

**A Human-in-the-Loop Methodology For Applying Topic Models to Identify Systems Thinking  
and to Perform Systems Analysis**

---

**A Thesis**

**Presented to  
the faculty of the School of Engineering and Applied Science  
University of Virginia**

---

**in partial fulfillment  
of the requirements for the degree**

**Master of Science**

**by**

**Ryan Culp Boyer**

**December**

**2016**

APPROVAL SHEET

The thesis  
is submitted in partial fulfillment of the requirements  
for the degree of  
Master of Science

*Ryan C. Boyer*  
\_\_\_\_\_  
AUTHOR

The thesis has been read and approved by the examining committee:

Dr. William T. Scherer

\_\_\_\_\_  
Advisor

Dr. Cody H. Fleming

\_\_\_\_\_  
Dr. Michael C. Smith

\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_  
\_\_\_\_\_

Accepted for the School of Engineering and Applied Science:

*C. H. Benson*

Craig H. Benson, Dean, School of Engineering and Applied Science

December  
2016

## **Abstract**

Systems thinking characterizes the paradigm needed to effectively design, analyze, maintain, and utilize systems. Prior work has shown that there is a language of systems thinking and that its presence can be quantified within text using supervised learning methods. Building on this foundation, this work presents an unsupervised, human-in-the-loop methodology that utilizes topic models to facilitate the identification of systems thinking within a corpus of unread documents. This methodology is then expanded for use in analyzing systems themselves with methods of visualization, summarization, and provision of research direction. The novel aspect of the methodology is the seeding of the corpus; the user encourages the topic model to reveal desired information in documents by seeding the corpus with several documents that demonstrate, discuss, or exhibit the information desired. For identifying systems thinking, documents exhibiting systems thinking are added and a systems thinking topic is encouraged to form; each document's topic proportion within this topic is used as a proxy measure for the potential of systems thinking. This method for identifying systems thinking is fundamentally an explorative methodology (not predictive or prescriptive) and requires no manual grading of documents, which makes it significantly faster than previous methods. A Tukey test on a graded corpus reveals that the top echelon of strong systems thinking papers have significantly higher mean topic proportions in the systems thinking topic than lower graded papers. Additionally, a case study on a corpus of Army documents related to the development, character, and management of soldiers demonstrates the methodology's effectiveness in overviewing a system, in providing research direction, and identifying systems thinking within a specific domain. Finally, the methodology is used to present analysis of the past nineteen years of transportation research, demonstrating that transportation research is growing more holistic by focusing on environmental, physical, and societal health. Additionally, this analysis shows that the growth of traditional infrastructure research (construction, bridges, pavement, and roads) is significantly outpaced by research in these more holistic areas. Overall, topic models and the human-in-the-loop methodology demonstrate value by pairing a computer's data processing and structuring power with human intuition's ability to make associations and process abstract concepts.

## Acknowledgements

I would like to sincerely thank Dr. William Scherer for his help, guidance, and patience as I completed this work. Dr. Scherer, your clever thinking and cunning intuition made this work much more manageable, though at times it was completely frustrating to see you quickly solve problems I had worked on for hours. Thank you for pushing me to publish, present, and exceed expectations. I would not have been able to finish this work without you.

I would also like to thank my committee members Dr. Cody Fleming and Dr. Michael Smith. Dr. Smith and Dr. Fleming, I am extremely grateful for your listening ear and fresh perspectives. Both of you helped me tremendously with your expertise, your advice, and your friendly discussion. I have strived for this work to encapsulate the systems thinking that you both embody as researchers.

MAJ Casey Connors, LTC Chris Smith, and all of the U.S. TRADOC Analysis Center Monterey, thank you for your support both academically and financially through this project.

To all of my various classmates, I'd like to thank you for your friendship. From discussions of our differing cultures to discussions of algorithm design in Python, I've learned so much from each of you. Thanks for keeping me sane, making me do fun things, and distracting me.

Finally, I'd like to thank both my wife, Kelly, and my Mom. Mom and KB, your support over the past year and a half (and longer) has been foundational for my success. You are so kind, so patient, and so loving, and I love you both so much. Kelly, you have made this research worthwhile, while simultaneously making me not want to research at all (I'd rather go play outside with you). Thanks for encouraging me, and thank you for making me and letting me work.



# Table of Contents:

<b>Abstract</b>	<b>I</b>
<b>Acknowledgements</b>	<b>II</b>
<b>Table of Contents:</b>	<b>III</b>
<b>List of Graphs and Figures</b>	<b>IV</b>
<b>List of Tables</b>	<b>V</b>
<b>1. Introduction</b>	<b>1</b>
<b>2. Background &amp; Literature Review</b>	<b>3</b>
2.1. <i>Language</i>	3
2.2. <i>Topic Models &amp; Machine Learning</i>	4
2.2.1. General Topic Models	5
2.2.2. Latent Dirichlet Allocation (LDA)	7
2.2.3. Using Topic Models	11
2.3. <i>Systems Thinking</i>	12
<b>3. Methodology</b>	<b>14</b>
3.1. <i>Methodology For Identifying Systems Thinking</i>	16
3.2. <i>Methodology For Understanding Systems and Systems Analysis</i>	17
3.3. <i>Methodology For Identifying Trends Through Time</i>	18
3.4. <i>Discussion of the Human-in-the-Loop Methodology</i>	19
3.4.1. General Advantages and Limitations of Topic Models and Seeding	19
3.4.2. Explorative Nature of the Methodology	19
3.4.3. Why Put A Human-in-the-Loop	20
3.4.4. Discussion Specific to Identifying Systems Thinking	21
<b>4. Validation Exercises and Case Studies</b>	<b>22</b>
4.1. <i>Overview of Validation</i>	22
4.2. <i>Evaluation With Graded Corpus</i>	22
4.3. <i>U.S. Army Human Dimension Corpus</i>	27
4.4. <i>18 Years of Transportation Research Board (TRB) Conferences</i>	29
4.4.1. TRB Analysis: Data Source and Topic Model	30
4.4.2. TRB Analysis Discussion of Results: Overview	33
4.4.3. TRB Analysis Discussion of Results: Transportation Research as a Whole	34
4.4.4. TRB Analysis Discussion of Results: Trends in Transportation Research	36
<b>5. Future Work and Conclusion</b>	<b>42</b>
<b>6. References</b>	<b>44</b>
<b>A. Appendix: Seed Document and Document Cleaning Information</b>	<b>48</b>
<b>B. Appendix: Supplemental Topic Model Tables</b>	<b>49</b>

## List of Graphs and Figures

Figure 1 Word Cloud Describing Physical and Psychological Resilience Topic from Topic Model of Army Human Dimension Programs.....	6
Figure 2 LDA Plating Diagram.....	8
Figure 3 Smoothed LDA in Plating Notation .....	10
Figure 4 Workflow of Human-in-the-Loop Methodology .....	15
Figure 5 Top Words from the Systems Thinking Topic of the IEEE Graded Corpus .....	23
Figure 6 Word Cloud of the Systems Thinking Topic of the IEEE Graded Corpus .....	24
Figure 7 Network Graph of IEEE Graded Corpus.....	25
Figure 8 Comparative Boxplot of Grades and Systems Thinking Topic Proportion from IEEE Graded Corpus in 25 Topic LDA Model .....	26
Figure 9 Network Graph of Army Human Dimension .....	28
Figure 10 Network Graph Of Committee Focuses Within TRB From Aggregated Topic Model.....	35
Figure 11 Network Graph Of First-Degree Connections Of Safety & Crashes Topic: Highlighting Interactions With Construction – Work Zones.....	36
Figure 12 Decreasing Trends In Topics Relating To Infrastructure From 1998 To 2016 .....	37
Figure 13 Decreasing Trends In Topics Relating To Pavements From 1998 To 2016.....	38
Figure 14 Trends In Topics Relating To Surveys And Human Behavior From 1998 To 2016 .....	39
Figure 15 Increasing Trends In Holistic And Global Topics From 1998 To 2016 .....	40

## List of Tables

Table 1	Sample Topic from Latent Dirichlet Allocation Topic Model .....	5
Table 2	Variables in Standard & Smoothed LDA.....	8
Table 3	The Dimensions of Systems Thinking .....	12
Table 4	Significance Values of Tukey Test on Graded IEEE Corpus for Systems Thinking Topic Proportion .....	26
Table 5	Number of Papers in TRB Annual Meeting Compendiums from 1998 to 2016 .....	31
Table 6	Summary Of Transportation Research Topics From LDA Model, Grouped By TRB Committee Structure .....	32
Table 7	Seed Documents Used in Training Topic Models .....	48
Table 8	Words Included in Systems Thinking Word List (Not the Dimensions of Systems Thinking) .....	48
Table 9	Words Removed From Graded IEEE Corpus Before Use In 25 Topic LDA Model .....	49
Table 10	Topics From Graded IEEE Corpus For Evaluating Methodology's Usefulness In Identifying Systems Thinking In 15 Topic LDA Model .....	49
Table 11	Topics From Army Human Dimension Corpus In 25 Topic LDA Model .....	51
Table 12	Transportation Research Topics from 19 Years of TRB Annual Meeting Papers in LDA Topic Model .....	53

## 1. Introduction

The ability to objectively identify good systems engineering and systems design would be useful to systems practitioners everywhere, allowing them to see examples of strong engineering and learn from them accordingly. However, systems engineering is still a field without clear boundaries, and good systems design is very subjective. An expert can often identify good design, but it is usually just because the system functions smoothly and effectively, not necessarily as a result of the design process. However, there is a language that characterizes systems thinking - the systemic, goal-driven, new-eyed thinking that enables practitioners to design innovative solutions and to not neglect important details. This “systems approach” should lead to improved system designs with fewer conflicts and minimal unexpected pitfalls.

Prior work has both defined a lexicon of systems thinking and used supervised learning methods to classify papers with strong application of systems thinking, generating an accuracy rate on the order of 70% [1], [2]. While this supervised method is relatively effective, it requires substantial human overhead to be implemented. This time consuming and subjective process is a significant obstacle for this methodology to be used in practice.

This work presents a human-in-the-loop iterative methodology of identifying documents with potential for systems thinking. The methodology uses topic models, an unsupervised text analytics technique that automatically structures text documents according to their themes. Topic models model a collection of documents by generating a set of topics that describe the documents. Each topic is a distribution of words present in the full collection of documents. Each document is then represented as a proportional mixture of these topics. By seeding the unread papers with a small subset of papers that demonstrate good systems thinking, a systems thinking topic can be induced. Each document is thus tagged with a proportion that belongs to this systems thinking topic, which serves as a proxy for the potential for systems thinking in each paper. This methodology can also be used to survey and visualize a corpus, providing quick information to practitioners to direct their research.

This proposed methodology for identifying systems thinking does not require extensive grading, making it more efficient than the prior supervised methodology. This also allows for seeds to be changed quickly as goals change. However, this methodology is designed for exploring the systems thinking of documents and not for outright prediction. Furthermore, the methodology is limited by the subjective nature of what defines good systems thinking and the subjective interpretation of topics within the topic models, though this is a general limitation that the prior methods suffered as well. However, this is what makes the human-in-the-loop iterative nature effective; people have intuition and expertise to handle abstract concepts more effectively than computers.

The goal of identifying systems thinking is to allow systems practitioners to learn from others quickly; this methodology can be expanded to allow systems practitioners to directly learn about their system of interest. The topic models and their structuring of the

corpus can be used to survey and visualize a corpus, providing quick information to practitioners to direct their research. Furthermore, relationships between documents and central themes within a system can be quickly identified.

After a brief discussion of language, the mathematics of topic models, and systems thinking, this work describes the human-in-the-loop methodology for utilizing topic models. Finally, the work presents several case studies and exercises demonstrating the effectiveness of the methodology for identifying the presence of systems thinking and performing systems analysis. First, a graded corpus from prior work is used to show its effectiveness at finding documents with strong systems thinking. A Tukey test demonstrates that the papers with the highest grades for quality of systems thinking, those scored with a 9 or 10 out of 10, have significantly higher mean proportions in the systems thinking topic than papers graded 1 through 7. Second, the work presents a case study demonstrating the utilization of the methodology to analyze and summarize a corpus related to human dimension programs in the military. This case study further supports the methodology's effectiveness at identifying systems thinking within a specific domain. Finally, the work presents an analysis of the past nineteen years of research in transportation via an analysis of the conference papers from the Transportation Research Board (TRB) Annual Meetings. This exploration shows that research related to construction of transportation infrastructure is decreasing as a proportion of the research presented at the TRB Annual meetings while human-focused research, especially research focused on health, attitudes, and the environment is growing.

## 2. Background & Literature Review

### 2.1. Language

Merriam Webster defines language as “a systematic means of communicating ideas or feelings by the use of conventionalized signs, sounds, gestures, or marks having understood meanings” [3]. Said another way, language is a commonly understood framework used for the expression of thought. Since language is a framework of understood rules (though there is inherent ambiguity), humans naturally use language to classify thoughts and information. Additionally, language conveys information indirectly through word choice and language itself might actually shape thought. However, the ambiguity of language is a hurdle that must be overcome by humans and computers alike for correct interpretation.

Speakers use language purposefully as a vessel for communicating thoughts, but language can also provide insights into how thoughts were formed and the nature of the thinker beyond semantic and pragmatic meaning of the words used. Word choice changes as knowledge of a subject and the authors intended audience change; understanding language in this way may offer clues of the quality, level, and type of analytical thought beyond the semantics and pragmatics of the words used to express it. A simple example can be seen in the following, roughly semantically similar expressions.

- A. “I love this song’s staccato sound!”
- B. “I love this song’s short, choppy sound!”

Absent of knowledge of the speaker, the casual reader would see the speaker of sentence A as likely more knowledgeable about music than speaker of sentence B as “staccato” is a purely musical word that is more esoteric than “short” and “choppy”. This doesn’t mean that either speaker is correct in calling a given song “staccato”, but the lexical choice does inform the listener, in a Bayesian sense, of the nature of the speaker. Throughout the remainder of the conversation the listeners might continue to gauge the speaker by performing more Bayesian updates on the language used. Furthermore, lexicons, that is “the words used in a language or by a person or group of people”, are not static but grow and change over time; as someone learns and masters new material, their lexicon should reflect this mastery [4].

Beyond active lexical choice, there is discussion within the linguistics community of whether words and language actively shape thought. This idea is known as the Sapir-Whorf hypothesis and argues that differences in language systems result in non-linguistic cognitive differences. More simply it argues that the native language of a speaker has a strong influence on their worldview [5]. It is important to note that the Sapir-Whorf hypothesis initially was framed for entirely different languages, not differing constructs, patterns, or vocabulary within a language. However, in recent years proponents have taken a weaker view of the Sapir-Whorf hypothesis, believing that language has the ability to affect thought, but that it does not in every context.

Boroditsky offers a survey of examples and studies in her article “How Language Shapes Thought” that argue in support of the modern view of the Sapir-Whorf hypothesis [6]. She

begins with an anecdote of an aboriginal tribe in Northern Australia whose language does not contain equivalent words for “right” or “left”. Instead they only use the cardinal directions. She asks a 5-year-old girl to point north and, without hesitation, she correctly does so. She then tells of asking a room full of her coworkers, English-speaking scholars at Stanford who respond diversely, pointing in every direction. Her argument is simple, English speakers are not forced to think only in terms of cardinal directions and generally find using “left” and “right” more clear; this comes at the tradeoff of lack of awareness of the cardinal directions [6] [7]. Similarly, Boroditsky studied how this tribe interprets time and found they identify east with the past and west with the future, independent of the direction they are facing [7]. To Boroditsky, their language shaped more than their understanding of direction; it shaped their communication of time.

Detractors of the Sapir-Whorf hypothesis further argue that it is impossible to isolate language from biological, physical, and other contextual stimuli. The Linguistic Society of America argues that language does not shape thought, but that culture shapes both language and thought. In the case of the Australian tribe, this this would be understood as their language expressing their culture’s lack of relative directions, which is also demonstrated in their understanding of time. Regardless, the Linguistic Society of America argues that language still does provide insight into a culture [8].

Despite the language foundationally being a framework of commonly understood rules, these rules are not absolute, have many exceptions, and can easily be confounded. One of the most common ways language is misconstrued is due to ambiguous word meaning. Groucho Marx’s famous two-line joke is a prime example of semantic and pragmatic ambiguity: “I shot an elephant in my pajamas; how he got in my pajamas I don’t know” [9]. The joke revolves around the phrase “in my pajamas”. Naturally and logically, English speakers understand this phrase to mean Marx shot an elephant while he [Marx] was wearing pajamas that he owned. However, the semantics of the words themselves do not clarify this, allowing Marx to turn the sentence upside down for humor. While this joke was intentionally set-up for its ambiguity, language everywhere is rife with unintentional and unavoidable ambiguity.

Language carries thoughts, and word choice can suggest the nature of a speaker to listeners. Furthermore, language may actively shape thought, allowing word choice to go beyond hinting to openly demonstrating the nature of the speaker. If computers are to correctly understand language and utilized language to understand a speaker, like humans they will be limited by the inherent ambiguity of language.

## **2.2. Topic Models & Machine Learning**

Topic models are “algorithms for discovering the main themes that pervade a large and otherwise unstructured collection of documents” [10]. An unsupervised machine learning process, topic modeling allows for a more goal-oriented approach to reading than search, links, personal recommendation, or blind exploration. In fact differences from these

methods are the primary motivation for the creation of topic models; users can intelligently look at text in the absence of any ideas, preconceived notions, or external direction.

### 2.2.1. General Topic Models

Topic models work by taking a collection of raw text documents, called a corpus, as input and generating topics and topic assignments for each document. Topics are typically probabilistically weighted lists of words; as the distribution is approximated, these random variables begin to differentiate topics. The top words from an example topic from a Latent Dirichlet Allocation topic model (discussed in further detail below) are in Table 1. The format of topic assignments varies between different topic models. Often, they are a multinomial distribution of proportions assigning a given document to all of the topics generated by the model.

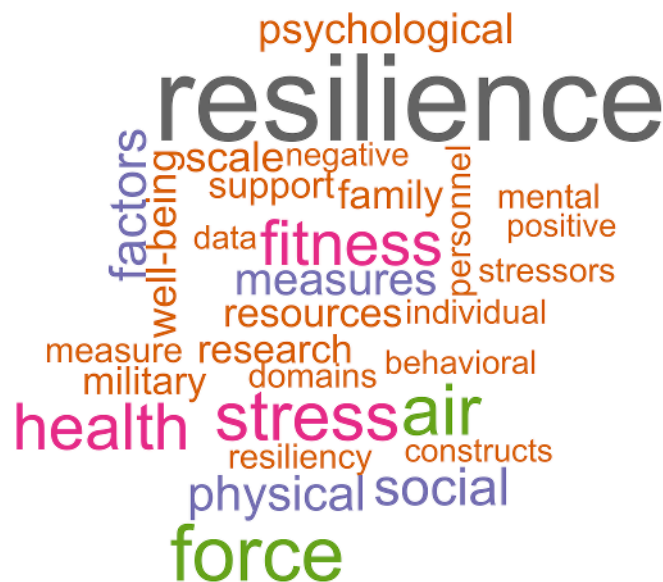
**Table 1 Sample Topic from Latent Dirichlet Allocation Topic Model**  
Topic can be summarized as “Physical and Psychological Resilience” and emerges from topic model of human dimension programs within the U.S. Army.

Word	Weight
resilience	520
stress	263
force	245
health	228
air	203
fitness	198
social	163
factors	157
physical	129
psychological	126
scale	115
stressors	101
resources	97
measures	94
well-being	92
research	89
mental	88

The topics can be visualized in word clouds for easier skimming and understanding. Within a word cloud, the size of the word is proportional to its importance and words of similar importance are given the same color. The corresponding word cloud for Table 1 is shown in Figure 1.



**Figure 1 Word Cloud Describing Physical and Psychological Resilience Topic from Topic Model of Army Human Dimension Programs**



Topic models work by leveraging the collocation of words within documents. By repeatedly sampling the documents and attempting to approximate a probabilistic distribution, topic models naturally find words that often occur together which results in the formation of topics. As the size of a corpus grows, the topic model has more data to find tighter and more salient topics [10].

Mathematically, topic models are mixture models that operate by assigning related stochastic random variables to elements of a text corpus. Often, the only observable random variable is the words in the documents; by attempting to approximate a joint probability distribution function the values of the hidden unobservable random variables can be found, which in turn define the topics. The actual joint probability distribution function of a topic model can often be written explicitly, but it is typically intractable to calculate exactly. To make up for this, topic models employ varying methods of distribution estimation.

Historically topic models have grown in complexity and have used several methods to approximate distributions. Latent Semantic Indexing (LSI) was one of the first information retrieval technologies to link related words together beyond stemming. However, LSI was not inherently probabilistic in nature and used singular value decomposition to reduce the size of the co-occurrence word space [11]. Probabilistic Latent Semantic Indexing (PSLI) was an evolution of LSI that was more statistically sound, but each document was only assigned one topic [12]. David Blei's Latent Dirichlet Allocation (LDA) is a generalization of PSLI that allows for documents to be seen as multinomial distributions over the topics [13]. Most new topic models are evolutions of LDA.

It is essential to remember that topic models are stochastic in nature. The same corpus run on the same topic model algorithm with the same parameters will often generate a

significantly different topic model purely as a result of random number generators. This does not diminish their usefulness, but any analyst using them should understand this and act accordingly.

### 2.2.2. Latent Dirichlet Allocation (LDA)

The most commonly used topic model is currently Latent Dirichlet Allocation, often referred to as LDA. David Blei, Andrew Ng, and Michael Jordan first published LDA in 2003 [13]. The key advantage that LDA has over its predecessors is that documents are tagged with multiple topics while using well-defined statistical modeling.

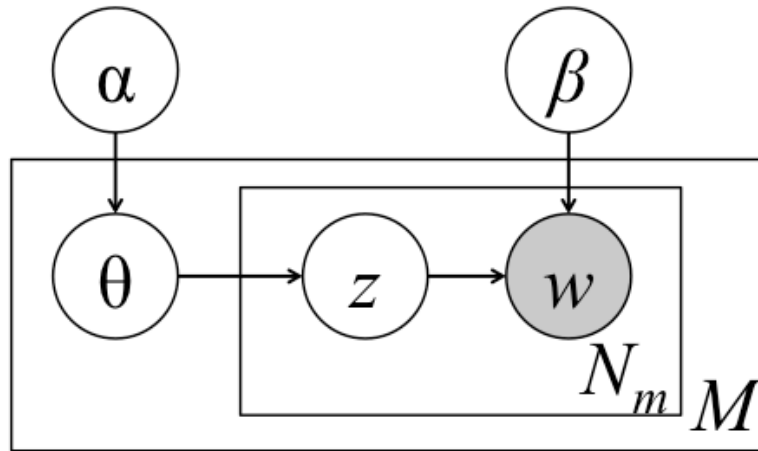
LDA is a mixture model (a type of generative probabilistic model) that models documents as mixtures of topics and models topics as probabilistically weighted lists of words. Given a set of topics, a new document can be generated by choosing proportions of these topics and drawing words from these topics according to these proportions [13]. Operating on the bag-of-words model of text, this document simulation portion is not actually used, but the distributions found by training the model provide keen insight to the nature of the corpus and the documents within the corpus. The generative process for regular LDA is shown in below, with variables shown in Table 2 [13]. Note that this is not the distribution estimation process, but the generative process.

1. Choose a topic proportion,  $\theta$ , from  $\text{dir}(\alpha)$ , where  $\alpha$  is the parameter vector for the dirichlet prior distribution
2. For each of the  $N$  words in the given document:
  - a. Choose a topic assignment,  $z_n$ , from  $\text{multinomial}(\theta)$
  - b. Choose a word,  $w_n$ , from  $p(w_n|z_n, \beta)$  which is a multinomial probability conditioned on the chosen topic assignment

The algorithm is more readily understood when shown in plating notation. Plating notation, developed originally by Gilks et al., clearly shows the dependence and repetition of stochastic random variables. Each circle represents a random variable or class of random variables. Shaded circles represent observable random variables while white circles represent latent or hidden random variables. The boxes represent repetition [14]. Figure 2 shows LDA in plating notation with additional notes about the random variables, similar to how it was presented by Blei in his original paper [13]. The random variables are further explained in Table 2.

**Figure 2 LDA Plating Diagram**

LDA algorithm is shown in plating notation, identifying the dependence and repetition of random variables.



**Table 2 Variables in Standard & Smoothed LDA**

Variable or Parameter	Type	Explanation	Quantity
$\alpha$	Positive, Real Number	Dirichlet Parameter, prior for the per-document topic distributions	K (one per topic, often all the same)
$\beta$	Positive, Real Number	Dirichlet Parameter, prior for the per-topic word assignments and distribution	W (one per unique word per topic, often all the same)
$\theta$	Multinomial Distribution, a vector of length K with values between 0 and 1 inclusive, summing to 1	Topic proportions for a given document	M (one distribution per document)
$z$	Integer between 1 and K	Word assignment; assigns a given word from a given document to a specific topic	The total number of words in the corpus, including duplicates of the same word
$w$	Word (often converted to positive integer)	A unique, specific word from a document	
$W$	Positive Integer	Number of Words in Corpus	
$M$	Positive integer	Number of documents in corpus	
$N$	Positive Integer	Total number of unique words in corpus	

$K$	Positive Integer	Number of Topics, assigned by user	
$\eta$	Multinomial Distribution, a vector of length $K$ with values between 0 and 1 inclusive, summing to 1	Word probabilities for a given topic	$W$ (one distribution per unique word)
$i, j, t$	Positive Integer	Indexing variables	

Effectively, LDA works by letting the hyper-parameters,  $\alpha$  and  $\beta$ , along with the observed words identify effective values for each random variable. The results are highly effected by these hyper-parameters. Setting them to small values,  $\alpha$  typically around 0.1 and  $\beta$  typically around 0.001, encourages fewer topics per document and fewer words per topic respectively.

The total probability of an LDA model can be explicitly represented through a joint probability distribution function, shown below in (1). However, this distribution is intractable to calculate exactly.

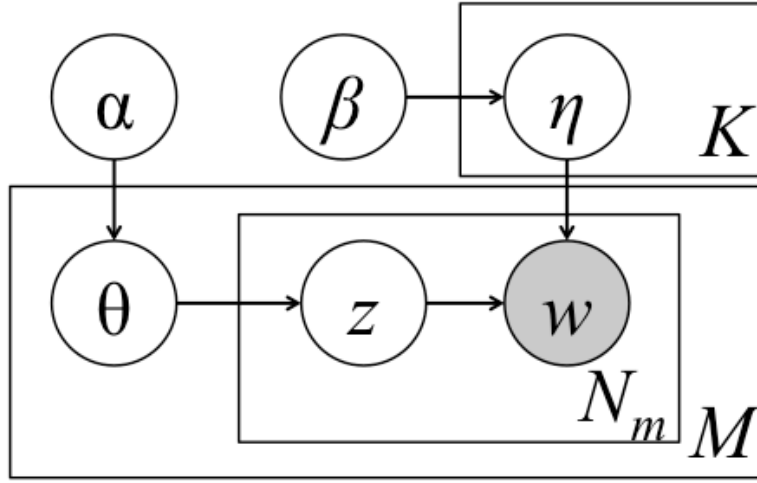
$$p(\boldsymbol{\beta}, \boldsymbol{\theta}, \mathbf{Z}, \mathbf{W}) = \prod_{i=1}^K p(\beta_i) \prod_{j=1}^M p(\theta_j | \alpha_j) \left( \prod_{t=1}^N p(Z_{j,t} | \theta_d) p(W_{j,t} | \beta_{1:K}, Z_{j,t}) \right) \quad (1)$$

Typically, this simplified version of LDA is not used when topic modeling a corpus. Instead LDA with smoothing is used, where the observable words depend on an additional set of multinomial distributions. This ensures that no word has a true 0% chance of being in a topic, effectively correcting for novel events during inference [13]. The new generative process for smoothed LDA is given below, Figure 3 shows the plating notation for smoothed LDA, (2) gives the total probability of smoothed LDA, and Table 2 shows the variables and parameters used in the model.

1. Choose a topic proportion,  $\theta$ , from  $\text{dir}(\alpha)$ , where  $\alpha$  is the parameter vector for the dirichlet prior distribution
2. Choose the word distributions for the topics,  $\eta$ , from  $\text{dir}(\beta)$ , where  $\beta$  is the parameter vector for the dirichlet prior distribution and  $\eta$  is the matrix of all  $\eta$ .
3. For each of the  $N$  words in the given document:
  - a. Choose a topic assignment,  $z_n$ , from  $\text{multinomial}(\theta)$
  - b. Choose a word,  $w_n$ , from  $\text{multinomial}(\eta_{z_n})$

**Figure 3 Smoothed LDA in Plating Notation**

The Smoothed LDA algorithm is shown in plating notation, identifying the dependence and repetition of random variables.



$$P(\mathbf{W}, \mathbf{z}, \boldsymbol{\theta}, \boldsymbol{\eta} | \boldsymbol{\alpha}, \boldsymbol{\beta}) = \prod_{i=1}^K P(\eta_i | \beta_i) \prod_{j=1}^M P(\theta_j | \alpha_j) \prod_{t=1}^N P(z_{j,t} | \theta_j) P(w_{j,t} | \eta_{z_{j,t}}) \quad (2)$$

As already mentioned, this total probability for both regular LDA and smoothed LDA is intractable to calculate exactly. Often variational Bayes, collapsed Gibbs sampling, or expectation propagation are used to approximate these distributions [13], [15], [16]. Furthermore, good choices for the  $\alpha$  and  $\beta$  hyper-parameters can be iteratively approximated [17], [18].

LDA does not perfectly reflect the way speakers and writers use language. It treats all documents under the bag-of-words model, which assumes that order of words does not matter. Furthermore, all topics are considered independent and uncorrelated under an LDA model.

Alternate topic models and evolutions of LDA address some of these assumptions and expand upon LDA's capabilities. Whye et al. developed a hierarchical topic model that allows for multiple levels of nested topics [19]. Chang and Blei developed a relational topic model that leverages meta-data of the corpus, such as authors, citations, journals, and links [20]. Chen et al. developed a topic model which can leverage topics generated by past topic models (or user-defined sets of words with the similar semantic meanings) so that topic models learn over time [21]. Jagarlamudi et al. developed a topic model that uses user-defined seed words that a user feels are descriptive of the corpus to improve topic coherence [22]. Andrzejewski et al. developed a topic model that allows users to specify "must-links" and "can't-links" between sets of words to allow for domain knowledge to guide the modeling process [23]. This is just a small sample of recent research.

### 2.2.3. Using Topic Models

Topic models are seeing growing utilization in many fields and are being used for explorative and descriptive purposes, as well as preprocessing for predictive learners. Often before the output of the topic model can be effectively utilized, additional calculations must be performed.

The similarity between documents within a corpus can be measured using their topic proportions. Several metrics have been proposed for this similarity measure, including the Kullback Leibler divergence, the symmetrized Jensen-Shannon (JS) divergence, Euclidian distance, the dot product, and cosine similarity [24]. This measurement can then be used in a predictive learner or for directed exploration of the corpus.

Similarly, LDAvis, an R package for visualizing LDA topic models, performs a principal component regression in order to visualize the topics. This allows the topic space to be displayed in two dimensions (the top two PCA components) for quick understanding of the topics and the potential for overlap between them. Furthermore, LDAvis provides a numerical metric for the relevance ( $r$ ) of each word to the topic given below in (3), where variables are as given in Table 2 with  $\lambda$  being a weight parameter and  $p_w$  being the marginal probability of term  $w$  being in the corpus [25].

$$r(w, k | \lambda) = \lambda \log(\eta_{kw}) + (1 - \lambda) \log\left(\frac{\eta_{kw}}{p_w}\right) \quad (3)$$

Topic models are seeing growing utilization. They are being used in an explorative way to understand trends through time and explore corpora. Blei et al. used topic models to explore the content of the journal *Science* [26]. Similarly, Livermore et al. used topic models to explore agenda formation within the Supreme Court and how its decisions differed from local courts [27]. All explorative uses of topic models suffer from the inherent subjectivity in the interpretation of the topics; however their structuring of information is still extremely helpful in gaining quick understanding in large data sets.

In a predictive way, topic models are being used to improve information retrieval processes and to tag documents appropriately for further sorting or cataloging. Wei et al. discuss how LDA topic models can be incorporated to database querying systems to improve retrieval over clustering based methods in a computationally efficient way [28]. Ramage et al. modified LDA to form tagged-LDA, a supervised algorithm that will tag documents with appropriate keywords or groupings based on the input of pre-tagged documents into the LDA model [29]. Additionally, topic models have seen usefulness for preprocessing data within the transportation realm. Zhang et al. used LDA topic models and geospatial data from twitter for traffic incident detection [30]. Brown paired random forests with topic models of the narratives of railway crashes to improve prediction of railway accident cost [31].

There are readily available implementations of LDA in most programming languages. The work in this paper utilizes Mallet, an open-source, Java-based package available from University of Massachusetts Amherst [18]. Mallet is one of the most popular implementations today as it is very fast, very light, can be run from the command line, and can be customized extensively. However, the methodology that is presented for utilizing topic models to identify systems thinking is not unique to Mallet or LDA.

### 2.3. Systems Thinking

The systems approach is a comprehensive method of designing and understanding systems, which looks at how the system functions as a whole and not how its individual pieces function [32]. This systems approach can contrast with modern understandings of systems engineering, which is sometimes seen as only a systematic process that should be walked through for a system to be designed effectively. However, systems thinkers both past and present feel that this is insufficient for good systems design and that it is the pairing of good processes with holistic views and well-defined goals that generate effective systems [33], [34].

However, as systems engineering is fundamentally multidisciplinary and interdisciplinary, the language of systems and the systems approach can quickly be confounded. Ackoff established a language to describe systems, systems of systems, and their parts [32]. Whitehead et al. extended Ackoff's linguistic foundation by creating a taxonomy of the systems approach, which they call the "dimensions of systems thinking" and is shown in Table 3 [1]. Both argue that clear communication and clear understanding are essential for systems research to thrive. Standing on the foundation of the systems approach, consensus between stakeholders is essential for a system's success; this consensus will be elusive if systems researchers and systems engineers cannot communicate effectively.

**Table 3 The Dimensions of Systems Thinking**

Descriptive Scenario
System Boundaries
System Stakeholders
Scope Of The Analysis
Type Of System
State Of System
Life Cycle Of System
Metrics
Axiological Components
Observer Effects
Normative Scenario
Objectives
Indices Of Performance
Develop Alternatives
Out Scope

---

---

Evaluate Rank Alternatives  
Interactions  
Iterate Analysis  
Leverage Points  
Recommendation

---

---

Whitehead et al. leveraged the dimensions of systems thinking to measure the quality of systems thinking in technical reports [2]. This work demonstrated a proof of concept that the language of systems thinking can be objectively classified using natural language processing and supervised learning methods. They began by grading a corpus of 295 documents, giving 62 life-cycle analysis papers either “top” for good systems thinking or “bottom” for poor systems thinking, and 233 IEEE papers between 1 (poor systems thinking) and 10 (excellent systems thinking). The 233 IEEE papers served as the training set while the 62 life-cycle analysis papers served as the test set. Whitehead et al. used two learners to evaluate the potential for measuring systems thinking. By converting all the documents to vectors via a term frequency-inverse document frequency (tf-idf) matrix, he identified the Rocchio centroid classification vectors between the good systems thinking documents in the training set (7 or higher) and poor systems thinking documents in the training set (4 or lower). This vector space model classified the test set with an accuracy of 61%. Additionally, Whitehead et al. utilized quadratic discriminant analysis on the tf-idf vectors, which gave an accuracy of 68%. These accuracy rates are not perfect, but do demonstrate that systems thinking can be identified computationally. The obvious caveats to this methodology is that grading a training corpus is a time-consuming and subjective process, and that even larger training sets are required for larger and more diverse corpora.



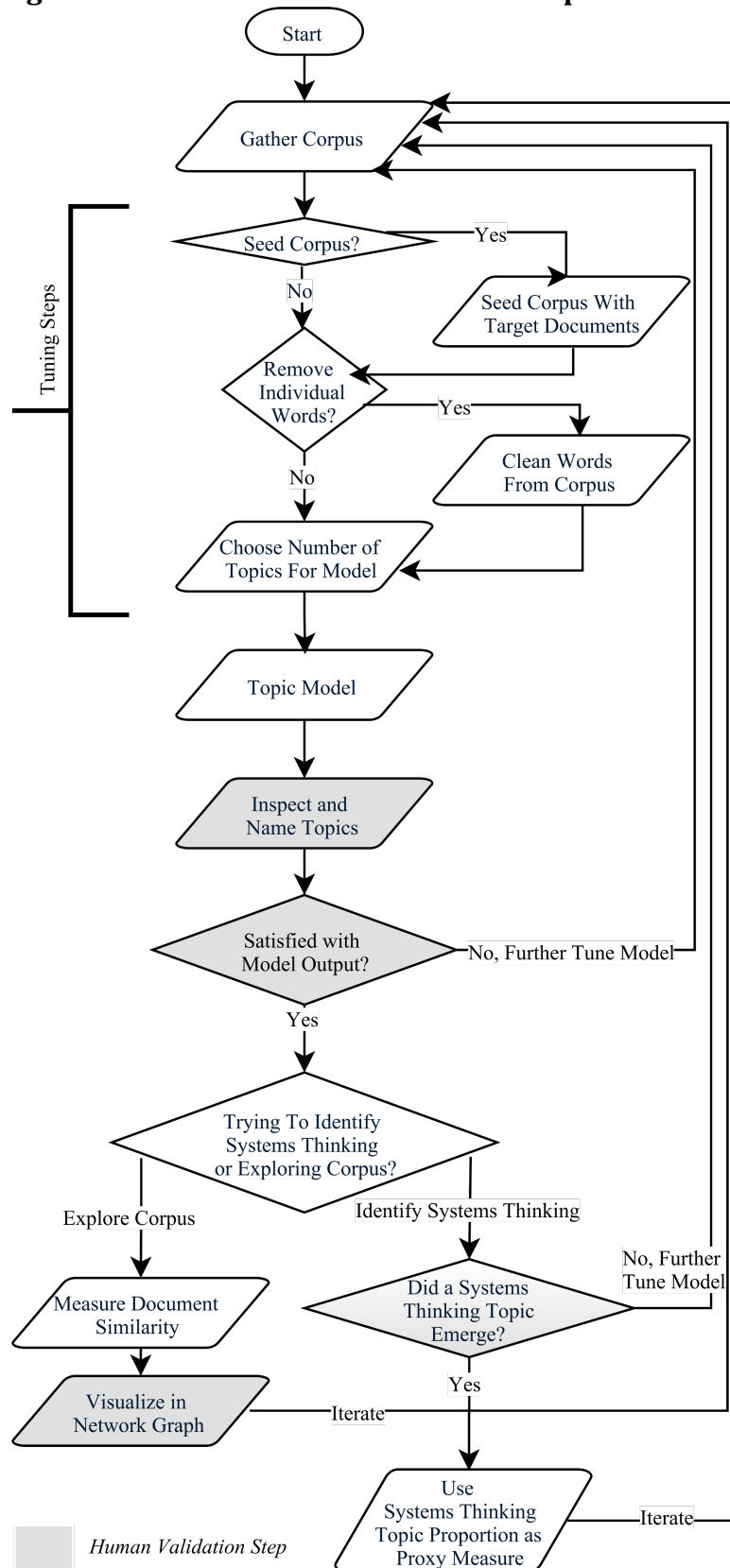
### 3. Methodology

The human-in-the-loop methodology for using topic models to identify documents with potential for systems thinking is fundamentally iterative and useful for exploring unread documents. At a high level, the user seeds the corpus with documents that they believe describe or exhibit strong systems thinking and then runs a topic model on the new corpus. Due to the seeds and iterative tuning, a systems thinking topic will emerge, which is filled with the dimensions of systems thinking or their domain-appropriate synonyms. The document's topic proportions within this topic serve as a proxy measure identifying the potential for systems thinking. This idea of seeding an unsupervised method to influence its structure is novel, unique, and effective for data exploration.

The human-in-the-loop methodology can also be used to visualize and understand a system, to identify relations between documents, or to find documents related to other ideas. This work first walks through and discusses the methodology for identifying systems thinking and then discusses its use in exploring a corpus for understanding. Figure 4 presents the flowchart for both of these processes. Three examples of the methodology's usage are provided afterwards to further clarify and explain the points of discussion within this section.

The most novel contribution from this methodology is the concept of seeding the corpus before performing a topic model. This technique provides some direction to the topic model while keeping it an unsupervised method. More importantly this allows the user to evaluate other documents within the corpus in terms of one they know about. This technique can be used to find documents addressing similar problems or using strong systems thinking [35].

**Figure 4 Workflow of Human-in-the-Loop Methodology**



### 3.1. Methodology For Identifying Systems Thinking

The user begins by gathering a corpus of unread documents that they are interested in, preferably with over 300 documents of at least a page in length, though smaller corpora may work depending on their content.

The user then “seeds” the corpus. He or she adds specific papers to the corpus that are representative of good systems thinking or actively describe good systems. This seeding methodology seeks to guide the topic model. The language within these seed papers will lead to the output of the systems thinking topic from the topic model, and, due to the collocation-based inference of topic models, other domain specific synonyms will naturally be found and brought into the systems thinking topic.

Despite Whitehead et al.’s proposal for a common language, the language of systems thinking will differ between domains in practice. Ideally the seed papers will discuss and use systems thinking by utilizing the equivalent dimensions of systems thinking within that domain. However, in our experiments with this methodology, a meta-paper discussing systems thinking from the systems thinking domain has been sufficient.

The number of seed papers required varies due to the number of papers in the corpus of interest and their thematic spread. However, a good rule of the thumb is one seed paper for every 40 to 60 papers of interest with a minimum of three seeds. For larger corpora, this ratio decreases dramatically. In practice getting the systems thinking topic to emerge can be difficult, and iteration is required.

The user then prepares the corpus for topic modeling by cleaning specific words from each document. This cleaning serves to focus the topics from the model on the valuable information. Typically the user removes stop words, which are words that have limited meaning on their own, such as common articles, prepositions, adverbs, and transitions words. On additional iterations of the human-in-the-loop methodology, a user might remove context specific words that provide limited information. For example, a corpus focusing on the canine diseases might remove the word “dog” as it will appear in every paper and provide no new information in any topic.

Next, the user picks a number of topics for the model (and any other model parameters if using a non-LDA model) and generates the topic model. The number of topics will tell the algorithm how many topics to look for within the corpus. Fewer topics result in more high-level themes being revealed, at the expense of detail and the possibility of convoluted ideas. More topics result in more detail and potential for specificity in each topic, but at the expense of possibly duplicated topics and the possibility for data overload. Several authors have suggested means of computationally identifying the ideal number of topics for an LDA topic model, but manual tuning by inspection and repetition provides the desired result with the added benefit of increased knowledge and understanding of the corpus [19], [36].

The user then inspects the topics produced by the model and gives them appropriate names. This provides insight into the makeup of the corpus and allows the user to see if a systems thinking topic emerged. A systems thinking topic would be one where the dominant words in the topic are made of the dimensions of systems thinking or their appropriate synonyms in the domain of the corpus.

If a system thinking topic emerged, it can be used as a proxy measure for the potential for systems thinking; the user can begin reading the documents with high proportions to see their analysis and learn from their results. However, if a system thinking topic did not emerge, the user must iterate through the human-in-the-loop methodology again. The topic model should be changed based on the outputs the user saw in the first iteration. For instance, if the topics are all very high-level and mixed, the number of topics may be increased. Or if there are several words appearing in many topics, they may be removed from the corpus. The user may also add or remove seed documents as necessary.

### **3.2. Methodology For Understanding Systems and Systems Analysis**

A pillar of systems thinking is the understanding of the system of interest; a natural extension of the human-in-the-loop methodology is to gain understanding of the system that the corpus describes. By changing model parameters, changing seed documents, and visualizing the output a user can gain understanding of how documents are related, identify central and pervasive themes, and find research direction.

The user may seed documents that demonstrate or discuss strong systems thinking, use seed documents related to another topic or concept, or use no seeds at all. Each of these options will unveil different aspects about the system described by the corpus. Just as systems thinking seeds will lead to a systems thinking topic, which provides information about the systems thinking in each document, alternative seeds may lead to alternative topics. The use of no seeds will describe the corpus as a whole, agnostic of any influence.

When the user reaches the step where he or she inspects and names the topics, he or she will naturally gain insight about the concepts within the corpus. This insight often leads to questions about how the topics are related, what ways the documents are interconnected, and how the ideas in the corpus are being used in practice. These questions can be answered through network graphs and through further iteration.

In creating a network graph of the topic model, each document is treated a node, and the similarities between their thematic content is calculated using the topic proportions from the model (a document-to-document network). Several metrics have been proposed for this similarity measure, including the Kullback Leibler divergence, the symmetrized Jensen-Shannon (JS) divergence, Euclidian distance, the dot product, or cosine similarity [24]. Similarly, the user can use the topic proportions to make a network graph of the documents and their relationships to topics, treating both documents and topics as nodes (a document-to-topic network). Both of these network graphs tend to display useful and informative information, especially when visualized in a clustering visualization, such as

those discussed in [37], [38], or [39]. However, the commonly used methods of Bayesian inference used to identify the probabilistic distributions of the topic models rarely give true zero values for any of topic proportions. This can result in an over-connected graph as all documents will have relationships with all topics though most of the relationships will be very weak (close to zero in topic proportion). This hurdle can be overcome by using a threshold function to keep only the most dominant edges.

Treating the topics as nodes, the user can make a similar network graph by identifying close relationships between topics through the documents (a topic-to-topic network). Topics that commonly have high proportions within one document are likely to be related (a topic-to-topic network). Similar metrics as discussed for the document-to-document similarity can be used, though the vectors would come from the transpose of the document-topic matrix. As different topics may correspond to differing proportions of the corpus, this type of graph is typically dominated by the more structural and prevalent topics. This can be normalized and provide a more holistic view of the relationships between topics by keeping only the strongest  $N$  edges for each node (where  $N$  is a positive integer between 1 and the number of topics - 1).

### **3.3. Methodology For Identifying Trends Through Time**

Given a corpus of documents with timestamps, a topic model can be used to visualize and understand the trends within the corpus through time. This work presents four methods of viewing these trends; two focus on the relative proportion of topics while two focus on the absolute quantity of papers in a topic.

First, the mean proportion of a topic can be used as a measure of relative prevalence of the topic within the corpus. This does not provide an absolute measure as the number of papers in a corpus for each time period may not be steady. For example, if there are an increasing number of papers for each time period, a decreasing mean topic proportion can still occur if the growth of other topics outpace it.

Additionally, mean topic proportions can be difficult to interpret as their significance varies due to model parameters. A mean topic proportion of 10% demonstrates a very dominant topic in a topic model with 100 topics but would represent a relatively rare topic for a model with only 5 topics. One method of providing the necessary context to the mean topic proportions is showing comparative graphs of the topics through time.

An alternative view of the relevance of a topic can be presented by looking at the percent change of the mean topic proportion. This measure loses all potential to compare the prevalence between topics, but does show how quickly topics are rising and falling within a topic model. This can alert an analyst quickly to emerging and falling trends.

In an attempt to provide more absolute context, indicator functions can be used to count the absolute number of papers focusing on different topics. Setting a threshold between 0 and 1, any paper with a proportion greater than or equal to the threshold in a given topic is

considered focused on that topic. The appropriate threshold value will vary due to topic model parameters. Additionally, this methodology has the drawback that certain papers may be equally spread through many topics and may not be counted anywhere while other papers may be counted multiple times.

This absolute lens still does not correct for the changing number of papers per time period. However, this measure can be normalized by simply dividing by the number of papers in a given time period. This normalized view is not as easily understood as the simple count of papers, but may provide a better understanding of the true balance of papers in the corpus. Together, these four measures can provide an appropriate view of trends of topics over time.

### **3.4. Discussion of the Human-in-the-Loop Methodology**

#### **3.4.1. General Advantages and Limitations of Topic Models and Seeding**

The use of topic models offers several distinct advantages and disadvantages over alternative methods. First, topic models are much more general than traditional methods of exploring textual information. While directly counting the frequency of the dimensions of systems thinking would provide some understanding of the potential for systems thinking in each document, the topic model utilizes the collocation of words. This allows the systems thinking topic to be corpus specific and to naturally find the domain-specific synonyms for the dimensions of systems thinking. Similarly, this generalized idea is useful for performing systems analysis as related words and concepts will emerge more quickly and clearly than classic term-frequency identification.

Additionally, the concept of guiding topic models by seeding is not unique to finding systems thinking. Alternative seeds can be used to find alternative concepts. This idea of seeding could be expanded to other unsupervised machine learning algorithms.

Like several other text mining and machine learning techniques, one direct limitation of topic models emerges from the bag-of-words model of text [40], [41]. A document can use the correct words and have a large tagged proportion in the systems thinking topic, but not say anything meaningful or provide poor systems analysis and thinking. For this reason, we refer to the topic proportion as a proxy measure for the potential for systems thinking.

#### **3.4.2. Explorative Nature of the Methodology**

The human-in-the-loop methodology for both systems analysis and identifying systems thinking is explorative and not predictive. This is due to several reasons. First, Topic models are stochastic methods that rely on random number seeds and approximations; values can change between iterations without changing any parameters or data. Additionally, there is not a “true” set of topics and proportions that best describe the corpus; all topic models are wrong as they oversimplify language. Furthermore, a certain document that exhibits strong systems thinking may not be tagged in this system thinking

topic at all if the focus and domain of the document is very different than the rest of the corpus.

Second, even if the stochastic nature and over-simplification of language is overlooked; there is no natural way to generate a traditional prediction of documents that exhibit or do not exhibit systems thinking. Traditional supervised methods identify a boundary between classes. In topic models no boundaries are identified between good and bad systems thinking documents; instead distributions are found. Occasionally unsupervised methods can be use for classification (such as K nearest neighbors), but in these cases each observation is assigned a value from a categorical distribution. However, in LDA and general topic models, each observation (a document) is assigned values from a multinomial distribution. This means that documents are assigned proportions to each topic and not simply in or out of a given topic. Due to the way these proportions are assigned, due to the fact that the values of these proportions can vary tremendously based on the number of topics in the model and the makeup of the corpus, and due to the fact that there is no clear non-systems thinking topic, it is difficult to truly classify the documents.

Collectively, this demonstrates that the methodology must be looked at from an explorative viewpoint; failure to do so can result in incorrect conclusions and overconfidence. Furthermore, the methodology is not designed to be predictive, but rather the use of the tool allows the user to quickly learn. In the process, documents with strong systems thinking may quantitatively identified, but the subjective nature of these concepts still requires human validation.

### **3.4.3. Why Put A Human-in-the-Loop**

Many modern applications of topic models and machine learning focus on removing the human from the prediction and decision making process. However, the human-in-the-loop methodology is different; people are included in the loop of the methodology.

Tremendous value comes from the process of performing the human-in-the-loop methodology as people tend to learn by doing. This inherent value in carrying out the process itself supersedes that of using the methodology as simply a “decision system”. By picking seeding documents, naming the topics, creating the visualizations, and iteration a user will naturally learn a tremendous amount about the content of the corpus. As they gain this knowledge, their questions about the domain will change; a human-in-the-loop allows this new knowledge to direct further exploration of the corpus via model tuning. Additionally, a human-in-the-loop allows the exploration to be directed towards the user’s level of prior knowledge about the corpus and the system the corpus describes.

Finally, the human is included in the loop to overcome the inherent subjectivity that fills language and abstract concepts. Computers and algorithms are great at handling data and binary ideas, but struggle to make sense of complex relationships and to identify associations that are obvious to humans. People are necessary to interpret the concepts that topics represent and to identify abstract ideas like systems thinking. The iterative

nature paired with human intuition serves to validate that the information coming from a topic model is reasonable and not a stochastic fluke, a poor association, or a misunderstanding. This reflects the idea that human computer collaboration can make better sense of the world together than either party on their own [42].

### **3.4.4. Discussion Specific to Identifying Systems Thinking**

The key idea behind the human-in-the-loop methodology is that as someone gains expertise in an area, his or her language changes to reflect it. More specifically, an expert in systems thinking will use language that reflects their expertise; thus the presence of the dimensions of systems thinking demonstrates strong systems thinking.

This methodology's use for identifying systems thinking has several advantages over the supervised method used by Whitehead et al.. First, this process does not require 60% to 80% of the papers of interest to be manually graded; instead it only requires a few additional papers that the practitioner feels represent good systems thinking. This makes the methodology rapid and fast, while allowing for system thinking seeds to be changed quickly if the output is unsatisfactory.

Getting the systems thinking topic to emerge can be difficult in practice, and not all systems thinking topics are equally effective for analysis. If the system's thinking topic is too tightly clustered, no non-seed documents will have significant proportions in it. However, if it is made too general, the proportions will contain no valuable information. Large corpora tend to require more seed documents and more fine-tuning to acquire effective systems thinking topics; the increase in available data allows the topic model to find more potential patterns and clusters. The best way to evaluate the effectiveness of a systems thinking topic is inspection. A user can begin by inspecting the document topic proportions as they quickly provide insight into the question of the topic spread. However, the ultimate validation of a systems thinking topic comes from inspection of the actual documents.

The user should be aware that not every paper *should* exhibit systems thinking. A report that is simply describing an event may not offer any analysis; it should not be surprising if it has little systems thinking. However, this does not make it useless in understanding the spread of the corpus or the system that is describing.

Finally, the human-in-the-loop methodology for identifying systems thinking is only useful if systems thinking is applied to the output. Identifying systems thinking is worthless if the analyst does not use systems thinking to better understand their problems and potential solutions. The human-in-the loop methodology of applying topic models to identify systems thinking does not solve any problems itself; it is simply a tool that can facilitate problem solving.



## **4. Validation Exercises and Case Studies**

### **4.1. Overview of Validation**

It is difficult to validate the effectiveness of the human-in-the-loop methodology as there is no baseline truth to measure success against. Furthermore, language and the interpretation of language are highly subjective. In an effort to demonstrate its usefulness this work presents several case studies and exercises. Each of these exercises seeks to address different goals and to demonstrate the usefulness of the methodology differently.

Whitehead et al.'s graded corpus is used in an attempt to establish baseline truth of what is and is not strong systems thinking [2]. Though this baseline is still highly subjective, the statistical analysis and visualizations of the topic model resulting from this corpus demonstrate that the methodology can quickly identify the documents with excellent systems thinking. The exploration of the corpus focusing on the "human dimension", that is the U.S. Army's framework and research around optimizing human performance and programs, encompasses both the goal of identifying systems thinking and performing systems analysis. The topic model resulting from this corpus shows that a corpus can be quickly summarized and organized for further systems analysis using the human-in-the-loop methodology, and that systems thinking can be identified with the methodology in a specific domain. Finally, the exploration of the past nineteen years of transportation research serves to demonstrate alternative ways of performing systems analysis with the human-in-the-loop methodology. In addition to summarizing the field of transportation research, the topic model is used to explore trends through time. These exercises are not an exhaustive demonstration of the methodologies usefulness in identifying systems thinking or performing systems analysis; they are simply examples.

### **4.2. Evaluation With Graded Corpus**

Application of the human-in-the-loop methodology on the graded corpus presented by Whitehead et al. demonstrates the effectiveness of the methodology for identifying systems thinking.

Initially this corpus consisted of 233 IEEE journal papers graded from graded from 1 (bad systems thinking) to 10 (good systems thinking). 5 systems thinking seeds (provided in Table 7 in the appendix) were added to the corpus and a topic models were run with 10, 12, 15, 20, 25, and 30 topics for 2000 collapsed Gibbs sampling iterations using the default stop words provided by Mallet. These initial topic models did not produce a clear systems thinking topic as they were skewed by the structural words of the paper, so the models were repeated but with the removal of additional words related to academic journals like "fig" "vol" "ieee" "model", "university, and "results" as well as every mention of "IEEE Systems Journal". Table 9 in the appendix provides the full list of additional removed words. These new models each produced an effective systems thinking topic, though an additional, less prevalent structural topic emerged too. The topics from the topic model are detailed in Table 10 in the appendix.

The model with 15 topics was used for analysis as its systems thinking topic seemed the most salient to the author; Figure 5 offers the top words and weights of the topic in table form and Figure 6 visualizes the topic in a word cloud. While the dimensions of systems thinking as enumerated by Whitehead et al. do not overly dominate this topic, the words it contains clearly hint at systems thinking and systems analysis. Additionally, it contains other strong words that supplement the dimensions of systems thinking. For example the words “factors” and “decision” may hint at the authors’ understanding of tradeoffs and traceability even though these words do not show up in the topic. These conclusions again highlight the role the human plays in the methodology.

**Figure 5 Top Words from the Systems Thinking Topic of the IEEE Graded Corpus**

<b>Word</b>	<b>Weight</b>
Systems	3590
System	2513
Engineering	1101
Sos	955
Analysis	853
Management	643
Complexity	595
Architecture	529
Level	524
Requirements	515
Cost	496
Development	491
Factors	449
Decision	427
Thinking	372

**Figure 6 Word Cloud of the Systems Thinking Topic of the IEEE Graded Corpus**

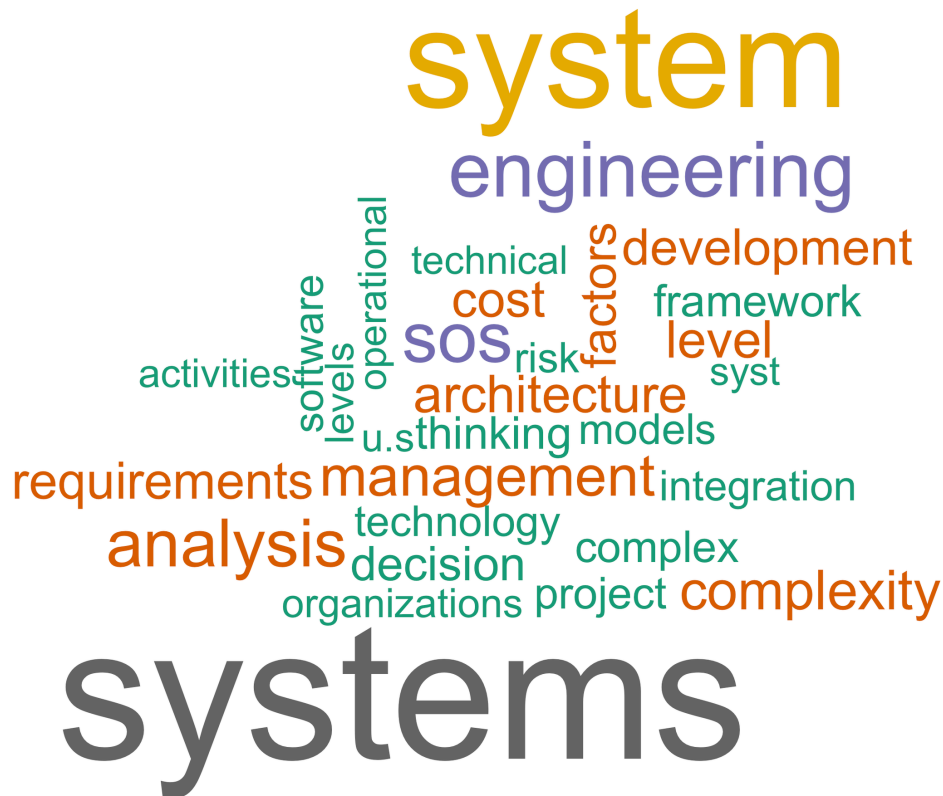


Figure 7 visualizes the modeled, graded corpus with seed documents in a network graph using the dot product as a measure of similarity between documents and the topic proportion as a measure of similarity between topics and documents. The network graph demonstrates the effectiveness of the seeding methodology to generate a systems thinking topic and to pull the documents with strong systems thinking (seeds and others) into the cluster of the visualization, which can be seen in the top right corner of the network graph.

No structural topics were removed for the network graph. Only edges with a similarity measure greater than 0.24 were included; this served to fix the overcorrected graph and was chosen manually and visually by the author. The golden nodes represent topics from the model and are labeled with their given name. The grey nodes represent regular documents, and the red nodes represent the seed documents that we felt demonstrated strong systems thinking. The names of the grey nodes are not displayed for readability, while the red nodes are only labeled with “TGT” for “target”. The graph was visualized using the OpenOrd algorithm with Gephi [38], [43].

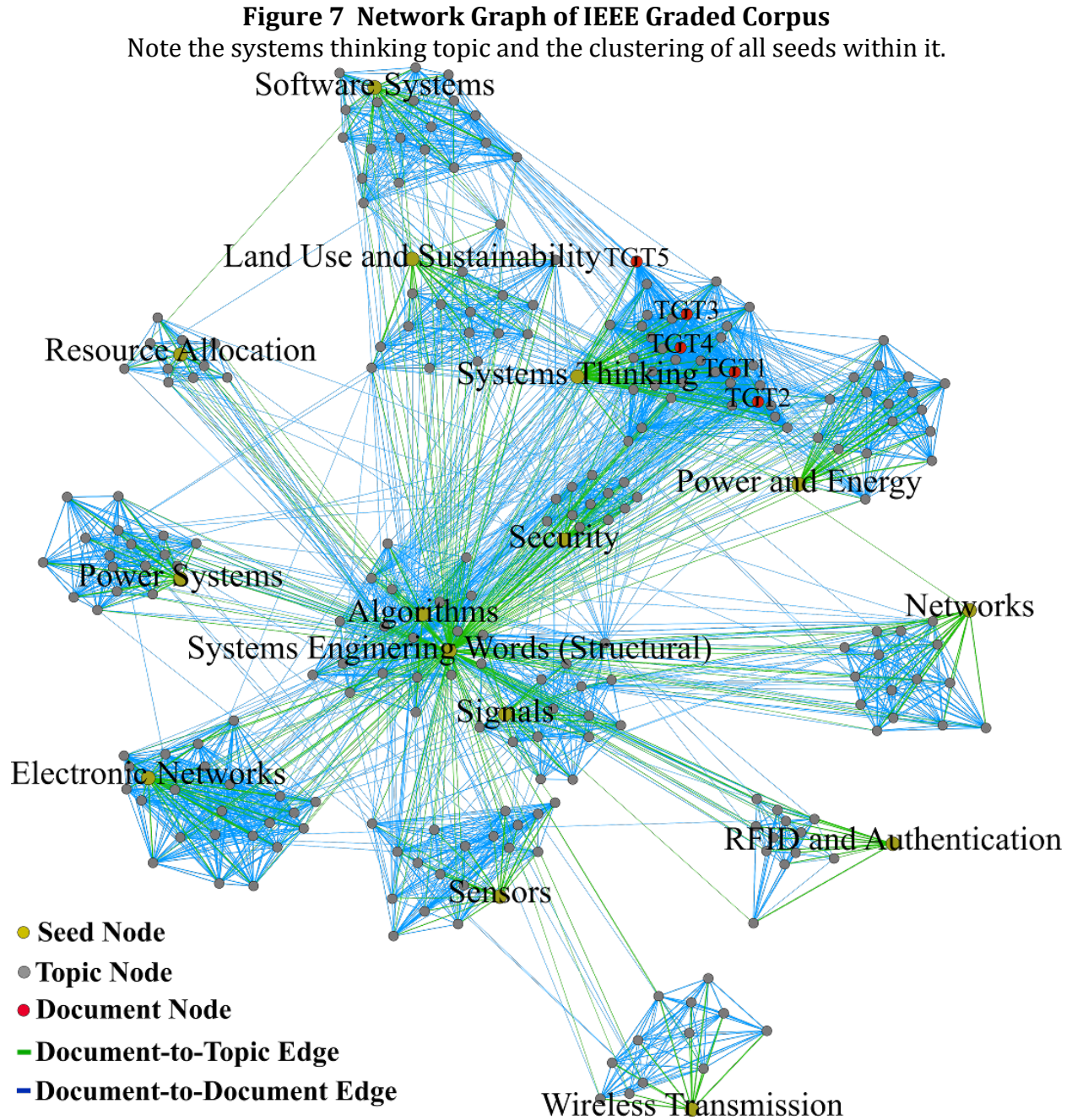
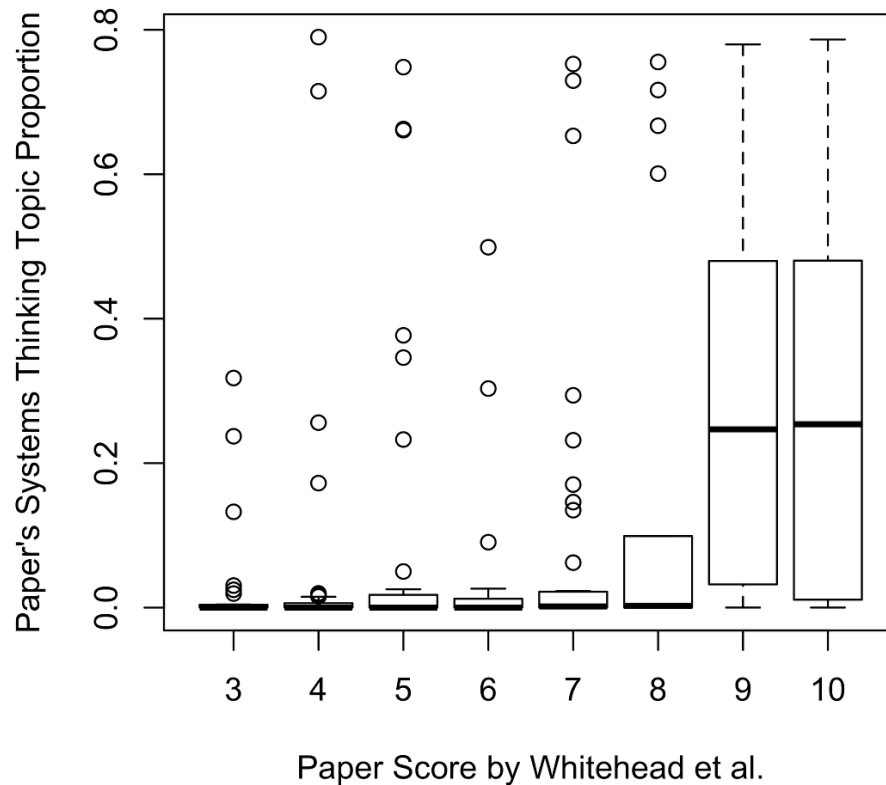


Figure 8 shows a comparative boxplot of the system thinking topic proportions grouped by the documents grade. Papers with a grade of 9 or 10 appear to have significantly higher proportions in the systems thinking topic when compared to the other scores, though there are outliers for each grade group.

**Figure 8 Comparative Boxplot of Grades and Systems Thinking Topic Proportion from IEEE Graded Corpus in 25 Topic LDA Model**



To test this statistically, Tukey test was performed on the papers grouped by their manual score. The mean topic proportion in the systems thinking topic for the groups of papers graded 9 and 10 was statistically higher than the groups of papers graded 3, 4, 5, 6, or 7 at the 0.05 level. There was no statistically significant difference at the 0.05 level between the means of the groups graded 3, 4, 5, 6, 7, or 8 or between the groups graded 8, 9, or 10. Table 4 details the p values for this test.

**Table 4 Significance Values of Tukey Test on Graded IEEE Corpus for Systems Thinking Topic Proportion**

Paper Grade		3	4	5	6	7	8	9	10
Paper Grade	3		1.00	0.88	1.00	0.91	0.24	<0.01	<0.01
	4	1.00		0.92	1.00	0.95	0.25	<0.01	<0.01
	5	0.88	0.92		0.95	1.00	0.88	0.02	0.02
	6	1.00	1.00	0.95		0.97	0.33	<0.01	<0.01
	7	0.91	0.95	1.00	0.97		0.82	0.01	0.01
	8	0.24	0.25	0.88	0.33	0.82		0.44	0.57
	9	<0.01	<0.01	0.02	<0.01	0.01	0.44		1.00
	10	<0.01	<0.01	0.02	<0.01	0.01	0.57	1.00	

The Tukey test and the comparative boxplot suggest that a practitioner would find strong systems thinking much more quickly using this methodology than by manual exploration of a corpus. The strongest system thinking documents have higher topic proportions in the emerging systems thinking topic. This validates the usefulness of the methodology for identifying top documents that exhibit strong systems thinking, but it does not support the use of the methodology as a prescriptive technique. The methodology cannot be used prescriptively for two reasons: first, the topics must be evaluated manually for each iteration and each corpus. Second, the topic proportions and their distributions are extremely dependent on the corpus content as well as the model parameters; there is not natural, consistent way to set a classification threshold that works for all corpora and topic models.

### **4.3. U.S. Army Human Dimension Corpus**

The “human dimension” refers to the U.S. Army’s framework for optimizing human performance. This framework encompasses the processes, programs, research, and operations needed to “assess, integrate, and synchronize [the Army’s] training and education, science and technology, holistic health and fitness, medical and personnel policies, programs, and initiatives in support of the Army Profession” [44]. This framework seeks to optimize human performance because the leaders of the U.S. Army believe that this is how America’s armed forces will remain the best military force despite the continually changing styles of warfare and the growing connectedness of the world [45].

The U.S. Army Training and Doctrine Analysis Center (TRAC) in Monterey, California provided a dataset of 180 documents related to human dimension projects, problems, and studies for a case study of the human-in-the-loop methodology. The Army was specifically interested in identifying where systems thinking was occurring within this domain and summarizing the system.

Several iterations of the human-in-the-loop methodology were performed on the corpus using 20, 25, 30, and 35 topics and seeding the corpus with the four documents that describe strong systems thinking, which are listed in the appendix in Table 7. Before topic modeling, the default stop words provided by Mallet were removed from the corpus. The emerging systems thinking topic was strongest in the model with 25 topics according to the judgment of the author, so it was used for all following analysis.

Each of the 25 topics was named by manually evaluating each topic’s word weights and the corresponding word cloud. Table 1 and Figure 1, from the introduction to topic models section, provide an example of the topic that was named “Physical and Psychological Resilience” as the words describe physical and psychological factors with a strong weight on the word “resilience”. The table lists the top 15 words in the topic with their corresponding weights while the figure shows the word cloud visualization of the topic. A similar table listing all of the topics and their top words is provided in the appendix (Table 11).

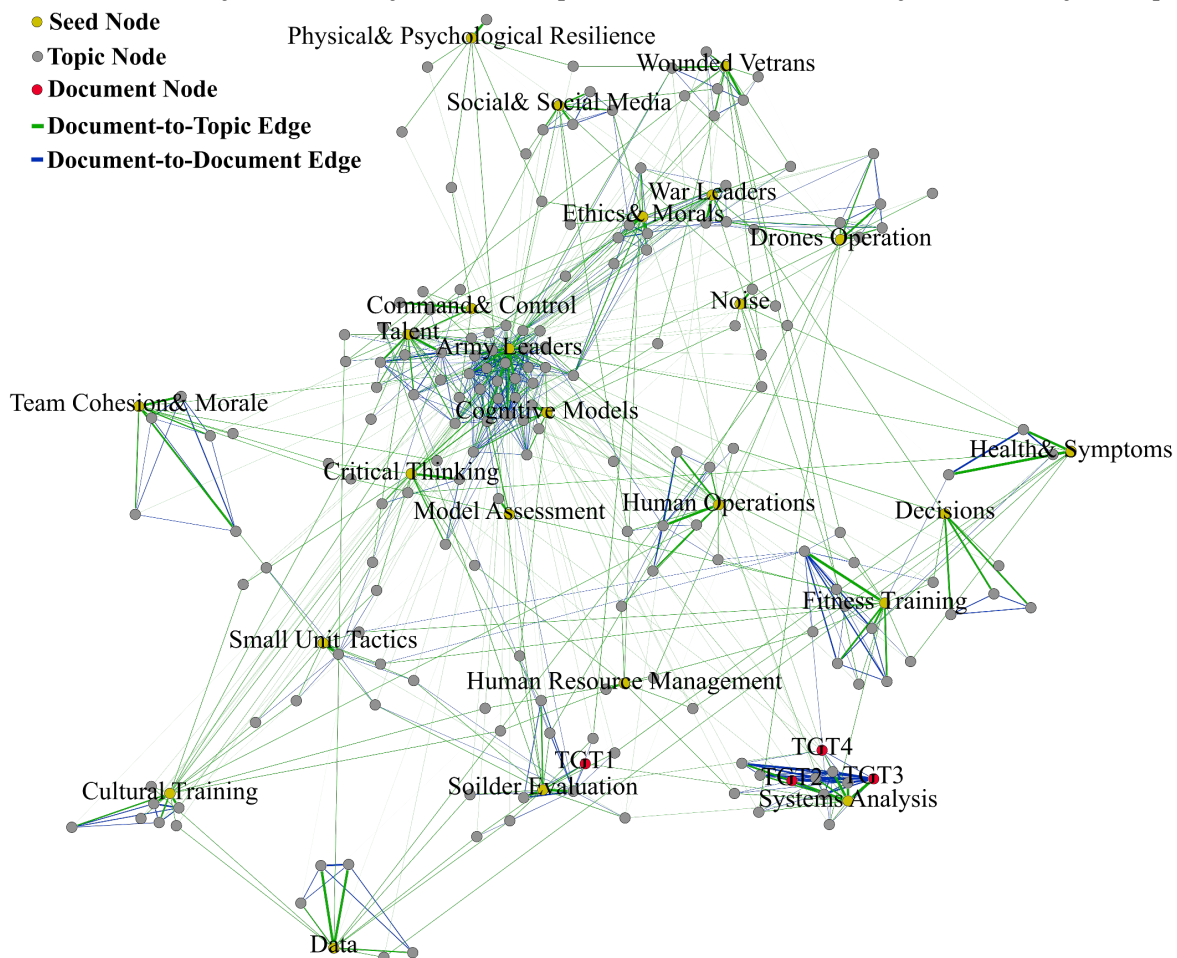


A network graph of the corpus was created, which is visualized in Figure 9, using the dot product as measure of similarity between documents and the topic proportion as a measure of similarity between topics and documents. In order to focus on only the dominant and informative relationships, the structure topic that was dominated by words like “research”, “university”, “table”, and “figure” was removed from the network. Additionally, only edges with a similarity measure greater than 0.05 were included in the graph. The node colors, edge colors, and visualization algorithm are the same as used for the prior network graph in section 4.2.

This network visualization provides a summary of the domain, and it can be further interpreted from the graph that everything the Army does in the human dimension is driven by the goal of making men and women leaders. This is supported by the centrality of the “Army Leaders” topic; almost all other clusters have connections to this topic. This leads to the interpretation that the other emerging clusters represent areas where the Army is trying to make men and women leaders or methods they are using to research and do this. Furthermore, it demonstrates the effectiveness of the seeding methodology to create a custom topic.

**Figure 9 Network Graph of Army Human Dimension**

Note the centrality of the Army Leaders topic and the seeds in the Systems Analysis topic.



This use of the human-in-the-loop methodology for summarization highlights several valuable features. First, the topic model is exhaustive and free from human bias. While a soldier offering an overview of human dimension research in the Army may have some bias towards certain themes or simply forget others, the topic model will not. Second, the summary it provides is effective and allows the user to quickly ask more questions to delve deeper into the material; documents can be read based on their position in the network graph or based on their proportion in topics of interest. Finally, the initial analysis and interpretation was done without the help of a military expert. However, MAJ Connors from TRAC Monterey agreed that the offered interpretation was accurate and insightful. This demonstrates that domain expertise is not required to benefit from the information provided by a topic model.

A systems thinking topic emerged from the methodology, bringing in three of the four seed documents and around nine other documents in the cluster visualization. Despite the subjectivity of strong systems thinking, a group of four members (two from the Department of Systems and Information Engineering at the University of Virginia and two from the Army Training and Doctrine Analysis Center in Monterey, California) read the five papers with the largest proportion in the systems thinking topic and the five papers with the smallest proportion in the topic.

The group collectively agreed that the top five documents demonstrated good systems thinking; they were appropriately scoped, identified important stakeholders, discussed decisions in terms of tradeoffs, and provided traceable arguments. Additionally, they agreed that the bottom five papers did not demonstrate strong systems thinking. About half of these bottom papers did not have potential for systems thinking; they were simply recounting an event or presenting information. However, they felt that the few papers in the bottom five that provided analysis lacked strong systems thinking.

Most of the papers in the top five and bottom five were heavily related to command and control, hinting that there is no clear area where systems thinking is especially strong or weak within the human dimension in the Army. A possible interpretation of this is that the Army needs system thinking the most in the people working within command and control.

Overall, this case study demonstrates that human-in-the-loop methodology can be extremely useful for understanding a system, and can be effective in alternative domains with alternative, domain-specific vocabularies.

#### **4.4. 18 Years of Transportation Research Board (TRB) Conferences**

In an effort to validate topic models' usefulness in exploring trends through time and to explore the use of topic models in overiewing a system, a topic model with 100 topics was created using 19 years of Transportation Research Board (TRB) Annual Meeting papers (~32,000 papers). This model was used to explore transportation research trends through topic models, to see if these trends lined up with our expectations, and to understand if



there is a clear way to aggregate trends and topic proportions. The four methods of viewing trends through time are discussed in section 3.3.

Overall, the topic model revealed that most areas of research within the TRB Annual Meetings have increased in volume over the past nineteen years due to increased volume of accepted papers. However, the topic model suggests that transportation research is becoming more holistic and more global. Research on energy and fuel; alternative transportation modes such as carsharing, bicycles, and busses; accessibility and health; and mobile technology are growing much faster than the increase in accepted papers. Additionally, more international research has been accepted, with a large number of papers coming from China. Research on construction and infrastructure, particularly pavement, bridges, pipes, signs and markings, and barriers and guardrails, has increased in volume, but is now a smaller proportion of the body of TRB papers accepted. One of the reasons for this may be the high cost of research laboratories and facilities required to support infrastructure research relative to that required to support research in some of the more information intensive emerging and rapidly growing research fields.

This effort built on a prior paper submitted to the 2016 TRB Annual Meeting: Text Mining and Topic Modeling on Compendium Papers from Transportation Research Board Annual Meetings by Das et al. [46] They summarize the TRB Annual Meetings between 2008 and 2014 using topic models, and show that “model”, “traffic”, “data”, and “vehicle” are the most used terms in a sample of 3000 TRB abstracts, and that the Texas Transportation Institute and the University of Texas are the top two contributors to the TRB by volume. Additionally, Das uses LDA to topic model the titles of all transportation papers from this period and provides an overview of the top topics and shows their trends over time. However, the Das et al. paper considers only titles and abstracts rather than the entire document. Additionally, the paper does not suggest any causal analysis of these trends. The project presented here builds on the Das et al. foundation by using papers from a time frame three-times longer (1998-2016), by using finer detail through modeling the entirety of each paper (not just the abstract or title), and by providing more insight into relationships and potential causes between transportation research trends and the changing world. This work was submitted to the 2017 TRB Annual Meeting and is in the review process [47].

#### **4.4.1. TRB Analysis: Data Source and Topic Model**

Nineteen years of TRB Conference Papers from 1998 to 2016 compiled the corpus used for this work. All of them were taken from either the CD-ROMs of the compendium of papers (1998 to 2012) or the compendium download available on the TRB website (2013 to 2016). Unfortunately, a CD-ROM for the 2000 conference was unavailable; a copy of the TRB 2000 Compendium Preprint CD-ROM served as a surrogate. This resulted in a collection of 32,946 conference papers. Of these, 856 papers had issues and could not be included in the corpus; the majority of papers with errors were marked as “unavailable for the CD-ROM” (555 out of 856), though others were withdrawn from the conference, had corrupted PDFs, or only included an abstract. This resulted in a final corpus of 32,090

conference papers; a breakdown of the distribution of these papers is shown below (Table 5). This breakdown reveals that the volume of accepted papers has grown significantly over the past 19 years.

**Table 5 Number of Papers in TRB Annual Meeting Compendiums from 1998 to 2016**

Year	Total Number of Papers in Compendium	Number of Papers Included in Corpus	Number of Papers Excluded Due to Errors
1998	1,000	791	209
1999	913	910	3
2000	920	916	4
2001	1,019	1,014	5
2002	1,186	1,183	3
2003	1,376	1,308	68
2004	1,507	1,483	24
2005	1,620	1,573	47
2006	1,713	1,663	50
2007	1,861	1,804	57
2008	1,765	1,694	71
2009	2,011	1,882	129
2010	2,149	2,009	140
2011	2,153	2,149	4
2012	2,518	2,509	9
2013	2,751	2,745	6
2014	1,534	1,525	9
2015	2,550	2,539	11
2016	2,400	2,393	7
<b>Total</b>	<b>32,946</b>	<b>32,090</b>	<b>856</b>

These removed papers and the use of the 2000 preprint instead of the official CD-ROM do change the makeup of the corpus and may result in the differing output from the topic model. However, the 2000 Compendium Preprint CD-ROM is assumed to be sufficiently close in size, content, themes, and language to serve as the data from the year 2000 and the 856 removed papers are assumed not to alter the composition of the corpus significantly.

Mallet, an open-source Java based topic-modeling package, modeled the corpus using a smoothed LDA topic model with 100 topics [18]. 100 topics were chosen after several iterations as it provides a balance between detail and general trends. The entirety of each paper was included in the corpus for the topic model. Prior to modeling, Mallet removed 524 stop words, which are common English articles, adverbs, conjunctions, pronouns, and prepositions. Removing these words results in topics focused more on the content of the papers than the supporting words. The model ran for 4000 iterations while optimizing hyperparameters every 10 iterations, as well as 10 iterative conditional modes iterations, which further smooth the topics after the model is finished.

The top 30 words of the resulting topics were then manually inspected, and each topic was given a name. The named topics were then further grouped by where they loosely fall within the TRB Committee structure. This grouping facilitates understanding of the output of the model and interpretation of trends through time. Note that the topics from the topic model, as aggregations and restructurings of the words within the corpus, do not directly correspond to a committee even though each of the original papers is associated with a committee. Many topics could belong in multiple committee groupings and some do not inherently belong in any. Furthermore, the naming and grouping of topics is inherently subjective, and identifying the uniting feature of a group of words can be difficult. Thus, the topics' names represent a starting point; expertise can be employed to refine the names. These assigned topic names and their groupings are summarized in Table 6; the full table identifying the manually-assigned grouping, topic name, prevalence, and top words in each topic is available Appendix A (Table 12).

**Table 6 Summary Of Transportation Research Topics From LDA Model,  
Grouped By TRB Committee Structure**

<b>Committee Grouping</b>	<b>Topics (From LDA Topic Model)</b>
Aviation	Passenger Logistics & Infrastructure; Airspace & Flight
Highway	Vehicles; Carsharing & Taxi; Motorcycles & Scooters; Driving Experiments & Simulators*
Marine Transport	Ports & Marine Freight
Motor Carriers	Trucks
Pedestrians & Bicycles	Pedestrian & Pedestrian Crossings; Bicycles;
Pipelines	Pipes
Public Transportation	City Rail Stations*; Public Transit Services; Busses
Rail	Rail
	General Construction: Projects; Load & Stress Tests; Work Zones; ????
Design & Construction	Pavement: General Maintenance & Assessment; Cracks, Joints, & Stress; Design & Performance; Concrete; Asphalt Surface & Application; Asphalt Mixes Roads: General Roads; Lane Features; Lane Types; Road Features; Road Geometry Other: Urban Design; Parking; Bridges; Signs & Markings; Barriers & Guardrails; Materials
Operations & Preservation	Emergency Evacuation & Security; Traffic: General Traffic; Traffic Simulation & Models; Traffic Flow & Capacity; Intersections - General; Intersections - Timing & Simulations; Intersections - Traffic Signals; Freeways & Freeway Traffic*

Planning & Environment	Drainage & Water; Sustainability; Weather; Emissions; Landscapes & Wildlife; Soil Testing; Energy & Fuel; Fares & Customers; Urban Areas & Cities; Noise & Sound
Policy & Organization	Development Agencies; Government Offices; Public Involvement; Decisions; Transportation Planning, Regional/Local; Land Development; Laws & Law Enforcement; Cost Benefit Analysis; Taxes & Pricing; Training & Education; Schools & Universities; Infrastructure Funding; Databases & Data Software*
Safety, System Components, & Users	Surveys: General Surveys; Travel Behavior - Distance & Mode; Behavior, Attitudes, and Perception; Income & Employment; Age & Gender; Transportation Alternatives; Travel Behavior - Patterns & Purpose; Safety: Speeds & Speed Limit; Safety & Crashes Tech: Advanced & Automated Tech; Equipment Installation, Power, and Operations; Video & Image Detection/Sensors; GPS, Cell Phones, and Mobile Tech; Other: Accessibility & Health; Employees & Commuting; Freight & Delivery
Paper Structure^	General Structure; TRB Words; Transportation; Data & Tables; Impact & Results; Frequency Words
Modeling & Simulation^	Performance Measures; Models & Variables; General Modeling; Path Planning & Networks; Probability & Machine Learning
Other ^	Europe & Australia; Texas & Florida; Asia; Canada; Travel Time; Park Visitors

\*Some topics could be binned in other committee groupings.

^Some topics did not have a clear committee grouping; a new grouping was created.

These topics provide an overview of what topics have been used in TRB papers over the past 19 years and provide a basis for identifying and visualizing trends in these topics, which is discussed in the following section.

#### 4.4.2. TRB Analysis Discussion of Results: Overview

Network graphs can effectively visualize the space of transportation research, demonstrating clusters and interactions between topics; however, with such a large number of topics and edges these are difficult to view without interactive network graph software or aggregating topics.

Trends in the topics can be seen by calculating the mean topic proportion for a given year or by counting the number of papers with a topic proportion above some threshold. These trends show that the volume of research in almost every topic has grown due to the tremendous increase in the number of papers accepted at TRB (~250% increase from 1998 to 2016). However, the balance in transportation research is changing; transportation

research is becoming more holistic, personal, technological, and global at the expense of traditional infrastructure and construction research.

#### **4.4.3. TRB Analysis Discussion of Results: Transportation Research as a Whole**

The topics provided in Table 6 offer an overview of what transportation research focuses on while the interactions between the topics can offer a deeper perspective. These interactions can be viewed in network graphs, which can include nodes for both the topics and the documents. However, with a large corpus these graphs are computationally expensive to visualize and difficult to understand without using interactive network graph software. For this reason, this work looks at network graphs using only the topics as nodes (topic-to-topic graphs as described in section 3.2).

Aggregating the topics by their committee grouping (that is summing the topic proportions together), as shown in Table 6, allows the relative bonds between committee groupings to be found. Using a graph clustering visualization generates Figure 10, which offers interesting insights into the nature of transportation research in general and, more specifically, the linkages between and among the topic areas. Note that the LDA topics that mapped into modal topics (e.g., Aviation, Marine transport, Highway, etc.) are on the “points” of the star created by the clustering visualization, indicating that mode-oriented research efforts typically focus exclusively on a single mode; the “inner circle” (shaded in Figure 10) includes topics whose keywords map into topics that are inherently multi-modal or intermodal and, not surprisingly, the thicker lines that link these topics indicate the larger volume of papers in these areas. Finally, the “kite” shape in the interior of the shaded circle illustrates the close ties between “Policy and Organization,” “Planning and Environment,” “Design and Construction,” and “Safety System Components and Users,” all representing major cross-cutting elements of transportation research.

**Figure 10 Network Graph Of Committee Focuses Within TRB From Aggregated Topic Model**

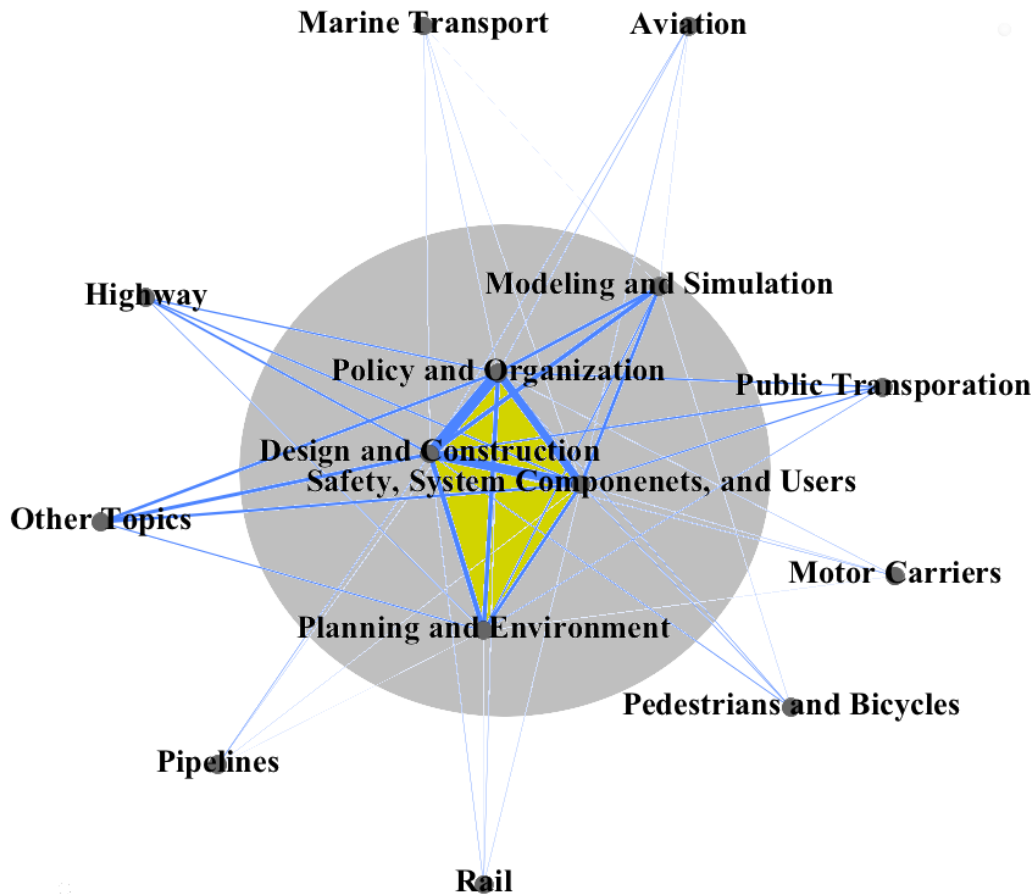
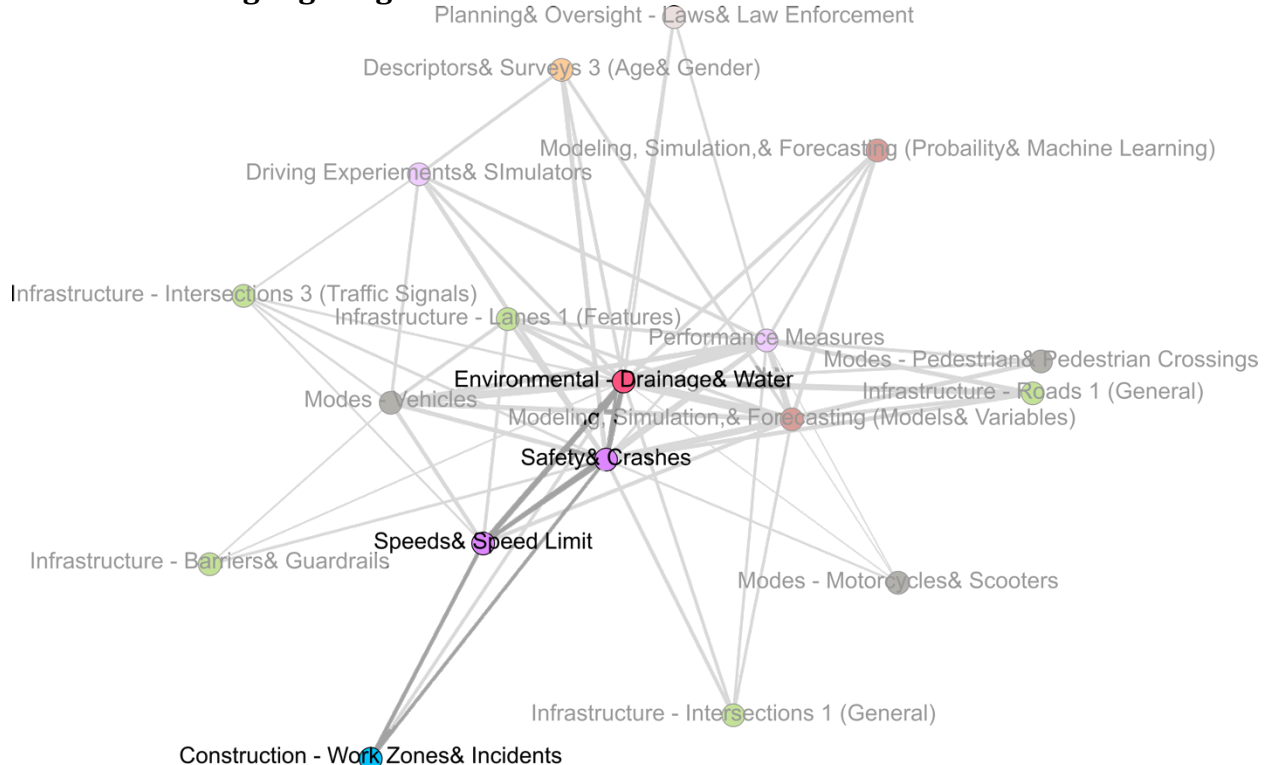


Figure 11 shows the sub-graph found by retaining only the 1st degree connections of the “Safety and Crashes” topic, after removing the structure topics and several of the policy and oversight topics. (The 1st degree connections refer to the topics that “Safety and Crashes” is most strongly connected to directly, not through another topic.) The emphasized sub-graph within Figure 11 highlights the shared connections between the “Safety and Crashes” topic and the “Construction – Work Zones” topic. They interact directly and through the “Environmental – Drainage and Water” and “Speed and Speed Limit” topics. This makes sense as water and drainage can quickly change a work zone, and rain decreases drivers’ visibility. Also, excessive speed is often seen as a factor in work zone crashes. This provides an example of how the transportation body of research can be visualized and understood at a high level with this methodology.

**Figure 11 Network Graph Of First-Degree Connections Of Safety & Crashes Topic: Highlighting Interactions With Construction – Work Zones**



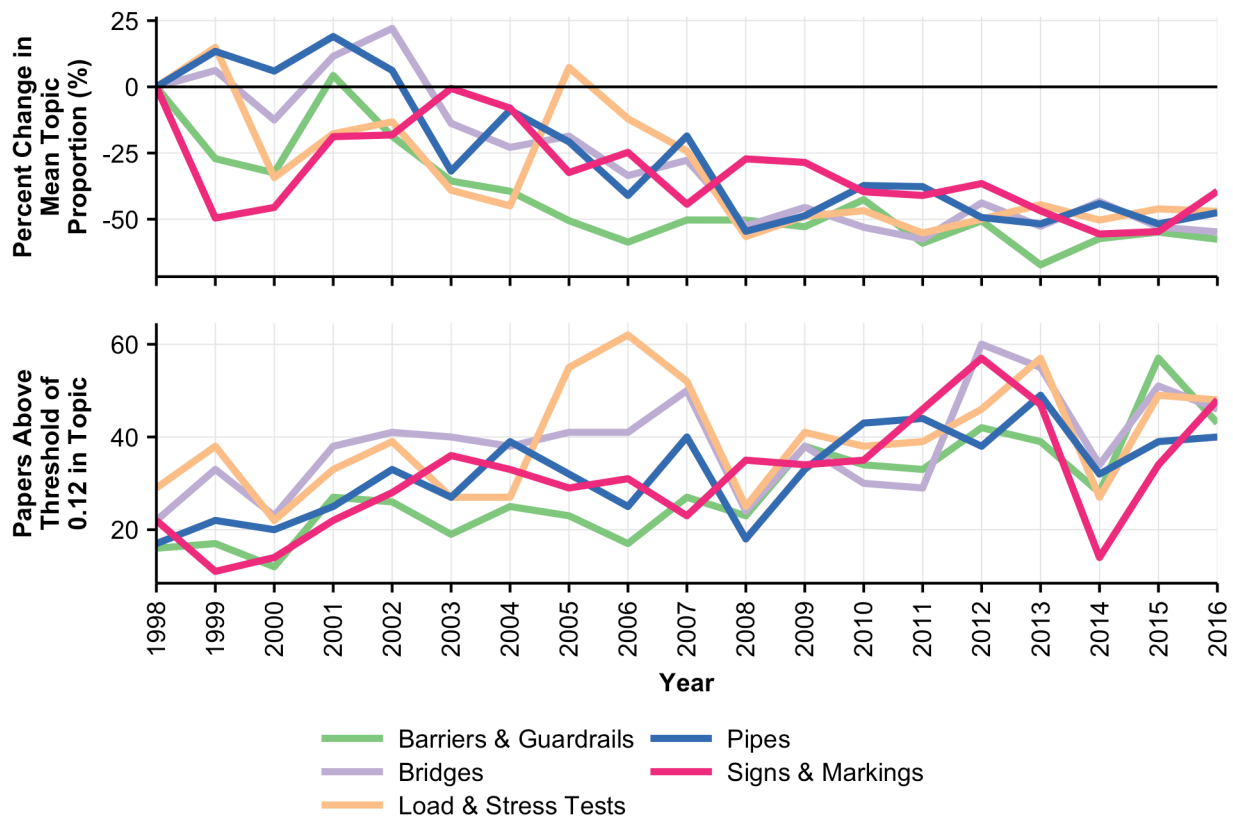
#### 4.4.4. TRB Analysis Discussion of Results: Trends in Transportation Research

All trends in this discussion of the TRB topic model are based on aggregations of topic proportions for each year as discussed in section 3.3; only two of the four methods are used in the visualizations. First, the percent increase of the mean proportion of a topic is used as it provides an easily digestible understanding of the trends in the topics prevalence. Second, the absolute count of papers in topics is used with a threshold value of 0.12. After trying multiple threshold values, 0.12 was chosen for visualization as it corresponds with the average paper having between 1 and 2 key topics that are not related to the paper structure topics (the average paper is 25.3% structure, and there are usually almost zero, but still non-zero, proportions in all topics). This methodology has drawbacks; certain papers may be equally spread through many topics and may not be counted anywhere. Together, these measures provide an appropriate view of the trends in TRB research. Finally, it is worth noting explicitly that neither of these methodologies addresses the quality of research being presented at TRB.

The trends reveal that research on pavement, infrastructure, and construction has generally fallen in popularity as a total proportion of TRB work over the past 19 years. Of the mean topic proportions with the biggest percent decrease (using the 2016 and 1998 topic proportions), the entire top 10 are from these three related sectors.

Of the infrastructure related topics with decreases in mean topic proportion, bridges, barriers and guardrails, pipes, signs and markings, and load and stress tests all see falls of near 50% (Figure 12) and have statically significant negative trends at the 0.05 level. This significance was calculated using a t-test on a basic linear model constructed with the mean topic proportion depending on the year. The decrease in research on bridges is surprising as there have been many news stories reporting on America's decaying bridge infrastructure and several high-profile bridge collapses in this span of time, including the 2007 collapse of the I-35 bridge in Minneapolis, MN that killed 13 and injured 145 people [48], [49]. There is an increase in the total volume of papers on these topics, but it is significantly less than the rate the papers are increasing.

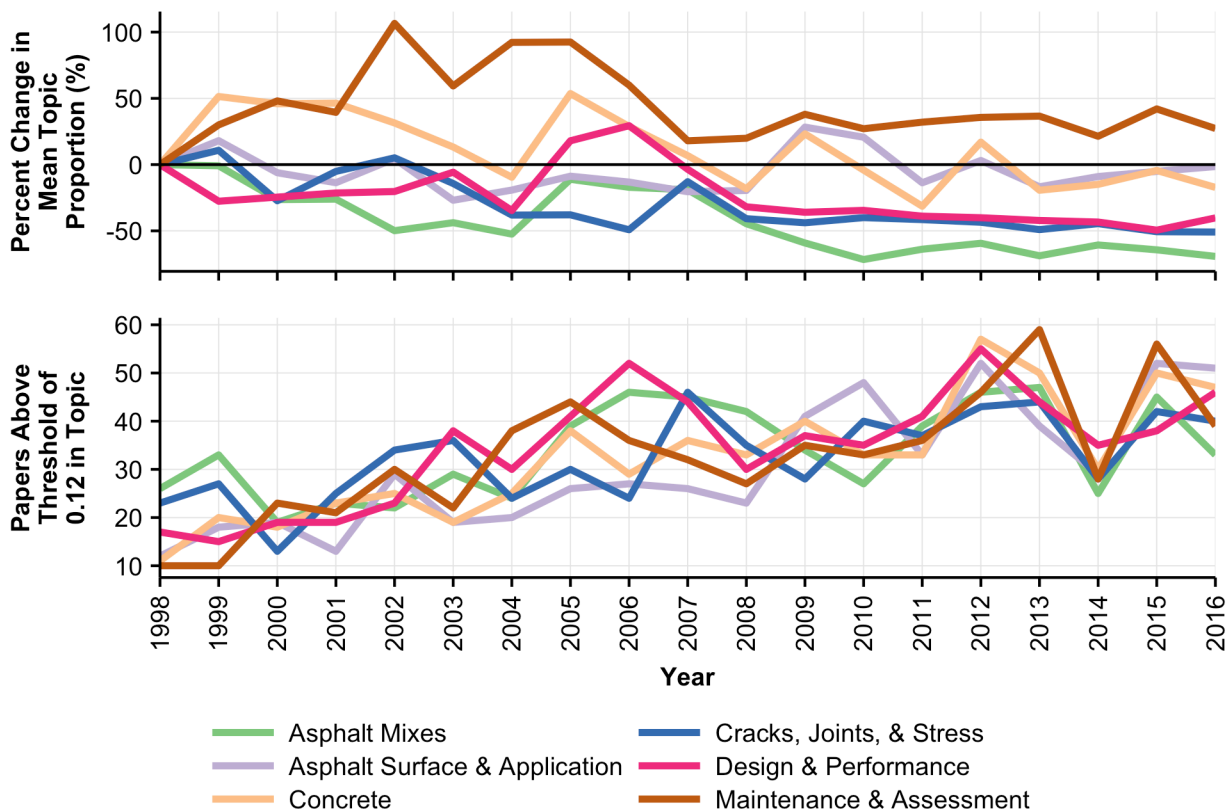
**Figure 12 Decreasing Trends In Topics Relating To Infrastructure From 1998 To 2016**



All topics focusing on pavement see similar behavior as shown in Figure 13; by 2010 all pavement topics' mean proportions have settled significantly lower than they were in the years prior with the exception of the asphalt surface characteristics topic and the concrete topic which stay relatively flat. Additionally, the topic related to maintenance and assessment sees initial strong growth but falls to new, steady lows by 2007. Once again, the volume of papers focused on these areas is growing but at a lower rate than the total volume of papers.



**Figure 13 Decreasing Trends In Topics Relating To Pavements From 1998 To 2016**



A possible explanation for the pavement decreases is that the Long Term Pavement Performance Committee's 20-year milestone occurred in 2009, and a large portion of their studies had projected 20-year time horizons [50]. This timing could imply that many pavement studies were finishing around 2007 to 2009, where the pavement related topics seem to reach a new normal. However, it is more likely that research on construction related issues is simply bounded by facilities and funding. A personal computer can run traffic simulations, but it takes heavy investment to run crash tests on barriers and guardrails.

Surveys and topics relating to human behavior generally saw increases in mean topic proportions, as well as many specific topics such as bicycles, carsharing and taxis, busses, accessibility and health, Asia, GPS and mobile technology, and energy and fuel. Collectively these paint a picture of transportation becoming more holistic and more global.

Figure 14 shows the increased emphasis on demographic specific travel behavior and opinions. Of the seven survey related topics, the three topics focusing on general surveys, how often and how far people travel, and why people travel were relatively flat over this time horizon and are not included in the figure. However, the four topics focused on demographic specific opinions and behaviors had statistically significant positive trends in the mean topic proportions at the 0.05 significance level, with the mean topic proportion of the topic focusing on perception, social behavior, and social attitudes (purple) increasing around 100% over the past 19 years. (These were calculated using a t-test on a linear

model as before.) These topics also see a tremendous increase in the volume of papers focused on them, significantly outpacing the increase in accepted papers.

Potentially, this increased focus can be explained by the ease and availability of this survey data. The widespread usage of mobile technology has made surveys cheap and easy to distribute widely; the use of mobile technology during transportation (for directions, fares, and scheduling) further eases this burden for transportation related work. While this barrier of access has been lowered, this does not explain why it has significantly outpaced the overall increase of papers. One explanation is that there is a growing value in tailoring transportation to specific people and groups, stemming from an increasingly holistic view of transportation research and practice.

**Figure 14 Trends In Topics Relating To Surveys And Human Behavior From 1998 To 2016**

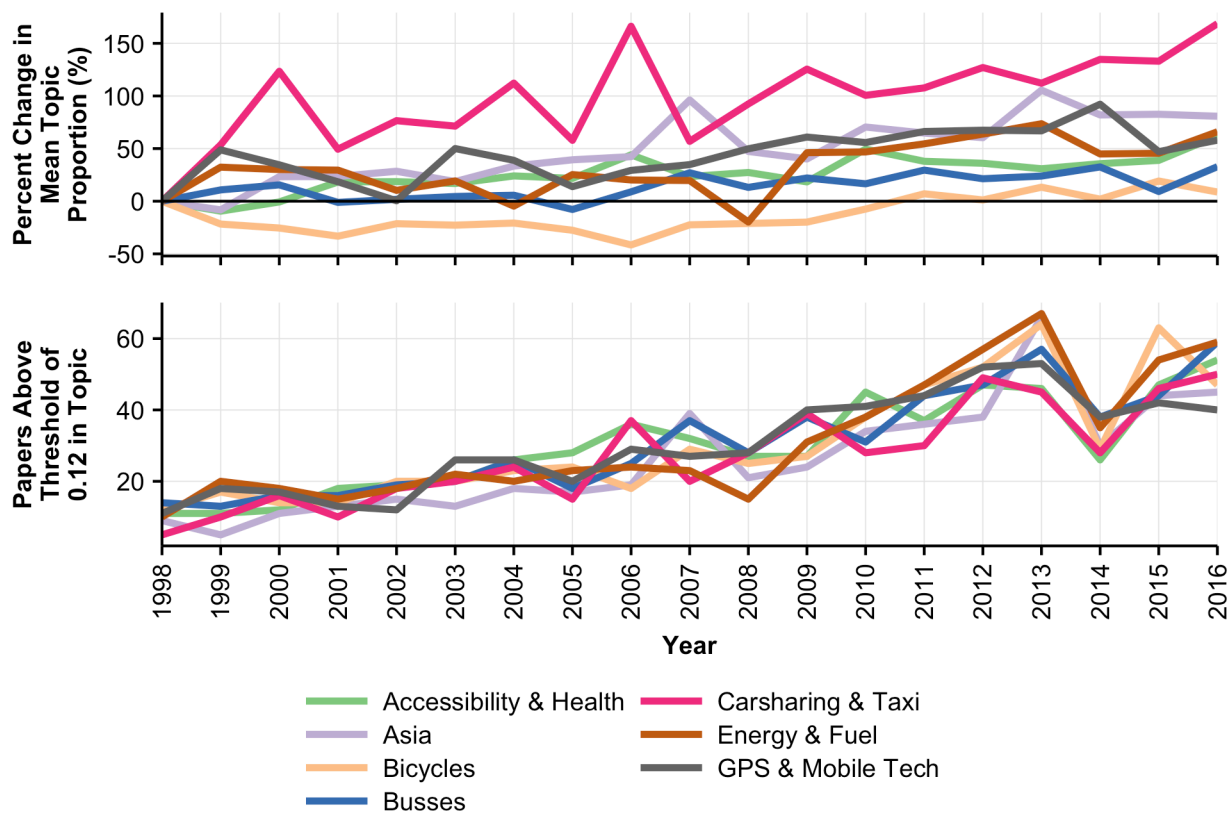


The increased focus on bicycles, carsharing and taxis, busses, accessibility and health, Asia, GPS and mobile technology, and energy and fuel is shown in Figure 15; all of these topics had statistically significant positive trends at the 0.05 level in the mean topic proportion (using the basic linear model and t-test as before). The bicycles topic increased 86% from its low in 2006, while the carsharing and taxi topic increased 168% and the busses topic 33% over the 19 years in their mean topic proportions. All three of these are viewed as environmentally responsible or health-conscious or both means of transportation; this points towards the growing movement towards alternative and holistic forms of

transportation. These trends are related to the increases in energy and fuel (66%) and health and accessibility (62%), which demonstrate that consumers and researchers desire to protect the earth and the people on the earth.

The growth in the Asia-related topic's mean proportion (81%) demonstrates the increase of research being presented at TRB from Asia and about Asia combined with TRB's efforts to embrace a global community. The top words in the topic suggest the majority of this Asian research is coming from China and Japan. This trend can be partially explained by the economic growth of China over the past 19 years; Chinese scholars have a greater ability to attend and present at TRB meetings. Finally, the growth in GPS and mobile technology topic's mean proportion (58%) demonstrates the effect that the recent globalization of technology has had on transportation; GPS and mobile technology is now being used to monitor traffic, plan routes, and report incidents, which was not possible just 10 years ago. Once again, all of these topics also see increases in the volume of papers focused on them, significantly outpacing the increases in accepted papers.

**Figure 15 Increasing Trends In Holistic And Global Topics From 1998 To 2016**



There are other, less systemic trends emerging from the topic model. Vehicles have seen a steady resurgence in focus in recent years, exhibiting close to a 50% increase in mean topic proportion since their low in 2005. Emergency evacuation and crisis related papers tend to see a brief resurgence after major events; there was a 40% increase in the two years following 9/11 and smaller increases after 2008 (bad US tornado year, Sichuan earthquake,

Hurricane Ike) and after 2012 (Hurricane Sandy). Finally, several topics saw bumps in research focus between 2006 and 2008, directly after the US passage of SAFETEA-LU. These include topics related to vehicles, roads, city rail, busses, speed limits, regional planning, and infrastructure funding.

This discussion paints a picture of where transportation has been, but not where it is going. The topic model is descriptive of the field, but is not designed to be predictive or prescriptive. It is up to the TRB Committees and, in some cases, Federal agencies, to decide the directions for future transportation research. All of the trends discussed demonstrate that transportation research's direction reflects a changing world. Ease of access and a societal value of holism have broadened the research agenda within the TRB. As TRB Annual Meeting participation continues to grow, so will the diversity of the accepted papers. Transportation research will anticipate and adapt to this growing and changing world.

## 5. Future Work and Conclusion

There is opportunity to continue building on the human-in-the-loop methodology; potential future work includes performing a true designed experiment with analysts to validate and quantify the effects of the methodology on workflow as well as further modification of the methodology and the topic model algorithm for use in alternative ways.

While the case studies and exercises discussed in section 4 demonstrate the value of the methodology for system analysis, they do not directly quantify or qualify the performance of the methodology versus alternatives. A true designed experiment could be run with the methodology to quantify and qualify this. A collection of analysts could be given a corpus that they have not seen and asked to find specific information within it or to summarize it effectively. A portion could be given a tool that facilitate the human-in-the-loop methodology with training, a portion could use computers and whatever methods they prefer, and a portion could be forced to simply read and skim the corpus.

Additionally, more research could be done on modifying the actual topic model algorithm to increase its effectiveness in identifying systems thinking. Perhaps a hybridization of traditional smoothed LDA and Ramage's labeled LDA could be use to improve the generation of the systems thinking topic [29]. Additionally, this work has only used LDA with the human-in-the-loop methodology; research is necessary to understand best practices in using other topic models such as hierarchical Dirichlet processes and the pachinko allocation model, which allows for correlation between topics [19], [51].

The methodology could be integrated into or expanded into a full-fledged database and information retrieval system, similar to [52]. Allowing users to run a more traditional search by providing a document instead of a search term, may add additional value to the research and exploration process. This project would require careful consideration to balance the non-predictive nature of topic models with instant answer objective of information retrieval.

Finally, there are a few lingering questions regarding the methodology. First, how does the methodology change based on the length of a document? Is there a limit to the length of the document that can be used effectively in the methodology (too long or too short)? Additionally, how does the methodology change when many documents of varying lengths are supplied? Second, is there a method of comparing the similarity between two topic models? As discussed prior, the human-in-the-loop methodology requires many topic models to be run throughout the process. Since these are stochastic models relying on random number seeds, it would be extremely beneficially to have a measure of how similar models are between iterations.

Overall, the human-in-the-loop methodology shows tremendous ability to leverage topic models to identify systems thinking and to assist in the exploration and understanding of a system. This can be valuable to system practitioners everywhere as it will facilitate their systems analysis and allow them to more easily learn from those who came before them.

There is still tremendous subjectivity in the definition of strong systems thinking and the interpretation of the topics, but this is why the human-in-the-loop methodology is so effective. Computers can handle the processing of the data while human intuition handles the subjective nature and tailors the results to focus the most important questions facing analysts

## 6. References

- [1] N Peter Whitehead, William T Scherer, and Michael C Smith, "Systems Thinking About Systems Thinking A Proposal for a Common Language," *IEEE Systems Journal*, vol. 9, no. 4, pp. 1-12, August 2014.
- [2] N Peter Whitehead, William T Scherer, and Michael C Smith, "Use of Natural Language Processing to Discover Evidence of Systems Thinking," *IEEE Systems Journal*, vol. PP, no. 99, p. 2015, May 2015.
- [3] "Language", Merriam-Webster, 2016.
- [4] "Lexicon", Merriam-Webster, 2016.
- [5] Paul Kay and Willet Kempton, "What Is the Sapir-Whorf Hypthesis?," *American Anthropologist*, vol. 86, no. 1, pp. 65-79, Mar 1984.
- [6] Lera Boroditsky, "How Language Shapes Thought," *Scientific American*, pp. 63-65, February 2011.
- [7] Lera Boroditsky and Alice Gaby, "Remembrances of Times East: Absolute Spatial Representations of Time in an Australian Aboriginal Community," *Psychological Science*, vol. 21, no. 11, pp. 1625-1639, November 2010.
- [8] Betty Birner. Does the Language I Speak Influence the Way I Think? [Online]. <http://www.linguisticsociety.org/content/does-language-i-speak-influence-way-i-think>
- [9] The Marx Brothers, *Animal Crackers*, 1930.
- [10] David Blei, "Probabilistic Topic Models," *Communications of the ACM*, vol. 55, no. 4, April 2012.
- [11] Scott Deerwester, Susan T Dumais, George W Furnas, Thomas K Landauer, and Richard Harshman, "Indexing by Latent Semantic Analysis," *Journal of the American Socitey for Information Science*, vol. 41, no. 6, pp. 391-407, September 1990.
- [12] Thomas Hofmann, "Probabilistic Latent Semantic Indexing," *Proceedings of the 22nd Annual Interanation ACM SIGIR Conference on Research and Developemnt in Information Retrieval*, pp. 50-57, 1999.
- [13] David M Blei, Andrew Y. Ng, and Michael I. Jordan, "Latent Dirichlet Allocation," *Journal of Machine Learning Research*, pp. 993-1022, January 2003.
- [14] Walter R Gilks, Andrew Thomas, and David J Spiegelhalter, "A Language and Program for Complex Bayesian Modelling," *Journal of the Royal Statistical Society. Series D (The Statistician)*, vol. 43, no. 1, pp. 169-177, 1994.
- [15] Stuart Geman and Donald Geman, "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. PAMI-6, no. 6, pp. 721-741, November 1984.
- [16] Thomas P Minka, "Expectation Propagation for Approximate Bayesian Inference," *Proceedings of the Seventeenth Conference on Uncertainty in Artificial Intelligence*, pp. 362-269, August 2001.
- [17] Hanna M Wallach, David Mimno, and Andrew McCallum, "Rethinking LDA: Why Priors Matter," *Advances in Neural Information Processing Systems 22*, pp. 1973-1981, 2009.

- [18] Andrew Kachites McCallum. (2002) MALLET: A Machine Learning for Language Toolkit. [Online]. <http://mallet.cs.umass.edu>
- [19] Yee Whye Teh, Michael I Jordan, Matthew J Beal, and David M Blei, "Hierarchical Dirichlet Processes," *Journal of the American Statistical Association*, vol. 101, no. 476, pp. 1566-1581, 2006.
- [20] Jonathan Chang and David M Blei, "Hierarchical relational models for document networks," *Ann. Appl. Stat.*, vol. 4, no. 1, pp. 124-150, 2010.
- [21] Zhiyuan Chen et al., "Leveraging multi-domain prior knowledge in topic models," *IJCAI International Joint Conference on Artificial Intelligence*, pp. 2071-2077, Dec. 2013.
- [22] Jagadeesh Jagarlamudi, Hal Daumé, and Raghavendra Udupa, "Incorporating lexical priors into topic models," *Proceedings of the 13th Conference of the European Chapter of the Association for Computational Linguistics*, pp. 204-213.
- [23] David Andrezejewski, Xiaojin Zhu, and Mark Craven, "Incorporating Domain Knowledge into Topic Modeling via Dirichlet Forest Priors," *Proceedings of the 26th Annual International Conference on Machine Learning*, pp. 25-32, 2009.
- [24] Mark Steyvers and Tom Griffiths, "Probabilistic Topic Models," in *Latent Semantic Analysis: A Road to Meaning*. Mahwah, New Jersey, USA: Laurence Erlbaum, 2007.
- [25] Carson Sievert and kenneth E Shirley, "LDavis: A Method for Visualizing and Interpreting Topics," *Proceedings of the Workshop on Interactive Language Learning, Visualization, and Interfaces*, pp. 63-70, June 2014.
- [26] David M Blei and John D Lafferty, "A Correlated Topic Model of Science," *The Annals of Applied Statistics*, vol. 1, no. 1, pp. 17-35, June 2007.
- [27] Michael A Livermore, Allen B Riddel, and Daniel Rockmore, "Agenda Formation and the U.S. Supreme Court: A Topic Model Approach," *Available at SSRN*, 2016.
- [28] Xing Wei and W. Bruce Croft, "LDA-Based Document Models For Ad-hoc Retrieval," *Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*, pp. 178-185, 2006.
- [29] Daniel Ramage, David Hall, Ramesh Nallapati, and Christopher D Manning, "Labeled LDA: A Supervised Topic Model for Credit Attribution in Multi-Labeled Corpora," *Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing*, pp. 248-256, 2009.
- [30] Shen Zhang, Jinjun Tang, Hua Wang, and Yinhai Wang, "Enhancing Traffic Incident Detection by Using Spatial Point Pattern Analysis on Social Media," *Transportation Research Record: Journal of the Transportation Research Board*, vol. 2528, pp. 69-77, 2015.
- [31] Donald E Brown, "Text Mining the Contributors to Rail Accidents," *IEEE Transactions on Intelligent Transportation Systems*, vol. 17, pp. 346-355, February 2016.
- [32] Russell L Ackoff, "Towards a System of Systems Concepts," *Management Science*, vol. 17, no. 11, pp. 661-671, 1971.
- [33] C West Churchman, *The Systems Approach*. New York, USA: Delacorte Press, 1968.
- [34] John E Gibson, William T Scherer, and William F Gibson, *How To Do Systems Analysis*. Hoboken, New Jersey, USA: John Wiley & Sons, Inc., 2007.



- [35] Ryan C Boyer, William T Scherer, Cody H Fleming, Casey D Connors, and N Peter Whitehead, "Using Topic Models to Understand Systems and to Identify Strong Systems Thinking," *Submitted for Publication to IEEE Systems Journal*, 2016.
- [36] Thomas L Griffiths and Mark Steyvers, "Finding Scientific Topics," *Proceedings of the National Academy of the Sciences*, vol. 101, no. 1, pp. 5228–5235, April 2004.
- [37] Thomas M. J. Fruchterman and Edward M. Reingold, "Graph Drawing by Force-Directed Placement," *Journal of Software: Practice and Experience*, vol. 21, no. 11, pp. 1129–1164, November 1991.
- [38] Shawn Martin, W. Michael Brown, Richard Klavans, and Kevin Boyack, "OpenOrd: An Open-Source Toolbox for Large Graph Layout," *Proceedings of SPIE - The International Society for Optical Engineering*, January 2011.
- [39] Yifan Hu, "Efficient, High-Quality Force-Directed Graph Drawing," *The Mathematica Journal*, vol. 10, no. 1, pp. 37–71, 2006.
- [40] Alexander S Yeh, Lynette Hirschman, and Alexander A Morgan, "Evaluation of Text Data Mining for Database Curation: Lessons Learned from the KDD Challenge Cup," *Bioinformatics*, vol. 19, no. Suppl. 1, pp. i331–i339, February 2003.
- [41] Andreas Hotho, Andreas Nürnberger, and Gerhard PaaB, "A Brief Survey of Text Mining," *LDV Forum - GLDV Journal for Computation Linguistics and Language Technology*, vol. 20, pp. 19–62, 2005.
- [42] Mary Cummings, "Man versus Machine or Man + Machine?," *IEEE Intelligent Systems*, vol. 29, no. 5, pp. 62–69, September 2014.
- [43] M Bastian, S Heymann, and M Jacomy, "Gephi: An Open Source Software for Exploring and Manipulating Networks," *International AAAI Conference on Weblogs and Social Media*, 2009.
- [44] United States Army Combined Arms Center, "The Human Dimension White Paper: A Framework for Optimizing Human Performance," United States Army, White Paper October 9, 2014.
- [45] United States Army Combined Arms Center. (2016, March) Force 2025 and Beyond Human Dimension Publications. [Online]. <http://usacac.army.mil/pubs/Force-2025-and-Beyond-Human-Dimension>
- [46] Subasish Das, Xiaoduan Sun, and Anandi Dutta, "Text Mining and Topic Modeling on Compendium Papers from Transportation Research Board Annual Meetings," *95th TRB Annual Meeting Compendium of Papers*, 2016.
- [47] Ryan C Boyer, William T Scherer, and Michael C Smith, "Trends Over Two Decades of Transportation Research: A Machine Learning Approach," *96th TRB Annual Meeting Compendium of Papers (Submitted)*, 2017.
- [48] National Transportation Safety Board, "Collapse of I-35W Highway Bridge Minneapolis, Minnesota August 1, 2007," National Transportation Safety Board, Washington, D.C., Report 2008.
- [49] Steve Kroft, *Falling Apart: America's Neglected Infrastructure*, 2014.
- [50] TRB Long-Term Pavement Performance Committee, "Fulfilling the Promise of Better Roads," Transportation Research Board of the National Academies, Washinton, D.C.,

ISBN 0-309-07247-6, 2001.

- [51] Wei Li and Andrew McCallum, "Pachinko Allocation: DAG-Structured Mixture Models of Topic Correlations," *Proceedings of the 23rd International Conference on Machine Learning*, pp. 577-584, 2006.
- [52] Commonwealth Computer Research Institute. (2016) Ketos: Use Neural Networks to Find the Documents You Need. [Online]. <http://www.ccri.com/case-studies/ketos/>
- [53] C West Churchman, "Perspectives of the Systems Approach," *Interfaces*, vol. 4, no. 4, pp. 6-11, August 1974.
- [54] Anthony M Hodgson, "Hexagons for Systems Thinking," *European Journal of Operational Research*, vol. 59, no. 1, pp. 220-230, May 1992.

## A. Appendix: Seed Document and Document Cleaning Information

This appendix includes supplemental tables describing and listing the seed documents used in prior sections and analysis as well as cleaning lists used in training topic models. . These are included to provide a more comprehensive view of the topic modeling process while not compromising readability of the thesis.

**Table 7 Seed Documents Used in Training Topic Models**

Document	Used in		Description or Citation
	Graded Corpus	Human Dimension Corpus	
Dimensions of Systems Thinking	TGT1	TGT1	List of the dimensions of systems thinking from [1], See Table 3
Chapter 10 – How To Do Systems Analysis	TGT2	TGT2	[34]
Systems Thinking about Systems Thinking: A Proposal for a Common Language	TGT3	TGT3	[1]
Systems Thinking Word List		TGT4	Custom list of potential synonyms for the dimensions of systems thinking, see Table 8
Perspectives of the Systems Approach	TGT4		[53]
Hexagons for Systems Thinking	TGT5		[54]

**Table 8 Words Included in Systems Thinking Word List (Not the Dimensions of Systems Thinking)**

Words				
Tradeoff	Backgrounds	Values	Metric	Goals
Tradeoffs	Environment	Belief	Metrics	Priorities
Trade-Off	Environments	Beliefs	Measurement	Priority
Trade	Context	Principle	Measure	Need
Exchange	Contexts	Principles	Index	Needs
Rank	Framework	Understanding	Indices	Requirement
Rankings	Frameworks	Stakeholders	Assess	Requirements
Compare	Situation	Stakeholder	Assessment	Require
Comparison	Conditions	Shareholder	Assessments	Requires

Comparisons	Circumstance	Shareholders	Indicator	Intent
Evaluate	Circumstances	Participant	Indicators	Intention
Evaluation	Foundation	Participants	Objective	Intentions
Evaluations	Foundations	Recipient	Objectives	Target
Background	Value	Recipients	Goal	Targets
Tradeoff	Backgrounds	Values	Metric	Goals

**Table 9 Words Removed From Graded IEEE Corpus Before Use In 25 Topic LDA Model**

Model		
Word		
Approach	Order	January
Based	Paper	February
Case	Performance	March
Data	Problem	April
Degree	Proc	May
Due	Process	June
Fig	Research	July
Function	Results	August
IEEE	Section	September
IEEE Systems Journal	Senior Member IEEE	October
Information	Set	November
Introduction	Shown	December
Journal	Table	
Life Fellow IEEE	Time	
Member IEEE	University	
Model	Vol	
Number	Work	

## B. Appendix: Supplemental Topic Model Tables

This appendix includes supplemental tables describing topic models addressed in prior sections and analysis. These are included to provide a more comprehensive view of the topic models' results while not compromising readability of the thesis. More detailed tables and complete data files can be acquired by contacting the author.

**Table 10 Topics From Graded IEEE Corpus For Evaluating Methodology's Usefulness In Identifying Systems Thinking In 15 Topic LDA Model**

Name	Drc. Prm.	Top Eight Words
Systems Engineering Words (Structural)	0.902	systems, design, system, simulation, engineering, values, high, analysis
Systems Thinking	0.190	systems, system, engineering, sos, analysis,

Software Systems	0.125	management, complexity, architecture service, services, system, analysis, architecture, software, components, design
Networks	0.100	node, network, nodes, networks, wireless, packet, mobile, routing
Sensors	0.088	sensor, system, control, sensors, formation, robot, tracking, decision
Power Systems	0.085	power, system, control, output, voltage, network, wind, current
Signals	0.084	signal, interference, frequency, signals, detection, filter, receiver, satellite
Power and Energy	0.076	energy, wind, power, load, market, cost, system, generation
Land Use and Sustainability	0.073	water, geoss, fire, global, earth, sustainability, learning, land
Security	0.071	network, supply, resilience, node, nodes, power, networks, hurricane
Algorithms	0.064	system, biometric, fuzzy, fingerprint, network, feature, algorithm, minutiae
Resource Allocation	0.060	grid, resource, resources, job, scheduling, user, application, jobs
Electronic Networks	0.058	network, traffic, bandwidth, frame, networks, delay, packet, transmission
Wireless Transmission	0.051	channel, system, spectrum, feedback, image, scheme, channels, networks
RFID and Authentication	0.042	rfid, tag, tags, reader, system, product, agent, readers

**Table 11 Topics From Army Human Dimension Corpus In 25 Topic LDA Model**

Name	Drc. Prm.	Top Eight Words
Structure Topic	1.408	research, military, number, report, time, work, analysis, provide
Army Leaders	0.478	army, leaders, training, development, leader, leadership, command, mission
War Leaders	0.274	war, military, leadership, u.s, press, leaders, strategic, public
Human Operations	0.232	human, force, operations, information, dynamics, military, u.s, cultural
Critical Thinking	0.146	thinking, critical, creative, education, army, innovation, students, simulation
Systems Analysis	0.136	systems, system, information, analysis, data, complex, decision, cognitive
Ethics & Morals	0.134	professional, army, ethic, ethics, ethical, profession, moral, military
Noise Topic	0.132	research, military, social, environments, context, organizational, army, factors
Decisions	0.125	war, cognitive, military, decisions, people, clausewitz, biases, decision
Soldier Evaluation	0.117	category, stimulus, stimuli, client, learning, responses, visual, problem
Model Assessment	0.099	model, assessment, human, validation, models, simulation, performance, behavior
Small Unit Tactics	0.099	building, opfor, team, blufor, squad, back, scenario, fire
Talent	0.098	talent, management, army, officer, strategy, strategic, contracting, accessed
Physical and Psychological Resilience	0.091	resilience, force, air, stress, health, fitness, social, factors
Wounded Veterans	0.086	research, wounded, defense, bureaucracy, warriors, warrior, department, ict
Human Resource Management	0.084	management, program, nsps, talent, apg, engineering, employees, army
Cultural Training	0.073	training, cross-cultural, culture, cultural, competence, army, learning, military
Team Cohesions & Morale	0.064	agility, morale, individual, motivation, factors, cohesion, team, learning
Social & Social Media	0.063	media, military, social, army, egyptian, study, chapter, ghana
Drones Operation	0.060	training, drones, cognitive, bias, game, biases, human, stress

Command & Control	0.056	command, army, information, cer, control, cers, staff, center
Data	0.055	data, big, hfs, session, practice, strong, community, xwh
Cognitive Models	0.033	human, army, dimension, mindset, concept, sketch, cognitive, performance
Fitness Training	0.028	fitness, physical, training, army, body, program, strength, test
Health & Symptoms	0.027	ptsd, smoking, participants, health, study, hurricane, cessation, symptoms

**Table 12 Transportation Research Topics from 19 Years of TRB Annual Meeting Papers in LDA Topic Model**

Committee Grouping	Name	Drc. Prm.	Top Eight Words
Paper Structure^	General	0.563	provide, case, important, time, potential, part, large, current
	TRB Words	0.445	paper, annual, meeting, trb, original, revised, submittal, figure
	Transportation	0.342	transportation, research, study, washington, studies, university, national, found
	Data & Tables	0.298	data, analysis, table, results, number, based, average, total
	Impact & Results	0.247	percent, average, increase, change, total, table, rate, increased
	Frequency Words	0.101	day, traffic, peak, period, hours, hour, days, daily
	Development Agencies	0.183	management, process, development, agencies, program, implementation, issues, agency
	Government Offices	0.159	state, highway, states, department, federal, dot, administration, county
	Cost Benefit Analysis	0.104	cost, costs, benefits, analysis, total, economic, table, savings
	Public Involvement	0.086	public, participants, group, community, groups, focus, project, people
Policy & Organization	Databases & Data Software	0.083	data, information, user, system, gis, database, map, software
	Decisions	0.081	social, decision, policy, planning, approach, decisions, knowledge, analysis
	Transportation Planning, Regional/Local	0.072	transportation, planning, regional, local, plan, projects, funding, region
	Infrastructure Funding	0.068	public, private, government, infrastructure, sector, services, market, financial
	Land Development	0.053	land, development, urban, area, areas, density, residential,



Modeling & Simulation	Taxes & Pricing	0.038	neighborhood pricing, tax, congestion, price, fuel, vmt, revenue, fee
	Laws & Law Enforcement	0.03	enforcement, law, driver, inspection, safety, program, carriers, motor
	Training & Education	0.029	training, students, transportation, engineering, program, learning, knowledge, education
	Schools & Universities	0.026	school, children, students, campus, schools, university, parents, child
	Performance Measures	0.121	performance, measures, level, quality, evaluation, based, measure, criteria
	Models & Variables	0.101	model, variables, models, variable, significant, regression, effect, effects
	General Modeling	0.061	model, models, demand, scenario, modeling, scenarios, developed, development
	Path Planning & Networks	0.058	network, link, problem, demand, number, links, path, algorithm
	Probability & Machine Learning	0.042	prediction, cluster, probability, approach, estimation, input, clusters, parameters
	Travel Time	0.105	time, travel, times, minutes, arrival, min, reliability, departure
Other ^	Europe & Australia	0.074	transport, european, public, london, policy, urban, countries, local
	Texas & Florida	0.045	texas, florida, utility, txdot, fdot, austin, houston, utilities
	Asia	0.031	china, japan, shanghai, city, zhang, beijing, chen, policy
	Canada	0.03	canada, toronto, ontario, canadian, city, region, montreal, vkt
	Park Visitors	0.029	information, route, park, national, traveler, routes, users, atis
Highway	Vehicles	0.097	vehicle, vehicles, cars, passenger, car, type, types, fleet
	Driving Experiments &	0.039	drivers, driving, driver, simulator,

Operations & Preservation: Traffic	Simulators		task, participants, performance, visual
	Carsharing & Taxi	0.016	car, carsharing, taxi, members, number, cars, mobility, sharing
	Bicycles & Scooters	0.009	motorcycle, kim, motorcycles, riders, dublin, motorcyclists, hawaii, riding
	General	0.096	traffic, corridor, congestion, highway, roadway, area, existing, segment
	Simulation & Models	0.053	simulation, traffic, model, models, parameters, behavior, modeling, based
	Flow & Capacity	0.037	capacity, delay, flow, queue, length, hcm, volume, lane
	Intersections - General	0.034	intersection, intersections, traffic, turn, left, lanes, volume, signal
	Freeways & Freeway Traffic	0.03	freeway, traffic, flow, ramp, loop, detector, speed, lane
	Intersections - Timing & Simulations	0.026	signal, traffic, control, priority, phase, delay, cycle, green
	Intersections - Traffic Signals	0.023	red, traffic, intersection, green, signal, light, stop, driver
Operations & Preservation	Emergency Evacuation & Security	0.027	emergency, evacuation, security, risk, response, events, event, fire
Safety, System Components, & Users: Tech	Advanced & Automated Tech	0.081	system, systems, technology, information, technologies, control, applications, communication
	Equipment Installation, Power, and Operations	0.074	system, equipment, power, prt, installation, operation, systems, maintenance
	Video & Image Detection/Sensors	0.044	detection, video, image, sensor, camera, sensors, images, system
	GPS, Cell Phones, and Mobile Tech	0.035	data, gps, location, collection, mobile, phone, cell, accuracy
Safety, System Components, & Users: Surveys	General Surveys	0.087	survey, respondents, surveys, responses, sample, questions, response, reported
	Travel Behavior - Distance & Mode	0.081	travel, trips, trip, mode, distance, modes, household, auto
	Behavior & Attitudes	0.046	behavior, car, factors, people, attitudes, analysis, environment, perceived

Safety, System Components, & Users: Safety	Income & Employment	0.043	income, households, household, population, census, urban, employment, areas
	Age & Gender	0.039	age, drivers, older, driving, women, group, driver, gender
	Transportation Alternatives	0.035	choice, car, van, alternative, utility, alternatives, attributes, preference
	Travel Behavior - Patterns & Purpose	0.027	activity, activities, household, time, work, travel, duration, tour
	Speeds & Speed Limit	0.064	speed, speeds, mph, limit, km/h, limits, percentile, reduction
Safety, System Components, & Users: Other	Safety & Crashes	0.059	safety, crashes, crash, accident, accidents, number, risk, injury
	Freight & Delivery	0.032	freight, transportation, industry, goods, delivery, logistics, commodity, supply
	Accessibility & Health	0.028	health, accessibility, mobility, access, people, services, walking, physical
	Employees & Commuting	0.023	work, employees, commute, program, programs, commuters, employers, employee
Design & Construction: Roads	General Roads	0.07	road, traffic, roads, rural, local, highways, south, users
	Road Geometry	0.057	design, curve, distance, curves, horizontal, highway, vertical, roadway
	Lane Features	0.053	lane, site, sites, lanes, shoulder, width, median, passing
	Lane Types	0.023	lanes, toll, lane, hov, hot, pricing, peak, express
	Road Features	0.014	gap, roundabout, tunnel, roundabouts, exit, entry, gaps, approach
Design & Construction: Pavement	Pavement - General Maintenance & Assessment	0.035	pavement, maintenance, condition, asset, management, life, system, years
	Pavement - Cracks, Joints, & Stress	0.026	pavement, cracking, crack, joint, cracks, concrete, slab, overlay
	Pavement - Design & Performance	0.023	pavement, design, sections, layer, section, base, pavements, thickness

Design & Construction: Other	Pavement - Concrete	0.021	concrete, strength, cement, test, aggregate, ash, results, air
	Pavement - Asphalt Surface & Application	0.021	surface, pavement, friction, tire, test, asphalt, texture, surfaces
	Pavement - Asphalt Mixes	0.018	asphalt, aggregate, mix, test, binder, mixtures, mixes, mixture design, street, streets, city, urban, community, sidewalk, environment
	Urban	0.054	community, sidewalk, environment
	Parking	0.046	parking, san, francisco, city, spaces, area, bay, bart
	Bridges	0.031	bridge, bridges, deck, steel, concrete, design, girder, span
	Signs & Markings	0.025	sign, signs, markings, pavement, marking, retroreflectivity, lighting, distance
	Materials	0.025	density, construction, asphalt, material, materials, quality, recycled, mix
	Barriers & Guardrails	0.015	barrier, test, impact, vehicle, post, guardrail, system, rail
Design & Construction: General	Projects	0.065	project, projects, construction, design, contractor, contract, work, quality
	Load & Stress Tests	0.036	load, test, stress, strain, loading, analysis, results, failure
	Work Zones	0.036	zone, work, traffic, incident, zones, incidents, message, closure
	????	0.021	profile, figure, iri, test, surface, wave, gpr, testing
Planning & Environment Environment	Drainage & Water	0.062	water, flow, drainage, runoff, quality, d"t, hydraulic, treatment environmental, sustainability, sustainable, impacts, climate, transportation, change, impact
	Sustainability	0.037	weather, temperature, conditions, snow, winter, maintenance, rain, ice
	Weather	0.033	emissions, emission, fuel, air, ghg, consumption, carbon, reduction
	Emissions	0.031	energy, fuel, electric, oil, hybrid, gas, electricity, economy
	Energy & Fuel	0.024	energy, fuel, electric, oil, hybrid, gas, electricity, economy
	Fares & Customers	0.022	fare, shopping, customer,

			customers, card, smart, business, online
	Urban Areas & Cities	0.021	transport, brt, urban, cities, city, system, countries, developing
	Landscapes & Wildlife	0.02	landscape, wildlife, mitigation, areas, roadside, species, area, environmental
	Soil Testing	0.018	soil, moisture, content, test, cement, strength, water, base
	Noise & Sound	0.012	noise, sound, levels, level, figure, barrier, measurements, dba
	City Rail Stations	0.058	station, stations, rail, line, city, york, chicago, commuter
Public Transportation	Public Transit Services	0.058	transit, service, ridership, services, bus, system, riders, agencies
	Busses	0.035	bus, buses, stop, passengers, stops, route, passenger, routes
Motor Carriers	Trucks	0.035	truck, trucks, axle, weight, class, load, heavy, trailer
	Pedestrian & Pedestrian Crossings	0.032	pedestrian, pedestrians, crossing, crossings, crosswalk, traffic, walking, cross
Pedestrians & Bicycles	Bicycles	0.024	bicycle, bike, cycling, cyclists, facilities, bicyclists, bicycling, bicycles
Rail	Rail	0.027	rail, train, track, trains, railroad, railway, passenger, freight
Pipelines	Pipes	0.023	pipe, soil, pile, design, wall, figure, rock, piles
Marine Transport	Ports & Marine Freight	0.019	port, container, ports, terminal, containers, vessel, ferry, barge
	Passenger Logistics & Infrastructure	0.018	airport, airports, air, passengers, passenger, airlines, airline, access
Aviation	Airspace & Flight	0.011	aircraft, airport, runway, flight, traffic, aviation, air, delay