GROUP BRAINSTORMING SUPPORT BY CHATGPT & AYOA AT THE DESIGN OF REGIONAL DEVELOPMENT PLAN

FRANC LAVRIČ, ANDREJ ŠKRABA

University of Maribor, Faculty of Organizational Sciences, Cybernetics & Decision Support Systems Laboratory, Kidričeva cesta 55a, Kranj, Slovenia
franc.lavric@student.um.si, andrej.skraba@um.si

Abstract The possibility to apply modern AI tools such as ChatGPT and Ayoa at the design of regional development plan of Gorenjska region is considered. With both systems 496 ideas were generated in the scope of human resources development. Generated ideas were categorized using Latent Dirichlet Allocation, Non-negative Matrix Factorization, human experts and ChatGPT. The entropy of the generated ideas was determined in order to compare ChatGPT and Ayoa. The hypothesis of diversity was confirmed by Mann-Whitney U test. In order to analyze the categories the word cloud was generated with programming language python for all four generative sessions. Python's pyLDA tool was applied for drill down analysis of indicated categories. The methodology has been proposed to integrate the human and AI actors to enhance brainstorming and decision-making processes.

Keywords: ChatGPT, brainstorming, ideation, categorization, regional development planning, NLP

University of Maribor Press

DOI https://doi.org/10.18690/um.fov.3.2023.43
ISBN 978-961-286-722-5

1 Introduction

The first phase of the design of the regional development plan for Gorenjska region is traditionally supported by the initial brainstorming sessions with expert groups (Škraba & Filipič 2009). Since new technologies have emerged, it is our intention to examine the possibility to augment brainstorming sessions with AI tools such as ChatGPT (Radford, 2019), Ayoa (Ayoa 2023) and python NLP libraries. Here we consider the generation of ideas as well as ranking. Our case study will target human resources development in Gorenjska region.

The ChatGPT technology has been successfully applied for the purpose of brainstorming in a bid to get better insight when faced with nebulous and diagnostically challenging cases (Kung et al., 2022). This is quite remarkable since medicine (Jeblick et al., 2022) is one of the most challenging areas of human edavor as well as stem cell research (Cahan & Treutlein, 2023). There have been significant concerns about the future after ChatGPT (Castelvecchi, 2022). One could predict the development of human endeavors similarly to the introduction of robots. Today no one complains that there are too many robots, for example in the car industry. Behind each robot, there are numerous new jobs which provide opportunities for entrepreneurship. With some fortune, the development, the positive effect might be similar. By analyzing the 10,732 tweets from early ChatGPT users (Haque et al., 2022) the in-depth qualitative sentiment analysis was performed. The results (Haque et al., 2022) show that most of the early adopters have expressed positive sentiments related to topics such as "Disruptions to software development, Entertainment and exercising creativity". It has been shown that only a limited percentage of users expressed concerns about issues such as the potential for misuse of ChatGPT (Grimaldi, 2023), especially in the education process (Susnjak, 2022).

2 Methodology

Figure 1 shows the structure of the system for electronic brainstorming. Here the expert group of subjects marked with s_1, s_2, … s_n is supported by the tools of Artificial Intelligence (AI), (Lavrič et al., 2022). In our case we have applied ChatGPT (Radford, 2019) and Ayoa (Ayoa, 2023) for generating ideas. ChatGPT could also be applied at the categorization as well as evaluation phase. ChatGPT

(Radford, 2019) or Chat Generative Pre-Trained Transformer, is a chatbot launched by OpenAI in November 2022.

Ayoa is a mind map app used powered by methods of Artificial Intelligence. Ayoa enables innovative brainstorming methods by providing the possibility to generate ideas on particular topic. We have successfully applied Ayoa for generating the innovative ideas for the design of regional development plan of Gorenjska region.

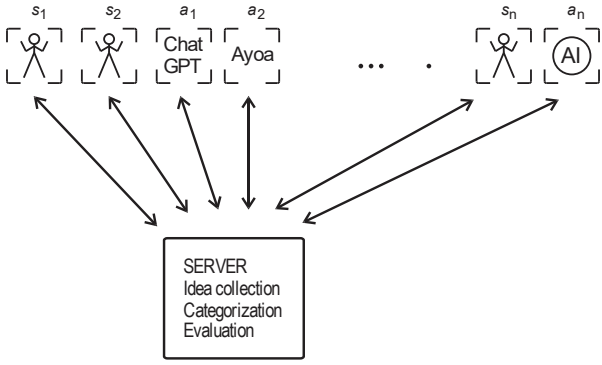

Figure 1: Electronic brainstorming with participants aided by the tools of Artificial Intelligence. Here we have considered two tools, ChatGPT and Ayoa. One could apply more tools as they emerge on the market.

It is important that there are diverse tools available, which might provide better results, as we have indicated in our previous research (Kološvari et al., 2019). Here the algorithms for ranking and integration of the results will be of crucial importance (Škraba et al., 2018). Therefore, we will test hypothesis **H1**: "*There is significant difference in generated ideas' entropy between ChatGPT and Ayoa*". It is important to provide a proper environment for different, independent AI tools which would provide significant improvement in performance as indicated in Kološvari et al., 2019. Detailed description of the methodology, that is behind ChatGPT is described in (Redford et al., 2019). At AI tools, one should be aware of the importance of user feedback, which contributes to the improvement of the system performance. The principle could best be described as the feedback in neural networks (Kofjač et al., 2003) or within decision group (Škraba et al., 2003). This simple click might not seem to be important, but one could imagine the information flow when this is multiplied by billions of users and use cases. The entropy of the generated ideas by ChatGPT and Ayoa will be determined according to the Shannon (1948):

$$H(\alpha) = -\sum_{x \in \alpha} p(x) \log p(x)$$

where α represents the particular idea generated by AI system an $p(x)$ the probability for the occurrence of particular symbol in the idea string set.

For the categorization of generated ideas, the method of Latent Dirichlet Allocation (LDA) (Blei et al., 2001, 2003) was applied with the help of the python libraries (Sievert et al., 2016; Mabey, 2021) as well as Non-negative Matrix Factorization (NMF) (Lee & Seung, 1999) implemented in python (Portilla, 2023). Here we will identify proper categories. The proposed categories will then be properly named by experts. The number of categories will be set to seven (7). This should be the user defined input. After the performed analysis, interpretation of particular categories will be user defined. We will apply Non-negative Matrix Factorization at the categorization together with Term Frequency - Inverse Document Frequency (TF-IDF) algorithms that use the frequency of words to determine how relevant those words are to a particular category. NMF is an unsupervised algorithm that performs dimensionality reduction and clustering. As the base, the document-term matrix (DTM) will be applied. DTM is a matrix that describes the frequency of terms that occur in one, generated idea. Here rows correspond to ideas in the generated idea set and columns correspond to terms. Latent Dirichlet Allocation will also be applied at the determination of categories (Blei et al., 2001, 2003).

In the first step, the vector space model for the ideas will be generated. Here we will also perform stopword filtering resulting in DTM matrix **A**. On matrix **A** the TF-IDF term weight normalization will be performed. Factors will be initialized by non-negative double singular value decomposition (NNDSV). Projected gradient NMF will be applied on matrix **A**. Basis vectors will provide the categories of the generated ideas. Coefficient matrix will provide the category membership weights for ideas (clustering).

2.1 Ranking capabilities of ChatGPT

The ChatGPT system addresses also the issue of criteria. Here it proposes three different criteria by which the ideas might be ranked:

"*It is difficult to provide a definitive ranking of these ideas as the importance of each one will* ***depend on*** *a variety of factors, including the* ***current state of technology and innovation in Slovenia****, the* ***resources available to implement the ideas****, and the* ***specific goals and priorities of the government or organization in question****. However, here is a possible ranking based on the general importance of each idea: …*"

3 Results

The ChatGPT has been entered with the following query: "Please provide 31 ideas on human resources development in Gorenjska according to the principles of brainstorming." and "Please provide another 31 ideas on the same topic.". With Ayoa, the following query was posed: "Human resources development in Gorenjska". Both systems generated 2x124 ideas in two separate sessions (named: ChatGPT1, ChatGPT2, Ayoa1, Ayoa2). For one, that is working with the group decision support systems for more than 25 years and conducting hundreds of brainstorming sessions, the results generated by ChatGPT were remarkable. The whole brainstorming session should be performed as a continuous chat, first we generate ideas, then we rank them and at the end we categorize them according to the similarity. It's important to erase previous chats or start a new chat since ChatGPT remembers the previous chatting. Ayoa is regionally better aware than ChatGPT, at least for Gorenjska region. For example, Ayoa proposed, on the topic "Development of tourism in Gorenjska region" the idea "Advertise cultural attractions in Kranj." Here one could observe spatial awareness of the tool that correctly determines the city Kranj, which is by the way the capital city of Gorenjska region. Correct spatial awareness could also be noted when considering ideas on the topic "Development of tourism in Slovenia." Ideas generated here referring to the Slovenia peculiarities are, refereeing Slovenia as well as Slovenian city Bled:

- Explore Slovenian vineyards and sample local wines.
- Uncover traditional Slovenian recepies in rural villages.
- Tour the capital city's food markets and restaurants.

- Cruise Lake Bled for an unforgettable dining experience.

Besides the categorizing, Chat GPT also provides a detailed description of particular category, which is nice and informative. In addition, Chat GPT comments after categorization: "Some of the ideas could fit into multiple categories, and I have tried to put them in the most relevant one. Also, these categories are not mutually exclusive, and some overlaps may occur." The system also abstracts particular category, which is rather handy. The system also explains it's ranking in particular category:

Please rank ideas in the category Waste management and recycling from most to least important.

1. Developing sustainable waste management policies to reduce pollution and promote recycling.
2. Investing in waste management and recycling infrastructure.
3. Developing sustainable resource management policies to reduce waste and conserve resources.

This ranking is based on the idea that developing sustainable waste management policies to reduce pollution and promote recycling is the most important aspect of waste management and recycling, followed by investing in waste management and recycling infrastructure and then developing sustainable resource management policies to reduce waste and conserve resources. Having a comprehensive waste management and recycling policy in place is crucial for the long-term sustainable development of Slovenia, reducing pollution and conserving resources. Investing in infrastructure and developing sustainable resource management policies are also important but to a lesser extent compared to the first one.

Categories were determined by the application of LDA and NMF method. The sets of extracted words are different, however in particular topic there is a good coherence. Here, the category labels are provided by the expert. From the expert point of view, the provided top 5 words provide good starting point for determination of categories. Table 1 shows the ideas categorized with the LDA and NMF method and two human "experts" Exp1 and Exp2, authors of the present paper. The labels of the particular category were also determined by humans, as said. By comparing LDA and NMF categorization one could see, that some categories

differ but in some, for example #1, #5, #6 … #9 the categories are harmonized. Here #0 represents idea in the first row, #1 idea in the second row etc. Based on this small sample, NMF might be slightly better semantically, but one should note that there are several parameters that determine the clustering. If we also consider the human categorization by experts Exp1 and Exp2 one could see, that in some cases, for example idea #1, the categorization is harmonized. On the left side of the Figure 2 is boxplot which enables us to check data spread of the entropy of the 248 ideas generated by ChatGPT and Ayoa. One could observe that the entropy values vary around 4 bit. Here we should mention that the entropy of letters in the English language is 4.11 bits (McCarthy, 1973). On the right side of the Figure 2 is frequency histogram of the entropy.

Table 1: Ideas categorized with the LDA and NMF method and Exp1 and Exp2

Idea	Lab. LDA	Lab. NMF	Exp1	Exp2
#0 Develop a comprehensive onboarding program for...	promotion	mentoring	Human resource planning	job
#1 Offer ongoing training and development opportu...	education	mentoring	Learning and development	edu
#2 Create a mentorship program to connect experie...	recognition	mentoring	Employee participation and communication	ment
#3 Implement a performance management system to t...	education	recognitionFeedback	Performance management	hr
#4 Establish a clear career development...	recognition	carerDev	Career planning	hr
#5 Create a rewards and recognition program to ac...	recognition	recognitionFeedback	Rewards	hr
#6 Encourage team building and collaboration...	network	teamNetwork	Employee participation and communication	hr
#7 Develop a strong corporate culture that promot...	workEnv	benefits	Employee participation and communication	hr
#8 Offer telecommuting options ...	rbenefits	workEnv	Personal wellbeing	worke
#9 Establish a wellness program to promote physic...	network	mentoring	Health and safety	worke

By performing the Shapiro-Wilk test ($p < 0.05$), we conclude that the data is not normally distributed, therefore, the Mann-Whitney U test (Bedre, 2023) is more appropriate for analyzing two samples than parametric counterparts.

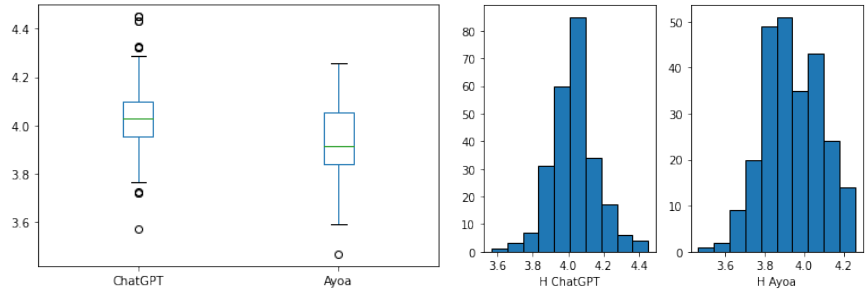

Figure 2: Box plot of the entropy for ideas generated by ChatGPT and Ayoa (left) and Frequency histrogram (right)

As the p value obtained from the Mann-Whitney U test is significant ($U = 41678.5$, $p < 0.05$), we conclude that the entropy of the generated ideas by ChatGPT is significantly higher than the entropy of ideas generated by Ayoa (confirm H1). Figure 3 shows word cloud for 4x124 ideas generated by ChatGPT (left column) and Ayoa (right column).

Figure 3: Word cloud for 4x124 ideas generated by ChatGPT (left column) and Ayoa (right column)

One could observe the highlighted words. For the ChatGPT these are: [employee, program, system, develop, create, offering] and for the Ayoa: [job, career, training, create, organize]. Inspection of the word cloud could improve one's ability to properly categorize the generated ideas. Figure 4 shows pyLDAvis (Sievert et al., 2016; Mabey, 2021) visualization that helps us explore the ideas and properly categorize them. This could be a difficult task if we have a large number of generated ideas. At the right, the frequency of words in particular category, in our case this is category 1, might be observed. According to the distribution, category 1 might be named as "careerDevelopmentSystem". The distance between the circles on the left, that represent different categories, represents semantic relationship between different categories.

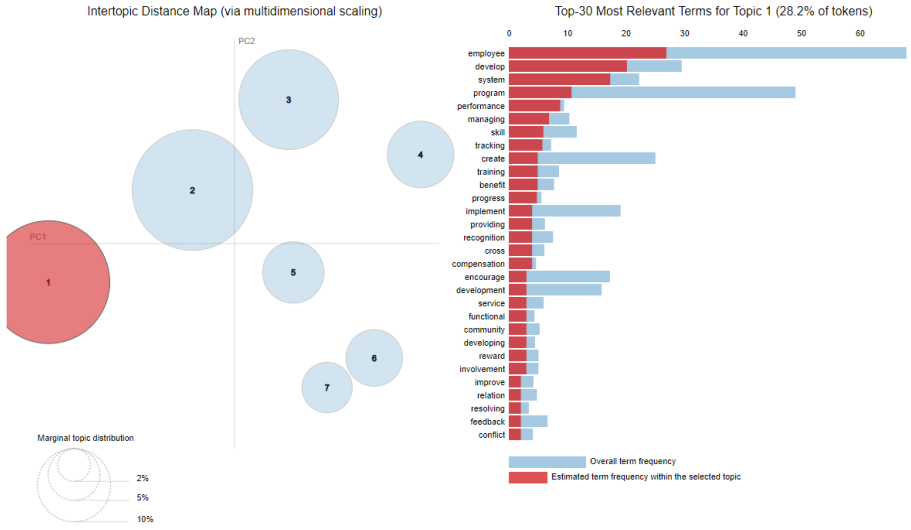

Figure 4: Categories associated with 124 ideas generated by ChatGPT set 1

The pyLDAvis tool is interactive and can help in understanding of the relation between categories and ideas.

Figure 5 shows the categories associated with 124 ideas generated by Ayoa set 1. The category 1 is marked in red circle. This category could be named »careerMentorshipProgram« One could observe different Intertopic distance on the left. The categories of the ChatGPT are semantically more dispersed than in the case of Ayoa.

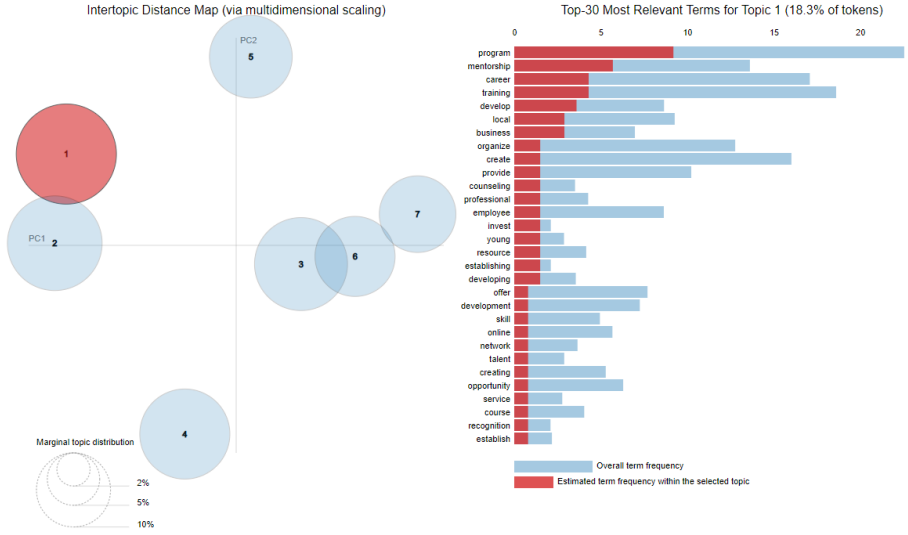

Figure 5: Categories associated with 124 ideas generated by Ayoa set 1

The pyLDAvis tool is also suitable for analysis of large data. This might come handy in the case of national gathering of ideas, where one could expect that several thousand proposals would be collected.

4 Discussion

New technologies raise the question, how does it fit into the human creativity and decision-making process. The issue might be similar to questioning how the car fits into the process of going from point A to B, where alternatives are: alt1) by foot, alt2) by car. Certainly, new means provide completely new possibilities and enhance our ability to build a better world. This idea might be questionable from Europe's point of view though, finding itself amidst the war in the year 2022 raging on from 2014 (Malhotra, 2023). With the new technology the brainstorming sessions could be significantly improved. However, here we also change the role of ideators and decision makers. The following new tasks could be considered in AI augmented brainstorming: a) Selecting good ideas from the set of generated by ChatGPT, b) Selection of the proposed criteria for evaluation, c) Categories naming, deep semantics analysis.

After generating hundreds of ideas with AI tools one could note that most ideas might be quite general, like reading the titles of the newspapers. Rarely, one could stumble on truly creative ideas that would be pinpoint focused on the posed brainstorming question. However, this might also happen as for the case of "Enterpreneurship" in Slovenia, the idea "Robotics incubator for aspiring entrepreneurs" makes a lot of sense, since Yaskawa Inc. has a strong production facility located in Ribnica.

We have observed that ChatGPT repeatedly generated exactly the same ideas. This is in contradiction with the issued instructions, to apply the brainstorming principle. Humans would rarely generate completely the same ideas. However, multiple generations of the same string of text could indicate the higher importance of the idea. Besides, the ideas might be ethically questionable as for example on the topic of "Human resources development in Slovenia", the ayoa AI generated the idea: "Automated attendance with facial recognition". On the one hand, the idea is again quite good and innovative, however here some ethical questions could arise. One could conclude that modern AI tools and techniques of NLP are indispensable tools for enhancing group brainstorming activity in the phase of generation of ideas as well as categorizing and ranking. This will probably change several creative workflows in organizations, government institutions and enterprises in the future.

Acknowledgements

This work was supported by the Slovenian Research Agency (ARRS) (program No.: UNI-MB-0586-P5-0018, Project "Bionics, Pilot Organizational Models of Professional Education and Systems Thinking for Digital, Green and Sustainable Development" within NextGenerationEU NRP No: 3330-22-3515, NOO No: C3330-22-953012 and Erasmus+ Project: 2021-1-MK01-KA220-HED-000027646.

References

Ayoa. Ayoa Mind Mapping, whiteboards & tasks. Retrieved January 27, 2023, from https://www.ayoa.com/

Bedre, R. (2021, November 7). Mann-Whitney U Test (Wilcoxon rank sum test) in python [Pandas and scipy]. Data science blog. Retrieved February 3, 2023, from https://www.reneshbedre.com/blog/mann-whitney-u-test.html

Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent dirichlet allocation. Journal of machine Learning research, 3(Jan), 993-1022.

Blei, D., Ng, A., & Jordan, M. (2001). Latent dirichlet allocation. Advances in neural information processing systems, 14.

Cahan, P., & Treutlein, B. (2023). A conversation with ChatGPT on the role of computational systems biology in stem cell research. Stem Cell Reports, 18(1), 1-2.

Castelvecchi, D. (2022). Are ChatGPT and AlphaCode going to replace programmers?. Nature.

Grimaldi, G., & Ehrler, B. (2023). AI et al.: Machines Are About to Change Scientific Publishing Forever. ACS Energy Letters, 8, 878-880.

Haque, M. U., Dharmadasa, I., Sworna, Z. T., Rajapakse, R. N., & Ahmad, H. (2022). " I think this is the most disruptive technology": Exploring Sentiments of ChatGPT Early Adopters using Twitter Data. arXiv preprint arXiv:2212.05856.

Jeblick, K., Schachtner, B., Dexl, J., Mittermeier, A., Stüber, A. T., Topalis, J., ... & Ingrisch, M. (2022). ChatGPT Makes Medicine Easy to Swallow: An Exploratory Case Study on Simplified Radiology Reports. arXiv preprint arXiv:2212.14882.

Kofjac, D., Skraba, A. and Kljajic, M., 2003. Neural network modeling by system dynamics methodology (Vol. 1, pp. 424-428). IEEE.

Kološvari, A., Stojanović, R., Zupan, A., Semenkin, E., Stanovov, V., Kofjač, D., & Škraba, A. (2019). Speech-recognition cloud harvesting for improving the navigation of cyber-physical wheelchairs for disabled persons. Microprocessors and Microsystems, 69, 179-187.

Kung, T. H., Cheatham, M., Medinilla, A., Sillos, C., De Leon, L., Elepano, C., ... & Tseng, V. (2022). Performance of ChatGPT on USMLE: Potential for AI-Assisted Medical Education Using Large Language Models. medRxiv.

Lavrič, F., Semenkin, E., Stanovov, V., Škraba, A. Andrej. (2022) Pregled odorij in metod za učinkovito skupinsko zbiranje idej in odločanje = Review of tools and methodology for efficient group idea generation and decision making. in Šprajc et al., ed., 41th International Conference on Organizational Science Development, University of Maribor, University Press.

Lee, D. D., & Seung, H. S. (1999). Learning the parts of objects by non-negative matrix factorization. Nature, 401(6755), 788-791.

Mabey, B. (2021). pyLDAvis documentation.

Malhotra, D. (2023, January 19). I ask chatgpt for advice to Zelenskyy on ending the war in Ukraine. what follows is fascinating... Medium. Retrieved February 3, 2023, from https://professordeepakmalhotra.medium.com/i-ask-chatgpt-for-advice-to-zelenskyy-on-ending-the-war-in-ukraine-what-follows-is-fascinating-818c8baa5736

McCarthy, J. P. (1973) "Automatic File Compression". International Computing Symposium. pp. 511-516.

Portilla, J. (2023) NLP - Natural Language Processing with Python [MOOC]. Udemy. Retrieved January 27, 2023 from https://www.udemy.com/course/nlp-natural-language-processing-with-python/

Radford, A., Wu, J., Child, R., Luan, D., Amodei, D., & Sutskever, I. (2019). Language models are unsupervised multitask learners. OpenAI blog, 1(8), 9.

Shannon, C. E. (1948). A mathematical theory of communication. The Bell system technical journal, 27(3), 379-423.

Sievert, C., Shirley, K., & Sievert, M. C. (2016). Package 'LDAvis'.

Škraba, A., & Filipič, B. (2009). Z informacijsko tehnologijo podprta izvedba sestankov regionalnih ravzojnih odborov v fazi zbiranja idej. In J. Nared & D. Perko (Eds.), Razvojni izzivi Slovenije [Slovenian development issues] (pp. 241–250). Založba ZRC.

Škraba, A., Kljajić, M., & Leskovar, R. (2003). Group exploration of system dynamics models—is there a place for a feedback loop in the decision process?. System Dynamics Review: The Journal of the System Dynamics Society, 19(3), 243-263.

Škraba, A., Stanovov, V., Semenkin, E., Kološvari, A., & Kofjač, D. (2018). Development of algorithm for combination of cloud services for speech control of cyber-physical systems. Int. J. Inf. Technol. Secur., 10(1), 73-82.