Usage of Artificial Intelligence Tools for Improvement of Expert Formulations During Construction of Knowledge Bases of Decision Support Systems

Oleh Andriichuk[1,2,3], Sergii Kadenko[1] and Anna Florek-Paszkowska[4]

[1]Institute for Information Recording of the National Academy of Sciences of Ukraine, 2, Shpak str., Kyiv, 03113, Ukraine
[2]Taras Shevchenko National University of Kyiv, 64/13, Volodymyrs'ka str., Kyiv, 01601, Ukraine
[3]National Technical University of Ukraine "Igor Sikorsky Kyiv Polytechnic Institute", 37, Prospect Beresteiskyi, Kyiv, 03056, Ukraine
[4]CENTRUM Católica Graduate Business School Pontificia Universidad Católica del Perú, Jirón Daniel Alomía Robles 125, Santiago de Surco, 15023, Lima, Perú

Abstract
The work is dedicated to the issue of applying artificial intelligence tools to enhance expert formulations in constructing knowledge bases for decision support systems. The challenges arising in constructing a knowledge base for weakly structured subject areas are considered, and corresponding approaches to the use of artificial intelligence tools are proposed. An appropriate experimental study has been conducted, the results of which indicate that such application of artificial intelligence tools in practice is advisable only in an automated version involving a group of experts.

Keywords
artificial intelligence, natural language processing, expert formulations improving, decision support system, knowledge base

1. Introduction

The activities of any manager are closely related to the need to make decisions on a daily basis. Decision-making is a specific type of activity that involves forming a set of decision options (alternatives) and then evaluating their relative effectiveness and resource allocation among the decision options, based on their assessments. The simpler types of decisions include accepting or rejecting an alternative, choosing the best alternative from a given set, and ranking alternatives.

In making complex decisions, there is often a need to consider numerous (tens or hundreds) interrelated factors that interact in complex ways. To ensure high-quality decision-making, the integration of knowledge from many expert specialists is necessary. However, due to psychophysiological limitations, humans are only capable of processing around 7-9 objects simultaneously [1]. Decision Support Systems (DSS) are used to overcome this limitation (Figure 1).

In solving problems in weakly structured domains, where DSS are increasingly utilized, the task of enhancing the adequacy of the domain model to improve the reliability of recommendations produced by DSS becomes relevant. An essential component of DSS is the information obtained from experts in the form of object names formulated in natural language. Therefore, it is crucial to unambiguously identify these objects in knowledge bases (KB) of DSS to adequately consider the collective expertise of experts when establishing relationships among these objects. Large KBs are created to describe complex subject areas, raising the issue of unequivocal object identification in these KBs.

ITS-2023: Information Technologies and Security, November 30, 2023, Kyiv, Ukraine
andriichuk@ipri.kiev.ua(O. Andriichuk); seriga2009@gmail.com(S. Kadenko); aflorekpaszkowska@pucp.edu.pe (A. Florek-Paszkowska)
0000-0003-2569-2026(O. Andriichuk); 0000-0001-7191-5636 (S. Kadenko); 0000-0002-1249-5069 (A. Florek-Paszkowska)

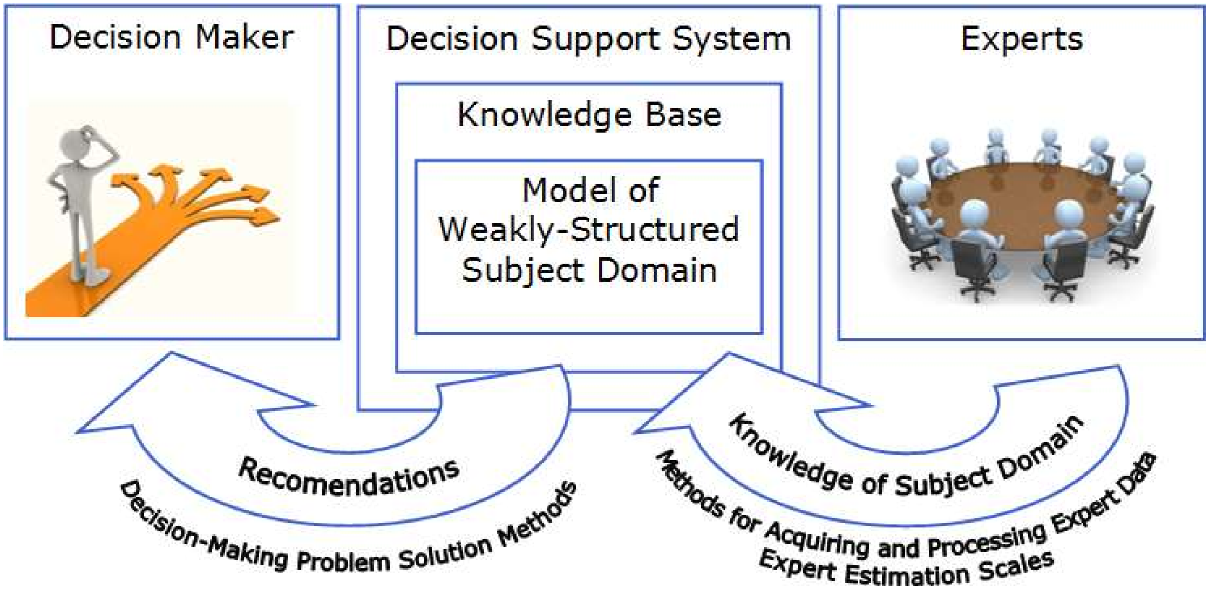

Figure1: Functional scheme of adecision support system

Currently, when constructing domain models in DSS, researchers focus significantly on methods for obtaining and processing expert evaluations, which are extensively covered in scientific publications. Typically, attention is given to evaluating objects from already constructed KBs. As the number of objects in DSS KBs grows, especially when using group methods for KB construction, the issue of error-free identification becomes increasingly important.

During the construction and maintenance of KBs, the most accurate representation of integrated expert opinions can be achieved by eliminating the erroneous repeated input of content-identical (but differently formulated) object names. To achieve this, it is advisable to search for content-related formulations. Enhancing the efficiency of using knowledge obtained from experts is enabled by the reuse of previously constructed KBs, but the process of merging KBs needs to be automated. In sufficiently large KBs, especially those built by expert groups of various profiles, there is a significant likelihood of errors.

The application of DSSs results in recommendations (see Figure 1) for decision-makers [2-4]. This involves modeling weakly structured subject areas [5, 6], as shown on Figure 2, using DSS tools based on the constructed corresponding KB. One of the important characteristics of such subject areas is the incomplete description of objects, which makes it difficult to create a quality training sample for machine learning. Therefore, expert knowledge needs to be utilized. Expertise, including group expertise, requires significant time and financial investments.

The properties of weakly structured subject domains outlined in Figure 2 include the absence of a formalizable functioning goal, the lack of an optimality criterion, uniqueness, dynamics, incomplete description, the presence of a subjective human factor, the inability to construct an analytical model, the absence of benchmarks, and high dimensionalities.

Let's delve deeper into the characteristics of these weakly structured subject domains. Objects within such domains are inherently unique. Management systems designed for these domains are typically tailored to address specific real-world problems, making replication on other entities costly or unfeasible.

In systems not engineered by humans, like biological systems, formalizing a functioning goal is often unattainable. While these systems aim for efficiency and parameter maintenance within defined limits, articulating a specific criterion for their functioning proves challenging due to the intricate and numerous factors at play, with complex and obscure connections that resist easy categorization.

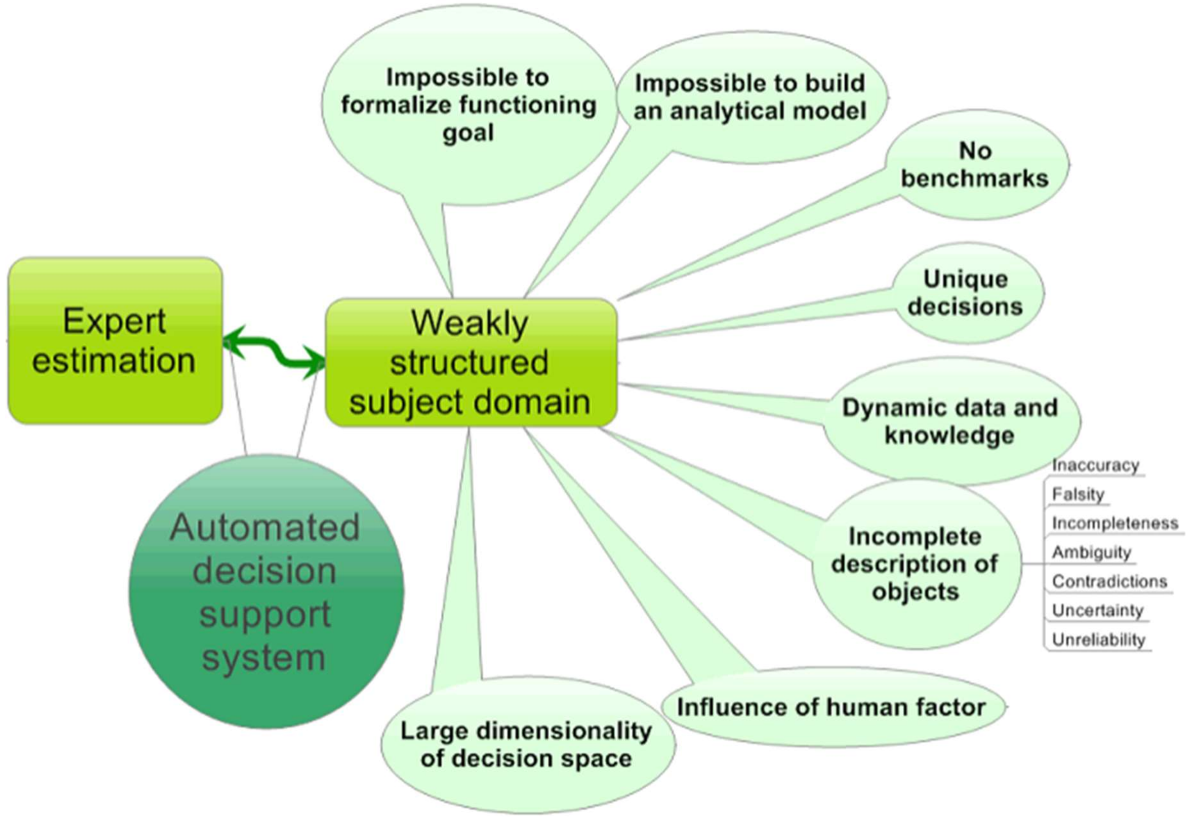

Figure 2: Features of weakly structured subject domains

The absence of a formalizable functioning goal precludes the construction of an optimization function that could dictate the ideal operational state of an object. Optimization becomes complex as factors are interdependent, disrupting the delicate balance necessary for system stability amidst environmental changes, potentially resulting in catastrophic outcomes and irreversible system alterations.

Without a formalizable functioning goal or an achievable optimization function, constructing an analytical model for these subject domains becomes unviable. Dynamism stems from the evolving nature of these objects over time, necessitating adaptive management strategies that mirror the object's changes.

Incompleteness in description arises from data inaccuracies, incompleteness, falsities, ambiguities, contradictions, uncertainties, and unreliabilities surrounding the object, making quantitative characterization challenging. Benchmarks for object characteristics within weakly structured subject domains are inappropriate due to these issues.

The vast decision space dimension results from the multitude and heterogeneity of criteria defining the subject domain. Human subjects with free will further complicate management efforts, as predicting and controlling human behavior within a system is intricate given individual goals and interests influencing actions.

The properties of weakly structured subject domains underscore the reliance on experts as the primary, and sometimes sole, source of information within these domains.

Research was conducted in the United States to examine the distribution of knowledge types utilized in the daily operations of specific organizations [6]. The findings from the Delphi Group's investigation, as depicted in Figure 3, revealed that a significant portion (42%) of the knowledge employed is not formally documented or stored on any data mediums. This type of knowledge is exclusively held by skilled experts, posing a challenge for AI tools to leverage it effectively for providing recommendations [7, 8]. Consequently, while AI serves as a valuable tool in Natural

Language Processing (NLP), it cannot completely supplant the expertise of a subject matter specialist. Despite AI's proficiency in processing vast datasets and recognizing semantic relationships, it may struggle to account for the distinctive context and nuances inherent in individual scenarios.

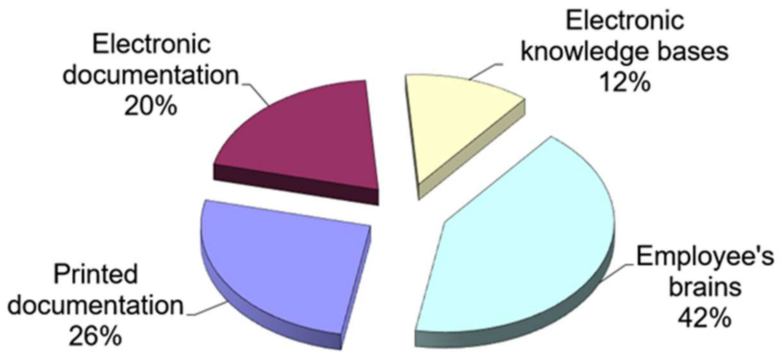

Figure 3: Delphi Group Research: Major Repository of Knowledge Organization

Recently, artificial intelligence (AI) tools based on large-scale linguistic models implemented using artificial neural networks with transformer architecture [9, 10] have gained significant popularity due to their convenience and accessibility. These tools have rapidly and widely spread among users, finding direct applications in the practice of preparing and writing various types of works, including academic papers, particularly English texts, as well as in publishing [11]. Some researchers even use these AI tools to construct networks of keywords as conceptual models for specific subject areas [12]. However, developers of such tools caution that large language models do not guarantee the truthfulness or reliability of the output data [13].

The above considerations necessitate and highlight the importance of researching the possibilities of using AI tools in DSS to build models for weakly structured subject areas [6, 11], as shown on Figure 2.

2. Construction of Knowledge Bases for Weakly-structured Subject Domains

Since the main component of DSS is the KB, let us consider its structure, which is shown in Figure 4. The main elements of the BR are objects and relations between them. KB objects can be goals and projects. The object of the BR is named in the form of a short formulation. A tuple of keywords is used to define the meaning of a KB object. The object of the KB can be quantitative or qualitative, threshold or quasi-linear. For projects, duration of execution and resources are specified. The relationship between BR objects can: be positive or negative, have a time delay, compatibility groups. It is also characterized by a private influence coefficient.

On Figure 5, tasks arising when constructing KBs for weakly structured subject areas are shown [14]. When obtaining knowledge from expert groups, the decomposition of goals and the determination (evaluation) of their impact levels are conducted. When formulating goals, each expert in the group provides texts of their individual sub-goal formulations. Figure 6 shows this stage in the Consensus-2 system [15]. Subsets of content-matching formulations are identified. The best formulation is selected in each subset. When establishing influences between goals, in the KB, goals that affect a specific goal and goals affected by a specific goal should be sought. Each expert determines the importance of each sub-goals by pairwise comparisons. To achieve higher reliability and consistency of expert evaluations, it is necessary to take into account the order of alternatives in pairwise comparisons [16, 17]. After group expert assessment of impact levels, aggregation of relevant assessments occurs upon reaching a sufficient level of consensus. When aggregating knowledge obtained from different expert groups, the search and consolidation in the KB of content-matching goals formulated by different groups take place. Additionally, "pre-establishing" influences on the combined set of goals with re-scaling the relevant impact level values occurs. In this process, goals affecting a specific goal and goals affected by a specific goal should be sought in the KB.

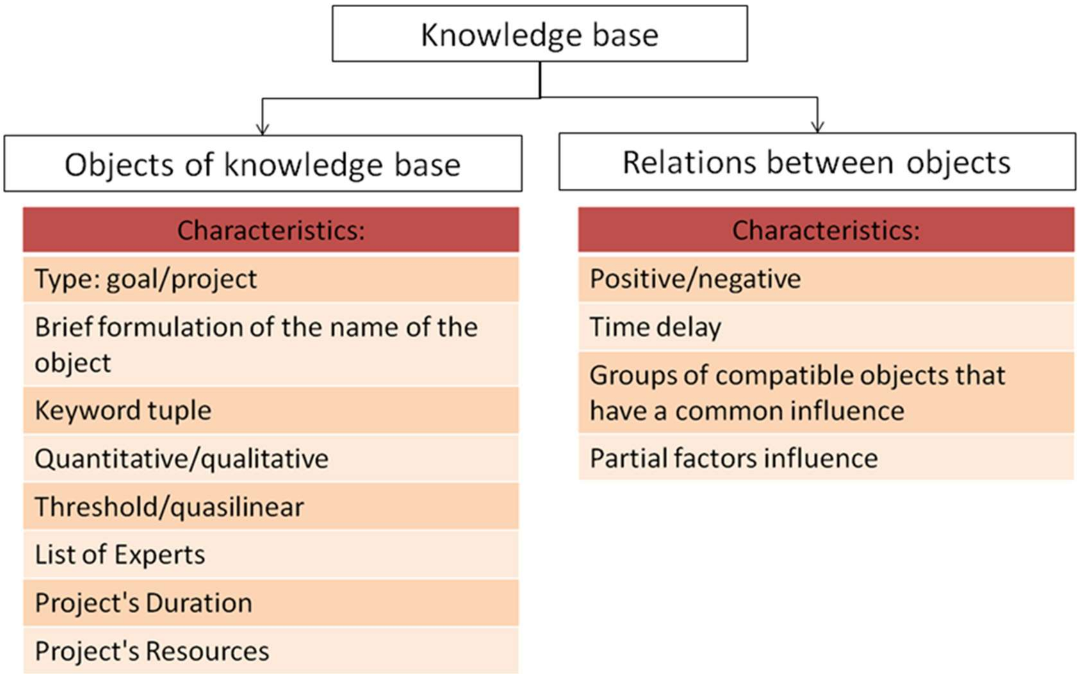

Figure 4: Structure of decision support systems knowledge base

When building KBs, the following issues arise: a high level of detail in the KB leads to a deterioration in the adequacy of models in weakly structured subject areas, namely: redundancy, ambiguity, and the presence of contradictions in the KB.

Figure 7 illustrates the factors that influence the quality of recommendations in DSSs. One of the most important among them is the "Adequacy of the domain model." In turn, this factor is influenced by the aforementioned redundancy, ambiguity, and the presence of contradictions in the KB of DSS.

For almost all the aforementioned tasks in constructing KBs for weakly structured subject areas, it is advisable to use AI tools. The question arises: can one fully rely on AI tool recommendations and perform these tasks in automatic mode? Below are the results of an experimental study on the accuracy of AI tool recommendations regarding improving expert formulations. These results indicate that such use of AI tools in practice is only advisable in an automated mode with the involvement of expert groups. It is not advisable to rely entirely on AI tool recommendations at this stage of its development.

During the construction of such KBs, knowledge engineers, analysts, and multidisciplinary experts consistently decompose a particular object of the subject area, step by step. This approach is applied to group modeling of subject areas within the framework of the "Consensus-2" system for distributed collection and processing of expert information for decision support systems [15].

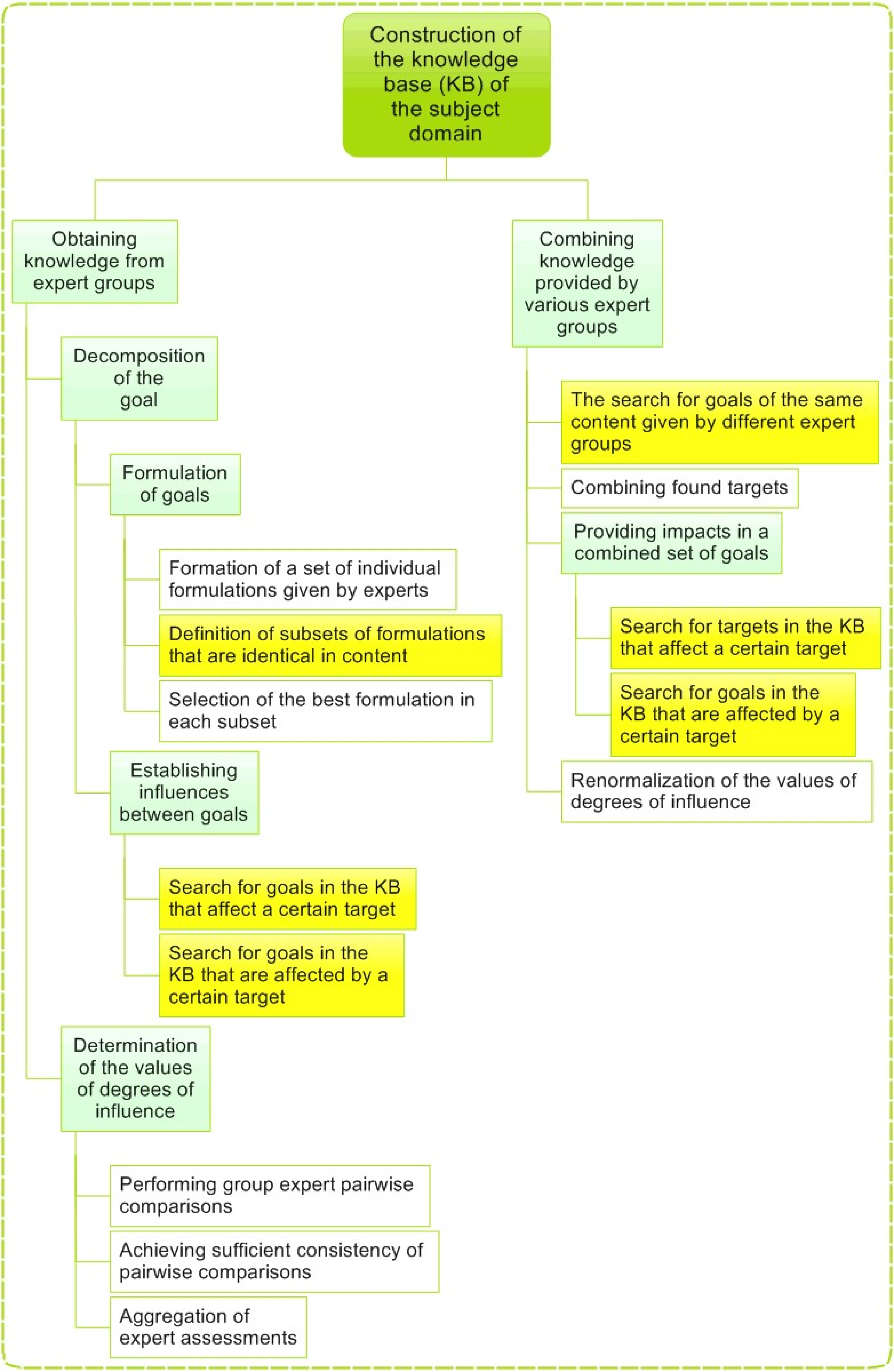

Figure 5: Tasks at the stage of building a subject domain knowledge base

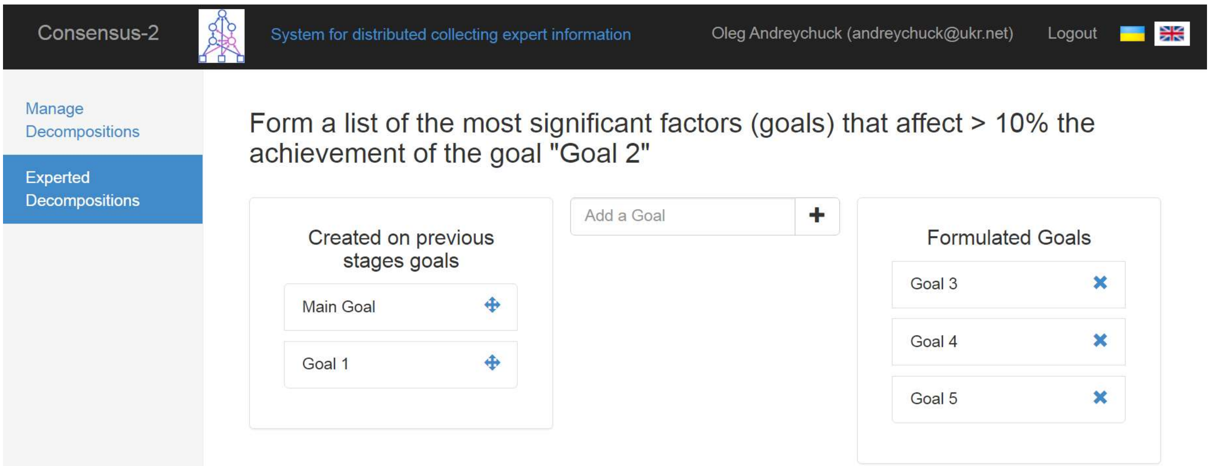

Figure 6: Factors affecting the quality of DSS recommendations

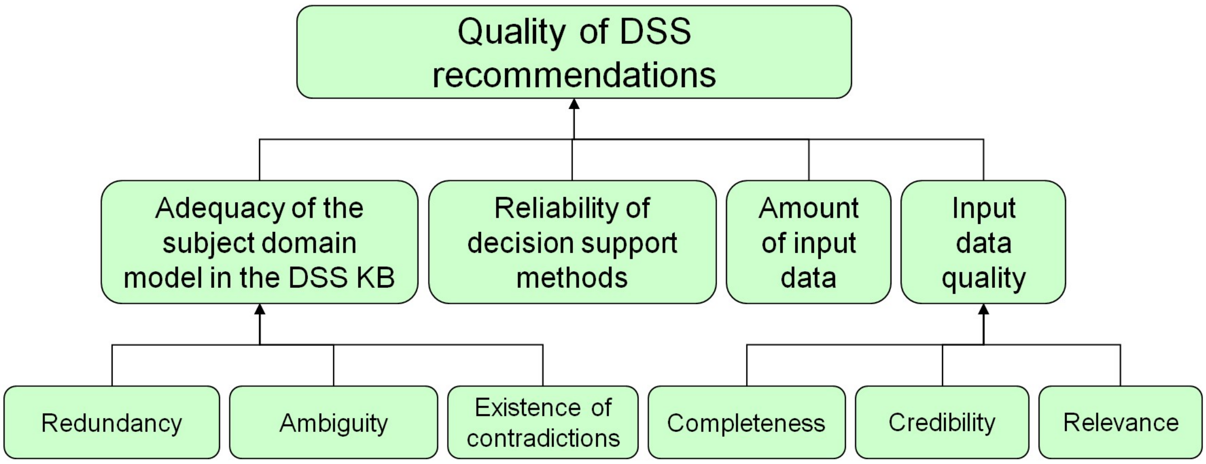

Figure 7: Factors affecting the quality of DSS recommendations

3. Experimental Study on the Accuracy of AI Tool Recommendations for Improving Expert Formulations

The conducted experimental study on the accuracy of AI tool recommendations for improving expert formulations consisted of three described above stages.

1. A group of expert respondents provided texts of their English-language formulations. Most of these texts were translated from Ukrainian and Russian using machine translation and qualified English translators. Each expert in this group specializes in the field of information technology and security. The English texts containing expert formulations belong specifically to this field. It is also worth noting that English is not the native language for these experts.
2. Recommendations for improving the quality of English texts containing expert formulations were obtained using AI tools with the architecture of GPT-3.5. Special prompts were used with the AI tool for this purpose.
3. A group of expert validators assessed the accuracy and quality of the English texts containing expert formulations, as well as the recommendations obtained from the AI tool. This group of experts has competence in the field of information technology and security and is proficient in English. Additionally, this group consulted with a group of experts specializing in English

language in general, translation, philology, linguistics, and English in the field of computer science.

The questionnaire for expert validators contained the following questions:

"Please compare the quality of the AI tool's recommendation with the quality of the original English-language expert opinion:

- AI tool recommendation is unreliable;
- the AI tool's recommendation is reliable but of equivalent quality;
- the recommendation of the AI toolkit is reliable and requires minor adjustment;
- the recommendation of the AI tool is reliable and does not require adjustment."

In order to ensure statistical credibility of the research, we calculated the necessary number of experiment instances. Evaluation of statistical credibility was conducted based on the central limit theorem [18]. If we set the confidence probability value at P_β = 0.95 (i.e., the probability that the random variable value falls within confidence interval β), and confidence interval size for the given experimental study is β = 0.05, the minimum necessary number of experiment instances can be calculated based on the following inequality:

$$n \geq \frac{p \cdot (1-p)}{\beta^2} \left(F^{-1}(P_\beta)\right)^2$$

where F^{-1} is the inverse Laplace function [19]; p is the frequency of repetition of value of the random characteristic under consideration.

We select the value of p based on previously obtained experiment results as the "worst" probability/frequency (i.e. the one closest to 0.5). As a result of test (initial) experiment series, we collected 61 ranking of alternative pair sequences. After filtering (screening), the remaining set of test experiment series constituted 33 rankings of alternative pair sequences. The results of test experiment series are presented in table 1.

Table 1
A test series of the experiment

Indicator name	Quantity (frequency)
total number of recommendations received	797
unreliable recommendations	285
reliable recommendations that are equivalent in quality to the original wording	10
reliable recommendations that require minor adjustments	76
reliable recommendations that do not require adjustment	426

Among the frequencies, defined based on the table {*285/797≈0.358; 10/797≈0.013; 76/797≈0.095; 426/797≈0.535*}, the worst one according to the specified criterion is frequency p = 0.535, which we will input into the formula for calculation.

After inputting all the respective values into the formula, we get:

$$F^{-1}(0.95) \approx 1.96,$$

then:

$$\left(F^{-1}(0.95)\right)^2 \approx 3.84,$$

$$n \geq \frac{0.535 \cdot (1-0.535)}{(0.05)^2} 3.84 = 382.12,$$

and, finally, $n \geq 382.12$. It means, that in order to draw credible conclusions based on the experiment results, it is sufficient to perform at least 383 instances of the experiment.

The results of the calculation of the accuracy of AI tool recommendations for improving expert formulations obtained during the experimental study are presented in Table 2.

Table 2
Results of an Experimental Study of the Credibility of AI Recommendations for Improving Expert Formulations

Indicator name	Quantity (frequency)
total number of recommendations received	1104
unreliable recommendations	365
reliable recommendations that are equivalent in quality to the original wording	17
reliable recommendations that require minor adjustments	108
reliable recommendations that do not require adjustment	614

Results of the experimental study on the credibility of AI tool recommendations for improving expert formulations:

- 56% of the recommendations provided by AI tools were accurate and did not require corrections.
- 65% of the recommendations improved the quality of formulations.
- 67% of the recommendations did not worsen the quality of formulations.
- 1.5% of the recommendations were found to be futile, adding no value.
- 33% of the recommendations were considered harmful, distorting the content of the formulations.

During the analysis of harmful recommendations, it was found that the main causes of distortions lie in the insufficient adaptability of the model to the specifics of the terminology and context used, as well as in the inadequate processing of terms that have specific meanings in the field of computer science.

4. Conclusions

It has been shown that it is beneficial to utilize artificial intelligence tools to address a range of tasks that arise during the construction of knowledge bases for decision support systems.

Ways of applying artificial intelligence tools in knowledge base construction have been proposed, particularly to enhance the quality of expert formulations.

An experimental study on the accuracy of recommendations from artificial intelligence tools for improving expert formulations has been conducted.

Corresponding empirical results have been obtained, indicating that the application of artificial intelligence tools in practice is advisable only in an automated mode, i.e., involving a group of experts. It is not advisable to rely solely on the recommendations of artificial intelligence tools at this stage of the technology's development.

References

[1] George A. Miller, The Magical Number Seven, Plus or Minus Two. *The Psychological Review*, 1956, vol. 63, pp. 81-97.

[2] T. L. Saaty, Principia Mathematica Decernendi - Mathematical principles of decision making - Generalization of the Analytic Network Process to neural firing and synthesis. RWS Publications, Pittsburg, 2010.

[3] F. Burstein, C. W. Holsapple, Handbook on decision support systems 1: Basic Themes, Springer-Verlag Berlin Heidelberg, 2008. doi:10.1007/978-3-540-48713-5.
[4] J. A. Zapata-Cortes, C. Sánchez-Ramírez, G. Alor-Hernández, J. L. García-Alcaraz, Handbook on Decision Making: Volume 3: Trends and Challenges in Intelligent Decision Support Systems, Springer, 2022. doi:10.1007/978-3-031-08246-7.
[5] D. T. Lee, Expert Decision-support Systems for Decision-making, *Journal of Information Technology* (3:2), 1988, pp. 85-94.
[6] Sergii V. Kadenko Prospects and Potential of Expert Decision-making Support Techniques Implementation in Information Security Area // CEUR Workshop Proceedings (ceur-ws.org). Vol-1813 urn:nbn:de:0074-1813-0. Selected Papers of the XVI International Scientific and Practical Conference "Information Technologies and Security" (ITS 2016) Kyiv, Ukraine, December 1, 2016, 8-14. URL: http://ceur-ws.org/Vol-1813/paper2.pdf]
[7] V. Shabrina, A. Silvianita, Factors Analysis on Knowledge Sharing at Telkom Economic and Business School (TEBS) Telkom University Bandung. In: Procedia - Social and Behavioral Sciences, vol. 169, 2015,198-206. doi:10.1016/j.sbspro.2015.01.303.
[8] Heather Smith, James Mckeen, Instilling a knowledge-sharing culture. In: Journal of Knowledge Management, vol. 15, no. 4, 2011, pp. 675-687. URL: https://www.researchgate.net/publication/228828294
[9] GhatGPT. URL: https://chat.openai.com
[10] Galactica. URL: https://galactica.org
[11] N. Punar Özçelik, A Comparative Analysis of Proofreading Capabilities: Language Experts Vs. ChatGPT / International Studies in Educational Sciences, ed. by Bülent Pekdağ, Serüven Publishing, June 2023, pp. 147-162.
[12] Dmitry Lande and Leonard Strashnoy, Concept Networking Methods Based on ChatGPT&Gephi (April 17, 2023). URL: https://ssrn.com/abstract=4420452 doi:10.2139/ssrn.4420452
[13] Aleksandr Koval, Safwan Al Salaimeh, Oleh Andriichuk, Usage of Expert Classification in Diagnostic Expert Systems' Knowledge Bases Construction // CEUR Workshop Proceedings (ceur-ws.org). Vol-2318. urn:nbn:de:0074-1813-0. Selected Papers of the XVIII International Scientific and Practical Conference "Information Technologies and Security" (ITS 2018) Kyiv, Ukraine, November 27, 2018, 139-149. URL: https://ceur-ws.org/Vol-2318/paper12.pdf.
[14] V. V. Tsyganok, O. V. Andriichuk, Experimental study of the method for determining the semantic similarity of objects of knowledge bases of decision support systems. *Data Recording, storage and processing*. 2014, Vol. 16, No. 4. P.65-76. (in Ukrainian)
[15] Certificate of copyright registration for work № 75023. Computer program "A system for distributed collection and processing of expert information for decision support systems – "Consensus-2"", Ukraine, registered on 27/11/2017. URL: https://dss-lab.org.ua/applications/consensus-2.
[16] Oleh Andriichuk, Vitaliy Tsyganok, Sergii Kadenko, Yaroslava Porplenko, Experimental Research of Impact of Order of Pairwise Alternative Comparisons upon Credibility of Expert Session Results // Proceedings of the 2nd IEEE International Conference on System Analysis & Intelligent Computing, (SAIC), 05-09 October, 2020, Kyiv, Ukraine, pp. 1-5. doi:10.1109/SAIC51296.2020.9239126
[17] V. Tsyganok, O. Andriichuk, S. Kadenko, Y. Porplenko, O. Vlasenko, An Approach to Reduction of the Number of Pair-Wise Alternative Comparisons During Individual and Group Decision-Making. In: Zgurovsky, M., Pankratova, N. (eds) System Analysis & Intelligent Computing. SAIC 2020. Studies in Computational Intelligence, vol 1022. Springer, Cham. URL: https://doi.org/10.1007/978-3-030-94910-5_9.
[18] Hans Fischer, A History of the Central Limit Theorem: From Classical to Modern Probability Theory (PDF). Sources and Studies in the History of Mathematics and Physical Sciences. New York: Springer, 2011. doi:10.1007/978-0-387-87857-7.
[19] A. M. Cohen, Inversion Formulae and Practical Results. Numerical Methods for Laplace Transform Inversion. Numerical Methods and Algorithms. 2007, Vol. 5. 23–44. doi:10.1007/978-0-387-68855-8_2.