ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

№2
2026

ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)

CENTRAL ASIAN ACADEMIC
RESEARCH CENTER

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE

2 (358)

APRIL – JUNE 2026

PUBLISHED SINCE JANUARY 1963
PUBLISHED 4 TIMES A YEAR

ALMATY, NAS RK

Academic Scientific Journal of Computer Science
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)
Owner: «Central Asian Academic Research Center» LLP (Almaty).
Certificate **№ KZ77VPY00121154** on the re-registration of the periodical printed and online publication of the information agency, issued on **05.06.2025** by the Republican State Institution «Information Committee» of the Ministry of Culture and Information of the Republic of Kazakhstan
Subject area*: information and communication technologies.*
Currently: *included in the list of journals recommended by the CCSES MSHE RK in the direction of «Information and communication technologies».*
Periodicity: *4 times a year.*
http://www.physico-mathematical.kz/index.php/en/

Academic Scientific Journal of Computer Science
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)
Меншіктеуші: «Орталық Азия академиялық ғылыми орталығы» ЖШС (Алматы).
Ақпарат агенттігінің мерзімді баспасөз басылымын, ақпарат агенттігін және желілік басылымды қайта есепке қою туралы ҚР Мәдениет және Ақпарат министрлігі «Ақпарат комитеті» Республикалық мемлекеттік мекемесі **05.06.2025** ж. берген **№ KZ77VPY00121154** Куәлік.
Тақырыптық бағыты: *ақпараттық-коммуникациялық технологиялар*
Қазіргі уақытта: *«ақпараттық-коммуникациялық технологиялар» бағыты бойынша ҚР БҒМ БҒСБК ұсынған журналдар тізіміне енді.*
Мерзімділігі: *жылына 4 рет.*
http://www.physico-mathematical.kz/index.php/en/

Academic Scientific Journal of Computer Science
ISSN 2518-1726 (Online),
ISSN 1991-346X (Print)
Собственник: *ТОО «Центрально-азиатский академический научный центр» (г. Алматы).*
Свидетельство о постановке на переучет периодического печатного издания, информационного агентства и сетевого издания **№ KZ77VPY00121154.** Дата выдачи **05.06.2025**
Тематическая направленность: *информационно-коммуникационные технологии.*
В настоящее время: *вошел в список журналов, рекомендованных КОКСНВО МНВО РК по направлению «информационно-коммуникационные технологии».*
Периодичность: *4 раза в год.*
http://www.physico-mathematical.kz/index.php/en/

CONTENTS

COMPUTER SCIENCE

МАЗМҰНЫ

КОМПЬЮТЕРЛІК ҒЫЛЫМДАР

СОДЕРЖАНИЕ

КОМПЬЮТЕРНЫЕ НАУКИ

ACADEMIC SCIENTIFIC JOURNAL OF COMPUTER SCIENCE
ISSN 1991-346X
Volume 2.
Namber 358 (2026). 105–124

https://doi.org/10.32014/2026.2518-1726.429

IRSTI 28.23.29
UDC 004.89:378.1

© Basheyeva Zh.[1], Tokesh A.[1]*, Bekish U.[2], Abdoldinova G.[3], 2026.
[1]Astana IT University, Astana, Kazakhstan;
[2]Academy of Public Administration under the President of the Republic of Kazakhstan, Astana, Kazakhstan;
[3]K. Kulazhanov Kazakh University of Technology and Business, Astana, Kazakhstan
E-mail: 255139@astanait.edu.kz

ARTIFICIAL INTELLIGENCE FOR ACADEMIC PROJECT MANAGEMENT: A BIBLIOMETRIC ANALYSIS AND SYSTEMATIC REVIEW

Basheyeva Zhuldyz — PhD, Astana IT University, Astana, Kazakhstan,
E-mail: zhuldyz.basheyeva@astanait.edu.kz, https://orcid.org/0000-0001-9605-2101;
Tokesh Alisher — Astana IT University, Astana, Kazakhstan,
E-mail: 255139@astanait.edu.kz, https://orcid.org/0009-0003-7538-2779;
Bekish Ulan — PhD, Academy of Public Administration under the President of the Republic of Kazakhstan, Astana, Kazakhstan,
E-mail: bekish_u@mail.ru, https://orcid.org/0000-0001-8979-4238;
Abdoldinova Gulsim — Candidate of Pedagogical Sciences, Kulazhanov Kazakh university of technology and business, Astana, Kazakhstan,
E-mail: abdoldinova@mail.ru, https://orcid.org/0000-0001-7891-7175.

Abstract. The review is a synthesis of the recent research on the implementation of artificial intelligence in managing academic projects in higher education institutions, specifically semantic approaches to track academic integrity, evaluation of the originality of the research topic, and supervisor selection. The previous research has often addressed plagiarism detection, novelty assessment, and the recommendation as separate features; therefore, the current paper is an attempt to unify the methodology, current trends, and the existing challenges related to these tasks. Based on the guidelines of the Preferred Reporting Items of Systematic Reviews and Meta-Analyses (PRISMA) 2020, we chose 49 peer-reviewed studies that were published between 2021 and 2025 and were found in the large scientific databases. The paper combines a systematic review and bibliometric analysis, including keyword co-occurrence, co-citation analysis, and network analysis of co-authorships through VOSviewer. Four major thematic clusters were identified:

(1) semantic integrity and plagiarism detection, (2) measuring the novelty and originality of the research topics, (3) intelligent recommendation and resource allocation systems, and (4) human-centred management and decision-support. In these clusters, lexical baselines are increasingly complemented or extended by embedding and transformer-based similarity models, and the most recent works use large language models to make higher-order reasoning and decisions. The literature is, however, fragmented, with an assortment of heterogeneous assessment procedures, limited use of real institution data, and scarce longitudinal validation. The results provide a comprehensive perspective on the methodology and research gaps, thus making it possible to build integrated academic management platforms, that combine integrity control, originality evaluation, and supervisory allocation.

Keywords: systematic review; bibliometric analysis; academic project management; natural language processing; semantic similarity; novelty detection; recommendation system

For citations: Basheyeva Zh., Tokesh A., Bekish U., Abdoldinova G. Artificial intelligence for academic project management: a bibliometric analysis and systematic review. Academic Scientific Journal of Computer Science, 2026. — No.2. – P. 105-124. DOI: https://doi.org/10.32014/2026.2518-1726.429

© Башеева Ж.[1], Төкеш Ә.[1]*, Бекіш Ұ.[2], Абдолдинова Г.[3], 2026.
[1]Astana IT University, Астана, Қазақстан;
[2]Қазақстан Республикасы Президентінің жанындағы Мемлекеттік басқару академиясы, Астана, Қазақстан;
[3]Қ. Құлажанов атындағы Қазақ технология және бизнес университеті, Астана, Қазақстан.
E-mail: 255139@astanait.edu.kz

АКАДЕМИЯЛЫҚ ЖОБАЛАРДЫ БАСҚАРУДАҒЫ ЖАСАНДЫ ИНТЕЛЛЕКТ: БИБЛИОМЕТРИЯЛЫҚ ТАЛДАУ ЖӘНЕ ЖҮЙЕЛІ ШОЛУ

Башеева Жұлдыз — PhD, Astana IT University, Астана, Қазақстан,
E-mail: zhuldyz.basheyeva@astanait.edu.kz, https://orcid.org/0000-0001-9605-2101;
Төкеш Әлішер — Astana IT University, Астана, Қазақстан,
E-mail: 255139@astanait.edu.kz, https://orcid.org/0009-0003-7538-2779;
Бекіш Ұлан — PhD, Қазақстан Республикасы Президентінің жанындағы Мемлекеттік басқару академиясы, Астана, Қазақстан,
E-mail: bekish_u@mail.ru, https://orcid.org/0000-0001-8979-4238;
Абдолдинова Гүлсім — педагогика ғылымдарының кандидаты, Қ. Құлажанов атындағы Қазақ технология және бизнес университеті, Астана, Қазақстан,
E-mail: abdoldinova@mail.ru, https://orcid.org/0000-0001-7891-7175.

Аннотация. Бұл шолу жоғары оқу орындарындағы академиялық жобаларды басқаруға жасанды интеллектіні енгізу бойынша заманауи зерттеу-

лердің, атап айтқанда, академиялық адалдықты бақылаудың семантикалық тәсілдерінің, зерттеу тақырыптарының бірегейлігін бағалаудың және ғылыми жетекшілерді таңдаудың синтезі болып табылады. Алдыңғы зерттеулер плагиатты анықтауды, жаңалықты бағалауды және ұсыныс беру функцияларын көбінесе бөлек сипаттамалар ретінде қарастырды; сондықтан бұл жұмыс осы міндеттерге қатысты методологияны, ағымдағы үрдістер мен қолданыстағы мәселелерді біріктіруге бағытталған талпыныс болып табылады. PRISMA 2020 (Жүйелі шолулар мен мета-талдауларға арналған таңдаулы есеп беру элементтері) нұсқаулығына сүйене отырып, біз ірі ғылыми деректер қорларынан 2021 және 2025 жылдар аралығында жарияланған 49 рецензияланған зерттеуді таңдап алдық. Жұмыс жүйелі шолу мен библиометриялық талдауды біріктіреді, соның ішінде түйін сөздердің бірлескен кездесуін, бірлескен дәйексөз талдауын және VOSviewer көмегімен авторлық бірлестіктің желілік талдауын қамтиды. Төрт негізгі тақырыптық кластер анықталды: (1) семантикалық тұтастық және плагиатты анықтау, (2) зерттеу тақырыптарының жаңалығы мен бірегейлігін өлшеу, (3) зияткерлік ұсыныс беру және ресурстарды бөлу жүйелері, және (4) адамға бағытталған басқару және шешім қабылдауды қолдау. Бұл кластерлерде лексикалық базалық модельдер эмбеддингтер мен трансформерлерге негізделген ұқсастық модельдерімен көбірек толықтырылуда, ал ең соңғы жұмыстарда жоғары деңгейдегі логикалық қорытындылар жасау мен шешім қабылдау үшін үлкен тілдік модельдер қолданылады. Дегенмен, әдебиеттер фрагменттелген, әртүрлі гетерогенді бағалау процедураларымен, нақты оқу орындарының деректерін шектеулі пайдаланумен және лонгитюдтік валидацияның тапшылығымен сипатталады. Нәтижелер методология мен зерттеулердегі олқылықтарға кешенді көзқарас береді, бұл адалдықты бақылауды, бірегейлікті бағалауды және жетекшілерді бөлуді біріктіретін интеграцияланған академиялық менеджмент платформаларын құруға мүмкіндік береді.

Түйін сөздер: жүйелі шолу; библиометриялық талдау; академиялық жобаларды басқару; табиғи тілді өңдеу; семантикалық ұқсастық; жаңалықты анықтау; ұсыныс беру жүйесі

© Башеева Ж.[1], Токеш А.[1]*, Бекиш У.[2], Абдолдинова Г.[3], 2026.
[1]Astana IT University, Астана, Казахстан;
[2]Академия государственного управления при Президенте Республики Казахстан, Астана, Казахстан;
[3]Казахский университет технологии и бизнеса имени К. Кулажанова, Астана, Казахстан.
E-mail: 255139@astanait.edu.kz

ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В УПРАВЛЕНИИ АКАДЕМИЧЕСКИМИ ПРОЕКТАМИ: БИБЛИОМЕТРИЧЕСКИЙ АНАЛИЗ И СИСТЕМАТИЧЕСКИЙ ОБЗОР

Башеева Жулдыз — PhD, Astana IT University, Астана, Казахстан,
E-mail: zhuldyz.basheyeva@astanait.edu.kz, https://orcid.org/0000-0001-9605-2101;
Токеш Алишер — Astana IT University, Астана, Казахстан,
E-mail: 255139@astanait.edu.kz, https://orcid.org/0009-0003-7538-2779;
Бекиш Улан — PhD, Академия государственного управления при Президенте Республики Казахстан, Астана, Казахстан,
E-mail: bekish_u@mail.ru, https://orcid.org/0000-0001-8979-4238;
Абдолдинова Гульсим — кандидат педагогических наук, Казахскийуниверситет технологии и бизнеса имени К. Кулажанова, Астана, Казахстан,
E-mail: abdoldinova@mail.ru, https://orcid.org/0000-0001-7891-7175.

Аннотация. *Актуальность.* Внедрение искусственного интеллекта в управление академическими проектами становится одним из значимых направлений цифровой трансформации высшего образования. Особую актуальность приобретают семантические подходы к мониторингу академической честности, оценке оригинальности тем исследований и подбору научных руководителей. В существующих исследованиях обнаружение плагиата, оценка новизны и рекомендательные функции часто рассматриваются как отдельные задачи, что ограничивает возможность формирования интегрированных платформ академического менеджмента. *Цель.* Систематизировать современные исследования по применению искусственного интеллекта в управлении академическими проектами, выявить основные тематические направления, методологические подходы и исследовательские пробелы в данной области. *Методы.* Исследование выполнено на основе руководства PRISMA 2020 для систематических обзоров и метаанализа. В выборку включены 49 рецензируемых исследований, опубликованных в 2021–2025 годах в крупнейших научных базах данных. Работа сочетает систематический обзор и библиометрический анализ, включая анализ совместной встречаемости ключевых слов, анализ социтирования и сетевой анализ соавторства с использованием программы VOSviewer. *Результаты и выводы.* В результате анализа выявлены четыре основных тематических кластера: семантическая целостность и обнаружение плагиата; измерение новизны и оригинальности тем исследований; интеллектуальные рекомендательные системы и системы распределения ресурсов; человекоцентричное управление и поддержка принятия решений. Установлено, что лексические базовые модели все чаще дополняются или заменяются моделями сходства на основе эмбеддингов и трансформеров, а в наиболее современных работах используются большие языковые модели для логического вывода и принятия решений более высокого уровня. Вместе с тем литература остается фрагментированной, характеризуется неоднородностью процедур оценки, ограниченным использованием реальных данных образовательных организаций и недостатком лонгитюдной валидации. Полученные результаты формируют комплексное представление о методологических подходах и исследовательских пробелах, а также могут быть использованы при создании интегрированных платформ академи-

ческого менеджмента, объединяющих контроль академической честности, оценку оригинальности и распределение научных руководителей.

Ключевые слова: искусственный интеллект, систематический обзор, библиометрический анализ, управление академическими проектами, обработка естественного языка, семантическое сходство, обнаружение новизны, рекомендательная система

Introduction. Over the past few years, the use of artificial intelligence (AI) technologies in higher education institutions has grown substantially, especially when it comes to the administration of student projects (Shimray and Subaveerapandiyan, 2025). Universities are operating under the conditions of greater competition and digitalization than ever before, so it is relevant to consider that AI-based technologies could become key aspects of automation, efficiency, and decision-making (Anghel and Popescu, 2022; Nurgaliyeva et al., 2025). In the broader framework of higher education management, decision-making processes are closely connected with the structured design of educational components, learning outcomes, and instructional content, which form the foundation of academic organization and governance (Bekish et al., 2020). The crucial place of AI technologies in the context of automating an academic process, especially recommendations for project topic selection, is discussed by Bibi et al. (2023). Despite rapid technological development, a number of relevant issues and problems have remained.

Literary review. The studies conducted so far have been devoted to AI-based approaches to the analysis of learning data and academic text processing. Lexical and metric-based baselines (e.g., TF-IDF and cosine similarity) can be limited in capturing deep semantic relationships (Hryn et al., 2025), while transformer-based contextual embeddings substantially improve STS-style similarity modeling (Hassan et al., 2025). Recent research increasingly highlights the benefits of embedding and transformer-based structures that are capable of capturing contextual semantics and enhance the detection of the reused texts as well as the evaluation of similarity at the sentence level (Korade et al., 2024). Meanwhile, novelty can be realized as recombinant novelty, when a document is considered new if it combines semantically distant cited references (Shibayama et al., 2021). Novelty has also been defined as a semantic distance in embedded idea spaces in similar studies, along with detailing comparisons between algorithmic estimates of novelty and human estimates of novelty, and SBERT-based systems that are most in line with human estimates (Just et al., 2024). However, most prior approaches have tended to address duplication control, novelty evaluation, and administrative tasks separately.

The current review centers its focus on AI-based approaches for project management in the academic domain, with respect to novelty detection, duplication and plagiarism, and also for the assignment of supervisors. The scope of this review is constrained within published and peer-reviewed literature between the period of 2021-2025, specifically for machine learning, natural language processing, and also

for hybrid approaches. The review combines systematic review procedures with bibliometric analysis.

The first objective of this review article is to compile existing studies, identify prevailing methodological patterns, as well as uncover gaps in linking semantic approaches and academic administration. In keeping with this objective, this study examines the following:

• RQ1: What are the main methodological trends in AI-based academic project management research?

• RQ2: What AI resources are the most frequently employed to detect novelty, control duplication, or assign a supervisor?

• RQ3: What are the limitations and unresolved problems in the existing research?

The rest of the article is organized as follows. The next section describes the methodology of systematic review and bibliometric analysis. The following sections discuss approaches based on artificial intelligence to ensure academic integrity, identify the novelty of research, and assign a supervisor. The final section discusses the implications of the findings and defines the directions for future research.

Methods. In this research, the mixed-methodological design is used, which combines a systematic review with a bibliometric analysis to integrate thematic and structural features of research on AI-based academic project management. The systematic component was reported in accordance with PRISMA 2020 guidelines to ensure transparency, replicability, and methodological rigor in study identification, screening, and inclusion procedures. At the same time, bibliometric mapping was conducted to analyze the quantitative trends in the field of research.

Search Strategy

A comprehensive literature search was conducted in Scopus and Google Scholar, selected for their broad coverage of computer science, artificial intelligence, and educational research. The final search was completed in December 2025.

The search strategy was structured around several conceptual categories reflecting the core dimensions of the review: (1) AI and semantic technologies, (2) academic integrity and duplication control, (3) research topic novelty assessment, (4) supervisor assignment and recommendation systems, and (5) the higher education academic project management context.

Within each category, the related terms were added together using the Boolean operator OR to represent the variations of the terms. The conceptual categories were then combined with operator AND to find studies that were located at the intersection of AI-driven semantic processes and academic project management processes in higher education. Where relevant, truncation symbols have been used in order to take into consideration the lexical variation.

Only peer-reviewed publications in English published between 2021 and 2025 were considered. Table 1 provides the detailed structure of the search terms and categorical grouping of search terms.

Table 1 – Search strategy

Category	Included Key Terms
AI and Semantic Technologies	("artificial intelligence" OR AI OR "machine learning" OR "deep learning" OR "natural language processing" OR NLP OR "transformer model*" OR "large language model*" OR LLM* OR semantic analys* OR semantic similarity OR semantic matching OR "text embedding*" OR "sentence embedding*" OR SBERT OR BERT OR "contextual embedding*" OR "deep semantic analys*")
Academic Integrity and Duplication Control	("academic integrity" OR plagiarism OR "semantic plagiarism" OR "intelligent plagiarism" OR paraphras* OR "text reuse" OR "similarity detection")
Research Topic Novelty Assessment	("research novelty" OR "topic novelty" OR originality OR "novelty detection" OR "originality assessment" OR "research originality" OR "innovation detection")
Supervisor Assignment and Recommendation Systems	("supervisor assignment" OR "supervisor matching" OR "supervisor allocation" OR "recommender system" OR "matching algorithm")
Higher Education and Academic Project Management Context	("higher education" OR universit* OR "academic project*" OR thesis OR dissertation OR "master* thesis" OR "diploma project*" OR "research management" OR "project management")

Inclusion and Exclusion Criteria

Inclusion criteria comprised peer-reviewed journal articles and review papers addressing AI or NLP applications in higher education contexts, particularly in relation to semantic similarity, novelty detection, plagiarism detection, text processing, or resource allocation within the academic project lifecycle. Exclusion criteria included non-peer-reviewed sources (e.g., editorials, blogs), purely technical AI studies without educational context, and articles lacking sufficient methodological transparency.

Quality Assessment and Data Extraction

The entire full-text assessment process included methodological selection to assess the clarity of the research objectives, the clarity of the data sets, the suitability of the assessment design, and sufficient detail to reproduce the study. Research papers that were not transparent in terms of methods were rejected. The systematic data analysis included author data, year of publication, aim of the study, methodology, participants data, main findings, limitations and keywords.

Study Selection Process

The dataset included 49 peer-reviewed articles published between 2021 and 2025. To attain a clear and reproducible means of selection, the literature on the topic was filtered with regard to the PRISMA 2020 protocol. In the first step of identifying relevant articles, a total of 1,829 articles were obtained from an extensive search of online databases such as Scopus and Google Scholar. After duplicate removal, a total of 418 duplicates were identified and eliminated, resulting in a preliminary list of 1,411 articles for further verification.

At the stage of selecting titles and abstracts, 1,176 entries were screened out because they did not correspond to the key areas of the research objectives. This resulted in 235 records being sought for full-text assessment. Of these, 223 full texts were successfully retrieved, whereas 12 records were excluded due to lack of access. The compliance of these 223 records with the requirements was carefully checked in accordance with a number of inclusion and exclusion criteria. At this stage, 174 reports were excluded for specific methodological and contextual reasons. The main grounds for exclusion included the absence of an AI-based methodological component (71 records), lack of focus on higher education settings (48 records), insufficient relevance to the core research themes of plagiarism detection, novelty evaluation, and supervisor assignment or recommendation (54 records), and inadequate methodological transparency preventing comparative analysis or reproducibility (1 record). Finally, 49 articles were peer-reviewed and included in the final systematic review process.

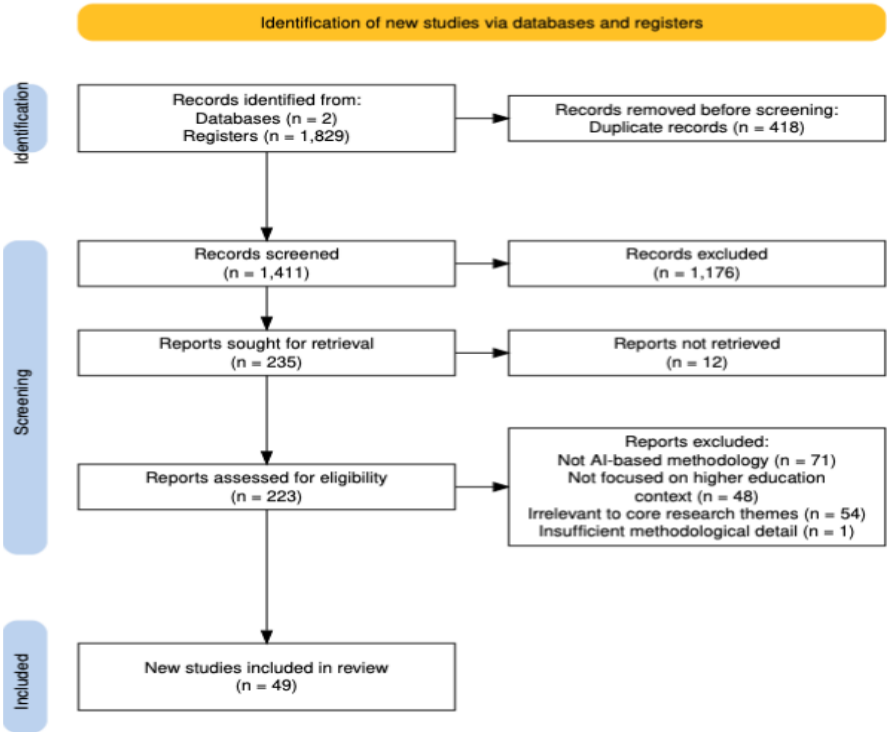

Figure 1 – PRISMA flowchart of the study selection process

Bibliometric and Thematic Analysis

Bibliometric analysis was conducted using VOSviewer to examine keyword co-occurrence patterns, co-authorship networks, and co-citation structures within the selected corpus. In parallel, a qualitative thematic synthesis was performed using an inductive coding approach. The included studies were iteratively coded and

grouped based on their methodological characteristics and functional application domains, resulting in four dominant thematic clusters.

Results. Results of the Bibliometric Analysis

Annual Publication Trends

Figure 2 illustrates the annual distribution of publications included in the review between 2021 and 2025. The number of studies demonstrates a consistent upward trend over the examined period. After a relatively moderate volume of publications in 2021, the output gradually increased in 2022 and 2023, followed by a more pronounced growth in 2024. The highest number of publications was recorded in 2025. Overall, the data indicate a growing scholarly interest in AI-driven academic project management and NLP-based administrative systems in higher education.

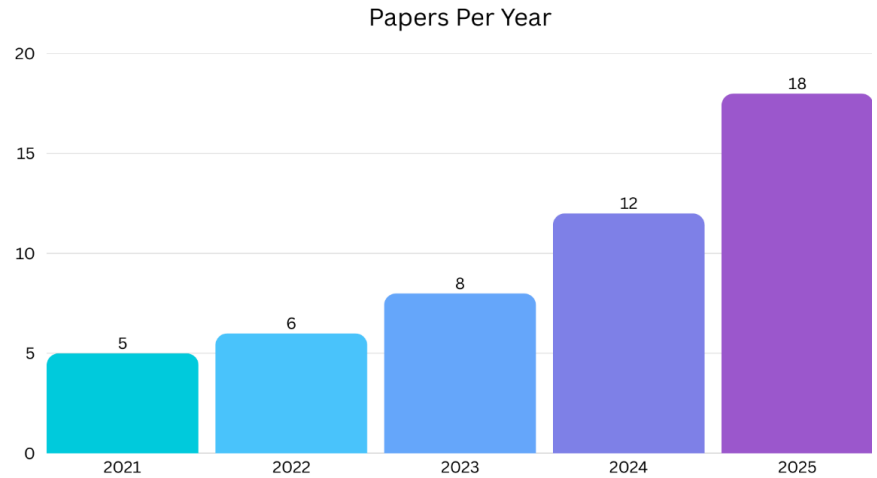

Figure 2 – Number of publications per year of inclusion (2021-2025)

Methodological Paradigms

In the given section, the quantitative comparison of the NLP similarity core methods used in the reviewed literature i.e. classical lexical methods, embedding-based semantic models, and LLM-based approaches is given. The 49 included studies were all analyzed and categorized based on the main analytical framework of the study. Since several studies employed hybrid architectures (e.g., combining TF-IDF baselines with embedding models or LLM components), the "primary" framework was determined based on the dominant method driving the final model performance or the central methodological contribution described by the authors. Classification in situations of hybrid implementations was done according to the predominant analytical paradigm.

According to the methodological descriptions of the respective articles, three major approaches became evident, including classical lexical and traditional machine learning techniques, transformer-based embedding models, and systems powered by large language models.

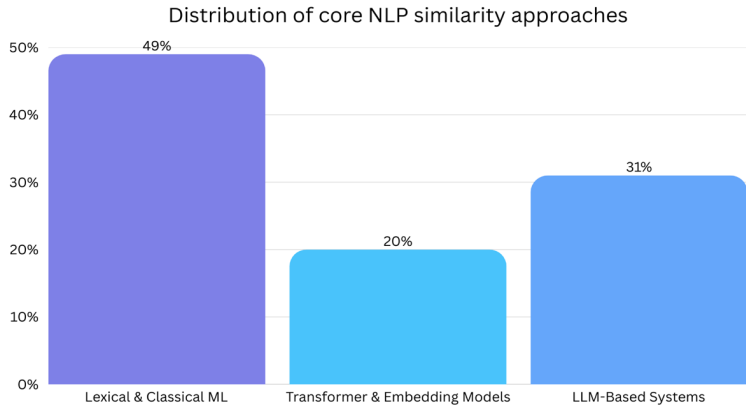

Figure 3 – Distribution of primary NLP similarity paradigms

The findings show that the studies based on lexical and classical machine learning capture 49% (n = 24). Such methods are TF-IDF, cosine similarity, supervised classification models, neural networks with no large-scale pre-trained models, and systems based on optimization. In most instances, the techniques are utilized as a form of baseline similarity or structured decision-supporting in recommendation and supervisor assignment models.

In the reviewed literature, 20% (n = 10) is transformer-based embedding. Such studies use mostly contextual sentence or document representations that were built based on architectures like BERT and SBERT. Similarity is commonly computed with dense vector embeddings and ranking or similarity measures and is generally used in semantic similarity measurement, paraphrase detection, and novelty measurement.

The approaches based on LLM represent 31% of the studies (n = 15). These are GPT-based, prompt-based interaction and zero-shot or few-shot inference strategies. These applications have been commonly used in AI-generated text detection and conceptual originality assessment tasks.

Overall, Figure 3 shows that lexical, embedding-based, and LLM-driven paradigms have a relatively high prevalence in the reviewed research body, which can be interpreted as a diversification of methodological approaches in current academic applications of NLP.

Conceptual Structure of the Research Field

The keywords co-occurrence network provides four significant thematic groups suggesting the conceptual framework of the field. Important nodes like artificial intelligence and natural language processing also act as connector concepts between multiple clusters.

The former cluster unites embedding-based and deep learning methods (e.g., LSTM, classification, text processing), which represent an expression of methodological improvement in automated text analysis.

The second cluster represents the domain of computational linguistics and semantically oriented studies, including research on computational linguistics, semantics, and novelty detection, with a focus on the theoretical foundations of language technologies.

The third cluster revolves around large language models and AI-based educational applications, such as language models, ChatGPT, machine learning, and students, which is a pointer to the applied aspect of generative AI in education.

The fourth cluster brings together semantic analysis, plagiarism detection and human-related evaluation, which display issues to do with assessment, integrity, and human-centered interpretation.

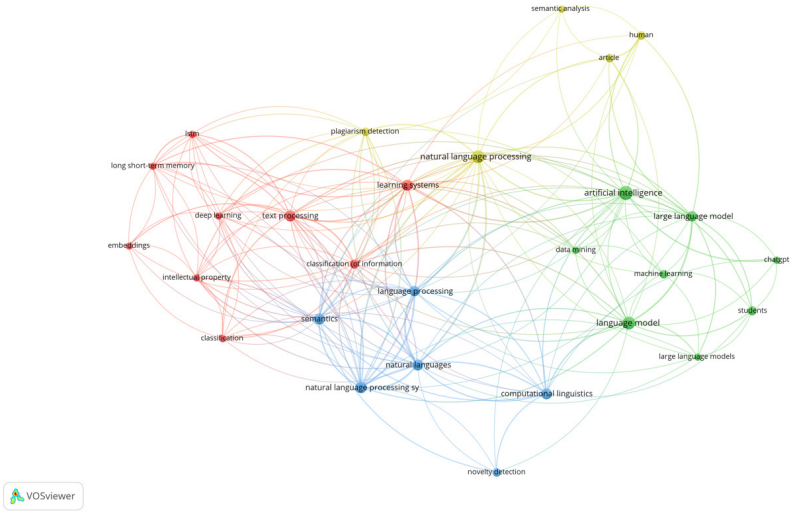

Figure 4 – Keyword co-occurrence network highlighting the shift toward generative AI and LLMs

Thematic Evolution of Author Keywords

In the overlay visualization shown in Figure 5, the co-occurrence of author keywords is presented over time in the analyzed literature, with natural language processing as the key term. Keywords that featured in the earlier publications predominantly 2022-early 2023 include deep learning, recurrent neural network, text plagiarism, and classification. Keywords that feature in the studies published between 2024-2025 include large language model, semantic analysis, academic writing, and AI innovation management.

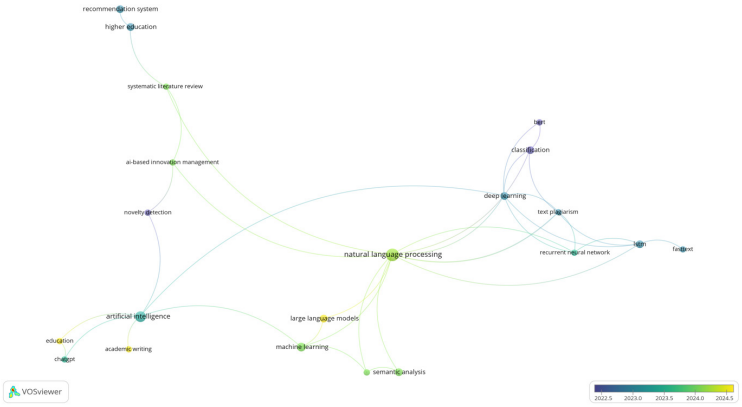

Figure 5 – Overlay-based thematic evolution of author keywords in AI-driven academic management research

Intellectual Sources and Publication Outlets

The source co-citation network shows six clusters of different intellectual foundations of the field. The core AI and NLP venues cluster (that is, the largest cluster by the number of sources) is centered around the Advances in Neural Information Processing Systems, ACL/EMNLP/NAACL proceedings, and ICML, showing the backbone of methodology in computational linguistics and deep learning. Conversely, the most significant hubs by cumulative link strength are seen in scientometrics and innovation-policy journals, including Scientometrics, Technological Forecasting and Social Change, and World Patent Information, so there is a strong evaluation-oriented research area. A second cluster will combine education-related journals, such as Computers and Education: Artificial Intelligence, British Journal of Educational Technology, and International Journal of Educational Technology in Higher Education. arXiv would be an exceptionally linked bridging source between conference and journal outlets.

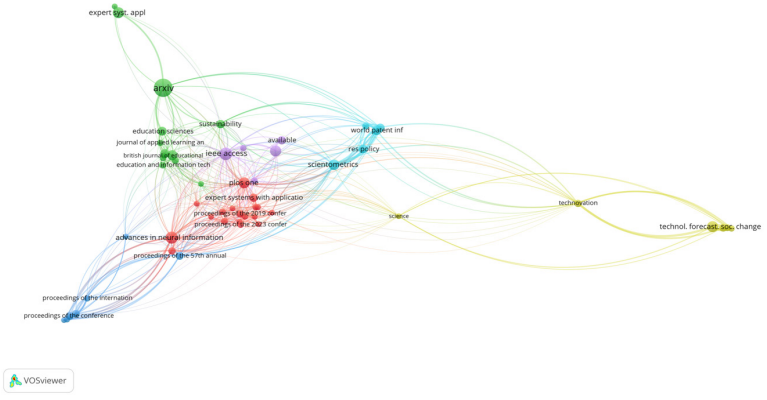

Figure 6 – Source co-citation network of publication outlets in AI-driven academic management research

Collaboration Patterns in AI Research

Figure 7 presents the co-authorship network among authors included in the review. The visualization reveals a fragmented structure consisting of several small groups with few connections among them, indicating relatively weak collaboration within the field.

Several clusters demonstrate strong internal connectivity indicating that specific research teams work together in the same manner. Meanwhile, some of the authors are shown as peripheral nodes with few co-authorship relationships revealing detached contributions to the field.

Thematic orientation of such clusters could be aligned with various research directions but the main outcome of the co-authorship network is a collaboration pattern as opposed to a thematic structure.

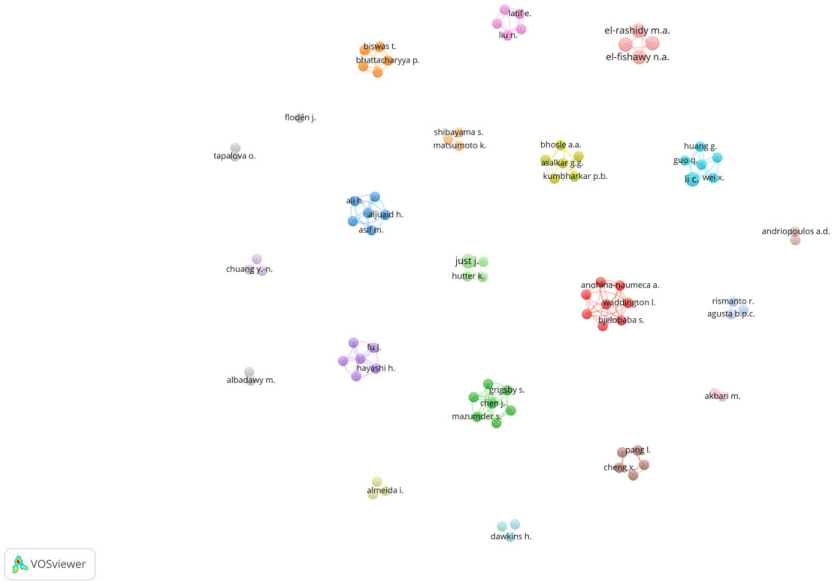

Figure 7 – Co-authorship network illustrating collaboration patterns among core authors

Intellectual Influences and Core Authors

The author co-citation network is organized into three broad clusters. The first cluster represents the core methodological foundations of Neural Natural Language Processing and Representation Learning, including highly co-cited authors such as Lee, K., Wang, Y., Chen, K., and Manning, C.D., which exhibit strong co-citation link strength.

The second cluster is centered on plagiarism detection, authorship analysis, and text forensics. Rosso appears as the central node in this group, with particularly strong co-citation ties observed between Potthast and Stein. Structurally, this cluster is clearly differentiated from the neural NLP cluster, with relatively limited connections between them.

A third, smaller cluster occupies an intermediate position between the neural NLP core and the plagiarism cluster, suggesting a bridging role within the intellectual structure of the field. The composition and spatial positioning of these clusters are illustrated in Figure 8.

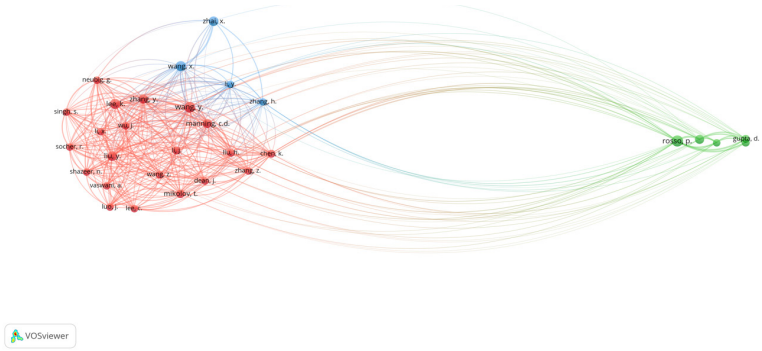

Figure 8 – Author co-citation network revealing the intellectual foundations of generative AI

Thematic Synthesis (Systematic Review)

Following the content analysis of the literature reviewed, four thematic directions emerged:

1. *Semantic integrity and plagiarism detection.*

This thematic area covers research issues related to the analysis of text similarity, plagiarism and paraphrasing, especially how they are applied in the academic environment (Amirzhanov et al., 2025; Darwish et al., 2023; Hryn et al., 2025; Sajid et al., 2025). The literature on this topic area included the use of classical algorithms to calculate the similarity between lexical texts, such as TF-IDF or cosine similarity, and the use of the latest transformer-based similarity models based on contextual embeddings to identify deep semantic connections (Amirzhanov et al., 2025; Hassan et al., 2025).

2. *Research topic novelty and originality assessment.*

Semantic novelty detection and redundancy identification studies are related to this thematic direction and define novelty as a change in topical context, in relation to an already existing or normal textual context (Ghosal et al., 2022; Just et al., 2024; Mat et al., 2021). The reviewed studies describe various applications of document embedding, contextual similarity and topic modeling techniques, where novelty is proved by semantic ranking or prioritized in comparison to a bibliographic context. Certain applications of the large language models in terms of assessing novelty or comparing ideas were observed (Wu et al., 2025).

3. *Intelligent recommendation and allocation systems.*

Supervisor allocation and recommendation in higher education constitute a distinct thematic direction within the reviewed literature, as they belong to this thematic direction (Anghel and Popescu, 2022; Dasri et al., 2025; Ramotsisi et al., 2022). Annisa et al. (2024) introduce a two-way recommendation system involving

the use of historical data of alumni and k-skyband view queries to make a broader span of recommendations than skyline objects, whereas Dasri et al. (2025) introduce a MapReduce-based implementation suitable to large academic datasets.

4. *Human-oriented management and decision-making.*

Research papers on the same topic that address transparency, explainability, and human interaction scenario within academic management systems through artificial intelligence fall within this thematic area (Khalifa and Albadawy, 2024; Ramotsisi et al., 2022; Shimray and Subaveerapandiyan, 2025). In parallel, methodological NLP research highlights the need for traceable and human-interpretable reasoning pipelines to support explainability (Fox et al., 2025).

Across all thematic areas, the reviewed studies report the use of embedding-based semantic representation, large language model-based methods, and lexical similarity-based approaches, as illustrated by various reports in the reviewed literature. The reported use of embedding-based and large language model-based methods was found in literature published in later years of the review period, while lexical similarity-based methods were illustrated as resource-efficient approaches.

Based on the reviewed literature, the data sources in this study, also indicate academic text data, including theses, research projects, and publications, as well as bibliographic, citation, and administrative data. Evaluation results based on real-world institutional deployments or long-term evaluations are rarely reported (Hryn et al., 2025).

The thematic distribution of the reviewed studies regarding semantic similarity, the evaluation of originality, recommendation systems, and the examined domains of governance is summarized in the next subsections.

Identified Research Gaps

An examination of the literature reveals several recurring gaps in the topics regarding the evaluation practices, data sources, methodological consistency, system integration in AI-based management research.

1. *Real-world institutional deployment and longitudinal evaluation.*

The reviewed studies report limited evidence of real-world institutional deployment or long-term evaluation. Many articles are based on autonomous experiments, environmental modeling, or retrospective analysis, with little or no evaluation of system performance in various academic terms (Ramotsisi et al., 2022). Only a small number of studies provide estimates based on real-world educational conditions and institutional project repositories (Luis et al., 2025).

2. *Use of primary institutional and administrative data.*

The reviewed literature indicates that frequent data sources include secondary data sources such as bibliographic databases, publication databases, and archival texts. Information related to academic texts is often described, while information related to institutional administration datasets is less frequently reported. Furthermore, datasets combining administrative data, texts, and evaluative data using a singular data analysis framework are rarely described.

3. *Management, transparency, and explainability mechanisms.*

The reviewed studies lack detailed information on interpretability, verifiability, and expert confirmation. This is consistent with broader NLP findings, according to which modern similarity pipelines usually do not guarantee explainability and the need to provide traceable hybrid reasoning procedures (Fox et al., 2025).

4. *Evaluation protocols and benchmarking practices.*

From the evaluated literature, numerous heterogeneous assessment schemes, performance criteria, and experimental designs are presented. Moreover, in the literature, various dataset creation methods, similarity boundaries, and assessment methods are recognized, yet the achievement of standard benchmarks is less emphasized.

5. *Integration of multiple academic management functions.*

The majority of reviewed works reported systems for individual functions like plagiarism detection, checking for novelty, or recommendation. Accounts of integrated system architectures for all or several academic management functions within a single system remain relatively scarce (Hryn et al., 2025; Ramotsisi et al., 2022; Kashkimbayeva et al., 2025).

Discussion.

This section interprets the findings that have been made and places these findings within the context of the wider body of research on AI-driven academic project management. Rather than reiterating the results presented in the systematic review and bibliometric analysis, the discussion focuses on explaining why the identified thematic structures, adopted methods, and observed collaboration patterns are relevant. The discussion is intended to align directly with the findings that have been made.

Key Trends Interpretation and Comparison with Previous Literature.

Considering the quantitative distribution of methodology paradigms, the development of AI-based academic project management can be discussed as evolutionary rather than disruptive. Even though classical lexical and traditional methods of machine learning are still a major part (49%), methods based on embedding and LLM are a narrow majority of recent works (51%). Such a balance implies the progressive change towards the semantically-based decision-support systems instead of replacing previous methods entirely.

The growing popularity of embedding and transformer-based models is an indicator of an appreciation of the weaknesses of superficial lexical similarity to identify paraphrased or semantically manipulated academic content (Amirzhanov et al., 2025; Darwish et al., 2023). In this regard, is increasingly conceptualized as conceptual positioning within an evolving research space, extending previous theoretical reflections (Ghosal et al., 2022).

Meanwhile, the emergence of the approaches based on LLM indicates growing interest in the abstraction-based reasoning and discourse-level assessment (Wu

et al., 2025). However, their coexistence with lexical and embedding approaches imply methodological diversification as opposed to paradigm supremacy. All these findings combine to present a growing yet structurally plural field of AI uses in academic integrity, originality assessment, and supervisor assignment.

Theoretical and Practical Contributions

Theoretically, the review is valuable for the literature because it combines both bibliometric analysis and systematic review to analyze how research streams are organized in relation to each other within a single methodological concept. The analysis of joint citations and keywords can also be used to determine that the current trend of research in the field of intellectual academic management is more related to progress in the field of basic NLP architectures than to the unification of theoretical traditions.

From a practical perspective, the results provide evidence-based guidelines that can be used in the design of more intelligent academic management systems. The reviewed literature shows that more effective systems are now being designed based on hybrid models that include semantic similarity modeling and rule systems or optimization systems to facilitate decision-making in academic institutions (Ramotsisi et al., 2022; Nurgaliyeva et al., 2025). Nevertheless, the low level of functional domain integration implies the need for unified systems that incorporate tasks like integrity, originality, and allocation in an integrated fashion. In addition to developing and harmonizing the results of previous research, the current study has the additional advantage that offers a conceptual model for developing an integrated academic project management system based on artificial intelligence, combining the main functional aspects established in the current review of scientific literature (Figure 9).

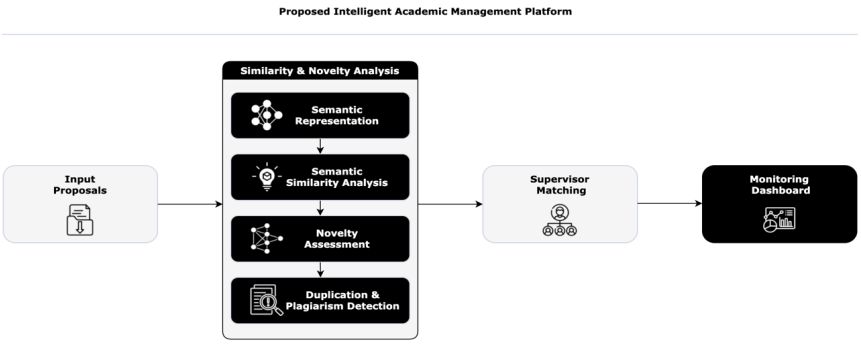

Figure 9 – Conceptual framework of an AI-based academic project management platform

In the proposed concept, semantic representation and similarity analysis constitute increasingly central structural components that serve as the basis for the implementation of various academic management operations, such as duplication control, novelty assessment, and resource allocation control. By illustrating how these components can be interconnected, this framework eliminates the

fragmentation of previous studies, in which systems for evaluating integrity, originality, and recommendations were considered largely as independent processes. Consequently, this framework provides a systematic view of how semantically sound artificial intelligence technologies can be used to facilitate the lifecycle of an academic project as a whole.

Methodological Observations and Limitations

There are certain trends in assessment methodology that need to be considered: first, there are still significant differences in assessment methods, with different datasets, similarity measures, and different assessment metrics, making it difficult to compare methods directly. Secondly, there is also a high level of use of secondary or bibliographic data with fewer reports on the use of administrative data, which partially affects the ecological validity of the assessment. Thirdly, the results of a real assessment are presented only in a limited number of studies. For example, Luis et al. (2025) evaluate the Augmented Search Generation System (RAG) to evaluate originality according to the instructors' judgments, while many other proposed systems remain at the conceptual or experimental stage.

These limitations indicate that the existing evidence is largely based on technical capabilities rather than institutional ones, and it is recommended that the former be taken into account when calculating accuracy or efficiency indicators.

Directions for Future Research

In light of the gaps mentioned above, the upcoming research should aim at empirical validation of an academic management system based on artificial intelligence, especially with regard to its effectiveness in the long term, across multiple academic cycles. Similar and convincing evidence will also be easier to collect through the introduction of standardized assessment procedures. In addition, additional research can be conducted on the architecture of the integrated system, which is a combination of a semantic integrity control scheme, a novelty assessment scheme, and a recommendation scheme.

Finally, research should be continued beyond the English-speaking space in order to generalize and make more practical academic management systems based on artificial intelligence, especially given that recent empirical evidence has shown widespread use of artificial intelligence tools in academic literature and other issues such as reliability, cost, and ethical management (Shimray and Subaveerapandiyan, 2025), as well as the growing complexity of new forms of plagiarism, such as paraphrased, interlanguage plagiarism and content developed by artificial intelligence (Amirzhanov et al., 2025).

Conclusion. The review is a summary of the literature on artificial intelligence in the field of academic project management in higher education, representing a comprehensive systematic analysis and bibliometric review. By consolidating evidence across semantic integrity control, research novelty assessment, and supervisor allocation systems, the study highlights a gradual shift from purely

lexical approaches toward embedding-based and large language model-driven paradigms. Rather than operating as isolated functions, these domains increasingly converge within semantically informed decision-support frameworks.

The results demonstrate a methodologically plural but structurally fragmented research landscape, characterized by heterogeneous evaluation protocols, limited institutional deployment, and scarce longitudinal validation. Meanwhile, the increasing use of transformer architecture and LLM-based logic points to the growth of higher-level abstraction and governance-oriented AI applications in academic field.

Future research should focus on empirical verification in real-world institutional environments, standardization of benchmarking processes, and modeling the architecture of an integrated system that combines integrity, originality, and allocation processes. With the continued development of artificial intelligence technologies, their application in academic project management will only expand, which requires complete technological progress, as well as clear, understandable and ethically justified implementation practices.

References

Anghel F., and Popescu E. (2022) Dynamic Graduation Project Allocation Based on Student-Teacher Profile Compatibility. *International Conference on Advanced Learning Technologies*. — P. 56–60. *https://doi.org/10.1109/ICALT55010.2022.00025* (in Eng.)

Annisa A., Adyatma M.R., Sampurno G I., and Li C. (2024) Two-way recommendation system for supervisor selection using historical data and skyband-view queries. *Journal of Advanced Research in Applied Sciences and Engineering Technology*, *34*(2). — P. 305–314. *https://doi.org/10.37934/araset.34.2.305314* (in Eng.)

Amirzhanov A., Turan C., and Makhmutova A. (2025) Plagiarism types and detection methods: a systematic survey of algorithms in text analysis. *Frontiers in Computer Science*, *7*. *https://doi.org/10.3389/fcomp.2025.1504725 (in Eng.)*

Bekish U., Nurgabyl D., Tulegulov A., Abdoldinova G., and Basheyeva Z. (2020) Methodology of designing the pithy components of elective courses on the higher mathematics of pedagogical profile. *International Journal of Engineering Research and Technology*, *13*(3). — 407 p. *https://doi.org/10.37624/ijert/13.3.2020.407-413* (in Eng.)

Bibi S., Trinanda K., Syamsudin M., Suheri Fitri N.A., Suharsono, Rigaux M., and Guarim R. (2023) *A Neural Network Recommendation Model for Final Project Topic Selection*. — P. 407–410. *https://doi.org/10.1109/ice3is59323.2023.10335346* (in Eng.)

Dasri D., Annisa A., and Haryanto T. (2025) Two-way thesis supervisor recommendation system using MapReduce k-skyband view queries. *JOIV International Journal on Informatics Visualization*, *9*(1). — 163 p. *https://doi.org/10.62527/joiv.9.1.2800* (in Eng.)

Darwish S. M., Mhaimeed I. A., and Elzoghabi A.A. (2023) A quantum genetic algorithm for building a semantic textual similarity estimation framework for plagiarism detection applications. *Entropy*, *25*(9). — 1271 p. *https://doi.org/10.3390/e25091271* (in Eng.)

Fox O.R., Bergami G., and Morgan G. (2025) Verified language processing with hybrid explainability. *Electronics*, 14(17). — 3490 p. *https://doi.org/10.3390/electronics14173490 (in Eng.)*

Ghosal T., Saikh T., Biswas T., Ekbal A., and Bhattacharyya P. (2022) Novelty detection: A perspective from natural language processing. *Computational Linguistics*, *48*(1). — P. 77–117. *https://doi.org/10.1162/coli_a_00429* (in Eng.)

Hassan E., Talaat A.S., and Elsabagh M.A. (2025) Intelligent text similarity assessment using RoBERTa with integrated chaotic perturbation optimization techniques. *Journal of Big Data*, *12*(1). *https://doi.org/10.1186/s40537-025-01233-3* (in Eng.)

Hryn O., Shevel A., Shcherbyna N., Kubrak,O., and Zadorozhnyi K. (2025) Implementation of artificial intelligence in the system for detecting academic dishonesty in Ukrainian secondary and higher education institutions. *Periodicals of Engineering and Natural Sciences (PEN)*, *13*(2). — P. 445–458. *https://doi.org/10.21533/pen.v13.i2.467* (in Eng.)

Just J., Ströhle T., Füller J., and Hutter K. (2024) AI-based novelty detection in crowdsourced idea spaces. *Innovation*, *26*(3), 359–386. *https://doi.org/10.1080/14479338.2023.2215740* (in Eng.)

Kashkimbayeva N., Bekish U., Abdoldinova G., Basheyeva Z., and Mitroshina A. (2025) Evaluating an integrated

user-feedbackapproach to software quality monitoring: enhancing accuracy and timeliness. *Eastern-European Journal of Enterprise Technologies*, 5(2 (137)). — P. 29–42. *https://doi.org/10.15587/1729-4061.2025.341516* (in Eng.)

Khalifa M., and Albadawy M. (2024) Using artificial intelligence in academic writing and research: An essential productivity tool. *Computer Methods and Programs in Biomedicine Update*, *5*. — 100145 p. *https://doi.org/10.1016/j.cmpbup.2024.100145* (in Eng.)

Korade N. B., Salunke M. B., Bhosle A.A., Kumbharkar P.B., Asalkar G.G., and Khedkar R.G. (2024) Strengthening sentence similarity identification through OpenAI embeddings and deep learning. *International Journal of Advanced Computer Science and Applications*, *15*(4). *https://doi.org/10.14569/ijacsa.2024.0150485* (in Eng.)

Luis S.Y., Reina D.G., and Marín S.T. (2025) Towards a retrieval-augmented generation framework for originality evaluation in projects-based learning classrooms. *Education Sciences*, *15*(6). — 706 p. *https://doi.org/10.3390/educsci15060706* (in Eng.)

Mat N., Politowicz A., Mazumder S., Chen J., Liu B., Robertson E., and Grigsby S. (2021) Semantic novelty detection in natural language descriptions. *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing.* — P. 866–882. *https://doi.org/10.18653/v1/2021.emnlp-main.66* (in Eng.)

Nurgaliyeva S., Amangali M., Basheyeva Z., Kashkimbayeva N., and Amantayev D. (2025) Design and evaluation of an intelligent waste monitoring system based on RGIS integration for smart cities. *Eastern-European Journal of Enterprise Technologies*, 4(9). — P. 70–78. *https://doi.org/10.15587/1729-4061.2025.337033* (in Eng.)

Ramotsisi J., Kgomotso M., and Seboni L. (2022) An optimization model for the student-to-project supervisor assignment problem: The case of an engineering department. *Journal of Optimization*, *2022.* — P. 1–11. *https://doi.org/10.1155/2022/9415210 (in Eng.)*

Sajid M., Sanaullah M., Fuzail M., Malik T. S., and Shuhidan S.M. (2025) Comparative analysis of text-based plagiarism detection techniques. *PLoS ONE*, *20*(4), e0319551. *https://doi.org/10.1371/journal.pone.0319551 (in Eng.)*

Shibayama S., Yin D., and Matsumoto K. (2021) Measuring novelty in science with word embedding. *PLoS ONE*, *16*(7), e0254034. *https://doi.org/10.1371/journal.pone.0254034 (in Eng.)*

Shimray S.R., and Subaveerapandiyan, A. (2025). Artificial intelligence in academic writing and research: Adoption and effectiveness. *Open Information Science*, *9*(1). *https://doi.org/10.1515/opis-2025-0026 (in Eng.)*

Wu W., Zhang C., and Zhao Y. (2025) Automated novelty evaluation of academic paper: A collaborative approach integrating human and large language model knowledge. *Journal of the Association for Information Science and Technology*, *76*(11). — P. 1452–1469. *https://doi.org/10.1002/asi.70005 (in Eng.)*

Publication Ethics and Publication Malpractice
in the journals of the Central Asian Academic Research Center LLP

For information on Ethics in publishing and Ethical guidelines for journal publication see http://www.elsevier.com/publishingethics and http://www.elsevier.com/journal-authors/ethics.

Submission of an article to the journals of the Central Asian Academic Research Center LLP implies that the described work has not been published previously (except in the form of an abstract or as part of a published lecture or academic thesis or as an electronic preprint, see http://www.elsevier.com/postingpolicy), that it is not under consideration for publication elsewhere, that its publication is approved by all authors and tacitly or explicitly by the responsible authorities where the work was carried out, and that, if accepted, it will not be published elsewhere in the same form, in English or in any other language, including electronically without the written consent of the copyright-holder. In particular, translations into English of papers already published in another language are not accepted.

No other forms of scientific misconduct are allowed, such as plagiarism, falsification, fraudulent data, incorrect interpretation of other works, incorrect citations, etc. The Central Asian Academic Research Center LLP follows the Code of Conduct of the Committee on Publication Ethics (COPE), and follows the COPE Flowcharts for Resolving Cases of Suspected Misconduct (http://publicationethics.org/files/u2/New_Code.pdf). To verify originality, your article may be checked by the Cross Check originality detection service http://www.elsevier.com/editors/plagdetect.

The authors are obliged to participate in peer review process and be ready to provide corrections, clarifications, retractions and apologies when needed. All authors of a paper should have significantly contributed to the research.

The reviewers should provide objective judgments and should point out relevant published works which are not yet cited. Reviewed articles should be treated confidentially. The reviewers will be chosen in such a way that there is no conflict of interests with respect to the research, the authors and/or the research funders.

The editors have complete responsibility and authority to reject or accept a paper, and they will only accept a paper when reasonably certain. They will preserve anonymity of reviewers and promote publication of corrections, clarifications, retractions and apologies when needed. The acceptance of a paper automatically implies the copyright transfer to the Central Asian Academic Research Center LLP.

The Editorial Board of the Central Asian Academic Research Center LLP will monitor and safeguard publishing ethics.

Requirements for articles design for publication in the journal are available on the websites:
www:nauka-nanrk.kz
http://physics-mathematics.kz/index.php/en/archive
ISSN2518-1726 (Online),
ISSN 1991-346X (Print)

Managing Editor: *A.Shormakova*
Editors: *D.S. Alenov, T. Apendiev*
Computer layout: *G.D. Zhadyranova*

Signed for print: June 15, 2026
Format: 70×90 1/16. 26.5 printed sheets. Order No. 2.

«Central Asian Academic Research Center» LLP
Almaty, Shevchenko Street, 28