

REVIEW

by **Assoc. Prof. Margarita Nikolaeva Terziyska, PhD**

University of Food Technologies – Plovdiv

of a dissertation submitted for the award of the educational and scientific degree **Doctor (PhD)**

Field of higher education: **4. Natural Sciences, Mathematics and Informatics**

Professional field: **4.6 Informatics and Computer Science**

Doctoral programme: **Informatics**

Author: Ivan Marianov Ivanov

Dissertation title: *Model and System for the Automated Generation and Verification of Test Questions Using Artificial Intelligence*

Scientific supervisors: Prof. Stanka Ivanova Hadzhikoleva, PhD, and Prof. Georgi Petrov Dimitrov, PhD

1. General description of the submitted materials

By Order No. RD-22-1413 of 17 July 2026 issued by the Rector of the University of Plovdiv “Paisii Hilendarski”, I was appointed a member of the scientific jury in the procedure for the defence of the dissertation submitted by Ivan Marianov Ivanov, a full-time doctoral student at the Department of Computer Informatics, Faculty of Mathematics and Informatics. At the scientific jury’s remote meeting held on 20–21 July 2026, I was assigned to prepare a review.

The set of materials submitted in hard copy by Ivan Marianov Ivanov complies with Article 36(1) of the Regulations for the Development of the Academic Staff of the University of Plovdiv and includes the following documents:

- application to the Rector of the University of Plovdiv for opening a dissertation defence procedure;
- curriculum vitae in European format;
- minutes of the Department Council concerning the report on readiness to open the procedure and the preliminary discussion of the dissertation;
- dissertation;
- abstract of the dissertation;
- list of scientific publications on the dissertation topic;
- copies of the scientific publications;
- list of identified citations;
- declaration of originality and authenticity of the submitted documents;
- statement of compliance with the specific requirements of the Faculty of Mathematics and Informatics.

The dissertation comprises 144 pages and includes an introduction, four chapters, a conclusion, a bibliography and an appendix. The exposition is illustrated with 50 figures and 19 tables. The bibliography contains more than 80 sources, predominantly in English.

2. Brief biographical information about the doctoral candidate

Ivan Marianov Ivanov was born in 1995. He completed a Bachelor's degree in Software Technologies and Design in 2018 and a Master's degree in Software Technologies and Design, specialising in Software Architectures and Tools, in 2019 at the University of Plovdiv "Paisii Hilendarski". From 2023 to 2026, he was a full-time doctoral student in the doctoral programme in Informatics.

Since 2018, he has worked in the field of information technology as an application software developer. The competencies stated in his curriculum vitae include C#, TypeScript, JavaScript, SQL and Java, as well as .NET, ASP.NET, Angular, Entity Framework Core, RESTful APIs, MS SQL Server and microservices. His professional profile is directly related to the software implementation and architectural aspects of the dissertation research.

3. Relevance of the topic and appropriateness of the stated aims and tasks

The topic is unquestionably relevant. The rapid adoption of large language models in education creates opportunities for the automated generation of learning and assessment content, while simultaneously raising issues concerning factual accuracy, pedagogical suitability, reproducibility, transparency and teacher oversight. The development of an integrated approach combining generation, independent automated verification, management of test content and adaptive assessment has clear scientific and applied significance.

The main aim—to develop a model, architecture and software implementation of an intelligent electronic testing platform—is clearly formulated. The five stated tasks encompass the analysis of existing solutions, the derivation of requirements, the creation of a conceptual model and architecture, the development of an adaptive module, prototype implementation and experimental evaluation. There is logical consistency between the aim, the tasks and the structure of the dissertation.

4. Knowledge of the research problem

The doctoral candidate demonstrates good knowledge of the current state and development of the research problem. The theoretical review covers the main areas related to the application of artificial intelligence in education, the automated generation and assessment of test questions, large language models, electronic and adaptive testing, and item response theory (IRT). Existing software platforms and technological solutions are also reviewed, and their functional characteristics and limitations are identified.

The bibliography comprises a total of 89 scholarly and information sources. Of these, 43 date from 2022–2026; therefore, approximately 48% of all sources used were published or updated within the last five years. This demonstrates an effort to account for current developments in an exceptionally dynamic scientific field. In addition to recent publications, earlier foundational studies in automated question generation, computer-based testing and adaptive testing are included. Most sources are in English, and only one source is in Bulgarian. Although this is understandable given the international nature of the topic and the leading role of English-language scientific literature, broader coverage of research and practical developments in the Bulgarian educational context would have enriched the analysis.

The literature review enables the doctoral candidate to identify an underexplored aspect: the absence of an integrated solution combining automated test-question generation, independent multi-model verification, test-content management and adaptive-testing mechanisms. On the basis of this analysis,

the need to develop the proposed model and software system is substantiated. Overall, it may be concluded that the doctoral candidate possesses the requisite knowledge of the scientific problem and of the current state of research in the field under consideration.

5. Research methodology

The methodology combines systems analysis, conceptual and architectural modelling, software design, prototyping and experimental evaluation. A human-in-the-loop approach is implemented, in which expert assessment is used as a reference against which the language-model assessments are compared. Four combinations of models acting as generator and evaluator are considered, and the generated questions are assessed according to predefined indicators and criteria.

The selected methodology is appropriate for achieving the stated aim and for examining the technical operability and applicability of the proposed model. The developed software prototype was experimentally tested using 72 generated questions based on the learning content of one discipline and evaluated by an expert. The results confirm the operability of the developed solution and indicate its potential for practical application.

The experimental validation is of a pilot nature and identifies opportunities for subsequent extension by including more academic disciplines, learning-content sources and independent experts, as well as by applying measures of agreement, statistical tests and analyses of result repeatability. These opportunities do not call into question the appropriateness of the selected methodology or the results obtained; rather, they define directions for future, larger-scale validation of the proposed model.

6. Characterisation and evaluation of the dissertation

The first chapter establishes the theoretical foundation of the research. It systematises the applications of artificial and generative intelligence in education, the evolution of electronic testing and the requirements for an intelligent platform. The chapter logically identifies the research niche: the absence of a comprehensive solution integrating generation, independent verification, management and adaptive testing.

The second chapter presents the conceptual model and architecture. User roles, functionalities, modules and use cases are defined. A notable strength is the consideration of multi-model verification, whereby the generator model and evaluator model may be different. The UML diagrams facilitate the tracing of processes and interactions.

The third chapter demonstrates the practical feasibility of the concept. It describes the technology stack, data model, REST API, real-time monitoring, integration with external and local AI models, robust processing of structured responses, and mechanisms for adaptive testing through IRT. The implemented prototype is a significant applied result, and the presented algorithms ensure traceability of the principal design decisions.

The fourth chapter contains a pilot experimental study. Four research questions are formulated, and model self-assessment, cross-assessment and expert assessment are compared. All final expert scores exceed 4.30 on the five-point scale; the highest expert score, 4.67, is assigned to questions generated by GPT. The results indicate highly rated question quality in the specific experimental setting, while also revealing dependence on the choice of evaluator model. The author appropriately distinguishes the risks and limitations.

Overall, the dissertation is logically structured, technically substantive and oriented towards a real problem. The credibility of the contributions is supported by the described architecture, the implementation of the prototype and the pilot experiments conducted. Some of the broader claims regarding increased efficiency, objectivity and personalisation, however, extend beyond what was directly measured in the experiment and should be regarded as prospectively supported rather than conclusively demonstrated.

7. Contributions and significance of the work for science and practice

I accept the principal contributions formulated by the author, namely:

Scientific contributions

- A conceptual framework for an intelligent electronic testing platform based on generative artificial intelligence has been created.
- An approach to the multi-model verification of test questions using independent AI models has been proposed.

Scientific and applied contributions

- A model of an intelligent platform for the automated generation of test questions using generative language models and parameterised prompts has been developed.
- A mechanism for the automated assessment of the quality of generated questions through independent AI verification has been proposed.

Applied contributions

- A software prototype of an intelligent electronic testing platform has been implemented, including functionalities for the automated generation, verification and management of test questions.
- An experimental study confirming the operability and initial applicability of the developed solution has been conducted.

The significance of the work lies primarily in integrating separate technologies and methods into a unified, practically implementable environment that provides for human oversight of generated and assessed content. The developed solution has potential for application in higher education and other forms of e-learning. Extending the empirical and psychometric validation would enable its effectiveness to be evaluated across different academic disciplines, groups of learners and conditions of use.

8. Assessment of the publications related to the dissertation

Six publications on the dissertation topic have been submitted. Three are indexed in Web of Science and/or Scopus. One article was published in *Applied Sciences* (Web of Science—Q2, IF 2.5; Scopus—Q2, SJR 0.52), and two further publications appeared in indexed proceedings of international scientific conferences. The remaining three publications appeared in national scientific outlets. The doctoral candidate is a co-author of all publications and the first author of three of them.

The publications thematically reflect the progressive development of the dissertation research—from the analysis of artificial-intelligence systems and approaches to automated test-question creation to practical implementation using large language models. According to the submitted statement, the

doctoral candidate has obtained 96 points under Group G indicators, compared with the required minimum of 30 points; thus, the minimum national requirements have been substantially exceeded.

A list of 29 identified citations of two publications related to the dissertation topic has been submitted. Of these, 25 refer to the article in *Applied Sciences*, and four to the publication devoted to the automated creation of educational questions.

9. Personal contribution of the doctoral candidate

The dissertation, the declaration of originality, the submitted software implementation and the publications provide grounds for concluding that the doctoral candidate made a substantial personal contribution to the development of the model, architecture, software prototype and experimental verification. His professional experience as a software developer is clearly evident in the quality and level of detail of the technical implementation. Although the individual contribution is not described separately in the co-authored publications, their thematic continuity and the candidate's authorship positions support the conclusion that he participated actively.

10. Abstract of the dissertation

The Bulgarian-language abstract is 32 pages long and presents the main components of the dissertation: relevance, aim, tasks, hypothesis, structure, conceptual model, software implementation, experimental results, contributions and dissemination of the results. An English-language abstract has also been submitted. Its content is consistent with the dissertation and provides a sufficiently clear account of the results achieved. The contributions are accurately reproduced.

11. Critical comments and recommendations

I have no substantial critical comments regarding the dissertation. As a recommendation for the doctoral candidate's future research activity, I would suggest that some subsequent results also be presented in sole-authored publications, which would more clearly highlight his personal contribution and scientific independence.

I would also like to pose the following question to the doctoral candidate: **What is the advantage of using different language models in the roles of generator and evaluator of test questions?**

12. Personal impressions

I do not know the doctoral candidate personally and therefore have no direct impressions of his educational, research or professional activities. The present assessment is based entirely on the submitted set of materials and the results presented in the dissertation.

13. Recommendations for the future use of the dissertation contributions and results

I recommend that the developed platform be piloted across a wider range of academic disciplines and educational institutions, involving teachers with different professional profiles and a larger number of learners. Promising directions for further development include integration with Moodle and other learning management systems, automated distractor analysis, and the combination of large-language-model-based assessment with classical psychometric indicators. This would broaden the applicability of the work and create conditions for a more comprehensive evaluation of its effectiveness in a real educational environment.

Conclusion

The dissertation contains scientific, scientific and applied, and applied results that constitute an original contribution to the field of intelligent electronic testing systems. The work meets the requirements of the Development of Academic Staff in the Republic of Bulgaria Act, the Regulations for its implementation, and the Regulations for the Development of the Academic Staff of the University of Plovdiv “Paisii Hilendarski”.

The work demonstrates that Ivan Marianov Ivanov possesses in-depth theoretical knowledge and professional skills in the doctoral programme in Informatics and is capable of independently conducting scientific and applied research, developing a complex software system and critically interpreting the results.

In view of the above, I give a positive assessment of the dissertation and propose that the esteemed scientific jury award Ivan Marianov Ivanov the educational and scientific degree **Doctor (PhD)** in Field of higher education 4. Natural Sciences, Mathematics and Informatics, Professional field 4.6 Informatics and Computer Science, Doctoral programme in Informatics.

17 August 2026

Reviewer:

(signature)

Assoc. Prof. Margarita Terziyska, PhD