Implementation of the Sequential Analysis of Variants Scheme using Custom GPTs for Verification of Student Scientific Competitions Documentation*

Iryna Domanetska[1,*,†], Vitaliy Tsyganok [1,2,†], Yaroslav Khrolenko[2,†]

[1] Taras Shevchenko National University of Kyiv, 64/13, Volodymyrska str., Kyiv, 01601, Ukraine
[2] Institute for Information Recording of NAS of Ukraine, 2, Mykoly Shpaka str., Kyiv, 03113, Ukraine

Abstract

The article investigates the use of custom large languages models to automate the checking of student research papers. A hybrid model is proposed, combining the method of sequential variant analysis with the capabilities of specialized GPTs. This approach allows automating the process of checking the compliance of documents with the formal requirements of the competition. Specialized GPTs is designed to automate document verification. The experiment involved comparing local implementations of GPTs on different models (Llama 3, LLaVA, Phi-3) and evaluating their accuracy, processing speed, and relevance of detected errors in order to select the best base model. Testing of the hybrid approach using custom GPTs developed from the Llama 3 model demonstrated a significant reduction in verification time compared to expert evaluation while maintaining high accuracy. The effectiveness and efficacy of the hybrid approach have been experimentally proven. The combination of the capabilities and advantages of LLM with the logic of sequential analysis within a single approach makes it perspective for digitalizing the competitive process, increasing its transparency and scalability.

Keywords

document verification automation, student research competitions, custom GPTs, LLM, sequential analysis of variants.

1. Introduction

In recent decades, there has been a significant rise in the popularity of student research competitions, which play a crucial role in fostering research skills, interdisciplinary approaches, and creative thinking. These events not only promote the generation of new ideas but also provide students with a platform for knowledge exchange, collaboration with experts, and professional recognition. However, the increasing number of participants, the growing demands for high-quality evaluation, and the need to ensure transparency in competition processes pose several challenges for organizers, particularly concerning the efficiency of document management and handling large volumes of information.

The widespread adoption of digital technologies has partially addressed these issues by transitioning application submission, evaluation, and communication between participants to online formats. Nevertheless, even with the use of existing platforms such as EasyChair or Submittable, a considerable degree of human intervention remains necessary for tasks such as validating submissions, assigning papers to reviewers, verifying academic integrity, and compiling final scores. Traditional approaches to managing competition documentation, which rely on manual verification

*International Workshop on Computational Intelligence, co-located with the IV International Scientific Symposium "Intelligent Solutions" (IntSol-2025), May 01-05, 2025, Kyiv-Uzhhorod, Ukraine

[1*] Corresponding author.
domanetska@knu.ua (I. Domanetska); vitaliy.tsyganok@gmail.com (V. Tsyganok); yaroskhr@gmail.com (Ya. Khrolenko)
0000-0002-8629-9933 (I. Domanetska); 0000-0002-0821-4877 (V. Tsyganok); 0009-0004-0641-827X (Ya. Khrolenko)

and administrative oversight, are becoming increasingly inefficient given the vast amount of data that must be processed.

One promising solution is the integration of intelligent systems, particularly generative pre-trained transformers (GPTs). These technologies have the potential to optimize key stages of competition document management: automating the preliminary verification of submissions against formal criteria, distributing papers to reviewers based on semantic topic analysis, conducting initial plagiarism checks, and even generating analytical reports for organizers. Furthermore, the integration of GPT-powered assistants into competition platforms could enhance participant communication by providing automated responses to inquiries and assisting with document formatting.

In this context, exploring the application of GPTs for automating competition document verification is of particular importance. The implementation of such technologies would not only enhance document processing efficiency and reduce human bias in evaluations but also make competition processes more transparent, accessible, and scalable. Given the ongoing digitalization of the academic environment and the increasing standards for research quality, the integration of intelligent technologies into the automation of student`s research competition document management is not merely desirable but essential for improving their efficiency and global accessibility.

2. Materials and Methods

Student research competitions play an important role in developing a research culture, stimulating innovative thinking, and preparing young scientists for the challenges of the academic environment. Due to the development of digital technologies and the growing number of participants, the issue of effective management of competition processes and document flow is becoming extremely relevant. Modern research shows that automation of such processes contributes to the efficiency of evaluation, transparency of competitions, and reduces the administrative burden on organizers.

Recent studies confirm that the transition to digital document management significantly improves the organizational aspects of scientific competitions [1]. The study by Markus [2] emphasizes that electronic document management minimizes the need for paper documents, increases the speed of processing applications, and ensures their safe storage.

The integration of artificial intelligence and text recognition plays a key role in improving the verification and analysis of submitted papers. Research [3] presents the OCR-D platform, which provides automated text recognition and pre-processing, which greatly simplifies the evaluation of applications.

Modern specialized competition management systems, such as EasyChair, Submittable, and others, significantly optimize the evaluation process. Paper [4] proposes a model of a specialized platform that allows automating certain stages of the competition. The platform is a comprehensive web application. The results of the study show that the use of specialized platforms can significantly reduce the time spent on organizational processes, which is especially important for large international competitions.

Additionally, the study by [5] demonstrates that the use of machine learning algorithms allows for thematic modeling of documents and automatic creation of metadata for digital archives, which improves the organization of competition information.

Automation of document management contributes to the transparency of tenders, in particular through the introduction of electronic signatures and access control technologies [6]. In addition, the introduction of algorithms for checking academic integrity, such as Turnitin and Copyscape, ensures compliance with the principles of scientific ethics and prevents plagiarism [7, 8].

Automating the workflow of student's research paper competitions is important for increasing the efficiency, scalability, and fairness of competitive processes. The use of digital platforms, artificial intelligence tools, and specialized competition management systems can significantly reduce administrative burdens, increase the speed and quality of evaluation, and ensure transparency and reliability of results. However, it is important to maintain a balance between technology and the human factor, as individual approach and expert judgment remain important elements of research evaluation.

2.1. Procedure for conducting a student research paper competition

The authors' analysis of modern technologies for organizing and conducting competitions [9] showed that most competitions that consider scientific creative works or projects are characterized by a unified technology for conducting the competition.

Let's consider the procedure of the competition on the example of the All-Ukrainian competition of student research papers in the fields of knowledge and specialties [10]. The generalized scheme of interaction between all participants of the competition is presented in Fig. 1.

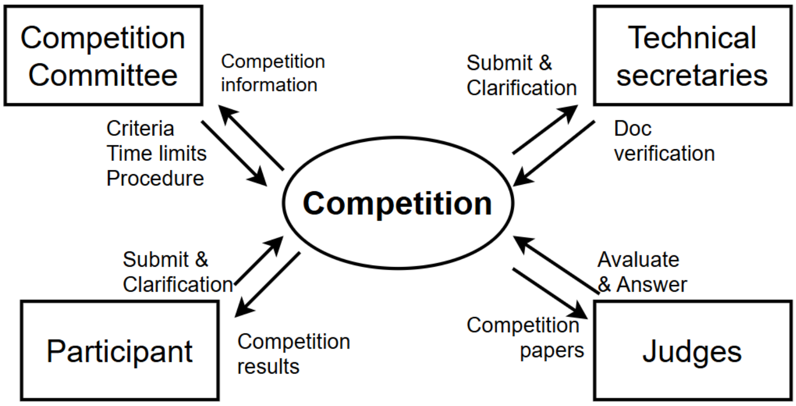

Figure 1: Actors of the competitive process

The Competition Committee performs a central function in the organization of the competition. It forms the basic rules:

- Evaluation criteria - for example, scientific novelty, methodology, practical value, quality of presentation.

- Time limits - deadlines for submitting papers, deadlines for checking documents, and the period of evaluation by judges.

- Procedure - the stages of the competition (submission, moderation, evaluation, appeals), methods of communication with participants.

- Information about the competition - publishing the rules on the website, sending them out via social media or emails.

The Committee accumulates all information about the competition. The raw data is collected, processed, analytical reports are generated (dynamics of participation compared to previous years, average scores by criteria, etc.), and visualization is performed: interactive graphs, heat maps of activity, score distribution charts.

Technical secretaries provide administrative support. Their tasks are to interact with participants to clarify details, organizational issues, and document participation in the competition.

Participants are students or researchers who submit papers to the competition. Participants of the competition submit documents through an online platform (for example, Google Forms, a specialized portal). After receiving confirmation of the correctness of the submitted documents, participants wait for the results of the competition.

The judges are industry experts responsible for the evaluation. The judges' responsibilities include checking the entries for compliance with the formal requirements of the competition, evaluating the entries according to certain criteria, formulating conclusions and comments on the advantages and disadvantages of the entries, and making recommendations.

The competitive selection technology includes several stages [Regulations on the All-Ukrainian]. Stage 1: "Submission of works". Students submit their research papers in compliance with the established conditions and deadlines.

Stage 2: "Initial verification". Technical secretaries check the works for compliance with formal requirements, such as compliance with the competition theme, volume, formatting, plagiarism check, etc. Works that do not meet the competition requirements are rejected.

Stage 3: Reviewing. Experts evaluate each entry according to the competition's criteria and review them. Each competition obviously has its own evaluation criteria that take into account its specifics, specialty, and focus.

Based on the results of the review, the Competition Committee forms and publishes a ranking list of scientific papers (hereinafter referred to as the ranking list).

The competition committee makes a decision on the selection of the best scientific papers, the authors of which will be invited to the final scientific and practical conference.

Stage 4: Presentation of works. At the final scientific and practical conference, students present their research, answer questions, and provide explanations.

Stage 5: Evaluation and determination of winners. The jury evaluates the papers and presentations, determining the winners.

Stage 6: Awarding the winners. Announcement of results and presentation of diplomas or awards.

The stage of initial verification of competition documents is essentially a kind of "weeding out" of works that violate the requirements of the competition. The main idea of this study is to implement the procedure for the initial verification of competitive documents as a scheme for sequential analysis of options, proposed by Ukrainian scientists Mykhalevych and Voloshyn.

2.2. Procedure for holding a student competition of research papers

This methodology is one of the most general approaches to solving multivariate problems, and its successful application to solving problems of research and design of complex systems formalized in classes of large-scale mathematical programming models has been widely confirmed [11, 12].

The method of sequential analysis of variants represents the process of finding a solution to a multivariate problem in the form of a multi-stage structure that resembles the structure of a complex experiment. Each step of the method is associated with checking the presence of certain properties in a subset of variants or individual variants and leads either to a direct reduction of the set of variants or prepares the possibility of reduction in the future [13]. Figure 2 shows a generalized scheme of the PAV method.

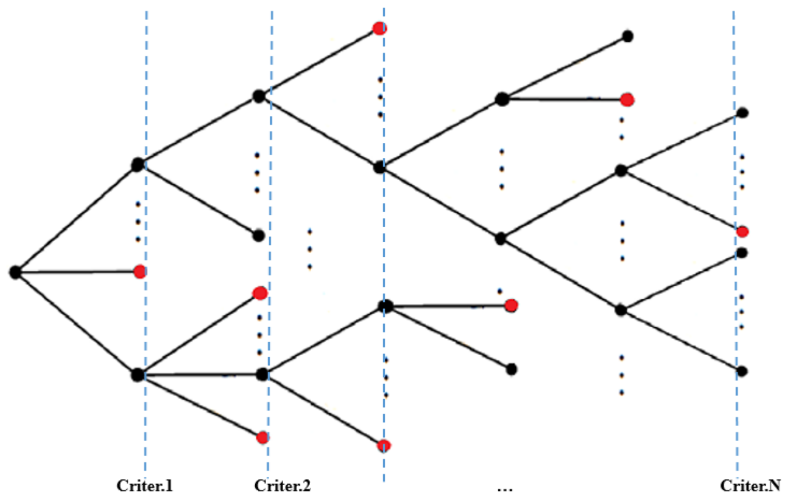

Figure 2: Scheme of the sequential analysis of variants method

Only variants that have certain properties (attributes) are selected for consideration in the next steps of solving the problem. The red vertices are the variants that turned out to not meet the criterion of the current step of the method (they do not have a certain feature defined by the criterion of the current step). It is due to such exclusion from the consideration that the dimensionality of the problem is reduced.

From the point of view of formal logic, the scheme of sequential analysis of variants is a development of A. Wald's sequential analysis and is reduced to repeating the sequence of actions

- dividing the set of decision variants and the problem into a family of subsets, each of which has additional specific properties;
- using these specific properties to find logical contradictions in the description of individual subsets;
- exclusion from further consideration of those subsets of variants whose descriptions contain logical contradictions.

To solve a specific problem based on its theoretical and practical analysis, it is necessary to formalize the properties that the desired variants should have. Then, it is necessary to identify as many features as possible to determine whether this variant is the desired one. Among them, choose those that are easily verifiable, as well as those that are inherent in as many variants as possible at the same time. Further, the choice of a calculation scheme consists in establishing a rational order of checking features, which allows to eliminate non-competitive options and find the optimal one [14].

Thus, the PAV methodology is based on an approach to forming a set of possible solutions and establishing criteria for their evaluation that allows for the elimination of unsuitable variants at an early stage without the need for their full consideration. The selection process occurs incrementally upon detecting non-compliance with the criteria, thereby preventing unnecessary computations. By eliminating all potential extensions alongside unsuitable variants, this approach drastically minimizes the computational workload required.

The PAV methodology can be used to implement the procedure for organizing and conducting competitions, since a number of requirements are imposed on the works submitted for the competition, which can be considered as criteria for eliminating options.

2.3. LLM for document processing

Large language models (LLMs) have become widely used in document automation because they are capable of performing tasks of analyzing, generating, and processing text with high accuracy. In recent years, researchers have been actively exploring ways to use LLMs in digital document management systems to increase efficiency, accuracy, and transparency of processes [15].

One of the key capabilities of LLM is automated text processing, which allows you to recognize, classify, and structure documents without manual intervention. For example, the LLM4Workflow system [16] offers automatic generation of workflow models, which can significantly reduce the time for processing large amounts of data in the document flow.

Another important aspect of LLM is the intelligent verification of documents. The study [17] demonstrates how such models can integrate into administrative processes by extracting key information, classifying documents, and identifying errors in the data submitted. The use of LLMs in document management enables automated verification of compliance with formal criteria.

A separate area is the automatic creation and summarization of documents. Paper [18] present a system that uses LLM to automatically abstract large text arrays, which greatly facilitates the process of reviewing documents and preparing reports. This is especially useful in academic document management, where experts often have to work with large volumes of scientific papers.

LLMs are often integrated with optical character recognition technologies process document images. For example, the ERPA system [19] combines LLM with OCR to process various document formats, allowing for automatic extraction of textual information and verification of its compliance with regulatory requirements.

Large Language Models now effectively address the task of automated plagiarism detection in academic and professional documents [20]. Unlike traditional systems such as Turnitin or Copyscape, which focus on exact textual matches, LLMs are able to analyze the content on a deeper level. They allow analyzing the semantics of the text, recognizing paraphrasing, finding conceptual coincidences, and assessing the compliance of research papers with the criteria of academic integrity. The conclusions of these studies show that LLM is a powerful tool for automating natural language processing and document management tasks.

Large language models are based on training on massive textual data, which allows them to analyze, generate, and interpret language with impressive accuracy. These models have universal skills, from writing creative texts to answering complex scientific questions. However, their “generality” often becomes a limitation for highly specialized tasks. That is why custom GPTs come in - adapted versions of basic models optimized for specific needs. Large models serve as a foundation: their “experience” gained from training on diverse data allows them to be quickly customized for specialized purposes. For example, the GPT-3 model can be fine-tuned using highly specialized datasets (medical records, legal documents, technical documentation) for better understanding the context of a particular industry. An alternative approach is prompt engineering, where the model is “taught” to perform tasks without changing its internal parameters through specially formulated questions or instructions [21]. Custom GPTs are widely used. The benefits of customization include increased accuracy, reduced query processing time, and lower costs compared to developing models from scratch. However, despite all the advantages, their use in document management has its challenges. The main problems include the need for large computing resources, the issue of trust in automatically generated answers, and the need to ensure data confidentiality. It is important to keep in mind the balance between automation and human control to ensure high quality and fairness of document evaluation.

2.4. Hybrid model of the task of automating the preliminary verification

The most time-consuming stage in the process of organizing and conducting f competition is the initial verification of the submitted entries. There is no point in expert evaluation if the work does not meet the requirements of the competition.

Combining the Sequential Analysis of Variants (SAV) method with the capabilities of custom GPTs for the task of preliminary verification of competition documents can significantly increase efficiency and automate one of the most labor-intensive stages of competition organization. The initial verification of documents is critical, it serves as a filtering mechanism, eliminating entries that do not meet the requirements, thereby saving experts’ time and improving the overall quality of the competition process.

Embedding GPT in a sequential option analysis scheme allows for a structured and automated step-by-step selection process, where each stage of analysis reduces the set of submitted documents by eliminating those that do not meet the established criteria.

The basic idea is that specialized GPTs act as smart filters that analyze documents in several stages, eliminating inappropriate options. This process can be represented as a multi-stage procedure, where each stage corresponds to a specific evaluation criterion.

Let D be the initial set of all documents submitted to the competition:

$$D=\{d_1, d_2, \ldots, d_n\}.$$

The set of selection criteria K consists of a set of mandatory checks that each document must pass:

$$K=\{k_1, k_2, \ldots, k_m\},$$

where

k_j is a specific criterion (e.g., checking for format compliance, academic integrity, presence of mandatory sections, etc.)

A set of GPTs modules:

$$GPTs=\{gpt_1, gpt_2, \ldots, gpt_m\},$$

where

gpt_j is a custom GPT that meets the criterion k_j and implements the verification of the document for compliance with this criterion.

The preliminary verification of competition documents consists of the following stages:

$$E=\{e_1, e_2, \ldots, e_m\},$$

where

e_j is the j-th stage of verification, which uniquely corresponds to the pair <criterion-GPTs>.

$$e_j = ¿\, k_j, gpt_j > ¿$$

The process of step-by-step verification can be represented as a sequence of rejection operators F_j, which are implemented by a custom GPT at each stage e_j:

$$F_j D_j \rightarrow D_{j+1}, D_{j+1} \subseteq D_j$$

where:

- D_j is the set of documents that have passed the jjjth filtering stage,
- F_j is a dropout operator that applies the criterion k_j to all elements of the set D_j, implemented by the corresponding GPTs gpt_j
- D_{j+1}- a subset of documents that have been checked according to the criterion k_j.

The process is completed at step m, after which the set D_m containing the documents admitted to further evaluation by experts remains:

$$D_m = F_m \circ F_{m-1} \circ \ldots \circ F_1(D)$$

Each operator F_j is implemented through a check function that returns a binary result for each document d_i, changing its current status $s_{i,j}$:

$$s_{i,j} = \begin{cases} F_j(d_i), \mathit{if\ di\ satisfies\ the\ criterion}\, k_j \\ 0, \mathit{if}\ d_i\, \mathit{does\ not\ meet\ the\ criterion}\, k_j \end{cases}$$

$\mathit{so}\ s_{i,j} \in \{0,1\} - \mathit{status\ of\ document\ di\ at\ stage}\ e_j$:

$$s_{i,j} = 1 : \mathit{the\ document\ has\ passed\ the\ stage}$$
$$s_{i,j} = 0 : \mathit{the\ document\ has\ not\ passed\ the\ stage}.$$

Feedback function $b_j : D \rightarrow R$, where R is the set of error reports.

$$b_j(d_i) = \mathit{GPTs\ report\ describing\ violations}, \mathit{if}\ F_j(d_i) = 0$$

Formally, the algorithm for passing through the verification stages is as follows:

Initialization: $s_{i,1} = F_1(d_i)$

Recurrent rule: For $j \geq 2$:

$$s_{i,j} = \begin{cases} F_j(d_i), \mathit{if}\ s_{i,j-1} = 1 \\ 0, \mathit{otherwise}. \end{cases}$$

Final state of the document:

$$S(d_i) = \prod_{j=1}^{K} s_{i,j}$$

Document d_i is allowed to participate in the competition, if $S(d_i) = 1$.

The model has the following properties.

Monotonicity: if $s_{i,j}=0$, then $s_{i,m}=0$ for all m≥ j.

Composability: Each stage independently processes the document, but the result depends on the previous stages.

Adaptability: Replacing F_j allows you to update the criteria without changing the architecture.

The principle of the model's functioning is shown in the figure below (Fig.3).

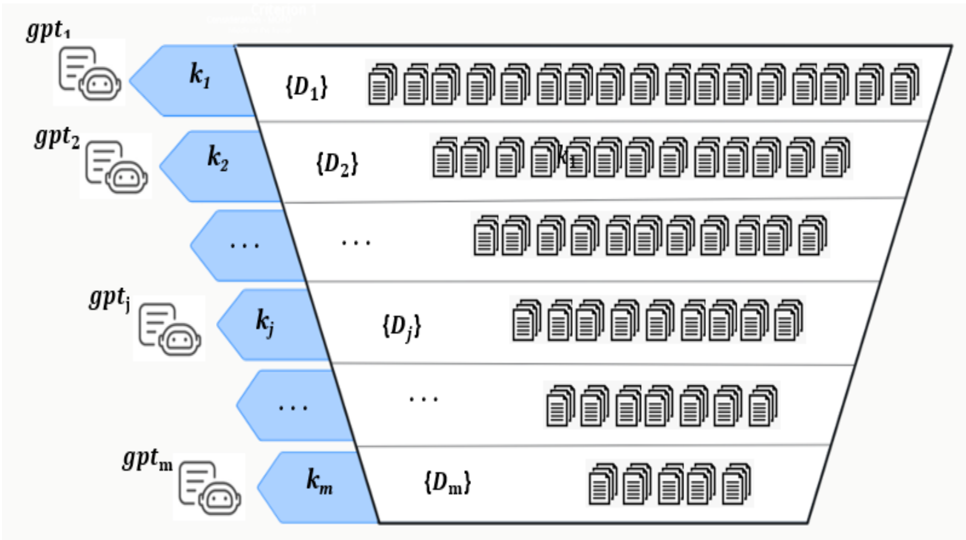

Figure 3: Generalized scheme of the hybrid method

2.5. Information model of a custom GPTs

Custom GPT models are complex adaptive systems tuned based on a set of parameters that determine their performance, functionality, and learning ability for specific tasks. The effectiveness of such a model depends on a coordinated choice of its parameters and configurations that affect accuracy, processing speed, security, and integration with external information environments.

To solve the problems of automated document processing, the approach of building custom GPTs based on industrial engineering is more appropriate, as it allows flexible and efficient adaptation of LLMs to specific tasks without the need for resource-intensive retraining.

The generalized information model of the custom GPTs can be represented as a system of sets:

$$GPTs=\langle Mb, Cs, Sc, Is \rangle,$$

where:

Mb – set of parameters of the base model;

Cs – set of contextual settings;

Sc – set of security constraints;

Is – set of integration settings.

The central element of a custom GPT is the base model M, which determines its ability to recognize, analyze, and generate text. The choice of the model determines the amount of knowledge and quality of the model's answers, contextual understanding of the text and its coherence, and performance in tasks requiring in-depth analysis.

Components of the basic model are

$$Mb=\langle Ua, Lcw, Ml \rangle,$$

where:

Ua – is the architecture of the underlying model (e.g., Llama 3, GPT-4, Mistral 7B, etc.) that determines the computational complexity;

Lcw — is the length of the context window that limits the maximum amount of text to be analyzed (e.g., 8K-16K tokens);

Ml — is a learning mechanism.

Contextual settings Cs are responsible for the behavior of GPTs when interacting with the user. These settings determine which task will be solved, how it will be solved, what the model receives as input, and how the should be presented.

$$Cs = \langle P_s, Lg, Fd, Om, Tg \rangle,$$

where:

Ps —system prompt that defines the behavior of the model;

Lg — language of the interface and reporting documents;

Fd — set of supported document formats (PDF, DOCX, TXT, etc.);

Om — mode of operation (generative or analytical);

Tg — generation temperature (determines the level of creativity of the response).

Contextual settings are key tools for controlling the behavior of custom GPTs. They determine how accurately and relevantly the model processes queries; what form of response it provides; how deeply it analyzes the content.

Security constraints are a critical component of specialized GPTs, as they ensure ethical, reliable, and confidential handling of textual data. In modern AI systems, security covers a wide range of measures aimed at preventing their malicious use, protecting confidential data, filtering unwanted content, and monitoring compliance with corporate standards.

$$Sc = \langle Qb, Cr, Mm \rangle,$$

where:

Qb — blocking incorrect requests (screening out manipulative, dangerous requests);

Cr — confidentiality restriction (prohibition of storing or transmitting confidential information);

Mm — moderation mechanism (automatic filtering or manual control).

The set of parameters Is (integration settings) is responsible for storage, automation, and interaction with other systems. Integration of GPT with external services significantly expands its capabilities.

$$Is = \langle API, Rg, Ds \rangle,$$

where:

API — API for collaboration;

Rg — requirements for automatic report generation;

Ds — parameters for connecting to databases or cloud services.

Thanks to its integration capabilities, the model can work dynamically, scalably, and efficiently, automating workflows, providing high-quality report generation, and providing access to up-to-date data in corporate or cloud environments.

The created model allows you to structure all the key parameters that affect the operation of the custom GPTs, identify the relationships between them, flexibly customize and adapt the model to specific tasks. This significantly reduces the cost of preparing new configurations, which is especially important for large-scale implementation.

3. Implementation

Creating custom GPTs (Generative Pre-trained Transformers) is the process of adapting large language models to specific tasks, which allows automating complex processes, from data analysis to content generation. Such models can be developed and implemented on different platforms, each of which offers unique capabilities, limitations, and target audience. A number of factors determine the choice of platform: from technical complexity to security and localization requirements.

Today, there are several key approaches to creating such models. Cloud-based solutions such as OpenAI GPTs or Microsoft Azure OpenAI offer intuitive interfaces for non-technical users [22]. They allow for quick model setup through text instructions and data uploads, but require constant access to the Internet and have privacy restrictions.

For organizations that work with sensitive information, local solutions such as Hugging Face Transformers or OLLAMA are more appropriate. These tools allow you to deploy models on your own hardware, which provides full control over the data, but requires technical expertise and resources. The features of platforms for creating custom GPTs are presented in Table 1.

Table 1
Platforms for creating custom GPTs

Criterion	OpenAI GPTs	Hugging Face	MS Azure OpenAI	OLLAMA	LangChain
Local Implementation	Cloud-based (None)	Possible (own servers)	Cloud-based with private instance	Fully local	Depends on integration
Customization Tools	Prompts + data upload	Model fine-tuning	Fine-tuning via Azure ML	Local model fine-tuning	Modular architecture
Complexity	Low (no-code interface)	High (requires ML/Python expertise)	Medium (Azure expertise needed)	High (DevOps skills required)	Medium (for developers)
API	REST API, Webhooks	Python/REST API	Azure Cloud integration	REST API (self-hosted)	Python/JS
Security	Cloud storage (OpenAI)	Full control (local)	High (GDPR, ISO compliant)	Maximum (data stays on-premises)	Depends on infrastructure
Ukrainian Support	Limited (via prompts)	Yes (via model fine-tuning)	Limited (similar to OpenAI)	Possible (with Ukrainian models)	Depends on model
Cost	From $0.03/1K tokens + $20.0 GPT Builder	Free (local) / from $0.005 (cloud)	From $0.03/1K tokens + Azure costs	Free (hardware costs)	Free library
Target Audience	Non-technical users	ML engineers, researchers	Corporations, government agencies	Technical experts	Developers

An important aspect for Ukrainian users is native language support. For those looking for a balance between flexibility and simplicity, hybrid solutions like LangChain are ideal, as they allow for the integration of different AI models into a single pipeline.

The cost also plays a crucial role. Free tools (OLLAMA, local Hugging Face model) are suitable for startups or educational projects, while cloud services (Azure, OpenAI) require constant investment.

The technological stack of the software implementation of the system provides flexibility and scalability. The Python language, which allows for the efficient integration of artificial intelligence libraries and data analysis tools, forms the basis of the server side. The client side is implemented using the React library. The server logic is based on the Django framework. The relational database MySQL was chosen to store and manage data [23].

The figure 4 shows a high-level workflow diagram of a part of an application program collaboration. The user initiates the interaction by uploading a document and choosing the type of verification through the interface. The user interface transmits data to the system core (model), which is responsible for processing business logic: it checks the correctness of the input data, formats the request, and prepares the context for further analysis. After that, the call controller uses the API to redirect the request to a custom GPTs - a specialized model trained to perform a certain type of verification. The custom GPTs interacts with the underlying large-scale language model, which generates a response based on the data received. The result of the LLM's work is returned through a chain of callbacks: first to the custom GPTs for additional processing, then to the controller that converts the data into a convenient format, then to the system kernel to generate a report file and enter the information into the database, and finally to the user interface.

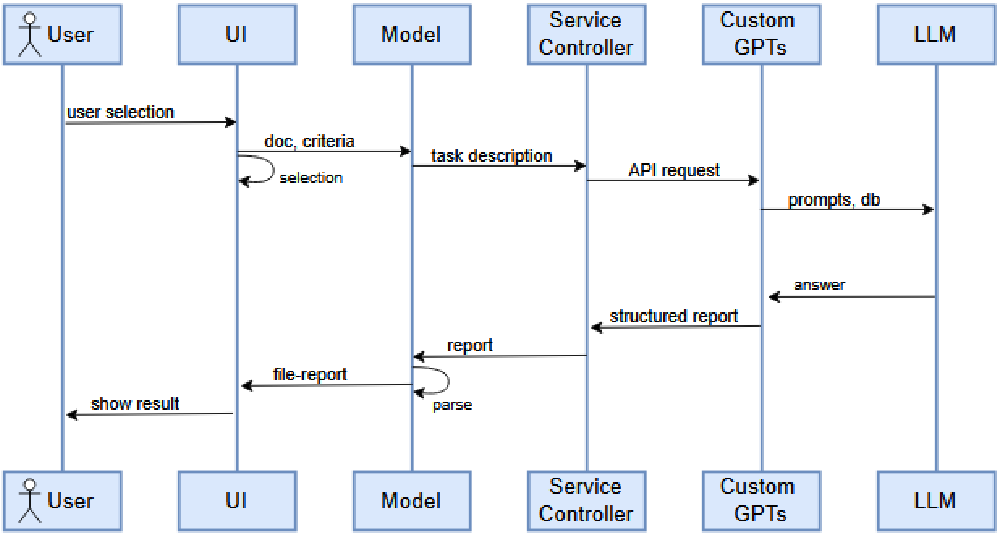

Figure 4: High-level diagram of the application workflow

4. Computational experiment

To evaluate the effectiveness of the developed system of custom GPTs, a series of experiments were conducted to verify the competition documentation of student research papers. The experiment involved student papers of the All-Ukrainian competition of scientific papers, in which one of the authors participated as a jury member (2021). Introducing certain deviations in the formatting of the papers allowed us to expand the set of initial data. The study compared the results of an automated check with an expert assessment. The experiment involved local implementations of GPTs based on the following models: Llama3 8B, LLaVA Llama3 8B, Phi-3 3.8B. The experiments conducted in three stages. At the first stage, a group of experts manually evaluated the documents, identifying errors in the design according to predefined criteria. At the second stage, the custom GPT analyzed the documents. The model automatically identified deviations and generated a report detailing the errors. At the third stage, the results of the automated check compared with expert opinions to assess the accuracy of error detection. To evaluate the model performance, metrics such as precision, recall, and F1 measurement were used. The speed of verification was analyzed in comparison with manual verification. The models used hints and regulatory documents. The experimental results are given in Table 2.

Table 2
Model performance indicators

Model	Precision	Recall	F1-score	Time
Experts	97%	89%	93%	900-1200s
Llama3 8B	95%	88%	90%	41.898s (8.66 tok/s)
LLaVA Llama3 8B	92%	85%	89%	57.920s (8.63 tok/s)
Phi-3 3.8B	87%	78%	89%	39.711s (8.88 tok/s)

A comparative analysis showed that custom GPTs models clearly outperform humans in terms of document processing time. The accuracy of their work is not much different from that of humans, although people also make mistakes. The efficiency of GPTs models depends significantly on the quality of the samples, which makes it possible to improve them further.

A hybrid model that combines the method of sequential analysis of variants with a cascade of specialized GPTs was also experimentally tested. The SAV scheme consisted of three stages: the first one was checking for structural compliance; the second one was checking for compliance with formatting requirements; and the third one was checking for correctness of the literature. The set of GPTs is based on LLM Llama3 8B (local version). The obtained results demonstrate that the method scheme works correctly, and the set of documents becomes smaller from stage to stage. Generalized performance indicators of the hybrid model:

- F1-mean of the hybrid model: 92% (compared to 89% in the manual test).
- Total processing time for 86 papers: 12 minutes (expert review - 29 hours at the rate of 20 minutes per review of one work by an expert).
- Errors of the second kind: papers with partial violation of the paper structure passed the filter; in the several cases, minor technical errors in the list of references caused the deviation of the works.

Thus, the hybrid approach proved to be viable and efficient, combining the advantages of artificial intelligence with the logic of sequential analysis, which makes it promising for scaling in large scientific competitions. To improve the model, it is necessary to expand the training data and optimize the custom GPTs.

5. Conclusion

The paper proposes an innovative approach to automating the management of student competition documents using a hybrid model that integrates the Sequential analysis of variants method with custom GPTs. The study demonstrates that the combination of these technologies allows for effective filtering of documents at the preliminary review stage, significantly reducing processing time compared to manual expert evaluation. The key advantage of the proposed approach is a structured multi-stage verification, where each stage corresponds to a specific criterion (formatting, structuredness, bibliography, etc.), which ensures objectivity, transparency and scalability of the process.

The experimental results confirm the rather high accuracy of custom GPTs, which makes them competitive alternatives to the existing technology.

The practical value of the study lies in the fact that the proposed approach opens up new opportunities for the digitalization of scientific competitions, making them more accessible, objective, and effective in a global context.

A promising direction for the development of the study is the development of a mathematical model of the competitive procedure for using an adaptive approach to determine the criticality of violations found in the competition documentation.

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] Yasinska, A. Problems and Prospects of Electronic Document Management in the Context of Digital Transformation. Young Scientist, 2022, pp. 128–134. doi: 10.32839/2304-5809/2022-11-111-27.

[2] Markus, M. L. Datification, Organizational Strategy, and IS Research: What's the Score? 2017.

[3] Neudecker, C., Baierer, K., Federbusch, M. OCR-D: An End-to-End Open Source OCR Framework for Historical Printed Documents. 2019.

[4] Jiugen, Y., & Ruonan, X. Design and Implementation of College Science and Technology Innovation Contest Management System Based on NET Platform. 2016.

[5] Glowacka-Musial, M. Applying Topic Modeling for Automated Creation of Descriptive Metadata for Digital Collections. 2022.

[6] Pho, H. T., & Tambo, T. Integrated Management Systems and Workflow-Based Electronic Document Management: An Empirical Study. 2014.

[7] Leong, Wai Yie, & Zhang, J. AI on Academic Integrity and Plagiarism Detection. ASM Science Journal, vol. 20, 2025. doi:10.32802/asmscj.2025.1918.

[8] Teremetskyi, V., Burylo, Yu., Zozuliak, O., Koshmanov, M., & Petruk, Olha. Academic Integrity in the Age of Artificial Intelligence: World Trends and Outlook for Ukraine from the Legal Perspective. Pakistan Journal of Life and Social Sciences (PJLSS), vol. 22, 2024. doi:10.57239/PJLSS-2024-22.1.00147.

[9] Khrolenko, Y. O. Analysis of Modern Technologies for Organizing and Conducting Competitions of Student Research Papers. URL: https://conference.ikto.net/public/static/about.html.

[10] Regulations on the All-Ukrainian Competition of Student Research Papers in the Fields of Knowledge and Specialties. URL: https://zakon.rada.gov.ua/laws/show/z0620-17#n101.

[11] Hnatiienko, H., Domanetska, I., Hnatiienko, O., & Khrolenko, Y. Tasks' Analysis of the Intelligent System for Conducting Project Competitions. Management of Development of Complex Systems, vol. 60, 2024, pp. 113–120. doi:10.32347/2412-9933.2024.60.113-120.

[12] Hnatiienko, H. M., Domanetska, I. M., Hnatiienko, O. H., & Khrolenko, Y. O. Scheme of Sequential Analysis of Options in Expert Technology for Analyzing Urban Transformation Proposals, in: Applied Information Systems and Technologies in a Digital Society: Proceedings of the 8th International Scientific-Practical Conference (Kyiv, October 1, 2024). pp. 63–74.

[13] Voloshyn, O. F., Hnatiienko, H. M., & Kudin, V. I. Sequential Analysis of Options: Technologies and Applications: Monograph. Kyiv: Stylos, 2013.

[14] Voloshyn, O. F., & Kudin, V. I. On the Technology of Sequential Analysis of Variants in Research and Design Problems, in: II International Conference Information Technologies and Interactions, Kyiv, November 3–5, 2015, Abstracts*. pp. 196–197.

[15] Mandvikar, Shreekant & Dave, Deep Manishkumar. Augmented intelligence: human-ai collaboration in the era of digital transformation. International Journal of Engineering Applied Sciences and Technology. International Journal of Engineering Applied Sciences and Technology, 2023, Vol. 8, Issue 06, pp. 24-33. doi: 10.33564/IJEAST.2023.v08i06.003.

[16] Xu, Jia & Du, Weilin & Liu, Xiao & Li, Xuejun. LLM4Workflow: An LLM-based Automated Workflow Model Generation Tool, in 39th IEEE/ACM International Conference on Automated Software Engineering. 2024. pp. 2394-2398. doi:10.1145/3691620.3695360.

[17] Gebreab, S. A., Salah, K., Jayaraman, R., Rehman, M. H. U., & Ellaham, S. (2024). LLM-Based Framework for Administrative Task Automation in Healthcare. In A. Varol, M. Karabatak, C. Varol, & E. Tuba (Eds.), in 12th International Symposium on Digital Forensics and Security, ISDFS 2024 IEEE. doi: 10.1109/ISDFS60797.2024.10527275

[18] Basyal, Lochan. (2023). Text Summarization Using Large Language Models: A Comparative Study of MPT-7b-instruct, Falcon-7b-instruct, and OpenAI Chat-GPT Models. 10.48550/arXiv.2310.10449.

[19] Abdellatif, Osama & Nader, Abdelrahman & Hamdi, Ali. ERPA: Efficient RPA Model Integrating OCR and LLMs for Intelligent Document Processing. 2024. 10.48550/arXiv.2412.19840.

[20] Pudasaini, Shushanta & Miralles Pechuan, Luis & Lillis, David & Salvador, Marisa. Survey on AI-Generated Plagiarism Detection: The Impact of Large Language Models on Academic Integrity. Journal of Academic Ethics. 2024. pp. 1-34. doi: 10.1007/s10805-024-09576-x.

[21] Zhao, Benjamin & Ikram, Muhammad & Kaafar, Dali. GPTs Window Shopping: An analysis of the Landscape of Custom ChatGPT Models. 2024. 10.48550/arXiv.2405.10547.

[22] Mohan, G. B., Prasanna Kumar, R., Vishal Krishh, P., et al. An Analysis of Large Language Models: Their Impact and Potential Applications. Knowledge and Information Systems, vol. 66, 2024, pp. 5047–5070. doi:10.1007/s10115-024-02120-8.

[23] Domanetska, I., & Khrolenko, Y. (2025). Analysis of the peculiarities of building a system for organising and conducting competitions of students' research papers. Management of Development of Complex Systems, 61, 226–231, dx.doi.org\10.32347/2412-9933.2025.61.226-231.