

AN ALGORITHM TO SUPPORT THE SCIENTIFIC MANUSCRIPT REVIEW PROCESS WITH THE ASSISTANCE OF CHATGPT

Assoc. Prof. Aldeniz Rashidov
Technical University of Gabrovo

Abstract. In this work, an algorithm is proposed and analyzed to support the peer review process of scientific manuscripts with the assistance of ChatGPT. The study showed that the use of ChatGPT can significantly improve the efficiency and objectivity of the peer review process. The algorithm includes quantitative and qualitative evaluation steps, using weighting and SWOT analysis for better accuracy and depth of evaluation. The algorithm can be useful for scientific editors, reviewers, and authors who are looking for a way to increase the efficiency and objectivity of the peer review process and to obtain preliminary evaluation and feedback on their manuscripts. The proposed algorithm has the potential to transform traditional peer review methods, leading to a more efficient, objective, and detailed scientific manuscript evaluation process. Further research and refinement of the algorithm can improve its accuracy and adaptability, providing even greater value to the scientific community.

Keywords: Peer Review; Artificial Intelligence; ChatGPT; evaluation; SWOT analysis; algorithm

1. Introduction

Peer review is critical to the scientific process as it provides an independent assessment of scientific manuscripts prior to publication. This process ensures that only high quality and reliable research reaches the wider scientific community, maintaining the standard of scientific publications and preventing the dissemination of inaccurate or misleading results. For example, in medical research, inadequately peer-reviewed publications can lead to inappropriate clinical practices, endangering the health and lives of patients. In engineering, peer review is equally important, as poorly conducted or unverified research can lead to the design of ineffective or even dangerous systems and technologies that could have serious consequences for the safety and functionality of engineering solutions.

Despite anonymity, the personal preferences and biases of reviewers in traditional review methods can influence the evaluation (Tennant & Ross-Hellauer 2020). Reviewers are often overworked, which can lead to delays in the review process and biased evaluations. Different reviewers may give different ratings for the same study. Lack of transparency in the process can lead to lack of trust and doubts about the objectivity of reviews.

With the development of information technology, the need for innovation in peer review is becoming more pressing. The volume of scientific publications is growing exponentially, making it difficult for reviewers to provide timely and quality reviews. In this context, automated peer review support systems can significantly improve the efficiency, objectivity and quality of this process (Schulz et al. 2022).

With the development of Artificial Intelligence (AI) and machine learning, various applications are starting to emerge to assist the peer review process. These technologies can be used to automatically analyze and evaluate scientific manuscripts, potentially addressing some of the challenges of traditional peer review. Some systems such as Turnitin and iThenticate use AI to detect plagiarism by comparing the text of manuscripts against huge databases of previously published material (Turnitin 2023; Turnitin 2024). Tools such as Semantic Scholar use AI to analyze citations in scientific publications, which can help assess the impact and significance of research (Irvine 2024). Algorithms are also being developed that can assess the quality of methodology, clarity of exposition, and innovativeness of research. Some AI systems can also analyze the tonality and emotional charge of text, which can help identify bias and unprofessional content (Nandwani & Verma 2021; Alrasheedy, Muniyandi & Fauzi 2022; Kusal et al. 2022). AI has the potential to play a pivotal role in the peer review of scientific manuscripts.

ChatGPT is a powerful AI tool that can be used to aid peer review (Hosseini & Horbach 2023; Gao et al. 2022). It can analyze the manuscript text, assess clarity, structure and methodology, and provide recommendations for improvements. An algorithm based on ChatGPT can support the peer review process by providing rapid and objective assessments, facilitating the work of reviewers and improving the quality of published research.

The aim of this study is to propose an algorithm involving a sequence of steps executed with the assistance of ChatGPT to support the review process. This algorithm will assist reviewers in creating a quality review of a scientific manuscript by fulfilling a number of specific conditions. With the help of ChatGPT, reviewers will be able to speed up the review process, increase the quality of reviews and provide more objective assessments. This study evaluates the capabilities and limitations of this approach and its impact on the efficiency and objectivity of the review process.

2. An algorithm to support the scientific manuscript review process with the assistance of ChatGPT

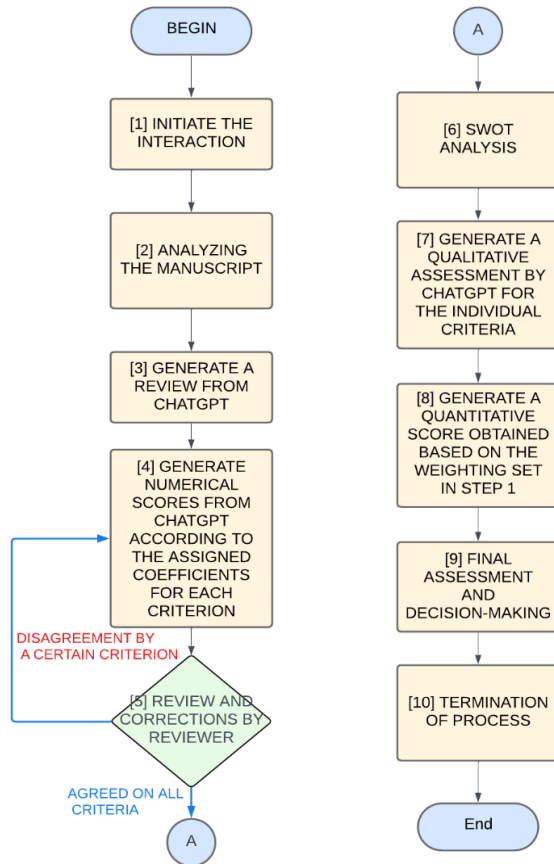


Figure 1. Flowchart of an algorithm to support the scientific manuscript review process with the assistance of ChatGPT

The peer review process requires a detailed and objective evaluation of multiple aspects of the manuscript's content and structure. As stated above, traditional methods of peer review often require considerable time and effort on the part of the reviewers, which can lead to delays in publication and differences in the quality of the

assessments. In this context, an algorithm to assist the review process using ChatGPT can significantly improve the efficiency and objectivity of the review process.

The proposed algorithm involves a sequence of steps that combine automated analysis generated by ChatGPT with the expert judgment of reviewers. This approach not only speeds up the process, but also provides an additional layer of objectivity and detail in the evaluation of scientific manuscripts. The block diagram of the algorithm is presented in Fig. 1.

A. Key steps of the algorithm to support the scientific manuscript review process with the assistance of ChatGPT:

- 1) *Initiating the interaction and asking a question:*
 - The manuscript is entered into the ChatGPT system for analysis;
 - The criteria for analysis are set (clarity, structure, originality, methodology, validity of results, etc.);
 - The weighting coefficients for each criterion required for the quantitative assessment of the content of the scientific manuscript are indicated.
- 2) *Analyzing the manuscript:*
 - ChatGPT analyzes the structure and text using the predefined criteria.
- 3) *Generate a review:*
 - ChatGPT generates a review that includes comments, recommendations, and ratings, sequentially for each criterion.
- 4) *Generate numerical estimates:*
 - ChatGPT generates numerical scores according to predefined weighting coefficients for each criterion based on the analyzed data.
- 5) *Review and corrections by a reviewer:*
 - The reviewer reviews each response and criterion score given by ChatGPT.
 - The reviewer checks the accuracy and precision of the comments and evaluations;
 - The reviewer adds his/her own comments and evaluations as needed and makes corrections to the ChatGPT-generated reviews.
 - The response may include disagreement on a criterion, discussion accordingly, and re-generation of a new score by ChatGPT on that criterion;
 - The reviewer can also use other trained AI models for additional opinions, creating a basis for a new discussion between ChatGPT and the reviewer until consensus is reached.
- 6) *SWOT analysis:*
 - Upon agreement on the criteria, ChatGPT generates a SWOT analysis of the manuscript content;
 - Assess the strengths and weaknesses of the work, as well as the opportunities and threats;
 - The reviewer may consult ChatGPT for any particular aspect of the SWOT analysis.

7) *Generating a qualitative assessment:*

– ChatGPT generates a qualitative assessment of the manuscript, including detailed comments and recommendations on individual criteria that help the author understand the strengths and weaknesses of the manuscript.

8) *Generating a quantitative assessment:*

– The quantification is based on the weighting coefficients set in step 1;
– ChatGPT generates a quantitative score that provides a numerical measure of the qualities of a scientific manuscript based on various criteria.

9) *Final evaluation and decision making:*

– The reviewer makes the final decision on the quality of the manuscript based on a pre-set threshold of quantitative assessment;
– The reviewer finalizes the evaluation and prepares the final review.

10) *Completion of the process:*

– Completion of the review process.

Using ChatGPT to support review offers several key advantages. The proposed algorithm covers both qualitative and quantitative aspects of peer review. The inclusion of weighting coefficients for the quantitative assessment and SWOT analysis for detailed content assessment adds additional levels of depth and objectivity to the review process (Rashidova 2023). The algorithm also highlights the importance of the human factor in vetting and finalizing evaluations, ensuring that automated comments and ratings from ChatGPT are relevant and accurate. Automated analysis can quickly identify key issues and provide objective assessments to serve as the basis for further review by reviewers. ChatGPT can generate detailed comments and recommendations that authors can use to improve the quality of the manuscript.

B. Sample criteria for evaluating the content of a scientific manuscript:

Table 1 presents a set of sample criteria that can be used to evaluate the content of a manuscript with the assistance of ChatGPT. These criteria are tailored to the functionalities of ChatGPT-4o and can be adapted depending on the specific needs of the reviewer and the characteristics of the text being evaluated. The criteria are designed to assist the reviewer in generating an objective and detailed assessment of the quality of the manuscript, covering different aspects of the scientific content.

Table 1. Sample evaluation criteria

Criterion	Sub-criteria	Criterion	Sub-criteria
1. Clarity and legibility	Comprehensibility of the text	5. Validity of results	Validity of results
	Clarity of objectives and hypotheses		Arguing the hypotheses
	Stylistic errors		Long-term relevance
	Terminological precision	6. Discussion	Interpretation of results
	Title relevance		Critical examination
	Abstract quality		Recommendations for future research
2. Structure	Logical structure	7. Originality and relevance	Originality of the study
	Connectivity of parts		Practical relevance
	Organization of information		Citation of relevant literature
3. Methodology	Description of methods	8. Ethics and integrity	Ethical considerations
	Consistency of methods		Completeness of the report
	Innovativeness of approach		Negative results
4. Data analysis	Objectivity and accuracy		
	Statistical significance		
	Graphical representation		

C. Sample table for quantitative assessment of the content of the scientific manuscript:

Table 2 provides indicative weightings for the different criteria against which the manuscript content will be assessed. These weightings reflect the relative importance of each criterion in the overall assessment and can be adapted to specific requirements. The table provides a structured framework for objective and quantitative assessment, ensuring clarity and consistency in the review process.

In the quantification table, the individual columns have the following meaning:

– *Criterion*: the criteria *represent* the different aspects of the manuscript that are being assessed.

– *Weighting coefficient*: The *weightings* reflect the importance of each criterion in the overall assessment. They are set in advance and summed should give 1 (or 100%).

– *ChatGPT score (1 – 5)*: the score that *ChatGPT* gives for each criterion based on the predefined evaluation criteria.

– *Weighted score*: The weighted score is calculated as the product of the weighting coefficient and the ChatGPT score. The weighted scores are summed to give the total quantitative score.

Table 2. Example weighting coefficients.

Criterion	Weighting coefficient	ChatGPT score (1-5)	Weighted score
Clarity and legibility	0.20		
Structure	0.20		
Methodology	0.20		
Data analysis	0.15		
Validity of results	0.15		
Discussion	0.10		
Originality and relevance	0.10		
Ethics and integrity	0.10		
Overall assessment			

D. Sample process of quantitative and qualitative assessment of manuscript content with the assistance of ChatGPT:

Given the functional capabilities of ChatGPT and the proposed algorithm to support the scientific manuscript review process, an example manuscript content evaluation process with the assistance of ChatGPT is presented. This process includes the following steps and elements that combine quantitative and qualitative evaluation:

- *Manuscript entry*: the reviewer enters the text of the manuscript into ChatGPT;
- *Initial discussion*: the reviewer starts a discussion with ChatGPT, asking questions on the different evaluation criteria;
- *Analysis and comments*: ChatGPT analyses the text and provides comments and recommendations for improvement on each criterion;
- *Criteria Score*: ChatGPT provides numerical scores on a scale of 1 to 5 for each criterion based on the data analyzed;
- *Review and corrections*: the reviewer reviews the ChatGPT comments and makes corrections where necessary;
- *SWOT analysis*: ChatGPT generates a SWOT analysis of the manuscript content;
- *Generate a qualitative score*: ChatGPT summarizes all comments and recommendations on the criteria in a tabular form;
- *Generate a numerical score*: ChatGPT adds up all the numerical scores and calculates the overall score for the manuscript;
- *Final evaluation*: the reviewer finalizes the quantitative and qualitative evaluation and prepares the final version of the review.

E. Example of formulating a question to ChatGPT to generate a review of a scientific manuscript:

When using ChatGPT to generate a review of a scientific manuscript, it is important to formulate a question that clearly defines the evaluation criteria and provides instructions for quantitative and qualitative evaluation. The question should be structured to include

all necessary aspects of review, such as clarity, structure, methodology, data analysis, validity of results, discussion, originality and significance, and ethics and integrity. This will allow ChatGPT to provide a detailed and objective review to assist the reviewer in assessing the quality of the manuscript. The following is an example of how a question to ChatGPT can be properly formulated to generate a review of a scientific manuscript:

Please generate review of the attached scientific manuscript using the following criteria:

1) *Clarity and legibility:*

– *Readability of the text:* is the text clearly written and easy to understand? Are there complex or unclear sentences that need to be rephrased?

– *Clarity of objectives and hypotheses:* are the research objectives and hypotheses clearly defined? Are there parts of the text where the objectives and hypotheses could be better stated?

– *Stylistic errors:* are there grammatical and spelling errors in the text? How often do they occur?

– *Terminological precision:* how accurately and consistently are scientific terms used?

– *Title relevance:* does the title reflect the content of the manuscript?

– *Abstract quality:* Does the abstract thoroughly reflect the ideology and content of the manuscript?

2) *Structure:*

– *Logical structure:* is the content coherently and logically ordered (introduction, methodology, results, discussion, conclusion)? Is there any part where the coherence of the presentation could be improved?

– *Connectedness of parts:* How do the different parts of the manuscript connect and complement each other? Are there transitions between sections that need to be improved?

– *Organization of information:* is there redundant information that can be removed or rearranged?

3) *Methodology:*

– *Description of methods:* Are the methods described in sufficient detail to be reproducible? Are there parts of the methodology that are insufficiently explained? Are the descriptions detailed enough for other researchers to reproduce the results?

– *Consistency of methods:* are the methods used adequate to achieve the objectives of the study? Are there methods that could be improved or replaced with more appropriate ones?

– *Innovativeness of approach:* to what extent are the methods used new and innovative for the relevant scientific field?

4) *Data analysis:*

– *Objectivity and accuracy:* is the data analysis objective and accurate? Are there biases or errors in the data analysis?

– *Statistical significance*: are the results statistically significant? Are the statistical tests correctly applied and results presented?

– *Graphical representation*: is the data presented visually in graphs, tables or charts and are they clear enough?

5) *Validity of results*:

– *Validity of results*: are the results clearly presented and easy to interpret? Are there any unsubstantiated claims or interpretations?

– *Arguing the hypotheses*: do the results argue for the propositions (hypotheses)? Are there inconsistencies between the results and the hypotheses?

– *Long-term relevance*: what are the long-term implications of the results for the scientific community?

6) *Discussion*:

– *Interpretation of results*: do the authors correctly interpret the results and relate them to the broader context of the study? Are there parts of the discussion that need to be improved or expanded?

– *Critical examination*: do the authors discuss the limitations of the study and potential sources of error? Are there gaps in the critical examination of the study?

– *Recommendations for future research*: do the authors provide specific recommendations for future research and how to address the identified limitations?

7) *Originality and relevance*:

– *Originality of the study*: Is the study original in the context of the existing literature? What is the contribution of the research to the relevant scientific field?

– *Practical relevance*: do the results have practical relevance and potential contribution to science or practice? Are there practical applications of the results that are not well represented?

– *Citation of relevant literature*: is the literature cited sufficiently relevant and what is the contribution to existing knowledge? At least 50% of the literature should have been published in the last 5 years?

8) *Ethics and integrity*:

– *Ethical considerations*: does the research comply with ethical standards, including ethics committee approval and informed consent? Are there potential ethical issues that have not been addressed?

– *Completeness of the report*: are all aspects of the study reported correctly and transparently? Are there any gaps or missing data in the report?

– *Negative results*: do the authors also report negative results or only positive results?

Quantitative assessment:

– Please generate me numerical scores for each of the above criteria and generate a table with weighting coefficients and weighted scores.

– The evaluation criteria have the following weights:

– Clarity and legibility – weighting coefficient 0.2;

- Structure – weight coefficient 0.2;
 - Methodology – weighting coefficient 0.2;
 - Data analysis – weighting coefficient 0.15;
 - Validity of results – weighting coefficient 0.15;
 - Discussion – weighting coefficient 0.1;
 - Originality and relevance – weight coefficient 0.1;
 - Ethics and Integrity – weighting coefficient 0.1.
- Each sub-criterion is scored on a scale of 1 to 5 as follows: 1. There are many gaps; 2. There are significant gaps; 3. There are small gaps; 4. There are small deviations; 5. There are no gaps.
- For each criterion, please form a numerical score from 1 to 5 and multiply this score by the weighting coefficient to obtain the weighted score. Present the scores in tabular form.
- Finally, add up all the weighted scores to get the total score.

3. SWOT analysis of the algorithm

In order to assess the effectiveness and feasibility of the proposed algorithm, a SWOT analysis was performed. This analysis examines the strengths, weaknesses, opportunities and threats associated with the use of the algorithm. The aim is to provide a reasoned assessment to help better understand the potential benefits and challenges associated with implementing the algorithm in the review process. Tables 3 shows a SWOT matrix of the proposed algorithm.

Table 3. SWOT matrix of the proposed algorithm

Strengths	Weaknesses
<i>Increased efficiency:</i> the algorithm significantly speeds up the review process by automating many of the steps that would otherwise take reviewer's considerable time.	<i>Technology dependency:</i> the algorithm is highly dependent on the functionality and quality of ChatGPT, which may lead to technical difficulties or limited accuracy of the analysis.
<i>Objectivity:</i> through ChatGPT, the algorithm provides more objective evaluations, reducing the influence of reviewers' personal preferences and biases.	<i>Limited adaptability:</i> despite the ability to adapt weightings, some specific needs of reviewers or requirements of publishers may not be fully met.
<i>SWOT Analysis:</i> The inclusion of a SWOT analysis provides additional perspective on the strengths and weaknesses of the manuscript as well as opportunities and threats that may arise.	<i>The human factor:</i> although the algorithm includes human intervention and corrections, automated comments and ratings may still require significant refinement by reviewers.

<i>Detailed analysis:</i> the algorithm includes both quantitative and qualitative assessments, providing a complete and detailed review of the manuscript.	
<i>Weighting:</i> The use of weighting for quantitative assessment adds an extra layer of precision and clarity to the assessment of different aspects of the manuscript.	
Opportunities	Threats
<i>Extension:</i> the algorithm can be extended and adapted to different scientific fields and publication types, which will increase its applicability and usefulness.	<i>Technological limitations:</i> limitations of current AI technologies and potential technical issues can affect the performance and accuracy of the algorithm.
<i>Algorithm improvement:</i> as AI technology evolves, the algorithm can be improved to provide even more accurate and relevant estimates.	<i>Ethical considerations:</i> the use of AI in review may raise ethical issues related to data confidentiality and reviewer autonomy.
<i>Training and adaptation:</i> the algorithm can be trained to recognize specific styles and requirements of different scientific journals, which will improve its accuracy and relevance.	<i>Reviewers opt-out:</i> Some reviewers may be resistant to the use of automated systems, preferring traditional review methods.

The SWOT analysis performed on the algorithm shows that the use of ChatGPT to support the review process has significant potential to improve the efficiency and objectivity of this process. Strengths include the increased efficiency, objectivity and detail of the analysis, as well as the ability to quickly identify key issues in the manuscript. The inclusion of SWOT analysis and weighting adds additional levels of precision and depth to the assessment. However, there are some weaknesses, such as the dependence on technology and the need for human verification and correction. Technological limitations and ethical considerations also pose threats that can affect the performance of the algorithm. On the other hand, the algorithm offers opportunities to extend coverage, improve accuracy as AI technologies evolve, and adapt to the specific needs of different scientific fields.

4. Conclusion

The study showed that the use of ChatGPT can significantly improve the efficiency, objectivity and detail of the peer review process. The proposed algorithm incorporates quantitative and qualitative evaluation steps, using weighting coefficients and SWOT analysis for better accuracy and depth of evaluation. The analysis showed that the algorithm can speed up the review process, reduce the influence of

personal preferences and biases, and provide detailed comments and recommendations.

Despite the significant advantages of the proposed algorithm, there are some limitations and disadvantages. The algorithm is highly dependent on the functionality and accuracy of ChatGPT, which may lead to technical difficulties or limited accuracy of the analysis. Furthermore, despite the ability to adapt the weighting coefficients, some specific reviewer needs or publisher requirements may not be fully met. Human verification and adjustments are still necessary, which may increase reviewers' time and effort.

The algorithm can be useful for a wide range of users in the scientific community. It is particularly needed for editors and reviewers who are faced with a large volume of manuscripts and are looking for a way to increase the efficiency and objectivity of the review process. In addition, the algorithm can be useful for authors who wish to obtain preliminary assessment and feedback on their manuscripts before formal submission for publication.

In conclusion, the proposed algorithm for supporting the review process with the assistance of ChatGPT offers significant advantages and has the potential to transform traditional review methods. Despite some challenges and threats, its implementation can lead to a more efficient, objective and detailed process of evaluating scientific publications. Further research and refinement of the algorithm can improve its accuracy and adaptability, providing even greater value to the scientific community.

REFERENCES

- TENNANT, J.P., ROSS-HELLAUER, T., 2020. The limitations to our understanding of peer review. *Res Integr Peer Rev*, vol. 5, no. 6. <https://doi.org/10.1186/s41073-020-00092-1>.
- SCHULZ, R.; BARNETT, A.; BERNARD, R., et al., 2022. Is the future of peer review automated? *BMC Res Notes*, vol. 15, no. 203. <https://doi.org/10.1186/s13104-022-06080-6>.
- TURNITIN, 2023. Turnitin turns on AI writing detection capabilities for educators and institutions (NA and EMEA). [Online]. Available: <https://www.turnitin.com/press/turnitin-turns-on-ai-writing-detection-capabilities-for-educators-and-institutions>. [Accessed: Jun. 28, 2024].
- TURNITIN, 2024. How does iThenticate work? Tools for advancing research integrity. [Online]. Available: <https://www.turnitin.com/blog/how-does-ithenticate-work-tools-for-advancing-research-integrity>. [Accessed: Jun. 28, 2024].
- IRVINE, A., 2024. We Review 'Semantic Scholar': An AI-Powered Literature Searching Tool. Thesis link, Auckland University of Technology. [Online]. Available: <https://thesislink.aut.ac.nz>. [Accessed: Jun. 28, 2024].

- NANDWANI, P.; VERMA, R., 2021. A review on sentiment analysis and emotion detection from text. *Soc Netw Anal Min*, vol. 11, no. 1, p. 81. <https://doi.org/10.1007/s13278-021-00776-6>.
- ALRASHEEDY, M.N.; MUNIYANDI, R.C.; FAUZI, F., 2022. Text-Based Emotion Detection and Applications: A Literature Review. 2022 International Conference on Cyber Resilience (ICCR), Dubai, United Arab Emirates, pp. 1-9. <https://doi.org/10.1109/ICCR56254.2022.9995902>.
- KUSAL, S.; PATIL, S.; CHOUDRIE, J.; KOTTECHA, K.; VORA, D.; PAPPAS, I., 2022. A Review on Text-Based Emotion Detection – Techniques, Applications, Datasets, and Future Directions. arXiv. [Online]. Available: <https://arxiv.org/abs/2205.03235>.
- HOSSEINI, M., HORBACH, S.P.J.M., 2023. Fighting reviewer fatigue or amplifying bias? Considerations and recommendations for use of ChatGPT and other large language models in scholarly peer review. *Res Integr Peer Rev*, vol. 8, no. 4. <https://doi.org/10.1186/s41073-023-00133-5>.
- GAO, C.A., HOWARD, F.M., MARKOV, N.S., DYER, E.C., RAMESH, S., LUO, Y., et al., 2022. Comparing scientific abstracts generated by ChatGPT to original abstracts using an artificial intelligence output detector, plagiarism detector, and blinded human reviewers. bioRxiv. [Online]. Available: <https://www.biorxiv.org/content/10.1101/2022.12.23.521610v1>. [Accessed: Jan. 31, 2023].
- RASHIDOVA, F., 2023. Determining priorities of indicators at the examination timetabling in a higher education institution for part-time students. *14th International Conference on Electrical and Electronics Engineering (ELECO)*, Bursa, Türkiye, pp. 1 – 4, Available from: <https://doi.org/10.1109/ELECO60389.2023.10415956>.

✉ **Dr. Aldeniz Rashidov, Assoc. Prof.**
ORCID iD: 0000-0003-4946-890X
Technical University of Gabrovo
Gabrovo, Bulgaria
E-mail: aldeniz@tugab.bg