Proceedings of the Design Society, Volume 5: ICED25
https://doi.org/10.1017/pds.2025.10182

Exploring Problem Framing activity using BERTopic

Gregory Litster[1,✉] **and Emily Moore**[1,2]

[1] *Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, Canada,*
[2] *Institute for Studies in Transdisciplinary Engineering Education & Practice, University of Toronto, Toronto, Canada*

✉ greg.litster@mail.utoronto.ca

ABSTRACT: Problem framing is a foundational aspect of the engineering design process, shaping how designers perceive challenges and potential solutions. Qualitiative methods, such as protocol analysis, have provided valuable insights about problem framing but are labor-intensive and time consuming. This study explores the use of a NLP technique BERTopic, to analyze framing in design conversations. BERTopic retains contextual nuances, offering a tool for uncovering the diversity and uniqueness of concepts explored by design teams while also making the analysis process more efficient. The results provide one representation of eight design group's processes, highlighting the different and changing topic representations that emerge throughout a design session. The findings highlight the potential of NLP tools for enhancing our understanding of framing in design cognition and team dynamics.

KEYWORDS: Design cognition, Problem framing, Research methodologies and methods, Design theory, Topic modelling

1. Introduction

Design problems are ill-structured and complex (Jonassen, 2000). As such, design teams must scope the problem in order for high-impact solutions to be generated. This process has been identified in the literature as problem framing (Dorst, 2015; Hey et al., 2007; Schön, 1984). Problem framing is an activity where teams will select from their environment the concepts they deem necessary for solution creation by assigning them significance and meaning, which helps to develop a deep understanding of the problem or challenge at hand (Dorst, 2015; Beckman, 2020). Problem framing has influence on the way solutions are sought out and has a huge impact on the design process overall (Stompff et al., 2016). The design community has been interested in understanding problem framing since Donald Schön's initial writing on reflective practice, which laid the foundation for early investigations of problem framing (Schön, 1984). As such, many different methods and techniques have been applied to understand framing, including interviews (Storm et al., 2019), surveys (Casakin, 2006; Silk et al., 2021), and most notably protocol analysis (Valkenburg and Dorst, 1998; Yang et al., 2019). Although these qualitative methods have provided valuable insight into framing activity, they are typically very time-consuming and laborious, having to spend time developing coding schemes and training coders. However, especially in the case of protocol analysis, advances in natural language processing (NLP) have helped alleviate some of those drawbacks by increasing the effciency of processing conversational text data.

Of particular interest in this work is the application of one NLP technique, topic modelling, which has a rich history in design research Dong (2005). Recent advances in Large Language Models (LLMs) architectures help to streamline and enhance topic modelling approaches which we believe can help uncover the complex process of problem framing. In the next section, we briefy review the literature on problem framing in more depth and explain how NLP, with a focus on topic modelling, has been used in design literature. We then present an exploration of a topic modelling algorithm - BERTopic

https://doi.org/10.1017/pds.2025.10182 Published online by Cambridge University Press

(Grootendorst, 2022) - applied to design transcripts to uncover characteristics of framing activity conducted by teams of designers.

2. Background

2.1. Framing in Engineering Design

The concept of problem or design framing in design theory is largely based on reflective practice (Schön, 1984). Reflective practice, or refection-in-action, is the process of designers grappling with a design situation that is uncertain (Schön, 1984). A good designer is said to have "reflective conversation" with the design situation as they weigh the consequences of design actions. When a designer is faced with the level of complexity inherent in design problems, they typically start with the identification of what they know about the problem and its elements. The selection of these known elements is said to help the designers create an appropriate mental model of the situation (Mathieu et al., 2000; Simon, 1969). Since Schön's initial writing of reflective practice, many definitions of problem frames - the products of framing activity - have been proposed. For example, in Cardoso et al. (2016) the authors state that a frame is "the perspective that is imposed by the designers on the design situation at a specific time during design activity" (pg. 67). This definition is fundamentally about how known problem and solution elements are characterized (Cardoso et al., 2016). Dorst (2015), describe frames as "a way of seeing" a problem or as a "complex thought tool". In this way, frames steer exploration and the perceptions of potential solutions because the concepts inside the frame have been given meaning (Dorst, 2015). Essentially, problem frames refer to alternative ways of communicating and understanding a problem, even when the underlying problem remains the same (Wright et al., 2015). More recently, Kelly and Gero (2022) reviewed the design framing literature and offered a definition: design frames are considered conceptual assemblages within cognition, which can then be made explicit (with speech or other mediums) as a representation of a design frame.

It is important to recognize that a problem can be reframed in multiple ways, influenced by an individual's interpretations of the situation (Silk et al., 2021). As such, problem frames are both personal and social concepts. Arguably more important though, is problem frames can change designer behaviour (Link et al., 2022). For example, Wright et al. (2015) found that by changing the way a design prompt is presented (neutral vs. innovative), participants will generate different types of ideas. Similarly, Silk et al. (2021) found evidence that when participants are provided an innovative problem frame, the participants perception of how creative their ideas are increased relative to a neutral problem frame. Given that problem frames have significant influence in the way design is conducted, investigations of framing activity are abundant but it remains a somewhat mysterious practice (Chandrasegaran et al., 2022).

2.2. Protocol analysis for measuring framing

Protocol analysis - recording and transcribing participant speech - studies have a rich history in measuring various aspects of design behaviour, including framing. The argument for the use of protocol analysis is that it provides one very powerful means for gaining information about cognitive processes (Ericsson and Simon, 1984). Most notably, Atman (2019) demonstrates the impact of this approach, offering considerable insight into the design processes of novices and experts using a developed coding scheme applied to many verbal protocols which culminated in useful design timeline representations. The design timelines have since been used to draw comparisons between groups with varying levels of expertise and inform new teaching methods in design. Similarly, Stumpf and McDonnell (2002) offer their own representation of framing by coding for different kinds of argumentation in transcripts to reveal instances of establishing and shifting frames as design activity progresses. Using the same dataset presented in this paper, Litster and Hurst (2024) drew parallels between design framing and systems thinking to create system map visualizations of framing activity in order to determine how system mapping can be useful for understanding framing.

The findings reported above are the result of considerable effort and resources, as protocol analysis studies are not straightforward to conduct. Protocol analysis requires signifcant investment of time and resources to create and verify coding schemes, train coders, and conduct final analyses. Fortunately, advancements in NLP methods have allowed for processing conversational data much more easily. For example, Gero and Milovanovic (2022) used NLP techniques on design transcripts – in conjunction with

https://doi.org/10.1017/pds.2025.10182 Published online by Cambridge University Press

the function-behavior-structure ontology – to map the design spaces of 19 different teams. Their results show promise in NLP's ability to correctly identify the first occurrence of concepts to map the design space of their participants. Similarly, Chandrasegaran et al. (2022) use n-gram analysis to identify common phrases associated with framing in their transcripts. They filtered these phrases on the basis of frequency and if the phrase was adopted by more than one person in the team. Their results indicate a powerful way to identify frames and other elements of design conversation. These examples demonstrate the increasing trend of applying NLP techniques for understanding different elements of design behaviour.

2.3. Topic modelling and BERTopic

Natural Language Processing (NLP) techniques have been used to understand design cognition for years. One NLP technique called topic modelling - an unsupervised model which learns a set of underlying topics from a set of documents (Nikolenko et al., 2017) - has been used to uncover underlying themes in a variety of research settings. Topic modelling approaches are diverse, with Latent Semantic Analysis (LSA) and Latent Dirichlet Analysis (LDA) being the most common implementations. LSA is a text analysis method that characterizes the semantic similarity between texts using a high dimensional semantic space (Dong et al.,2004). This method has been used by design researchers to measure shared understanding in design teams (Dong, 2005; Hill et al., 2001), develop models of a design (Dong and Agogino, 1997) and predict design team performance (Ball et al., 2020). LDA, on the other hand, focuses on modelling a corpus of documents into topics based on the distribution of words and documents in the corpus (Gyory et al., 2020). LDA has been used to understand the impact that managers have on engineering team cognition (Gyory et al., 2020), analyzing the performance of capstone design teams (Ball et al., 2020), and to discover high-quality design ideas from web scraped data (Ahmed et al., 2016). One of the assumptions of both LSA and LDA, is that the documents are of sufficient length (>100 words) (Ferguson et al., 2022). Recording and transcribing participant conversational speech often results in the length of each document to be quite short, as they are usually segmented based on turn taking in conversation, which makes the topic modelling algorithms less effective. Additionally, after removing stop words and punctuation in the transcripts, the data can become quite sparse. Additionally, these techniques typically represent topics as a 'bag-of-words' which disregards semantic relationships between words (Grootendorst, 2022). Fortunately, analyses can now take advantage of powerful Large Language Models (LLM's) to improve the effciency and effectiveness of analyzing verbal protocols, while retaining all contexts included in the short segments of speech. For example, Sakib et al. (2024) use OpenAI's GPT-4 model to determine the extent to which an LLM could accurately classify question utterances according to a question-asking taxonomy. They found that, despite some limitations, the LLM's are able to enhance the quality of the quantitative analysis (Sakib et al., 2024). One of Google's LLM, BERT (Bidirectional Encoder Representations from Transformers) is designed to pre-train deep bidirectional representations from unlabelled text by jointly conditioning on both left and right context on all layers (Devlin et al., 2019). BERTopic is a topic modelling algorithm that takes advantage of BERT's pre-trained language models to convert sentences or paragraphs into dense vector representations (Grootendorst, 2022). These vector representations, which do not depend on the length of the document, can be clustered to create interpretable topics. For example, creativity researchers interested in understanding the various topics associated with Do-it-yourself (DIY) YouTube videos used BERTopic to cluster groups of related transcripts taken from the videos and were able to identify 35 different topics. Their analysis was able to represent the overall landscape of creative endeavors, including activities like gardening and woodworking, which DIYers are working on (Ceh et al., 2023). In the case of BERTopic, researchers are able to effectively analyze large amounts of data quite quickly, provided the topics are representative of the phenomenon they are interested in. In our case here, uncovering the complexities of problem framing activity with manual qualitative analysis is time consuming and laborious, so we are interested in determining how BERTopic can aid in uncovering elements of problem framing in an efficient way.

2.4. Obndings highlight the potential of NLP tools for enhancing our understanding of framing in design cognition and team dynamicsjective

Recent literature has indicated that the way in which a problem is framed is one of the most important parts of the design process as it can dictate which kinds of solutions are available to the team (Stompff

https://doi.org/10.1017/pds.2025.10182 Published online by Cambridge University Press

et al., 2016). Provided the past success of topic modelling in design research, as well as the success of BERTopic in uncovering useful topics in other areas as described above, we wanted to explore the use of the LLM topic modelling algorithm BERTopic to determine what extent it might be helpful in characterizing or understanding problem frames used by design teams. Leveraging BERTopic to understand the semantic relationships between words we might be able to identify the diversity or uniqueness of concepts that are brought into the designers attention throughout their session. By doing this, we aim to draw some conclusions about how well the designers explored the design space.

3. Method

3.1. Design task and participants

In this study we use a set of transcripts collected by Dr. Carlos Cardoso at the Technical University of Delft. The dataset was used with permission and has been analyzed in other problem framing papers (Litster and Hurst, 2024). Eight groups of three Industrial Design master's students were tasked with generating solutions to the following open-ended problem and provide a sketch of their solution. The task was a part of a graduate course so the students were asked to audio record their own sessions which would be transcribed later. Each group was provided the following instructions:

> *Different people have different waking up experiences in the morning. However, a great number of people consider this process as unpleasant. How might you improve the morning waking up experience? As a team of three, generate new and useful ways (a product/system/service) that provide people with a positive waking up experience. If you generate several ideas, make sure you choose one final concept, and make a clear sketch of it. You should spend approximately 30 minutes on this activity.*

The 'How might you...' phrasing at the beginning of the design prompt is intentionally vague and open-ended which usually encourages further exploration of a given problem (Siemon et al., 2018). That is, designers tend to spend most of their time searching for elements of the problem that warrant a solution to be designed.

The students were randomly assigned to groups of three, usually with two students with industrial design backgrounds and one with either mechanical or civil engineering backgrounds. The age of the participants ranged from 22 to 26 years old. The transcripts from the videos were later generated by a research assistant and not the students themselves. Eight transcripts make up the dataset. The average duration of the sessions was 34 minutes.

3.2. BERTopic pipeline

In total, there were 2109 lines of data across all eight groups that made up the data and were analyzed using BERTopic. The transcripts were segmented based on turn-taking, so a new line was created when a different participant starting speaking. For the purpose of topic modelling, the data must be segmented into documents which we defined as lines of the transcript. We decided to learn the topics on the entire dataset rather than each group individually because all of the teams were working on the same problem, so we assumed that the topics would be similar between groups. Therefore, the transcripts were first combined together to learn document embeddings across the entire dataset. The document embeddings for each of the documents in the transcript were learned from a sentence-transformer model with approximately 22 million parameters[1]. These document embeddings were then used in the clustering task and the rest of the analysis was completed by following the quick start guide for BERTopic [2]. In total, 46 topics were identified, with one of the assigned topic labels as outliers. The BERTopic algorithm, and more specifically the clustering algorithm that is used, is a soft-clustering approach which allows for noise in the data to be modelled as outliers. After fitting the model, more than 30% of the documents in the dataset (n=781) were considered outliers. When examining some of documents that were considered outliers by the algorithm, we noticed that there were documents which we believe a human rater would

[1] https://huggingface.co/sentence-transformers/all-MiniLM-L6-v2

[2] https://maartengr.github.io/BERTopic/index.html

https://doi.org/10.1017/pds.2025.10182 Published online by Cambridge University Press

not consider as an outlier but rather might have been better assigned to one of the other topics. For example, one participant said:

> *"And also usually people don't think that waking up is important to the morning ritual, and I think it should be. Like, breakfast and uh washing your teeth, washing your face and stuff like that. It's part of your morning ritual"*

Given that there are relevant words representing concepts or actions associated with waking up (e.g., breakfast, morning ritual, washing your face) we wanted to ensure that each line in the transcript, which are acting as documents for topic modelling, was assigned to an appropriate topic. BERTopic has a built-in function to reduce the number of outliers, which can be done in a number of different ways. One strategy for reducing outliers is to use the embeddings of the outliers to find the closest topic cluster based on cosine similarity (https://maartengr.github.io/BERTopic/index.html). When implemented on this dataset, it reduced the number of outliers from the original 781 down to 0. As such, all documents in the transcript were assigned to a topic, and the total number of topics was reduced to 45.

4. Results

Our goal in generating the topic model was to identify useful representations of the design conversations in each of the groups. The 45 topics generated by the algorithm show promise in that endeavor. For example, Topic 17 focused on temperature with representative words including 'temperature', 'cold', 'warm', 'weather', and 'winter'; meaning that some groups considered temperature as part of their discussion of the waking up experience. We notice that most of the topics provided insight into the concepts that were discussed among the groups. The assignment of a topic to each of the documents signifcantly reduces the amount of time necessary to understand the entire transcript of a team. Although the algorithm did well at identifying coherent topics, not all topics provide insight into the concepts or ideas that were discussed by the participants. This is likely the result of the conversational nature of the sessions and how the transcripts were generated. That is, six topics were identified and tagged based on documents with a single word. For example, Topic 2 consists of 73 documents with "yeah" as the only text. In traditional NLP tasks, like latent semantic analysis, stop words like "yeah" or "okay" are removed in order to help in with the computation associated with dimensionality reduction. By using the LLM architectures, we no longer need to remove stop words (e.g., "yeah"); however, the downside is we end up with topics which include stop words as part of their representation, ultimately reducing the coherence, or topics that are strictly made up of single words. In the latter case, topics made up of single words were removed from analysis. A list of the remaining topics, the number of times that topic was identified in the dataset and representative words for each topic can be found in Table 1. We considered 'useful' topics to be those with documents that are longer than one word in length as well as referenced a minimum of a single concept. For the remainder of the paper, all of the analysis and discussion is focused only on those topics that provide meaningful representations of the design conversation.

Upon initial inspection of the topics, we noticed that most represented easily identifiable clusters of concepts or closely related concepts (e.g., topic 0: smell, topic 1: light, topic 3: alarms) which is consistent with the defnition of problem frames as "conceptual assemblages"explained in (Kelly and Gero, 2022). That is, we interpret topic representations as one model of the "conceptual assemblage"developed by the designers throughout the session. Again, using the LLM architecture does include some words in topic representations (e.g., "that", "it") which may have otherwise been removed using another topic modeling approach and may not seem directly related to the concept. We note that despite those words being included in the representation, the topics are still clear enough to discern a a single concept or cluster of related concepts.

We also noticed some (n=6) of the topics were more closely associated with actions of design process (denoted in Table 1 with an *). These topics are represented by words, for example in Topic 4, like 'ideas', 'concepts', 'brainstorm', 'think', and 'problem', which indicate that along with problem-related concepts, the team also explicitly discusses what they need to do in order to accomplish the goal of developing a solution to the design prompt. Table 2 includes four demonstrative examples of the documents that were categorized into 'concept' and 'process' categories. We also note that Topic 4 is associated with the most number of documents in the dataset (n=103) which may indicate that a signifcant portion of the participant conversations revolves around process.

https://doi.org/10.1017/pds.2025.10182 Published online by Cambridge University Press

Table 1. Description and distribution of topics

Topic, * denotes process	Document count	Associated words
0	86	smell, awakening, natural, senses, smells, coffee, sound, smelling
1	85	light, the, something, lights, like, sunlight, or, you, have, maybe
3	88	alarm, an, to, is, that, it, you, alarms, the, so
4*	103	we, ideas, concept, brainstorm, should, problem, solution, think
7*	80	minutes, we, five, ten, have, min, time, how, much, for
8	58	sound, nature, sounds, noise, like, of, in, the, select, novel
9	80	like, and, hours, time, at, sleep, my, you, to, when
10	63	phone, app, can, smart, this, an, have, it, you, the
11	58	morning, someone, message, call, you, nice, get, if, something, up
14*	41	sketch, draw, notes, looks, final , make, four, oh, room, drawing
15	37	pillow, shape, on, the, cushion, put, function, it, just, your
16	70	up, wake, waking, ways, to, what, so, good, morning, mindfu
17	31	temperature, cold, warm, weather, control, winter, dark, outside, the
18	42	touch, bump, pushing, moving, like, body, sort, position, matrass, of
20*	44	complete, choose, two, combination, one, do, think, wanna, these, pick
21	33	phone, your, my, sleep, use, it, apps, know, like, and
22	31	motivation, week, negative, motivational, motivating, motivate, loser
24	22	voice, record, recording, yeah, actor, particular, radio, sounds, or, should
25	30	boring, tv, activities, really, active, some, sort, stress, asleep, relaxing
26	52	sleep, sleeping, bed, going, night, you, probably, and, out, waking
27	57	wake, up, you, waking, to, but, slow, it, also, when
29	36	experience, waking, improve, how, good, up, we, ideal, morning, says
30	17	snooze, snoozing, button, down, for, my, others, alarm, notes, writing
31	40	product, service, expensive, package, whole, integrated, external, sell
32	26	music, listen, temperate, song, first, the, stephen, fry, if, mean
33	39	sleep, before, after, activities, sleeping, during, go, bed, to, you
34	23	cold, bed, warm, your, the, covered, is, hot, winter, out
35	22	hangover, alcohol, drink, broken, are, brush, throwing, teeth, eat, don
36*	30	write, down, writing, we, wrote, do, read, over, it, mean
38	17	story, stories, tells, telling, remember, only, someone, he, three, instance
40*	55	we, can, this, let, do, see, maybe, turn, what, got
41	24	curtain, open, window, food, eyes, screen, the, order, control, show
43	14	sound, asleep, music, way, light, up, is, fall, that, for
44	23	bed, worries, without, late, went, go, get, to, school, internet

Table 2. Topic representations

Topic	Associated topic words	Representative document
0 - Concept	smell, awakening, natural, senses, smells, coffee, sound, smelling	"It can be a waking up I guess. If you, if you smell like eh coffee or like nice breakfast, that will help." (G3P2)
1 - Concept	light, the, something, lights, like, sunlight, or, you, have, maybe	"Sometimes you really get annoyed by light. If you 're really sleepy or hungover then...[shakes head]"
4 - Process	we, ideas, concept, brainstorm, should, problem, solution, think	"Yeah five minutes on experience, and. ..brainstorm the problem?"
14 - Process	sketch, draw, notes, looks, final , make, four, oh, room, drawing	"Maybe I can do a storyboard. Maybe four pictures like this. It's one of the easiest ways to understand"

https://doi.org/10.1017/pds.2025.10182 Published online by Cambridge University Press

The topic labels (e.g., numbers 0-44) also allow for investigations of when particular topics appear more or less often in the design conversations. In particular, because we are interested in framing, we see value in demonstrating the number of different topics that are discussed throughout the session as this could be an indication of more or less sophisticated design behaviour Atman (2019). By using the topic labels alone, Figure 1 shows the diversity of topics over time based on a rolling window of the number of topics discussed by the participants. The values presented in the figure are identified by counting the number of unique topic values within a 10-line segment of the transcript and then dividing by the total number of unique topics found across that individual group. For example, in the transcript for group 1, in one window of ten documents (documents 32-41) the participants considered 5 unique topics which, when divided by the total number of topics for that group, represents the diversity value of the first window. In the next window (documents 33-42), the participants consider 6 unique topics, as one additional unique topic was added into the window, which causes the diversity in that window to increase. This means that darker areas of the heat map indicate a greater diversity of topics discussed within that window, while lighter areas indicate less diversity of topics. Again, because of our interest in framing, we only represent those topics that are relevant to concept development within the team. As such, the white areas of the graphic are instances in the transcript where a process topic (*) was assigned by the algorithm and thus do not contribute to the diversity of conceptual topics.

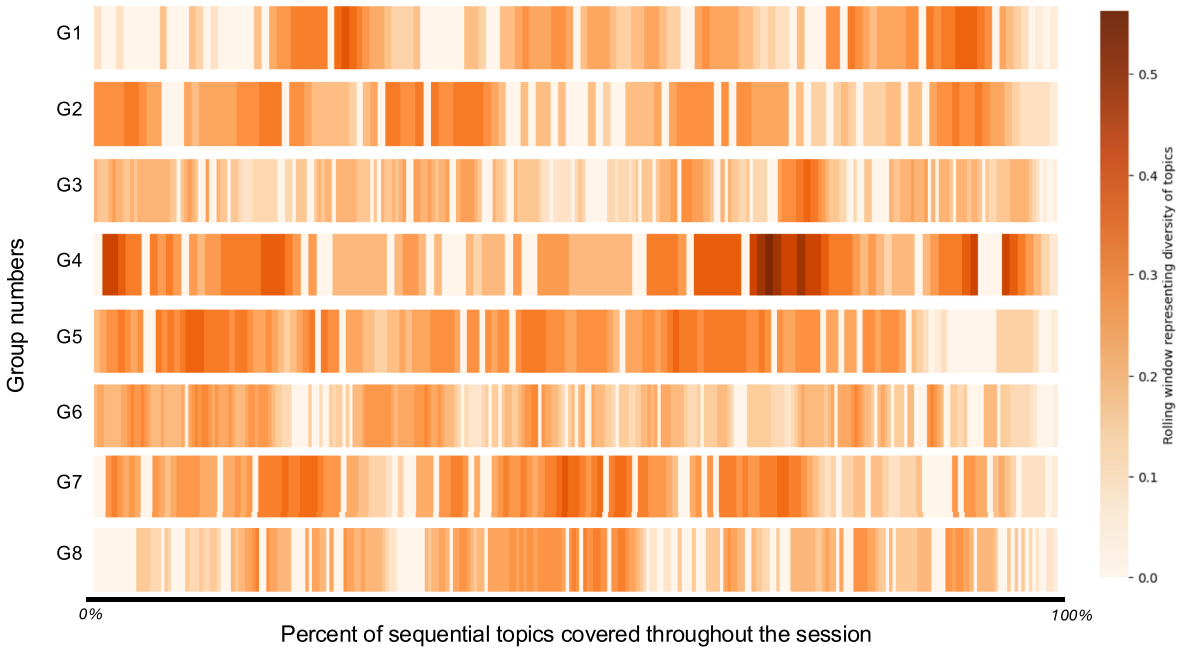


Figure 1. Diversity of topics over the duration of the session

This figure alone is not an ideal representation for the diversity of topics because the representation of the diversity of topics does not depend on the number of new topics identified, but simply the number of unique topics in the 10-document segment. It is possible that the topics repeated from the previous 10-document segment. A cumulative graph for the number of new topics for each sessions can provide some indication for when new topics appear throughout the session. Therefore, using Figure 1 in conjunction with a cumulative graph of new topics identified over the session provides some insight into when diversity of topics might be the highest.

We offer the cumulative graph of group four in Figure 2 to provide a feasible explanation for one of the patterns seen in Figure 1. We selected this fourth group because it includes the most extreme diversity values found at roughly 10% and 60% through the session. The slope found in the first 10% of the session for Group 4 found in Figure 2 demonstrates that they are discussing a number of new topics, with the slop quickly flattening indicating fewer new topics under discussion, something you might expect to see as designers explore their design space. We note that this signifcant increase in new topics can also be

https://doi.org/10.1017/pds.2025.10182 Published online by Cambridge University Press

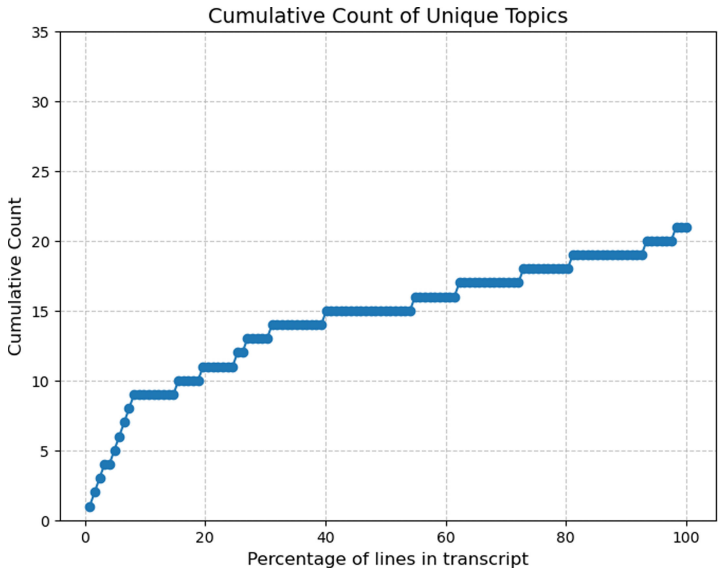


Figure 2. Group 4 cumulative count of new topics identified

seen in Figure 1 with the darker red chunk at the beginning. However, another similar dark section of the figure can be found at 60% of the way through the session. Comparing this to the same percentage point on the cumulative graph, we notice that not many new topics are being discussed at this time. This means that the group seems to no longer be introducing as many new topics but rather discussing many of, rather than a select few, previously identified topics. Although we can't say which of these interpretations represents the actual diversity of topics at those points in the design session from the graphs alone, the two visualizations together provide an easily identifiable part of the session that could be further investigated using a more qualitative approach.

5. Discussion

Design problem framing is a critical activity that occurs at the earliest parts of the design process. The way in which a design team frames a problem will influence their subsequent behaviour, especially with respect to the solutions that they develop (Link et al., 2022; Silk et al., 2021; Wright et al., 2015). In this paper, we were interested in exploring if topics generated from an LLM-enhanced topic model algorithm, BERTopic, can be useful for understanding framing activity.

The topics generated from eight transcripts, as seen in Table 1, demonstrate the algorithm's ability to capture high-level elements of design that were discussed by the groups. We noticed in particular that two types of topics were found, those associated with concepts related the problem and those related to process. Although no in-depth analysis was conducted on each of these types in this paper, we see value in being able to identify this type of speech in the transcripts quickly and with minimal effort and training. That is, we see the insight gained from the visualizations presented here from the topic modelling results as a supplemental tool which could lead to more targeted qualitative analyses. Gero and Milovanovic (2022) propose that their network analysis could be used in real-time of a design session to act as an external memory prompt for designers. We see the potential of real-time topic modelling as well, with the algorithm keeping track of or summarizing the different overall topics that the team has explored. For example, if we can identify potentially problematic problem frames early in the process based on the topics that are discussed and tagged as designers work, effort can be made to avoid them. Likewise, the identification of promising design concepts early in the design process will encourage more exploration of that concept and how it might inform the design frame.

The topics identified by the algorithm provide one way for us to look at how a design conversation evolves over time. The heat maps and cumulative topic graphics are useful for identifying what point in the session the groups are considering the most unique topics, potentially acting as a measurement for divergent or convergent thinking within the group (Goldschmidt, 2016). Problem framing as an activity focuses on understanding the different components of a problem situation and assigning them

https://doi.org/10.1017/pds.2025.10182 Published online by Cambridge University Press

significance and meaning. When teams are diverging onto a particular topic, we might speculate that they perceive these concepts as signifcant to the problem. This kind of analyses are useful for both getting a better understanding of design behaviour and, if presented as feedback to the designers themselves, can act as strong evidence for how their process has unfolded.

5.1. Limitations

Although the topic modelling algorithm provided useful representations of the design activity this work does not come without limitations. For example, framing involves designers giving meaning and significance to concepts they deem important to their problem context (Dorst, 2015). Simply assigning topics to each line in a transcript does not necessarily tell us what significance those concepts had to the designers. A future analysis would beneft from sharing the generated topics generated from the algorithm with the participants themselves to see how well it aligns with their perception of the session. If we could determine a level of agreement between the participants self-identified frames and those identified by the algorithm, a frame quality metric could be developed. Additionally, the algorithm assumes that each document (e.g., a line in the transcript) contains only one topic. It is possible that an individual participant is comparing two different topics in the same utterance which could reduce the effectiveness of the approach. This assumption may provide an explanation for why our dataset resulted in so many outliers as highlighted with the example in Section 3.2. Further investigations of topic quality would provide insight into how to mitigate this limitation of the approach as it stands.

6. Conclusion

In this paper, we have explored the use of BERTopic for understanding problem framing activity. Our qualitative analysis of the topics generated from the algorithm shows that topics are useful for understanding high-level concepts considered by each group. As such, we see value in using this technique to uncover elements of problem frames explored by design teams, with the hope that these insights can ensure the products, systems, and services they design are impactful.

Acknowledgements

We wish to acknowledge the participants of the study and Dr. Carlos Cardoso for the use of the data. Also Dr. Sharon Ferguson and Dr. Alison Olechowski provided valuable insight on implementing topic modelling algorithms on an earlier version of this manuscript.

References

Ahmed, F., Fuge, M., Gorbunov, L. D. (2016). Discovering Diverse, High Quality Design Ideas From a Large Corpus. *In ASME 2016 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 7.

Atman, C. J. (2019). Design timelines: Concrete and sticky representations of design process expertise. *Design Studies*, 65, 125–151.

Ball, Z., Bessette, J., Lewis, K. (2020). Who, What, and When? Exploring Student Focus in the Capstone Design Experience. *In ASME 2020 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*.

Beckman, S. L. (2020). *To Frame or Reframe: Where Might Design Thinking Research Go Next? California Management Review*, 62(2), 144–162.

Cardoso, C., Badke-Schaub, P., Eris, O. (2016). Infection moments in design discourse: How questions drive problem framing during idea generation. *Design Studies*, 46, 59–78.

Casakin, Hernan (2006). Assessing the Use of Metaphors in the Design Process. *Environment and Planning B: Planning and Design*, 33, 253–268.

Ceh, S., Putze, F., Benedek, M. (2023). D-I-WHAT? Identifying creative domains in do-it-yourself videos on YouTube. *Computers in Human Behavior*, 137, 107424.

Chandrasegaran, S., Lloyd, P., Salah, A. A. (2022). Identifying problem frames in design conversation. *In DRS2022*: Bilbao.

Devlin, J., Chang, M. W., Lee, K., Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv, 1810.04805*.

Dong, A., Agogino, A. M. (1997). Text analysis for constructing design representations. *In Artifcial Intelligence in Design*.

Dong, A., Hill, A. W., Agogino, A. M. (2004). A Document Analysis Method for Characterizing Design Team Performance. *Journal of Mechanical Design*, 126(3), 378–385.

https://doi.org/10.1017/pds.2025.10182 Published online by Cambridge University Press

Dong, A. (2005). The latent semantic approach to studying design team communication.*Design Studies*, 26(5), 445–461.
Dorst, K. (2015). Frame Innovation: Create New Thinking by Design. MIT Press, Cambridge, MA, USA.
Ericsson, K. A., Simon, H. A. (1984). Protocol analysis: Verbal reports as data. *The MIT Press*, 426.
Ferguson, S. A., Cheng, K., Adolphe, L., Van De Zande, G., Wallace, D., Olechowski, A. (2022). Communication patterns in engineering enterprise social networks: an exploratory analysis using short text topic modelling. *Design Science*, 8, e18.
Gero, J., Milovanovic, J. (2022). Creation and characterization of design spaces. *In DRS2022*: Bilbao.
Goldschmidt, Gabriela (2016). Linkographic Evidence for Concurrent Divergent and Convergent Thinking in Creative Design. *Creativity Research Journal*, 28(2), 115–122.
Grootendorst, M. (2022). BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv*. Available at: http://arxiv.org/abs/2203.05794.
Gyory, J. T., Kotovsky, K., Cagan, J. (2020). A Topic Modeling Approach to Study the Impact of Manager Interventions on Design Team Cognition. *In ASME 2020 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*.
Hey, J. H. G., Joyce, C. K., Beckman, S. L. (2007). Framing innovation: negotiating shared frames during early design phases. *J. of Design Research*, 1(1), 79.
Hill, A., Song, S., Dong, A., Agogino, A. (2001). Identifying Shared Understanding In Design Using Document Analysis. *In Proceedings of the ASME Design Engineering Technical Conference*, 309–315.
Jonassen, D. H. (2000). Toward a design theory of problem solving. *Educational Technology Research and Development*, 48(4), 63–85.
Kelly, N., Gero, J. S. (2022). Reviewing the concept of design frames towards a cognitive model. *Design Science*, 8, e30.
Link, S., Krishnakumar, S., Menold, J. (2022). Evaluating the Effect of Project Frame on Communicative Patterns in Capstone Design Pitches. *In ASME 2022 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, 4
Litster, G., Hurst, A. (2024). Analyzing problem framing in design teams: a systems mapping approach. *Artifcial Intelligence, Engineering Design and Manufacturing*, 38.
Mathieu, J. E., Heffner, T. S., Goodwin, G. F., Salas, E., and Cannon-Bowers, J. A. (2000). The influence of shared mental models on team process and performance. *The Journal of Applied Psychology*, 85(2), 273–283.
Nikolenko, S. I., Koltcov, S., Koltsova, O. (2017). Topic modelling for qualitative studies. *Journal of Information Science*, 43(1), 88–102.
Sakib, Ahmed Shahriar; Hurst, Ada; and Safayeni, Frank (2024). Towards AI-Assisted Protocol Analysis in Design Research: Automating Question Labelling with GPT-4 According to Eris' (2004) Taxonomy. In: Gero, John S. (Ed.), *Design Computing and Cognition*'24, 38–55. Springer Nature.
Siemon, D., Becker, F., and Robra-Bissantz, S. (2018). How Might We? From Design Challenges to Business Innovation. *Journal of Creativity and Business Innovation*, 4, 96–110.
Silk, E. M., Rechkemmer, A. E., Daly, S. R., Jablokow, K. W., McKilligan, S. (2021). Problem framing and cognitive style: Impacts on design ideation perceptions. *Design Studies*, 74, 101015.
Stumpf, S. C. McDonnell, J. T. (2002). Talking about team framing: using argumentation to analyse and support experiential learning in early design episodes. *Design Studies*, 23(1), 5–23.
Stompff, G., Smulders, F., Henze, L. (2016). Surprises are the benefts: reframing in multidisciplinary design teams. *Design Studies*, 47, 187–214.
Storm, R., Maanen, J. van, and Gonçalves, M. (2019). Reframing the Design Process: Integrating Goals, Methods and Manifestation into the Co-Evolution Model. *Proceedings of the Design Society: International Conference on Engineering Design*, 1(1), 359–368.
Simon, H. A. (1969). The sciences of the artifcial. MIT Press, Cambridge, Mass., 3rd ed.
Schön, D. A. (1984). The reflective practitioner: how professionals think in action. Ashgate, Farnham.
Valkenburg, R., Dorst, K. (1998). The reflective practice of design teams. *Design Studies*, 19(3), 249–271.
Wright, S. M., Silk, E. M., Daly, S. R., Jablokow, K. W., McKilligan, S. (2015). Exploring the Effects of Problem Framing on Solution Shifts: A Case Study. *In 2015 ASEE Annual Conference & Exposition.*
Yang, Y., Brik, N., de Jong, P., and Guerreiro Goncalves, M. (2019). Visualised Frames: How Sketching Infuences Framing Behaviour in Design Teams. *Proceedings of the Design Society: International Conference on Engineering Design*, 1(1), 419–428.

https://doi.org/10.1017/pds.2025.10182 Published online by Cambridge University Press