Chapter 40

Application of Topic Modelling for Literature Review in Management Research

Barsha Saha

Economics and Public Policy
Indian Institute of Management Shillong, Shillong, India

Abstract—The paper proposes topic modeling as an efficient method for an exploratory literature review for management research. The model has been built to analyze research contributions of Game Theory predominant in the area of management research. The study primarily caters to 2 objectives, first, to build a topic model using the Latent Dirichlet Allocation (LDA) method. Secondly, the study presents one of the optimization methods to build the improved topic model. The results for both models are compared in terms of statistical measures (coherence scores in this case), semantic coherence, and interpretability. A considerable improvement from the base model is observed in the optimized version of the base model. Solution concept of Nash Equilibrium, Market games and Market competition, Allocation games, and Network games are found to be predominant sub-themes in the extracted topics.

Keywords—Exploratory Literature Review, Game Theory, Text Mining, Topic Modelling, Optimization, Management Research

I. INTRODUCTION

Management research by nature is interdisciplinary as it draws in from multiple domains of research. To better understand the concepts of a research area, the interdisciplinary nature of management research calls for in-depth knowledge of contributions to the domain, i.e., through a review of the existing literature. An efficient way to initiate the study of such a vast body of contributions is through an exploratory literature review (ELR). The traditional ELRs had suffered from the biases of a researcher in terms of selection of articles under study and also have been very time-consuming to conduct [14]. It has also received criticism due to a lack of replicability and validity. To curb these short-comings and to ensure replicability and validity, the author proposes a systematic topic modeling based exploratory literature review. To illustrate, the author conducted a topic modeling based ELR on Game Theory. The reason for selecting Game Theory is primarily to exhibit its interdisciplinary background and multifaceted applications. Another reason for the selection was the author's expertise with the topic to better interpret the model results. Topic modeling is an efficient and scalable text mining tool that had been used in this paper to analyze articles to extract major research themes of Game Theory. These extracted themes are treated as the output of the ELR, which provides a general idea of the discipline [3]. These individual themes then could be subjected to further in-depth structured/systematic study [10].

The objectives of this study are twofold. The first is to conduct a proper exploratory literature review of an interdisciplinary topic through the Latent Dirichlet Allocation (LDA) based topic modeling method [1] [3]. The second objective is to build an optimized topic model for better interpretation. Contrary to the usual iterative optimization method, here in this paper, the optimized model had been built on the base model using a wrapper function. This approach aids in the efficient utilization of CPU resource

and time compared to the iterative method of tweaking model parameters to find the best fit.

The purpose of ELR is to obtain broader research themes of a domain. Since topic modeling operates with the bag of words, i.e., analyses the textual corpus based on frequency and co-occurrences of keywords, it helps to extract important words indicative of the aforementioned themes of an ELR. It has to be kept in mind that the topic modeling is an unsupervised machine learning method. The underlying computational mechanism can extract keywords for the desired number of topics. But they might not be semantically coherent or useful at all. Hence, it requires human judgment and the expertise of the researcher to interpret the results successfully [9].

II. METHOD

The paper builds a framework (Figure 1) to conduct topic modelling for an exploratory literature review. The study utilizes the open-source topic modeling library Gensim, available in Python, to create the program [15]. The entire analysis can be segregated into three stages: pre-processing, topic modelling, and post-processing [3].

The pre-processing handled the data clean-up and data preparation to build topic models in the next step. The preparation followed the standard process of tokenization followed by building N-grams, removal of Stopwords, and finally, Lemmatization.

The final corpus was then used to build topic models, beginning with the base model and then derive the optimized model. Here, the author employed the LDA method to build the topic models since it has several advantages over other topic modeling methods like LSA (Latent Semantic Analysis) [12]. LDA is a generative model, allowing to group the articles in the corpus into distinct themes [20]. Moreover, it is possible to employ the fitted model to analyze a new document/text [4]. Also, LDA is useful for large scale corpus, making it suitable for exploratory literature review on the bulk of research articles [8][11].

The post-processing catered to understanding and interpreting the latent themes from the keywords extracted in the topic modeling step. The objective of a researcher in a topic modeling based study is to obtain the best fit model which is statistically valid and semantically coherent.
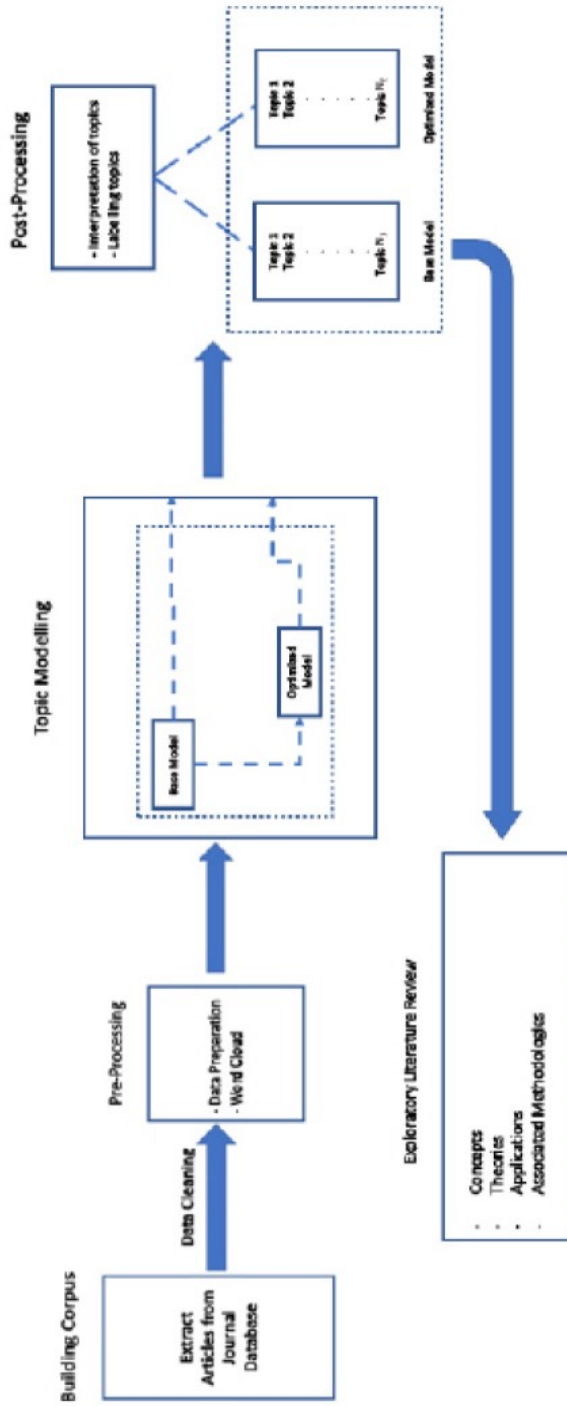

Fig. 1. Schematic representation of the topic modeling based exploratory literature review framework (Figure prepared by author)

III. RESULT

The result section is spread over in subsections to present the outcomes of pre-processing, topic modeling, and post-processing operations. The details of each subsection are as follows:

A. Pre-processing

Game Theory as a discipline deals with the strategic interaction between 2 or more agents. The agents have payoffs for each action taken by them and the other players. This interaction can be translated to represent cases in different disciplines. For

example, in 1944, game theory was formally introduced as a mathematical solution concept in economics [18]. However, researchers soon found its applications in areas other than Economics, like Social Science, Computer Science, Biology, etc. Management studies had incorporated learnings from Game Theory to study the interaction among players in the market, price and quantity competition, business strategy (M&A, expansion, internationalization).

Since the study focuses on the literature pertaining to management research, the search criteria to extract research articles were restricted to Economics, Business Management, Decision Science, and Social Science domain. Furthermore, the result was filtered for articles only in English. It is necessary to specify the source language since the program has to handle language-specific nuances, parts of speech, and default stop word list.

Using the aforementioned search criteria in the SCOPUS database, the search resulted in 15,103 articles. The extracted data contained information on the respective research article spread across 25 columns like Authors, Title, Year, ..., Abstract, ..., etc. A sample of the data headers is as shown below Table 1.
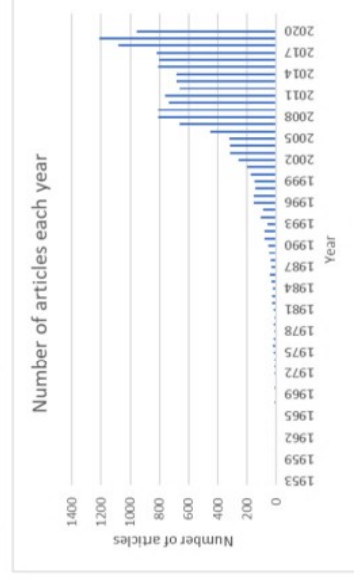
Table 1   Overview of the extracted data

Authors	Author(s) ID	Title	Year	Source title	Volume	Issue	Art. No.	Page start	Page end	Page count	Cited by	DOI
Leifer E.M	164210975	Trails of in	1988	Sociologic	3		4	499	524		8	10.1007/B
Brams S.J.	660373962	National s	1988	Synthese	76		2	185	200		5	10.1007/B
Marte A.L.	700338945	How does	1988	Journal of	19		4	166	168		21	
Judice J.J.	700489435	Reformul	1988	Journal of	57		1	123	149		13	10.1007/B
Pachter M	700501540	An optima	1988	Journal of	56		2	179	204		5	10.1007/B

Link	Abstract	Author Keyword	Index Keyword	ISSN	ISBN	CODEN	Document Type	Publication Stage	Access Type	Source	EID
https://ww	Game theory seeks solutions th			8848971			Article	Final		Scopus	2-s2.0-0007246863
https://ww	Issues that arise in using game			397857			Article	Final		Scopus	2-s2.0-0141844225
https://ww	[No abstract availabl		aged; agin	220124			Article	Final		Scopus	2-s2.0-0024043805
https://ww	A family o	absolute		223239		COMPUTE	Article	Final		Scopus	2-s2.0-0024000344
https://ww	We consid	Collision		223239		KINEMATI	Article	Final		Scopus	2-s2.0-0023964872

This raw data was then cleaned and arranged for further processing. First, the problems with missing data and erroneous data were handled, which resulted in the removal of 410 entries, thereby leaving 14,693 entries for the next step of pre-processing.

An overview of the selected articles is as follows:

A gradual increase in Game-Theoretic studies was observed over the years from 1953 to 2020 (till October), with a peak of 1203 in 2019, as shown in Figure 2. In 2020, by October 948 articles were published, with the possibility of more publications by the end of the year.

The top 10 most cited works are presented in Table 2. It can be observed that all these articles broadly fall under the scope of Game Theory. However, conceptually, they are from different backgrounds. For example, mathematical modeling, supply chain management, decision strategies (at the political and institutional level), behavioral or sociological phenomena (preference, reciprocity), and experimental studies were the most prominent research areas. Hence, there are ample sub-topics under the game theory that could be the focus of the next level of in-depth literature review. In order to find the major themes under Game Theory, this exploratory literature review is performed. This study can be considered as precedence to conducting an in-depth survey of the literature in the future.

From the extracted data (25 columns), only the fields Authors, Title, Year (of publication), Cited by, and Abstract are essential for the analysis in this paper. The abstracts of the papers act as the corpus for this study.

A corpus is essentially a collection of documents/ text which serves as the input for topic modeling. This corpus was further processed to derive the topic models to extract the latent themes of Game Theory.

Number of articles each year

Number of articles

Year

1953 1959 1962 1965 1969 1972 1975 1978 1981 1984 1987 1990 1993 1996 1999 2002 2005 2008 2011 2014 2017 2020

0 200 400 600 800 1000 1200 1400

Figure 2   Number of articles published in each year

Table 2 Top 10 highest cited article

Title	Authors	Year	Cited by
On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games	Dung P.M.	1995	2735
The digitization of word of mouth: Promise and challenges of online feedback mechanisms	Dellarocas C.	2003	1788
Veto players: How political institutions work	Tsebelis G.	2011	1786
Supply chain coordination with revenue-sharing contracts: Strengths and limitations	Cachon G.P., Larivière M.A.	2005	1544
Understanding institutional diversity	Ostrom E.	2009	1358
Democracy and redistribution	Boix C.	2015	1071
Predicting How People Play Games: Reinforcement Learning in Experimental Games with Unique, Mixed Strategy Equilibria	Erev I., Roth A.E.	1998	1009
A theory of reciprocity	Falk A., Fischbacher U.	2006	998
Direct marketing, indirect profits: A strategic analysis of dual-channel supply-chain design	Chiang W.-Y.K., Chhajed D., Hess J.D.	2003	898
Preferences, property rights, and anonymity in bargaining games	Hoffman E., McCabe K., Shachat K., Smith V.	1994	884

Continuing with the pre-processing step, the clean-up procedure of the corpus catered to the four following purposes:

1. Conversion of the corpus to lowercase,
2. Removal of punctuations,
3. Removal of URLs, and
4. Removal of STOPWORDS (includes default stop words in English and a Customized list).

A wordcloud was built to observe the most frequent words in the corpus. The program was designed to fetch the top 200 words (Figure 3).

Fig. 3. WordCloud of top 200 words in the corpus

The largest words are Game Theory (as expected), model, games, strategies, equilibrium, information. These are indicative of key terminologies in game theory. Looking closer, the smaller terms represent economic games, with terms like behavior, network, decision, price, conditions, management, competition, agents, product, price. It can be interpreted as narrowing down scope from a general broader perspective of game theory to application in the areas pertaining to Economics, consistent with the aforementioned search criteria.

B. Topic Modelling and Post-Processing

This processed data (i.e., corpus) was carried forward to the next stage to build the topic models. The topic models were built using the LDA (Latent Dirichlet Allocation) Method. It is an unsupervised and probabilistic modeling method that generates topics from an unlabelled corpus [1][2]. Following the bag of words approach, the frequency of occurrences of

the words is computed instead of analyzing the semantics and meaning of the sentences in the corpus. The built-in functionalities of the Gensim library were utilized to build the topic models.

The post-processing is the analysis phase, where topics are interpreted from the keywords derived from each model. These themes are represented with brief one- or two-word labels or with a single sentence description where required.

For ease of presentation and comparability, each model is explained below, followed by their post-processing themes and labels.

1. Base model

The base model was prepared using the Gensim library [15] and estimated with 10 topics. The coherence score was computed for both base and optimized models [19]. In the newer literature, this external model quality parameter is considered a better measure to evaluate model quality. There are multiple ways to compute coherence, for example, C_v, C_{UCI}, C_{UMass}, C_A, and C_P.

Here, coherence is measured by C_v, which is computed on a sliding window [16]. It is a single set segmentation of the most frequently occurring words and an indirect confirmation measure using NPMI and the cosine similarity [16]. The perplexity is computed as -18.6922972369299, and the coherence score is 0.39130963562193 6. The keywords for the 10 topics of the base model are presented in Table 3.

From an analyst's point of view, there is no fixed good or bad coherence (C_v) score since the score is highly dependent on the nature of the corpus. Nevertheless, an improvement in the coherence score is an indicator of better topic interpretability. The program is designed to fetch only the top 10 words for each topic. From the keywords, it is evident that some topics are interwoven, e.g., topic 3 and topic 7. Also, there are topics like 4, 6, and topic 9, with some redundant words, which compromises the interpretability of the themes. The labels for the interpretable topics are presented in Table 5.

The dynamic visualization of the model in Figure 4 shows that there are much overlap among the topics [17]. It represents the repetition of keywords in multiple themes, which results in closely related topics. Hence there exists a need for optimization of the model.

2. Optimized Model

From the computational point of view, it is assumed that an improvement in coherence score is an indicator of a better model. The base model has been iteratively evaluated for improved coherence score. But due to the nature of the corpus, the best fit model extracted only 2 topics with a better coherence score. These topics were very generic, which is also expected intuitively. The model quality parameters like perplexity (hence loglikelihood) and topic coherence are often found to be uncorrelated or, perhaps, negatively correlated [6]. Hence, at least for this type of corpus, a different optimization method is required. There are multiple optimization methods available to enhance a topic model [forthcoming working paper by author].

In this paper, an enhanced version of LDA is implemented using Mallet (Machine Learning for Language Toolkit) wrapper to build the optimized model [13].

The model is tested for 10 topics. The resultant coherence score is 0.46895355346569, which is an improved estimate compared to the previous models 1 and 2. The keywords for this model are given in Table 4.

The inter topic distance map in Figure 5 shows a comparatively balanced spread of topics across the plane. As the model has improved on coherence score and a good balance of topics as seen from the map, the semantic analysis is performed on the optimized model to identify the major themes hidden in the corpus [Table 5]. It was found that this model covers a set of 10 diverse themes. The themes cover aspects from market games in the context of economic theories, the study of social behaviour. It also proposes experimental methods to understand game-theoretic concepts. The relevance of these sub-themes was validated and confirmed from domain expertise.

Each of these themes is worth pursuing a focused study. Hence, this model serves the original purpose of conducting a literature review.

IV. DISCUSSION

It was found that both models extract at least one topic that is indicative of the central idea or definition of game theory. The topics in the base model cover the broader application area (topic 2, 5) of game theory. The other topics catered to the key terminologies and the popular solution concept of Nash Equilibrium. Strangely, the latter context did not surface in the optimized model. The base model's point of concern is that 3 topics are incomprehensible from the top 10 keywords extracted to identify the topic. Hence, the need for an optimized model arose.

The optimized model performed much better in extracting the most discussed sub-topics in Game Theory (Topics 2, 3, 5, 6, 7, 9, and 10). Management research fields such as business strategy, public policy, international relations, microeconomics, and macroeconomics heavily rely on these research contributions.

The point to be highlighted here is that this study has three parts for analysis. First, the python program efficiently extracts keywords from the corpus. Each set of keywords represent each topic. The second part of the analysis is the trickiest as it is subjective, depends on human judgement and the researcher's expertise [7]. The researcher first checks for the semantic coherence of the group of keywords assigned to a topic. If they made sense, then based on domain expertise, the researcher chose 'labels' to indicate the themes [8]. This two-step process completes the topic modeling part. The final part is about conducting the ELR. The purpose of an exploratory study is to extract themes from a broad discipline of research. It allows to obtain an overview of the contents of the discipline as well as familiarizes the researcher with the seminal works in the area. The researcher then can draw from these extracted themes to proceed with a more focused and systematic study. For example, taking a cue from this study, one could narrow down the research area to one of the topics extracted by the models [9].

V. CONCLUSION

The paper fulfilled its objectives by presenting topic modeling as an efficient tool for an exploratory study. It had also displayed one of the optimization methods to prepare a better model. The paper used LDA and its Mallet wrapper to build the models. The inherent advantage of LDA lies in its capability to handle large scale corpus efficiently [11]. Since the management research is interdisciplinary in nature, a full-fledged well-defined literature review requires analysis of many research contributions. The manual method of conducting a literature review is not devoid of the researcher's biases and suffers from a lack of replicability and validity, and also very much time-consuming [14]. Therefore, an LDA based method aids in handling these concerns.

A systematic study of the literature on the selected topic(s) from the topic modeling based ELR would complete the process of a rigorous literature review [forthcoming working paper by author].

The lack of a one-size-fits-all optimization method could be considered a drawback of this method. As the nature of the data, co-occurrences of words, key terminologies could affect the program output. One way to handle the disruption from unwanted keywords is by having a good custom stopword list. Alternatively, the base model can be tested with a different combination of hyperparameters to test for best fit.

In conclusion, this is a new and efficient way of conducting a literature review. Recent literature by authors from multiple domains has presented their findings and effectivity of similar methods to conduct a literature review [5]. The scope of this method is promising and would be benefitted from future contributions to the area.

FIGURES AND TABLES

The relevant tables and figures pertaining to topic models are presented as follows.

Table 3　Topics of the base model

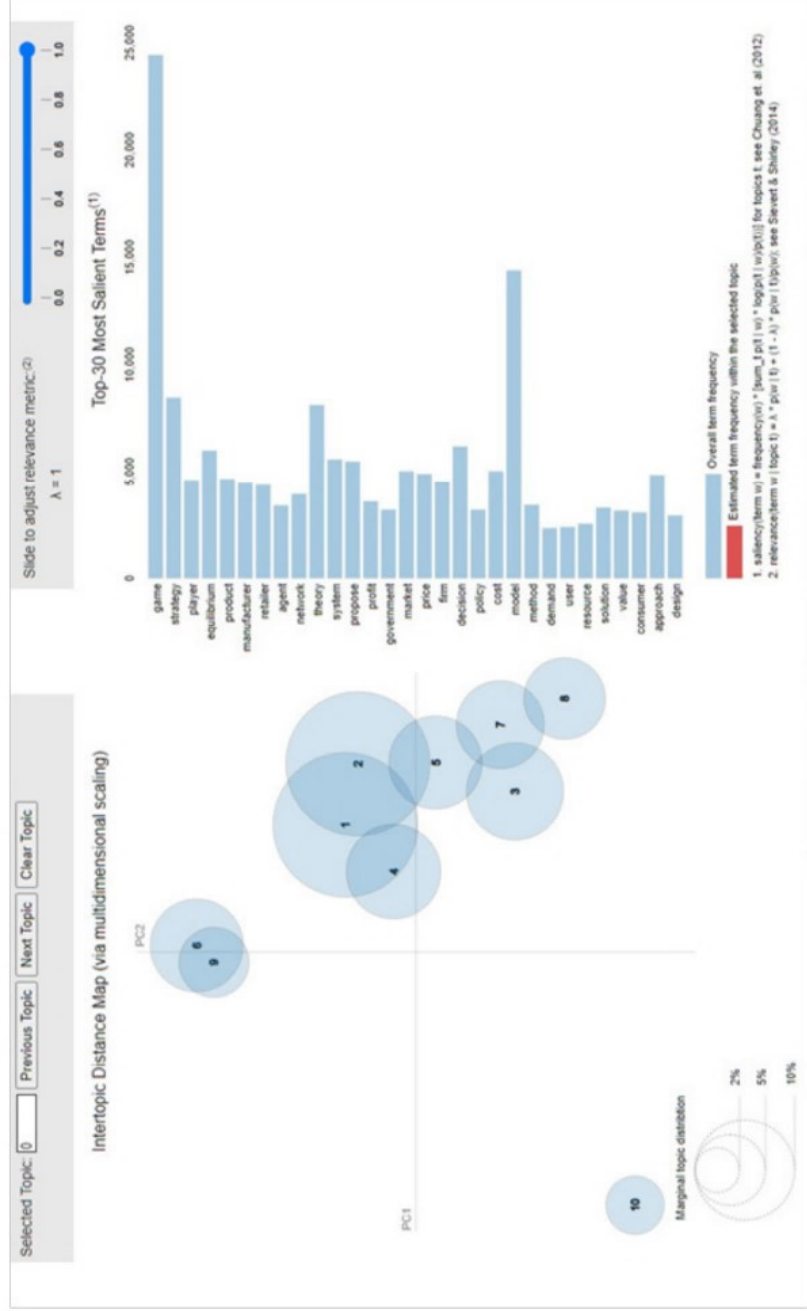| Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10
Game	Revenue	Game	Agent	Product	Game	Game	Network	Government	Model
Model	Innovation	Theory	Belief	Manufacturer	Cooperative	Player	Propose	Policy	Market
Theory	Regulation	Design	Communication	Retailer	Value	Equilibrium	System	Cooperation	Price
Decision	Provider	System	Build	Profit	Solution	Strategy	User	Finding	Firm
Approach	Participation	Approach	Group	Cost	Energy	Set	Resource	Incentive	Strategy
Risk	Urban	Method	Individual	Demand	Method	Payoff	Allocation	Subsidy	Consumer
Social	Land	Optimization	Trust	Decision	Model	Nash	Power	Relationship	Level
Process	Consumption	Application	Behavior	Channel	Algorithm	Dynamic	Scheme	Public	Increase
Behaviour	Heterogeneous	Present	Play	Contract	Member	Solution	Platform	Country	Information
Economic	Driver	Development	Path	Supplier	Capacity	Mode	Low	Reveal	Effect

Fig. 4.　Intertopic distance map of the base model

Table 4   Topics of the optimized model

Topic 1	Topic 2	Topic 3	Topic 4	Topic 5	Topic 6	Topic 7	Topic 8	Topic 9	Topic 10
Model	Cost	Theory	Game	Policy	Market	Game	Strategy	Network	Price
Decision	Cooperative	Social	Solution	Group	Firm	Design	Game	System	Product
Process	Mechanism	Economic	Set	Conflict	Information	Learn	Equilibrium	Propose	Retailer
Management	Cooperation	Individual	Function	Country	Increase	Knowledge	Player	User	Manufacturer
Risk	Game	Behaviour	Method	Level	Model	Present	Agent	Resource	Profit
Theory	Allocation	Approach	Introduce	Political	Strategic	Play	Model	Time	Demand
Enterprise	Share	Interaction	Solve	State	Effect	Experiment	Dynamic	Control	Consumer
System	Order	Explain	Player	Government	Investment	Test	Payoff	Performance	Quality
Government	Benefit	Make	Present	Public	Competition	Subject	Behaviour	Approach	Contract
Develop	Distribution	Trust	Concept	International	Technology	Application	Time	Power	Supplier

Selected Topic: 0   Previous Topic   Next Topic   Clear Topic

Intertopic Distance Map (via multidimensional scaling)

Slide to adjust relevance metric λ = 1

Top-30 Most Salient Terms

Overall term frequency
Estimated term frequency within the selected topic

1. saliency(term w) = frequency(w) × [sum_t p(t | w) × log(p(t | w)/p(t))] for topics t; see Chuang et. al (2012)
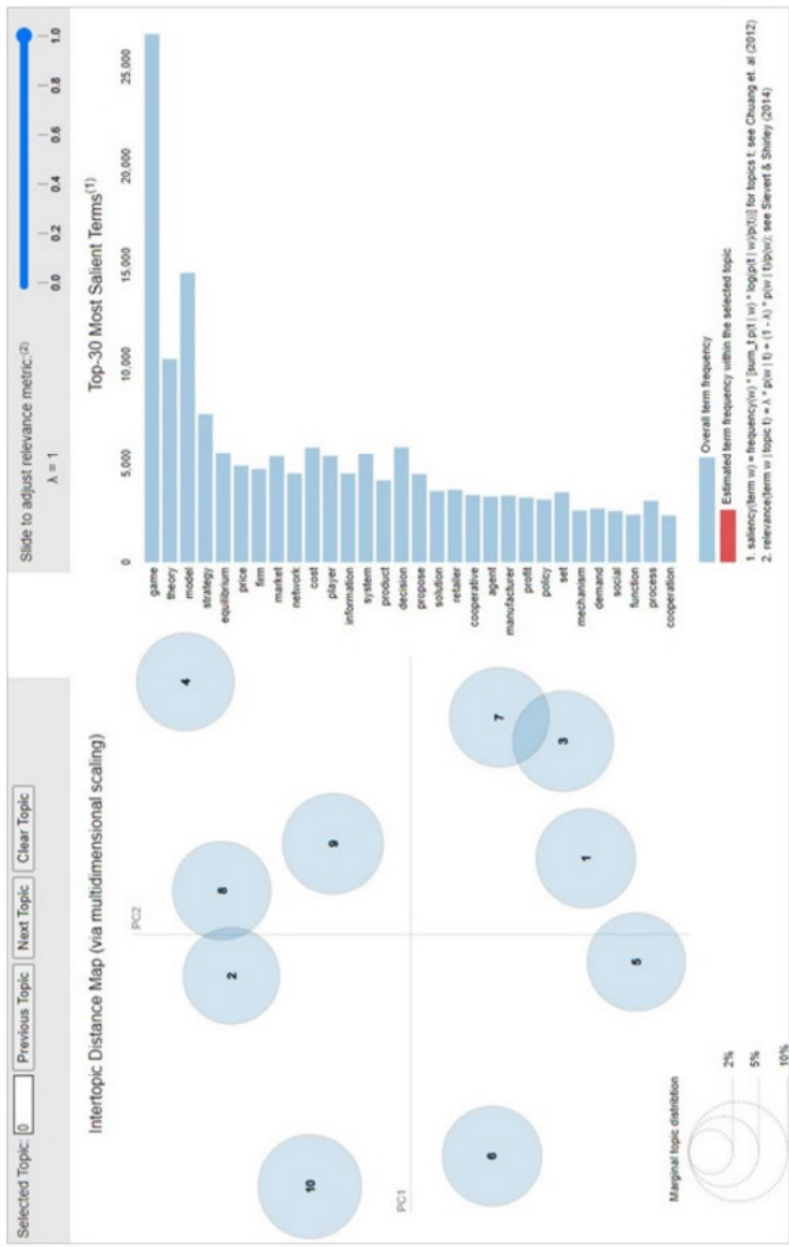2. relevance(term w | topic t) = λ × p(w | t) + (1 − λ) × p(w | t)/p(w); see Sievert & Shirley (2014)

Fig. 5.   Intertopic distance map of the optimized model

Table 5   Topic labels of base and optimized model

		Top 10 keywords									Topic Label
Base Model											
Topic 1	Game	Model	Theory	Decision	Approach	Risk	Social	Process	Behaviour	Economic	Key terminologies of Game Theory
Topic 2	Revenue	Innovation	Regulation	Provider	Participation	Urban	Land	Consumption	Heterogeneous	Driver	Application in regulations
Topic 3	Game	Theory	Design	System	Approach	Method	Optimization	Application	Present	Development	Related terminologies in Game Theory
Topic 4	Agent	Belief	Communication	Build	Group	Individual	Method	Behaviour	Play	Path	Incoherent*
Topic 5	Product	Manufacturer	Retailer	Profit	Cost	Demand	State	Channel	Contract	Supplier	Market Games
Topic 6	Game	Cooperative	Value	Solution	Energy	Method	Model	Algorithm	Member	Capacity	Incoherent*
Topic 7	Game	Player	Equilibrium	Strategy	Set	Decision	Nash	Dynamic	Solution	Mode	Solution concept of Nash Equilibrium
Topic 8	Network	Propose	System	User	Resource	Payoff	Power	Scheme	Platform	Low	Allocation Game
Topic 9	Government	Policy	Cooperation	Finding	Incentive	Allocation	Relationship	Public	Country	Reveal	Incoherent*
Topic 10	Model	Market	Price	Firm	Strategy	Consumer	Subsidy	Level	Increase	Effect	Related terminologies in Game Theory
Optimized Model											
Topic 1	Model	Decision	Process	Management	Risk	Theory	Enterprise	System	Government	Develop	Managerial decision making
Topic 2	Cost	Cooperative	Mechanism	Cooperation	Game	Allocation	Share	Order	Benefit	Distribution	Allocation Games
Topic 3	Theory	Social	Economic	Individual	Behaviour	Approach	Interaction	Explain	Make	Trust	Game theory to study social Behaviour
Topic 4	Game	Solution	Set	Function	Method	Introduce	Solve	Player	Present	Concept	Concept of Game Theory
Topic 5	Policy	Group	Conflict	Country	Level	Political	State	Government	Public	International	International Policy
Topic 6	Market	Firm	Information	Increase	Model	Strategic	Effect	Investment	Competition	Technology	Market Competition
Topic 7	Game	Design	Learn	Knowledge	Present	Play	Experiment	Test	Subject	Application	Experimental game design
Topic 8	Strategy	Game	Equilibrium	Player	Agent	Model	Dynamic	Payoff	Behaviour	Time	Dynamic games
Topic 9	Network	System	Propose	User	Resource	Time	Control	Performance	Approach	Power	Network Games
Topic 10	Price	Product	Retailer	Manufacturer	Profit	Demand	Consumer	Quality	Contract	Supplier	Market Games

REFERENCES

[1] Blei, D. 2011. Probabilistic Topic Models.

[2] Blei, D., and Lafferty, J. 2007. A correlated topic model of Science. The Annals of Applied Statistics, 1

[3] Blei, D., Ng, A., and Jordan, M. 2001. Latent Dirichlet Allocation.

[4] Blei, D., Ng, A., Jordan, M., and Lafferty, J. 2003. Journal of Machine Learning Research 3 (2003) 993–1022 Submitted 2/02; Published 1/03 Latent Dirichlet Allocation.

[5] Boye Asmussen, C., and Møller, C. 2019. Smart literature review: a practical topic modelling approach to exploratory literature review. Journal of Big Data, 6.

[6] Chang, J.; Boyd-Graber, J. L.; Gerrish, S.; Wang, C. & Blei, D. M. (2009), Reading Tea Leaves: How Humans Interpret Topic Models., in 'Neural Information Processing Systems', pp. 288–296 .

[7] DiMaggio, P., Nag, M., and Blei, D. 2013. Exploiting Affinities Between Topic Modeling and the Sociological Perspective on Culture: Application to Newspaper Coverage of U.S. Government Arts Funding. Poetics, 41, p. 570–606.

[8] Grimmer, J. 2010. A Bayesian Hierarchical Topic Model for Political Texts: Measuring Expressed Agendas in Senate Press Release. Political Analysis, 18.

[9] Jockers, M., and Mimno, D. 2013. Significant Themes in 19th-Century Literature. Poetics, 41, p. 750–769.

[10] Levy, Y., and Ellis, T. 2006. A Systems Approach to Conduct an Effective Literature Review in Support of Information Systems Research. International Journal of an Emerging Transdiscipline, 9.

[11] Lu, Y., Mei, Q., and Zhai, C. 2011. Investigating task performance of probabilistic topic models: An empirical study of PLSA and LDA. Inf. Retr., 14, p. 178–203.

[12] M. Girolami, and A. Kabán 2003. On an equivalence between PLSI and LDA. In SIGIR '03.

[13] McCallum, A. 2002. MALLET: A machine learning for language toolkit.

[14] Quinn, K., Monroe, B., Colaresi, M., Crespin, M., and Radev, D. 2010. How to Analyze Political Text With Minimal Assumptions and Costs. American Journal of Political Science, 54, p. 209–228.

[15] Řehůřek, R., and Sojka, P. 2010. Software Framework for TopicModelling with Large Corpora.

[16] Röder, M., Both, A., and Hinneburg, A. 2015. Exploring the Space of Topic Coherence Measures. WSDM 2015 - Proceedings of the 8th ACM International Conference on Web Search and Data Mining, p. 399–408.

[17] Sievert, C., and Shirley, K. 2014. LDAvis: A method for visualizing and interpreting topics.

[18] von Neumann, J., and Morgenstern, O. 1944. The Theory of Games and Economic Behavior, Princeton: Princeton University Press.

[19] X, Quan., C, Kit., Y, Ge., and S, J Pan. 2015. Short and sparse text topic modeling via self-aggregation. Twenty-Fourth International Joint Conference on Artificial Intelligence

[20] Y, Sun., K, Loparo., and R, Kolacinski., 2020. Conversational Structure Aware and Context Sensitive Topic Model for Online Discussions. IEEE 14th International Conference on Semantic Computing (ICSC), San Diego, CA, USA, 2020, pp. 85–92, doi: 10.1109/ICSC.2020.00019.