International Journal of Internet, Broadcasting and Communication Vol.12 No.4 204-211 (2020)
http://dx.doi.org/10.7236/IJIBC.2020.12.4.204

IJIBC 20-4-25

A Study on Socio-technical System for Sustainability of the 4th Industrial Revolution: Machine Learning-based Analysis

Jee Young Lee

Senior Researcher, Graduate School of Information, Yonsei University, Korea
J.Ann.Lee@yonsei.ac.kr

Abstract

The era of the 4th industrial revolution is a complex environment in which the cyber world and the physical world are integrated and interacted. In order to successfully implement and be sustainable the 4th industrial revolution of hyper-connectivity, hyper-convergence, and hyper-intelligence, not only the technological aspects that implemented digitalization but also the social aspects must be recognized and dealt with as important. There are socio-technical systems and socio-technical systems theory as concepts that describe systems involving complex interactions between the environmental aspects of human, mechanical and tissue systems. This study confirmed how the Socio-technical System was applied in the research literature for the last 10 years through machine learning-based analysis. Eight clusters were derived by performing co-occurrence keywords network analysis, and 13 research topics were derived and analyzed by performing a structural topic model. This study provides consensus and insight on the social and technological perspectives necessary for the sustainability of the 4th industrial revolution.

Keywords: *Socio-technical System, The fourth industrial revolution, Co-occurrence keywords networks, Structural topic model*

1. Introduction

The era of the 4th industrial revolution is the era of hyper-connection, hyper-convergence, and hyper-intelligence. In other words, it is a society where the cyber world and the physical world are integrated to realize automation and intelligence [1, 2]. Therefore, technology in the cyber world and objects, humans, and society in the physical world are integrated and interacted [3]. In order for the era of the Fourth Industrial Revolution to be successfully implement and sustainable, not only the technological aspects that implement digitalization but also the social aspects must be recognized and dealt with as important. Among the 4th industrial revolution, there was a general consensus that Industry 4.0, which focuses on the manufacturing sector, should consider both technological and social factors in order to be sustainable [4]. In the field of complexity science, the concept of socio-technical systems is introduced to solve the problem of complexity from a technical and social perspective [5]. Socio-technical systems describe systems that involve complex interactions between human, mechanical and environmental aspects of tissue systems. The

Manuscript Received: October. 2, 2020 / Revised: October. 8, 2020 / Accepted: October. 11, 2020
Corresponding Author: J.Ann.Lee@yonsei.ac.kr
Tel: +82-2-2123-4187, Fax: +82-2-2123-4180
Senior Researcher, Graduate School of Information, Yonsei University, Korea

socio-technical systems theory is a method of optimizing system design that can cope with environmental complexity, dynamism, new technology and competition, taking into account the characteristics of open systems for various stakeholders [6]. Socio-technical systems theory has been successful in the design of new technologies and in design-driven change [7, 8].

Recognition and understanding of the Socio-technical System is essential for the successful implementation and sustainability of the 4th Industrial Revolution. However, there have been few such studies yet. Accordingly, this study attempted to confirm how the Socio-technical System was handled in the research literature for the last 10 years. To this end, we collected research literature from SCOPUS and analyzed keywords that appeared simultaneously with the Socio-technical System. And the research topics covered in the literature related to the Socio-technical System were analyzed by applying a machine learning-based text mining method. The results of this study provide consensus and insight on social and technological perspectives necessary for sustainability in the era of the 4th industrial revolution.

2. Research methods

In order to find out the Socio-technical System dealt with in the research literature, this study uses "Socio-technical System" as a search word for journal articles and conference papers, and 2,469 research documents written in English until October 25, 2020 were collected in Scopus. The purpose of this study is to identify the social and technological perspectives necessary for system design in the era of the 4th industrial revolution. Therefore, 1,890 items, filtered from 2010 onwards, were used as the research data.

2.1. Co-occurrence keywords network analysis

Co-occurrence keywords analysis was performed to identify keywords related to the socio-technical system perspective. First, create a keyword co-occurrence matrix. Next, a similarity matrix is obtained by normalizing the keyword co-occurrence matrix using the association strength [9]. Keywords clustering is performed based on the similarity matrix. This study used the VOSviewer tool as a tool to perform and visualize co-occurrence keywords network analysis [10]. We set the minimum number of occurrences of a keyword to 10 in the parameter setting of VOSviewer. The minimum cluster size was set to 5 [11].

2.2. Analysis of research topics related to socio-technical system

In this study, topic modeling was performed among text mining based on machine learning in order to analyze studies on the Socio-technical System in terms of topics. This study used structural topic model (STM), the latest algorithm among topic modeling algorithms. STM is a modified and extended model of latent Dirichlet allocation (LDA), the most widely known topic modeling method [12]. As shown in Figure 1, the process of STM is conceptually composed of three components; (1) a topic prevalence model, which controls how words are allocated to topics as a function of covariates, (2) a topical content model, which controls the frequency of the terms in each topic as a function of covariates, and (3) a core language model, which combines these two sources of variation to produce the actual words in each document.

In this study, STM was performed after text preprocessing (converting to lower case, tokenization, removing punctuation, removing stopwords, removing numbers, stemming) for the abstract of collected research literature. R package was used to implement STM [13].

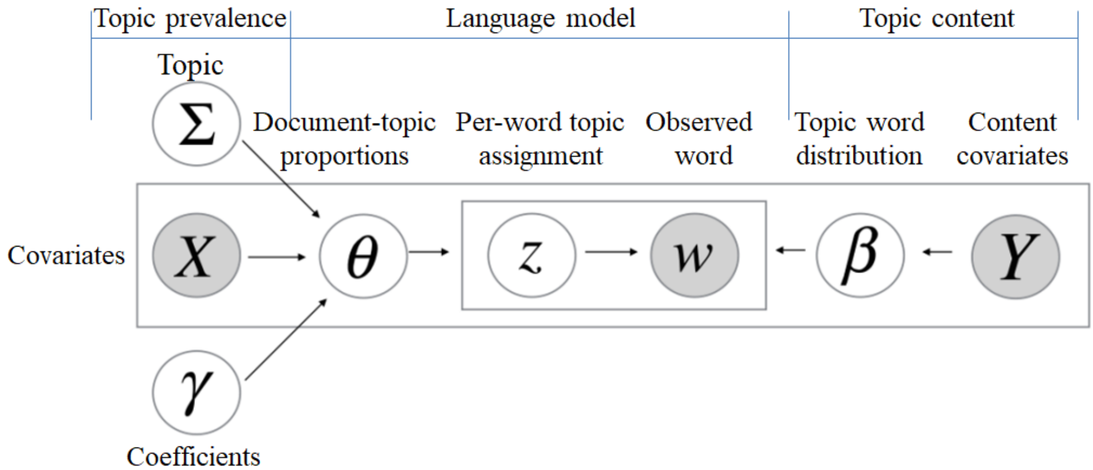

Figure 1. Conceptual structure and process of the STM

3. Results and discussion

3.1. Results and discussion on co-occurrence keywords network analysis

As shown in Figure 2, a co-occurrence network map and 5 clusters were derived. The co-occurrence network map was derived from 5,148 keywords. There were 61 keywords with at least 10 simultaneous occurrences, and 5 clusters were identified. Table 1 summarizes the top 5 keywords based on 5 clusters and weight of links. Cluster 1, shown by the red nodes in the lower right of Figure 2, consists of 14 keywords, including sustainability, governance, innovation, infrastructure, and transition. This cluster was summed up as 'Transition for sustainability'. Cluster 2, represented by green nodes in the upper left, has 13 keywords such as human factors, security, privacy, internet of things (IoT), and technology. This cluster was interpreted as being related to 'IoT'. Cluster 3, shown as blue nodes across the center and upper right area, consists of 13 keywords such as socio-technical systems, resilience, design, information systems, and knowledge management. This cluster was interpreted as 'Information systems design for resilience'. Cluster 4, shown as yellow nodes in the lower center, has 11 keywords such as systems engineering, requirements engineering, modeling, agent-based modeling, and simulation. This cluster was summarized as 'Systems engineering modeling'. Cluster 5, represented by purple nodes in the lower left corner, consists of 10 keywords such as complex systems, complexity, risk, resilience engineering, and safety. This cluster was interpreted as 'Complex systems'.

Through co-occurrence network analysis, it was confirmed that the STS concept was applied not only to the complex systems research area, but also to sustainability, IoT, information systems design for resilience, and systems engineering modeling. This demonstrates the use of the STS concept in system design for complexity, sustainability, and resilience in the Fourth Industrial Revolution represented by IoT.

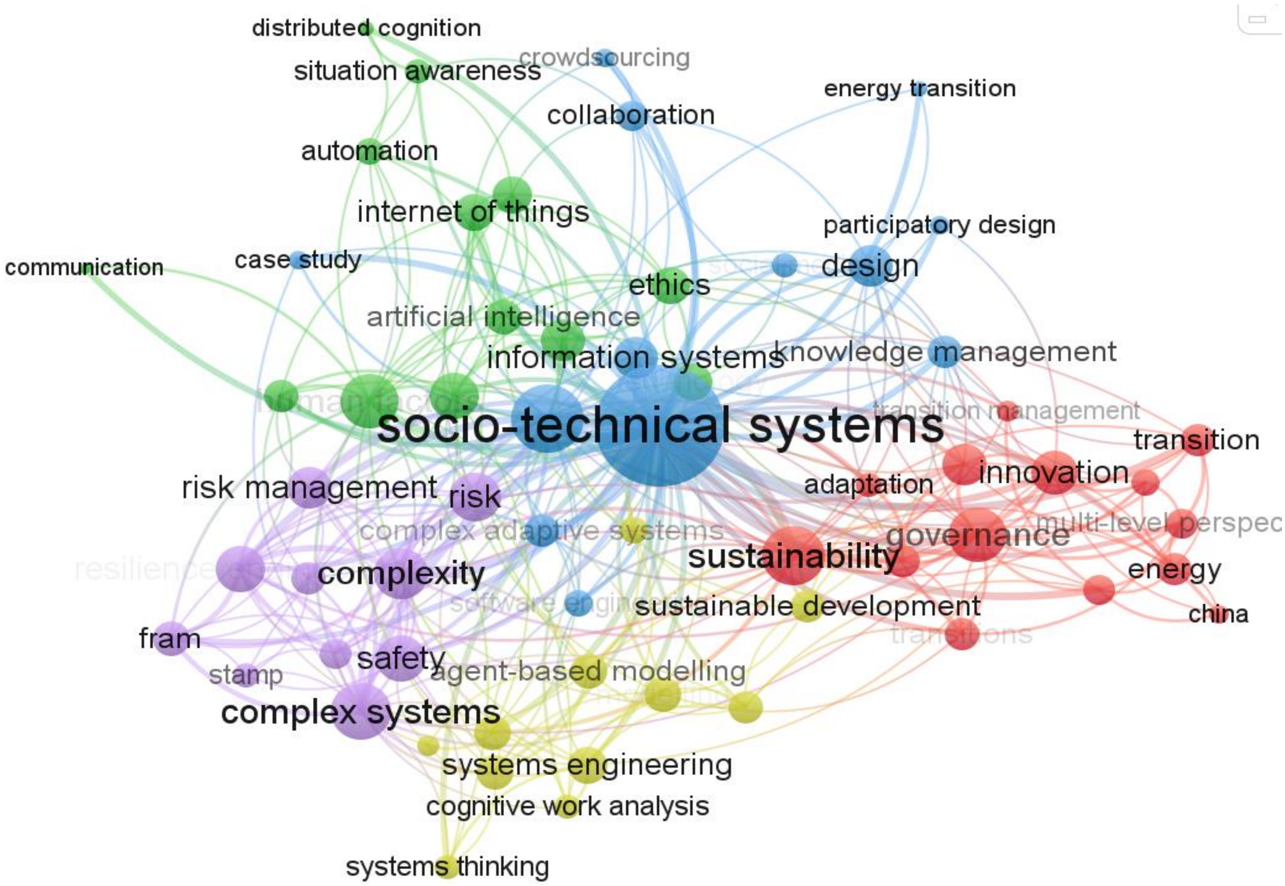

Figure 2. Keyword co-occurrence network map using VOSviewer

Table 2. Keyword co-occurrence based cluster extraction and top keywords

Cluster	Top 5 keywords based on weight of links				
1. Transition for sustainability	sustainability	governance	innovation	infrastructure	transition
2. IoT	human factors	security	privacy	IoT	technology
3. Information systems design for resilience	socio-technical systems	resilience	design	information systems	knowledge management
4. Systems engineering modeling	systems engineering	requirements engineering	modelling	agent-based modelling	simulation
5. Complex systems	complex systems	complexity	risk	resilience engineering	safety

Figure 3 is an overlay network showing the simultaneous occurrence keywords superimposed with the publication year. You can see the evolution of research in terms of co-occurring keywords over the last 10 years. In early 2010, many keywords related to distributed systems such as information systems, situation awareness, communication, transition, and distributed awareness appeared. In the mid-2010s, keywords such as collaboration, crowdsourcing, participatory design, sustainability, complex systems, and safety increased.

This period can be inferred that there were many studies on sustainability in a complex systems with various participants.

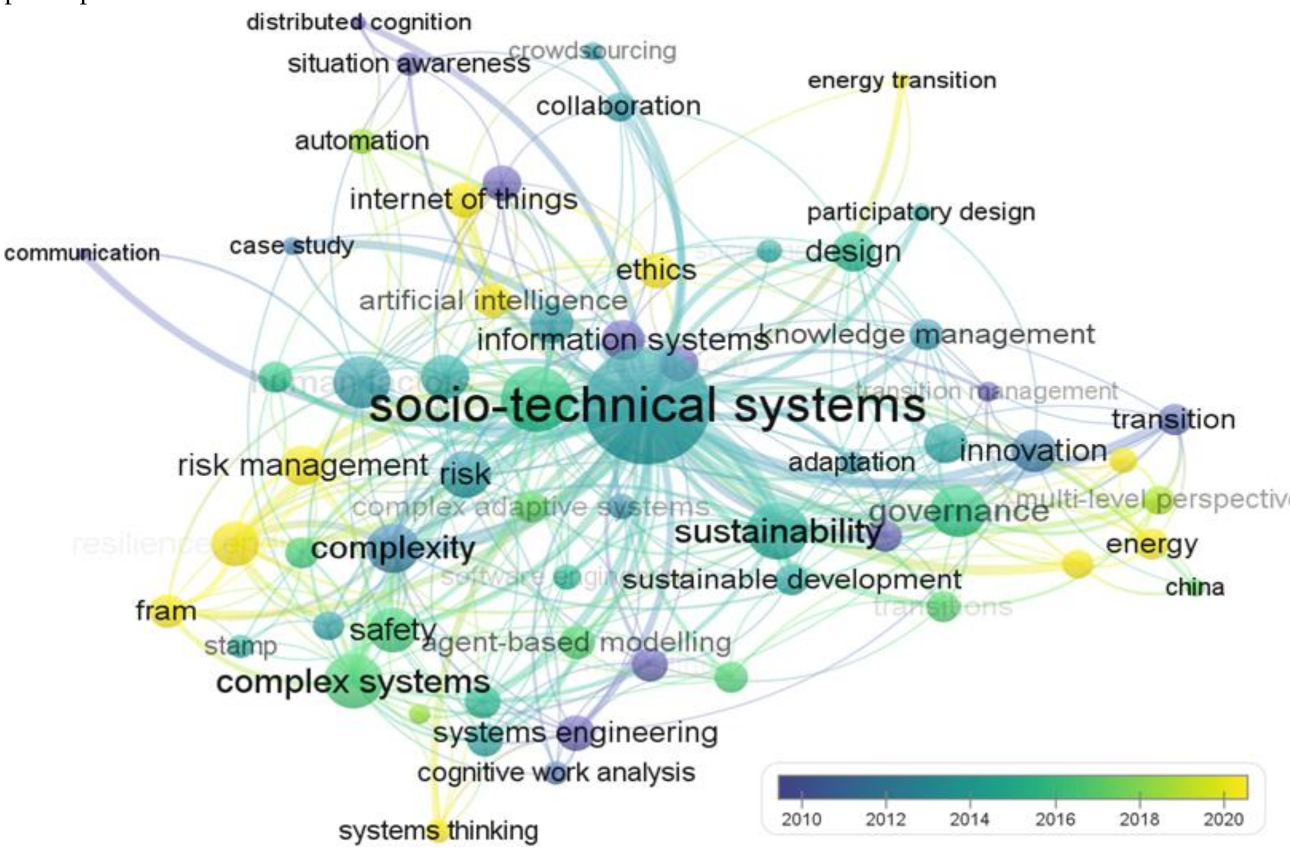

Figure 3. Keyword co-occurrence network map using VOSviewer

3.2. Results and discussion on analysis of research topics related to socio-technical system

To determine the optimal number of topics K to be used for STM, we gradually increased the number of topics from 5 to 20 and identified four indicators: held-out likelihood, residual, semantic coherence, and lower bound. As a result, the optimal number of topics was determined to be 13. Table1 summarizes the five high-level words for each topic derived from the performance of STM. The label of the topic was determined by analyzing the top words by topic and 5 documents with high weight.

Table 2. STM based topic extraction and top words

Topic	Top words					Topic label
1	design	technology	value	ethic	practice	Computing ethic
2	energy	water	efficience	resource	consumption	Energy saving
3	data	open	ecosystem	share	decentration	Computing evolution
4	human	society	complex	machine	ann[1]	Complex systems

[1] ANN : Artificial Neural Networks

5	work	team	cognition	group	rbps[2]	Innovation management
6	safety	risk	system	operation	accident	Safety engineering
7	transition	sustainability	policy	change	climate	Sustainability transitions
8	social	interact	community	network	behavior	Collective behavior
9	require	engine	secure	model	software	Security requirements engineering
10	practice	implement	lean	service	paradigm	Lean service
11	complex	resilience	simulation	fram[3]	functional	Resilience quantification
12	energy	smart	electron	power	grid	Smart grid
13	product	innovation	knowledge	industry	holacracy	Co-evolution

Topic 1 (Computing ethic) was a study to apply ethic to computing such as technology and design process. Computing should also have a sense of social responsibility, and these are studies that approached the social impact of computing from a socio-technical perspective. Topic 2 (Energy saving) was about applying the socio-technical system from a broad and systematic perspective to analyze the institutional, cultural, and behavioral reasons for the inconsistency of energy saving. Topic 3 (Computing evolution) is a study on evolution such as pervasive computing, open system, and cloud ecosystem, and introduced a socio-technical system for the interaction of all physical entities. Topic 4 (Complex system) was to apply the socio-technical system to the study of the characteristics and problem solving methods of complex systems of interconnected cyber physical systems. Topic 5 (Innovation management) was a study on responding to environmental changes and managing enterprise-wide innovation. Topic 6 (Safety Engineering) was a study on risk assessment and management in interconnected systems. Socio-technical system theory was applied to analyze deterministic and nondeterministic causal relationships for risk assessment. Topic 7 (Sustainability transitions) was a study on the transitions of sustainability due to the mixture and change of various policies. Topic 8 (Collective behavior) was about how people connected with each other interact through social media, social networks, and online communities, and what collective behavior patterns appear. Topic 9 (Security requirements engineering) was a study on how to derive, analyze, and manage requirements for security in an interaction network between actors. Topic 10 (Lean service) was a study on implementing lean and understanding lean as a paradigm of a Socio-technical system to establish correlations and impacts on organizational performance. In Topic 11 (Resilience quantification), securing and managing Resilience for socio-technical systems in complex systems can lead to system innovation. For this, it was a study on a method to quantify resilience. Topic 12 (Smart grid) was a study that applied a socio-technical perspective to investigate the expectations of actors and control uncertainty in the smart grid, one of the important complex systems. Finally, Topic 13 (co-evolution) was about co-evolution in a complex interconnected system. Adopted to respond to rapid change, the agile paradigm promotes knowledge management and collaboration, helping to develop product innovation and an organization's innovative culture. This was a study that introduced and used the socio-technical system as an analytical approach to co-evolution and co-creation of value due to changes in manpower and organization.

---

[2] RBPS : Responsible Boundary Production System

[3] FRAM : Functional Resonance Analysis Method

3.3. STM based topic trend estimation

To estimate the trend of the topic, a regression where topic-proportions are the outcome variable was performed. As shown in Figure 4, the hot topics for which research is increasing recently were analyzed as Topic 5 (Innovation management) and Topic 8 (Collective behavior).

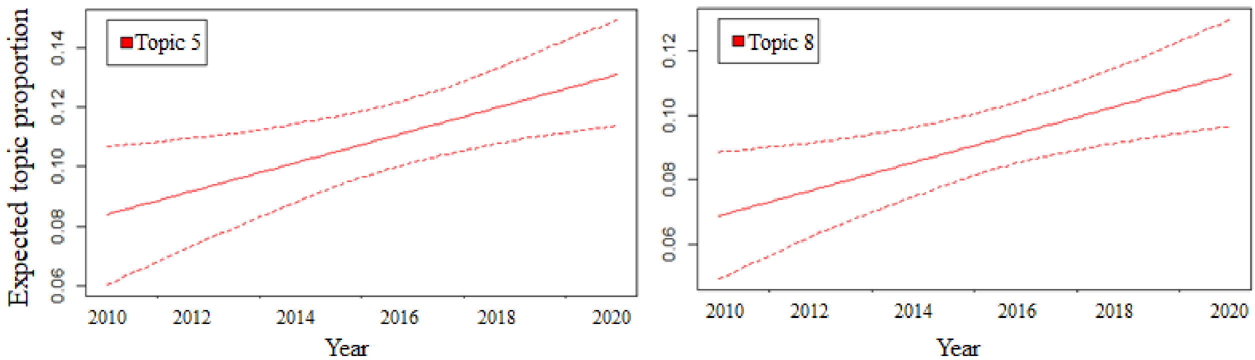

Figure 4. Topic trend estimation over year

4. Conclusion

In order for the era of the 4th industrial revolution to be successfully realized and sustainable, not only the technological aspects that implemented digitalization, but also the social aspects must be recognized and dealt with as important. This study confirmed the necessity of a socio-technical system that should be introduced from a research perspective for the successful realization and sustainability of the 4th industrial revolution. In addition, research literature and research topics related to the socio-technical system over the past 10 years were checked.

The results of this study are as follows.

- Identification of the field of introduction of the STS concept: It was confirmed that the STS concept was applied not only to the complex systems research field, but also to sustainability, IoT, information systems design for resilience, and systems engineering modeling.
- Checking the evolution of research involving STS: The evolution of research can be confirmed in terms of co-occurring keywords over the last 10 years.
- Classification of research topics including STS: Research topics were derived and classified through STM-based topic analysis.
- Research trend confirmation: Hot-topic and cold-topic were identified through topic trend estimation.

This study was analyzed by performing machine learning-based text mining on research literature. The limitation of this study is that specific case studies on the research field derived from topic analysis were not conducted. This is a future study and is planned to be carried out as a multidisciplinary research method. Despite these limitations, the contribution of this study is as follows.

- Consensus on the necessity of STS was sought in designing Internet broadcasting and communication services in the era of the 4th industrial revolution.
- Research conducted by introducing STS was confirmed in terms of topics, and this will provide a reference that researchers can use in research design.
- Research literature that introduced STS showed the evolution of keywords and research topics, which will help determine future research projects.
- The work of deriving and classifying research keywords and research topics using the machine learning-based method performed in this study can be applied to other researches.

References

[1] J. Lee, "A Study on Research Trend Analysis and Topic Class Prediction of Digital Transformation using Text Mining," *International journal of advanced smart convergence*, vol. 8, no. 2, pp. 183-190, 2019. DOI: https://doi.org/10.7236/IJASC.2019.8.2.183

[2] J. Y. Lee, "A Study on Agile Transformation in the New Digital Age," *International Journal of Advanced Culture Technology*, vol. 8, no. 1, pp. 82-88, 2020 DOI: https://doi.org/10.17703/IJACT.2020.8.1.82

[3] L. D. Xu, E. L. Xu, L. Li, "Industry 4.0: state of the art and future trends," *International Journal of Production Research*, vol. 56, no. 8, pp. 2941-2962, 2018. DOI: https://doi.org/10.1080/00207543.2018.1444806

[4] R. Davies, T. Coole, A. Smith, "Review of socio-technical considerations to ensure successful implementation of Industry 4.0," in *7th International Conference on Flexible Automation and Intelligent Manufacturing, FAIM2017*, pp. 1288-1295, June 27-30, 2017. DOI: https://doi.org/10.1016/j.promfg.2017.07.256

[5] N. Zhao, J. Bao, N. Chen, "Ranking Influential Nodes in Complex Networks with Information Entropy Method," *Complexity*, vol. 2020, June 2020. DOI: https://doi.org/10.1155/2020/5903798

[6] G. H. Walker, N. A. Stanton, P. M. Salmon, et al., "A review of sociotechnical systems theory: a classic concept for new command and control paradigms," *Theoretical issues in ergonomics science*, vol. 9, no. 6, pp. 479-499, 2018. DOI: https://doi.org/10.1080/14639220701635470

[7] K. Eason, "Local sociotechnical system development in the NHS National Programme for Information Technology," *Journal of Information Technology*, vol. 22, no. 3, pp. 257-264, 2007. DOI: https://doi.org/10.1057/palgrave.jit.2000101

[8] G. Baxter, I. Sommerville, "Socio-technical systems: From design methods to systems engineering," *Interacting with computers*, vol. 23, no. 1, pp. 4-17, 2011. DOI: https://doi.org/10.1016/j.intcom.2010.07.003

[9] N. J. v. Eck, L. Waltman, "How to normalize cooccurrence data? An analysis of some well-known similarity measures," *Journal of the American Society for Information Science and Technology*, vol. 60, no. 8, pp. 1635-1651, 2009. DOI: https://doi.org/10.1002/asi.21075

[10] N. J. Van Eck, L. Waltman, "Software survey: VOSviewer, a computer program for bibliometric mapping," *Scientometrics*, vol. 84, no. 2, pp. 523-538, 2010. DOI: https://doi.org/10.1007/s11192-009-0146-3

[11] N. J. Van Eck, L. Waltman, "Citation-based clustering of publications using CitNetExplorer and VOSviewer," *Scientometrics*, vol. 111, no. 2, pp. 1053-1070, 2017. DOI: https://doi.org/10.1007/s11192-017-2300-7

[12] M. E. Roberts, B. M. Stewart, D. Tingley, et al., "The structural topic model and applied social science," in *Advances in neural information processing systems workshop on topic models: computation, application, and evaluation*, Harrahs and Harveys, Lake Tahoe, 2013.

[13] M. E. Roberts, B. M. Stewart, D. Tingley, "stm: R package for structural topic models," *Journal of Statistical Software*, vol. 10, no. 2, pp. 1-40, 2014. DOI: https://doi.org/10.18637/jss.v091.i02