ISSN: 1991-3494 (Print)
ISSN: 2518-1467 (Online)

SCIENTIFIC JOURNAL OF PEDAGOGY AND ECONOMICS

№3
2026

ISSN 2518-1467 (Online),
ISSN 1991-3494 (Print)

CENTRAL ASIAN ACADEMIC
RESEARCH CENTER

SCIENTIFIC JOURNAL OF PEDAGOGY AND ECONOMICS

PUBLISHED SINCE 1944

3 (421)

May – June 2026

ALMATY, 2026

Scientific Journal of Pedagogy and Economics
ISSN 2518-1467 (Online),
ISSN 1991-3494 (Print).
Owner: «Central Asian Academic Research Center» LLP (Almaty).
The certificate of registration of a periodical printed publication in the Committee of information of the Ministry of Information and Communications of the Republic of Kazakhstan
№ KZ50VPY00121155, issued on 05.06.2025
Thematic focus: *«publication of the results of new achievements inthe field of fundamental sciences»*
Periodicity: 6 times a year.
http://www.bulletin-science.kz/index.php/en/

Scientific Journal of Pedagogy and Economics
ISSN 2518-1467 (Online),
ISSN 1991-3494 (Print).
Меншіктенуші: «Орталық Азия академиялық ғылыми орталығы» ЖШС (Алматы қ.).
Қазақстан Республикасының Ақпарат және коммуникациялар министрлігінің Ақпарат комитетінде 05.06.2025 ж. берілген **№ KZ50VPY00121155** мерзімдік басылым тіркеуіне қойылу туралы куәлік.
Тақырыптық бағыты: *«іргелі ғылым салалары бойынша жаңа жетістіктердің нәтижелерін жариялау»*
Мерзімділігі: жылына 6 рет.
http://www.bulletin-science.kz/index.php/en/

Scientific Journal of Pedagogy and Economics
ISSN 2518-1467 (Online),
ISSN 1991-3494 (Print).
Собственник: ТОО «Центрально-азиатский академический научный центр» (г. Алматы).
Свидетельство о постановке на учет периодического печатного издания в Комитете информации Министерства информации и коммуникаций и Республики Казахстан
№ KZ50VPY00121155 выданное 05.06.2025 г.
Тематическая направленность: *«публикация результатов новых достижений вобласти фундаментальных наук».*
Периодичность: 6 раз в год.
http://www.bulletin-science.kz/index.php/en/

CONTENTS

PEDAGOGY

ECONOMICS

МАЗМҰНЫ

ПЕДАГОГИКА

ЭКОНОМИКА

СОДЕРЖАНИЕ

ПЕДАГОГИКА

ЭКОНОМИКА

SCIENTIFIC JOURNAL OF PEDAGOGY AND ECONOMICS
ISSN 1991-3494
Volume 3.
Number 421 (2026), 364-379

https://doi.org/10.32014/2026.2518-1467.1213

IRSTI 14.35.07
UDC 378.147.004.04

© Shyndaliev N.T., Serikkazy A.S., Nurbekova G.F.*, 2026.
L.N. Gumilyov Eurasian National University, Astana, Kazakhstan.
E-mail: gulnurfaz@mail.ru

ENHANCING STUDENTS' PROFESSIONAL COMPETENCIES IN BIG DATA PROCESSING THROUGH ARTIFICIAL INTELLIGENCE: RESULTS OF A PEDAGOGICAL EXPERIMENT

Shyndaliev Nurzhan — candidate of pedagogical sciences, Associate Professor, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,
E-mail: nurazhan0412@gmail.com, https://orcid.org/0000-0001-5284-3526;
Serikkazy Adilet — 1st year doctoral student of the specialty «8D01511 – Computer Science», L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,
E-mail: adiletserikkazy2@gmail.com, https://orcid.org/0009-0005-5387-6254;
Nurbekova Gulmira — PhD, Assistant Professor, L.N. Gumilyov Eurasian National University, Astana, Kazakhstan,
E-mail: gulnurfaz@mail.ru, https://orcid.org/0009-0005-8468-9523.

Abstract. The rapid development of Big Data and Artificial Intelligence (AI) technologies is imposing new requirements on higher education systems in terms of developing future specialists' data-handling skills, analytical thinking, and digital competencies. At present, Big Data and AI technologies are predominantly taught in higher education institutions as separate disciplines, while their integrated and comprehensive application remains insufficiently developed. In this regard, the development and implementation of project-based learning models grounded in the use of artificial intelligence tools represent an important scientific and pedagogical challenge. The aim of the study is to experimentally determine the pedagogical effectiveness of a competency-based project learning model implemented through the integration of artificial intelligence technologies for developing professional competencies in the field of Big Data. The study was based on the Design-Based Research methodology. The participants included 22 students enrolled in the 6B01511 – "Computer Science" educational program at the L.N. Gumilyov Eurasian National University. Students' professional competencies were assessed in the areas of data collection and preprocessing, machine learning, data interpretation, scientific visualization, and analytical report preparation.

Assessment was conducted using project artifacts evaluated through a rubric-based system. Quantitative data were processed in SPSS using descriptive statistics, paired-samples t-tests, and Cohen's d effect size indicator, while qualitative data were interpreted through the analysis of project reports. The research findings revealed statistically significant differences between pre-test and post-test results. The highest levels of competency development were identified in the areas of data preparation, visualization, and scientific interpretation. In addition, the development of students' analytical, research, and interpretive skills was observed. The proposed pedagogical model can be recommended for implementation in higher education teaching practices related to Big Data and artificial intelligence and may serve as an effective methodological tool for improving the professional training of future computer science specialists. This is of significant importance for the development of future specialists' professional competencies in the context of the digital transformation of the educational process.

Keywords: Big Data education, artificial intelligence, professional competencies, pedagogical model, projects, higher education

For citations: Shyndaliev N.T., Serikkazy A.S., Nurbekova G.F. Enhancing students' professional competencies in big data processing through artificial intelligence: results of a pedagogical experiment. Scientific Journal of Pedagogy and Economics, 2026. — No.3. — P. 364-379. DOI. https://doi.org/10.32014/2026.2518-1467.1213

© Шындалиев Н.Т., Серікқазы Ә.С., Нурбекова Г.Ф.*, 2026.
Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан.
E-mail: gulnurfaz@mail.ru

ЖАСАНДЫ ИНТЕЛЛЕКТ АРҚЫЛЫ СТУДЕНТТЕРДІҢ ҮЛКЕН ДЕРЕКТЕРДІ ӨҢДЕУДЕГІ КӘСІБИ ҚҰЗЫРЕТТІЛІГІН АРТТЫРУ: ПЕДАГОГИКАЛЫҚ ЭКСПЕРИМЕНТ НӘТИЖЕЛЕРІ

Шындалиев Нуржан — педагогика ғылымдарының кандидаты, қауымдастырылған профессор, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,
E-mail: nurazhan0412@gmail.com, https://orcid.org/0000-0001-5284-3526;
Серікқазы Әділет — «8D01511-Информатика» білім беру бағдарламасының 1 курс докторанты, Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,
E-mail: adiletserikkazy2@gmail.com, https://orcid.org/0000-0001-5284-3526;
Нурбекова Гульмира — PhD, қауымдастырылған профессор м.а., Л.Н. Гумилев атындағы Еуразия ұлттық университеті, Астана, Қазақстан,
E-mail: gulnurfaz@mail.ru, https://orcid.org/0009-0005-8468-9523.

Аннотация. Үлкен деректер (Big Data) мен жасанды интеллект (Artificial Intelligence) технологияларының қарқынды дамуы жоғары білім беру жүйесінде болашақ мамандардың деректермен жұмыс істеу, аналитикалық

ойлау және цифрлық құзыреттіліктерін қалыптастыруға жаңа талаптар қойып отыр. Қазіргі уақытта жоғары оқу орындарында Big Data және AI технологиялары көбіне жеке пәндер ретінде оқытылып, олардың өзара кіріктірілген және кешенді қолданылуы жеткілікті деңгейде жүзеге асырылмай келеді. Осыған байланысты жасанды интеллект құралдарына негізделген жобалық оқыту модельдерін әзірлеу мен енгізу өзекті ғылыми-педагогикалық мәселелердің бірі болып табылады. Зерттеудің мақсаты – жасанды интеллект технологияларын кіріктіру арқылы жүзеге асырылатын құзыреттілікке негізделген жобалық оқыту моделінің Big Data саласындағы кәсіби құзыреттіліктерді дамытудағы педагогикалық тиімділігін эксперименттік тұрғыдан анықтау. Зерттеу дизайн-бағдарланған зерттеу (Design-Based Research) әдіснамасына негізделді. Зерттеуге Л.Н. Гумилев атындағы Еуразия ұлттық университетінің 6В01511 – «Информатика» білім беру бағдарламасының 22 студенті қатысты. Студенттердің кәсіби құзыреттіліктері деректерді жинау және алдын ала өңдеу, машиналық оқыту, деректерді интерпретациялау, ғылыми визуализация және аналитикалық есептерді дайындау бағыттары бойынша бағаланды. Бағалау рубрикаға негізделген жобалық артефакттар арқылы жүзеге асырылды. Сандық деректер SPSS бағдарламасында сипаттамалық статистика, paired-samples t-test және Cohen’s d көрсеткіштері арқылы өңделді, ал сапалық деректер жобалық есептерді талдау негізінде интерпретацияланды. Зерттеу нәтижелері алдын ала және қорытынды тест көрсеткіштері арасында статистикалық мәнді айырмашылықтардың бар екенін көрсетті. Әсіресе деректерді дайындау, визуализациялау және ғылыми интерпретациялау бағыттары бойынша жоғары құзыреттілік деңгейі анықталды. Сонымен қатар студенттердің аналитикалық, зерттеушілік және интерпретациялық дағдыларының дамығаны байқалды. Ұсынылған педагогикалық модель жоғары оқу орындарында Big Data және жасанды интеллект пәндерін оқыту процесіне енгізуге ұсынылады және болашақ информатика мамандарының кәсіби даярлығын жетілдіруге бағытталған тиімді әдістемелік құрал ретінде қарастырылуы мүмкін. Бұл білім беру процесін цифрлық трансформациялау жағдайында болашақ мамандардың кәсіби құзыреттіліктерін дамыту үшін маңызды болып табылады.

Түйін сөздер: үлкен деректер бойынша білім беру, жасанды интеллект, кәсіби құзыреттілік, педагогикалық модель, жобалар, жоғары білім

© Шындалиев Н.Т., Серикказы А.С., Нурбекова Г.Ф.*, 2026.
Евразийский национальный университет имени Л.Н. Гумилёва,
Астана, Казахстан
E-mail: gulnurfaz@mail.ru

ПОВЫШЕНИЕ ПРОФЕССИОНАЛЬНОЙ КОМПЕТЕНТНОСТИ СТУДЕНТОВ В ОБРАБОТКЕ БОЛЬШИХ ДАННЫХ С ИСПОЛЬЗОВАНИЕМ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА: РЕЗУЛЬТАТЫ ПЕДАГОГИЧЕСКОГО ЭКСПЕРИМЕНТА

Шындалиев Нуржан — кандидат педагогических наук, ассоциированный профессор, Евразийский национальный университет имени Л.Н. Гумилёва, Астана, Казахстан,
E-mail: nurazhan0412@gmail.com, https://orcid.org/0000-0001-5284-3526;
Серикказы Адилет — докторант, Евразийский национальный университет имени Л.Н. Гумилёва, Астана, Казахстан,
E-mail: adiletserikkazy2@gmail.com, https://orcid.org/0000-0001-5284-3526;
Нурбекова Гульмира — PhD, и.о. ассоциированного профессора, Евразийский национальный университет имени Л.Н. Гумилёва, Астана, Казахстан,
E-mail: gulnurfaz@mail.ru, https://orcid.org/0009-0005-8468-9523.

Аннотация. Стремительное развитие технологий больших данных (Big Data) и искусственного интеллекта обусловливает необходимость совершенствования профессиональной подготовки будущих специалистов в области информатики и развития у них компетенций, связанных с анализом данных, машинным обучением и принятием решений на основе данных. В статье представлены результаты педагогического эксперимента, направленного на оценку эффективности компетентностно-ориентированной модели проектного обучения, основанной на интеграции инструментов искусственного интеллекта в процесс изучения технологий Big Data. Исследование выполнено в рамках методологии дизайн-ориентированного исследования (Design-Based Research). В эксперименте приняли участие 22 студента образовательной программы «6B01511 – Информатика» Евразийского национального университета имени Л.Н. Гумилёва. Оценка сформированности профессиональных компетенций осуществлялась по направлениям подготовки и предварительной обработки данных, машинного обучения, интерпретации результатов, научной визуализации и подготовки аналитических отчётов. Количественная обработка данных проводилась с использованием методов описательной статистики, парного t-критерия Стьюдента и коэффициента Cohen's d, а качественный анализ основывался на изучении проектных отчётов студентов. Результаты исследования выявили статистически значимые различия между показателями входного и итогового оценивания, а также подтвердили положительное влияние проектного обучения с использованием инструментов искусственного интеллекта на развитие аналитических, исследовательских и интерпретационных навыков

обучающихся. Наиболее выраженный рост компетентности зафиксирован в области подготовки данных, визуализации и интерпретации результатов анализа. Практическая значимость исследования заключается в возможности использования предложенной педагогической модели при реализации дисциплин, связанных с технологиями больших данных и искусственным интеллектом, в образовательных программах высших учебных заведений.

Ключевые слова: большие данные, искусственный интеллект, профессиональные компетенции, проектное обучение, педагогический эксперимент, цифровые технологии, высшее образование

Introduction. The rapid development of digital technologies is imposing new requirements on higher education systems, particularly in relation to the formation of students' information-technological and data-related competencies. In the context of ongoing digitalization, one of the key educational objectives is the systematic development of students' ability to operate in Big Data environments, their analytical thinking skills, and their proficiency in using artificial intelligence (AI) tools from the early stages of professional training. According to global analytical reports, the exponential growth of data volumes has positioned Big Data and AI-related competencies among the most in-demand professional skills in the contemporary labor market (Reinsel, Gantz and Rydning, 2018). Despite this growing demand, higher education practice often addresses these objectives in a fragmented manner. In many academic programs, Big Data, machine learning, and artificial intelligence technologies are taught as separate courses or isolated modules, while their interconnected and integrated application remains insufficiently developed. As a result, although students may acquire theoretical knowledge, they frequently experience difficulties in systematically performing the full data-processing cycle — including data collection, preprocessing, modeling, visualization, and interpretation — within authentic practical contexts (Chen, Chiang, and Storey, 2012; Blazquez and Domenech, 2018). This issue is also evident in the context of higher education in Kazakhstan. Although national digital transformation initiatives, including the "Digital Kazakhstan" program, identify the development of data literacy and competencies in artificial intelligence and Big Data as strategic priorities, scientifically grounded and practice-oriented pedagogical models for implementing these directions in the educational process remain insufficiently explored (Ministry of Digital Development of the Republic of Kazakhstan, 2021). As a result, professional training often remains limited to disciplinary knowledge, without providing sufficient opportunities for the comprehensive development and assessment of students' analytical capabilities, data-handling skills, and research competencies. Nevertheless, pedagogical models that directly target the preparation of future computer science teachers and integrate artificial intelligence and Big Data tools remain insufficiently explored. This situation underscores the need for empirically grounded instructional approaches capable of supporting the

systematic development of learners' professional knowledge and skills in data-intensive fields. In response to this research gap, the present study proposes an AI-integrated, competency-based instructional model implemented through project work and examines its pedagogical effectiveness in Big Data education. The aim of the study is to experimentally evaluate how the integration of artificial intelligence technologies into project-based learning influences students' professional preparedness for working with data. Particular attention is given to the development of competencies related to the full data-processing cycle, including data collection, preprocessing, modeling, visualization, and interpretation, as well as to students' analytical, research-oriented, and interpretative skills in authentic Big Data contexts. Accordingly, the study addresses the following research tasks:

1) To identify and conceptualize the structure of students' professional competencies in Big Data and artificial intelligence

2) To develop a pedagogical model for enhancing professional competencies in Big Data education through Artificial Intelligence integration.

3) To experimentally evaluate the pedagogical effectiveness of the proposed pedagogical model

Literary review. An analysis of studies conducted by Kazakhstani scholars indicates that a body of research aimed at integrating digital technologies and artificial intelligence into the preparation of future specialists is gradually emerging. In particular, Nurbekova G. F., in her doctoral dissertation "Theoretical and Practical Foundations of Training University Students in Big Data," examines methodological and instructional-organizational approaches to developing students' competencies in the field of Big Data within higher education. This research establishes the theoretical and practical foundations for teaching Big Data in university curricula.

In addition, the works of Serik M. and co-authors devote considerable attention to the development of students' digital and analytical competencies. By substantiating the importance of forming skills related to data handling, logical reasoning, and the use of information technologies, the authors analyze the content and teaching methods of Big Data courses, including approaches to cluster analysis (Serik et al., 2022; Serik, Abdrakhmanova and Zhanatbekova, 2023).

The theoretical and practical foundations for integrating artificial intelligence into the educational process are also addressed in the works of several scholars. In the context of preparing computer science teachers, the pedagogical potential of using programming environments and digital platforms is widely discussed in contemporary educational research. Zhunusbekova investigates students' perceptions of artificial intelligence technologies and their relationship to issues of academic integrity, while Orynbasar focuses on the challenges that arise during the introduction of AI into the educational process (Zhunusbekova, 2024; Orynbasar, 2024).

At the institutional level, Zhukabayeva et al. propose a conceptual model of

a digital university and describe the key steps for its implementation, whereas Nazyrova et al. highlight the role of technological and organizational changes in the process of AI-driven digital transformation (Zhukabayeva et al., 2025; Nazyrova et al., 2025).

Overall, the reviewed studies demonstrate increasing scholarly interest in the integration of Big Data and artificial intelligence into higher education. However, existing research primarily focuses on theoretical, technological, or organizational aspects, while pedagogical models aimed at the competency-based preparation of future computer science teachers through AI-integrated project learning remain insufficiently investigated.

Materials and methods. The research methodology is grounded in the logic of design-based research (DBR). Within the study, a pedagogical model integrating artificial intelligence and Big Data was developed and piloted in an authentic educational context. The effectiveness of the model was evaluated by analyzing changes in students' competencies using pre–post measurements. In addition, the study was based on the following methodological approaches: a competency-based approach, which was used to define the structure of professional competencies encompassing the full data-processing cycle (data collection, preprocessing, modeling, visualization, and interpretation); and a mixed-methods approach, which enabled the integrated analysis of quantitative data (SPSS results) and qualitative data (project reports).

The study was conducted in three main stages: a diagnostic stage, a formative stage, and a final stage. The participants were 22 undergraduate students from the Eurasian National University after Gumilyov who were preparing for a career in computer science in fields related to computer science.

To assess students' professional competencies, a rubric-based and artifact-based evaluation system was employed. The assessment system covered key competency areas, including data collection and preprocessing; exploratory analysis and visualization; machine learning and modeling; Explainable AI; qualitative data analysis; bibliometric and network analysis; data integration and synthesis; and data storytelling and scientific communication. Each competency was evaluated on a 0–4 scale.

As assessment evidence, students produced Python code, machine learning models, ATLAS.ti projects, VOSviewer maps, and analytical reports.

Quantitative data were processed using SPSS. Descriptive statistics (mean and standard deviation) and paired-samples t-tests were applied to identify pre–post differences for each competency. Cohen's d was calculated to estimate effect sizes. Qualitative data were analyzed through the evaluation of students' project reports using clearly defined assessment criteria, and formative feedback was provided to support competence development.

Results. *Results related to Objective 1: Identification of the Structure of Students' Professional Competencies in Big Data and Artificial Intelligence.*

Scholars highlight that the preparation of professionals in the fields of Big Data and AI requires the development of multiple competencies, with technical, analytical, and ethical competencies identified as the main categories. In our study, we substantiated the set of competencies necessary for improving the knowledge and developing the skills of future professionals in these two fields based on findings from existing academic research.

One of the most essential competencies is data processing and preprocessing. Gurcan and colleagues found that professionals devote a significant portion of their time to data preprocessing tasks. Their study highlights that activities such as error correction, handling missing values, data transformation, and standardization consume substantial time during the data analysis process. Similarly, Martins et al. identify data cleaning as a critical component of analytics, emphasizing its direct impact on the accuracy and reliability of subsequent machine learning models (Gurcan et al., 2025, Martins et al., 2025).

In this regard, it becomes evident that before initiating any analytical operations, professionals must be proficient in collecting, cleaning, and managing large datasets. This includes the ability to detect and correct errors in "messy" data, ensure data quality, and prepare datasets for large-scale analysis. Thus, data preprocessing is not merely a technical step but a foundational competency that underpins the success of all subsequent AI and Big Data processes.

Another crucial skill is the ability to perform data visualization and exploratory analysis. This entails the use of charts, graphs, and other interactive techniques for the analysis and communication of insights. It also entails the use of statistical analysis or machine learning techniques on Big Data and the representation of the findings through the use of data visualization tools. For instance, a good Big Data analyst should ideally know how to handle tools such as Spark's MLlib, R or Python (pandas) for analysis, and Tableau or matplotlib for creating visualizations. Although data visualization might seem as a smaller skill for a Big Data professional, studies indicate that it is a very important skill for the analysis and interpretation of findings from large datasets (Gurcan et al., 2025).

Machine learning and modeling is one of the competencies that is central and directly developed within the context of artificial intelligence. The area encompasses training professionals so that they understand and can apply various techniques of machine learning, such as supervised learning (regression, classification), unsupervised learning (clustering, dimensionality reduction), and reinforcement learning. The understanding of models learning from data and the judgment regarding its evaluation also depend on a sound knowledge of probability and statistics, which is why definitions of AI competencies always stress "machine learning algorithms, data analysis, and mathematics" as its necessity (Serik et al., 2022).

Although Big Data and AI have much in common when it comes to foundational skills, they also diverge in several specialized competencies. Table 1 below highlights these areas of convergence and distinction across the two domains.

Table 1 — Comparison of competencies in Big Data and AI

Competency	Big Data	AI
Programming and Software Skills (Python, Java, libraries)	Core	Core
Data Analysis and Visualization	Yes	Yes
Mathematics and Statistics Foundations	Yes (basic)	Yes (advanced)
Distributed Computing and Big Data Tools (Hadoop, Spark, NoSQL)	Core	Mostly not needed
Data Storage Architecture (databases, data lakes, ETL)	Core	Not a focus
Machine Learning Algorithms	Often utilized	Core
Deep Learning and Neural Networks	Not typical	Core
Algorithmic/Computational Thinking	Yes	Yes (critical)
Model Interpretation and Explainability	in analytics context	Core (XAI, etc.)
Ethics, Security and Privacy	Yes (data privacy)	Yes (AI ethics)
Domain/Business Knowledge	Important	Important
Communication and Teamwork	Important	Important

Source: compiled by the authors based on the analysis of existing studies.

As illustrated above, programming, data literacy, and analysis skills are some of the cross-disciplinary competencies that both Big Data and AI practitioners require (Persaud et al., 2020; Gurcan et al., 2025). This applies to knowledge of ethics or domain knowledge, in addition to problem-solving and communications skills, which are also required in both fields (Nastasa et al., 2025). This indicates that there is a close nexus between Big Data and AI, where Big Data gives access to the raw material in terms of data, and AI gives the intelligence to derive knowledge from this data.

On the other hand, there are clear domain-specific competencies. Big Data specialists focus more on the "plumbing" of data – building data pipelines, managing storage, and ensuring systems can handle volume and velocity (Gurcan et al., 2025; Miller, 2014). They are experts in distributed computing frameworks and database technologies that are less emphasized in pure AI roles. In contrast, AI specialists concentrate on the "intelligence" aspect – developing algorithms, models, and AI techniques (like neural networks) to make decisions or predictions (NJIT, 2024). For example, an AI engineer working on a neural translation model requires deep knowledge of sequence-to-sequence networks and NLP, which a Big Data engineer may not need (Serik et al., 2022). Conversely, configuring a Spark cluster for a data pipeline or optimizing a NoSQL database is typically outside the AI researcher's focus. There is also a nuanced middle ground. Some competencies might appear different but are related. Machine learning is a related field that encompasses both big data and artificial intelligence. Big data analysts use machine learning for predictive analytics, while artificial intelligence specialists focus on machine learning methods and research (Nastasa et al., 2025).

Finally, big data and artificial intelligence specialists share a common foundation in computational and data skills, but differ in their specializations: big

data tends to focus on data ecosystems and scalability, while artificial intelligence tends to focus on algorithms and model complexity.

Results related to Objective 2: Development of a pedagogical model for improving professional training in Big Data using AI tools.

In response to the identified gaps in integrating Artificial Intelligence (AI) and Big Data into professional training, this section presents the results of developing a pedagogical model aimed at improving students' competencies in these fields. The model was designed based on competency-based learning principles and implemented through an AI-supported project framework. The goal was to enable students to master the full data-processing cycle—from raw data acquisition to meaningful interpretation—while simultaneously developing both technical and analytical skills in realistic settings.

The proposed model (Figure 1) is built on a spiral eco-system logic, where students engage in iterative project work cycles centered on authentic data analysis tasks. It integrates AI and Big Data tools to support various stages of learning, modeling, and reflection. The model is structured into four key stages:

1. Input Block**:** Diagnostics and baseline data (student profiles, pre-tests, data types)
2. Integration Block**:** Four project life cycles combining AI and Big Data tools
3. Assessment Block**:** Multi-point evaluation system based on competencies
4. Output Block**:** Students' developed artifacts and demonstrated learning outcomes

Each block plays a specific role in building and assessing professional readiness.

In the Artificial Intelligence and Big Data Integration block, students are expected to work on the following four stages and do project work. The first stage of the learning process is focused on data collection and preparation. For this, the use of Python, Pandas, NumPy tools is aimed at improving the skills of data search, cleaning, and normalization. The second stage of the cycle is Machine Learning and Interpretation. The goal of this cycle is to develop students' abilities in classification, clustering, performance evaluation, and development of interpretable artificial intelligence using Scikit-learn, SHAP, and LIME. The next 3 stages are focused on qualitative interpretation. For this, ATLAS.ti is used, and practical work such as text coding, thematic analysis, and mapping of joint events is carried out. The last 4th cycle is dedicated to bibliometrics and scientific visualization. Network visualization and trend analysis are performed using the VOSviewer tool. In addition, at the end of each cycle, the main artifacts are confirmed by reflective analysis and iterative refinement.

The assessment and monitoring block indicates that students' work will be assessed at three levels. First, a formative test is organized to determine the initial level of competence. In the second stage, formative assessment is carried out at the end of each cycle. Its format is as continuous feedback. In the next stage, after completing the students' projects, their reports are assessed based on the assessment criteria. The assessment framework used in the experimental study

is summarized in Table 2. Finally, a formative test is conducted with questions to determine the levels of competence. As a result, a comparative assessment is carried out using pre-test and post-test measurements.

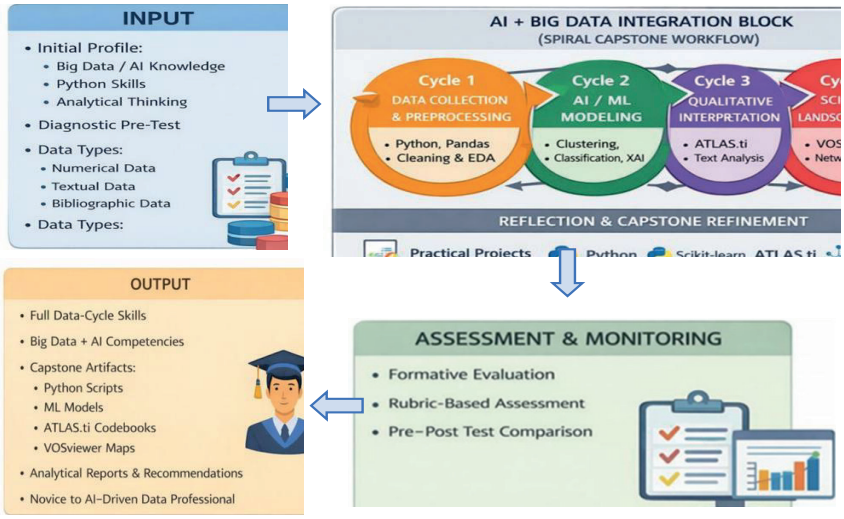


Figure 1 — AI and Big Data integrated pedagogical model for competency development through project reports

Table 2 — Assessment criteria and indicators for Big Data and AI competencies

Assessment criteria and key indicators	Low level	Medium level	High level
Cycle 1-Data collection and preparation			
Data source selection; handling missing values and anomalies; normalization and transformation; feature engineering	Performs isolated preprocessing steps without consistency	Applies multiple preprocessing techniques with partial justification	Independently completes the full data preparation pipeline with methodological justification
Cycle 2 - Machine learning and Explainable AI			
Model selection; parameter tuning; metric-based evaluation; use of SHAP/ LIME for explainability	Applies models without analytical justification	Evaluates models with limited interpretation	Develops, evaluates, and explains models with well-justified decisions
Cycle 3 - Qualitative data interpretation			
Systematic coding; hierarchical categories; co-occurrence analysis; codebook development (ATLAS.ti)	Performs basic coding without thematic integration	Constructs structured codes with partial interpretation	Conducts in-depth qualitative analysis with coherent thematic synthesis
Cycle 4 - Bibliometric and scientific visualization			
Data import; co-occurrence mapping; cluster identification; trend interpretation (VOSviewer)	Creates maps without analytical interpretation	Identifies clusters with limited explanation	Analytically interprets clusters and trends and integrates findings into the research context

Results related to Objective 3: Experimental evaluation of the pedagogical effectiveness of the proposed model.

Pre-test and Post-test Results. To examine changes in students' knowledge and competencies before and after the implementation of the proposed pedagogical model, a pre-test and post-test design was employed. Data were collected using a Google Forms questionnaire consisting of 10 structured questions, each targeting a specific aspect of Big Data and Artificial Intelligence knowledge. The same instrument was administered at both stages to ensure comparability.

Students' responses were analyzed by employing a dual analytical approach. First, students' responses were analyzed on a frequency basis (n, %), which allowed the detection of changes in the number of correct responses for each question. The responses were then coded dichotomously - 1 for correct responses and 0 for incorrect responses - and exported into the IBM SPSS Statistics data file to compute mean scores and standard deviations for statistical comparisons and further inferential analysis by paired samples t-tests, for example. The question-level analysis of pre-test and post-test results is presented in Table 3.

Table 3 — Question-level analysis of pre-test and post-test results

№	Questions	Pre-test correct n (%)	Pre-test Mean	Post-test correct n (%)	Post-test Mean
Q1	What is artificial intelligence, and how is it defined in the context of modern digital technologies?	19 (85.7%)	0.86	20 (100%)	1.00
Q2	What key characteristics define Big Data, commonly described by the 3Vs concept?	22 (100%)	1.00	20 (100%)	1.00
Q3	What is the main purpose of using Big Data technologies in data-intensive systems and organizations?	22 (100%)	1.00)	20 (100%)	1.00
Q4	How is machine learning defined as a field of artificial intelligence, and what role does it play in data analysis?	21 (95.2%)	0.95	20 (100%)	1.00
Q5	Why is data analytics considered an essential component of professional training for future specialists?	22 (100%)	1.00	20 (100%)	1.00
Q6	What are the main applications of artificial intelligence in higher education?	18 (81.8%)	0.82	19 (95.0%)	0.95
Q7	Which tools and technologies are most commonly used for Big Data processing and analysis?	14 (63.6%)	0.64	18 (90.0%)	0.90
Q8	How can artificial intelligence tools support the development of students' professional competencies?	17 (77.3%)	0.77	19 (95.0%)	0.95
Q9	What are the main stages of the Big Data processing cycle?	20 (90.9%)	0.91	20 (100%)	1.00

To visually illustrate the dynamics of students' learning outcomes, a comparative bar chart was constructed based on the pre-test and post-test mean scores. The graph (Figure 2) highlights the relative improvement across all ten questions and supports the numerical findings reported in Table 4.

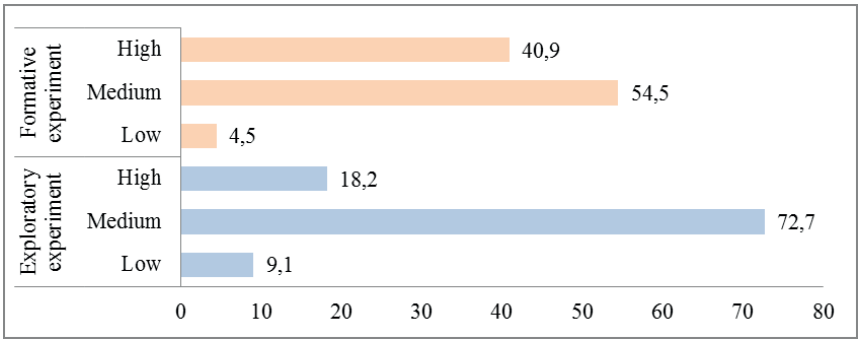
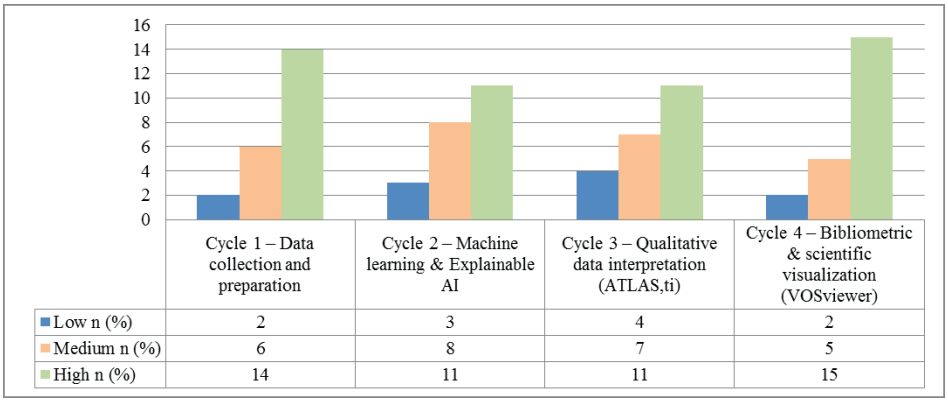


Figure 2 — Comparison of pre-test and post-test mean scores by question

The graphical comparison clearly shows that post-test scores exceeded pre-test scores for all questions, with the most pronounced gains observed in areas requiring practical and integrative knowledge. This visual evidence reinforces the conclusion that the pedagogical intervention positively influenced students' understanding of Big Data and Artificial Intelligence concepts.

The final project reports were evaluated based on the assessment criteria presented in Table 2, which are structured into four cycles. For each cycle, students' competency levels were classified as low, medium, or high. The evaluation was conducted using the final project artifacts produced by 22 students, including source code, machine learning models, analytical reports, and visualizations. As shown in Figure 3, the majority of students demonstrated high competency levels across all four project assessment cycles.



	Cycle 1 – Data collection and preparation	Cycle 2 – Machine learning & Explainable AI	Cycle 3 – Qualitative data interpretation (ATLAS,ti)	Cycle 4 – Bibliometric & scientific visualization (VOSviewer)
Low n (%)	2	3	4	2
Medium n (%)	6	8	7	5
High n (%)	14	11	11	15

Figure 3 — Distribution of students' competency levels across the four project assessment cycles

The results of the final assessment of the project reports demonstrate that the highest proportions of students at the high competency level were observed in data collection and preparation (Cycle 1) and bibliometric and scientific visualization

(Cycle 4), accounting for 63.6% and 68.2% of participants, respectively.

Although the application of machine learning and Explainable AI tools (Cycle 2) proved to be relatively more challenging for students, 50.0% of the participants nevertheless reached a high level of performance, demonstrating the ability to justify model selection and interpret results. These findings indicate that the use of artificial intelligence tools contributes to the development of students' analytical thinking, while also highlighting that the Explainable AI component requires targeted methodological support.

With respect to qualitative data interpretation (Cycle 3), a comparatively higher proportion of low-level outcomes (18.2%) was observed, which can be attributed to the analytical rather than technical complexity of systematic coding and thematic synthesis. Nevertheless, half of the students achieved a high level of competency in this cycle, successfully integrating quantitative and qualitative data.

Overall, the frequency analysis confirms that AI-integrated projects create effective conditions for the development of comprehensive professional competencies in Big Data and artificial intelligence, supporting students' ability to operate across the full data-processing and interpretation cycle.

Discussion. The findings of the study demonstrate that the integration of artificial intelligence tools into project-based learning positively influences the development of students' professional competencies in Big Data education. The results indicate that competency-based and AI-supported instructional approaches create favorable conditions for developing both technical and analytical skills required for working with large-scale data. The study confirmed that competencies related to data preprocessing, visualization, and machine learning play a central role in preparing future specialists for data-intensive environments. These findings are consistent with previous studies emphasizing the importance of data handling, analytical thinking, and machine learning competencies in Big Data and AI education (Gurcan et al., 2025; Serik et al., 2022). The comparatively higher performance observed in data preparation and scientific visualization may be explained by the practical and application-oriented nature of these tasks. Students demonstrated greater confidence when working with data cleaning, visualization tools, and bibliometric mapping. In contrast, comparatively lower results in qualitative interpretation and Explainable AI indicate that these areas require more advanced analytical reasoning and interpretative skills. This suggests the need for additional methodological support in teaching model explainability and qualitative analysis. Another important finding is that AI-integrated project work enabled students to combine quantitative and qualitative approaches within authentic educational contexts. The use of tools such as Python, Scikit-learn, SHAP, LIME, ATLAS.ti, and VOSviewer supported the development of interdisciplinary competencies and strengthened students' ability to complete the full data-processing cycle.

Overall, the study demonstrates that AI-supported project-based learning can

serve as an effective pedagogical framework for improving professional training in Big Data education. The proposed model may be applied in higher education curricula aimed at developing future specialists' professional competencies in the context of digital transformation.

Conclusion. The study was guided by three research questions, each of which was examined to address the stated aim.

The first research question concerned the identification and conceptualization of the structure of students' professional competencies in Big Data and artificial intelligence. The analysis of existing studies and educational practices made it possible to define a set of interconnected competencies covering the full data-processing cycle, including data collection and preparation, analysis and modeling, interpretation, and scientific communication. The results confirmed that professional training in data-intensive fields requires an integrated combination of technical, analytical, and interpretative competencies.

The second research question focused on the development of a pedagogical model for enhancing students' professional competencies through the integration of artificial intelligence tools. The proposed model, implemented through four iterative project life cycles, provided a structured learning environment in which students were able to apply data analysis and modeling tools in practice-oriented tasks. The model facilitated the systematic integration of quantitative, qualitative, and visualization methods within a single project workflow.

The third research question addressed the pedagogical effectiveness of the proposed model. The results of the final assessment of project reports showed that most students achieved medium to high levels of competency across all four assessment cycles. Higher levels of performance were particularly evident in data preparation and scientific visualization, while activities related to model interpretation and qualitative analysis required greater analytical effort. These findings indicate that the proposed instructional approach supports the development of applied professional competencies, while also revealing areas that benefit from additional methodological support.

In summary, the study demonstrates that competency-based projects provide an effective framework for developing students' readiness to work with large-scale data. The findings may be used to inform the design of Big Data and data analytics curricula in higher education.

Conflict of interest: *There are no conflicts of interests.*

Contribution of the authors: *All authors have equally contributed to the conceptualization, development and writing of the article.*

References

Blazquez D., and Domenech J. (2018) 'Big Data sources and methods for social and economic analyses', *Technological Forecasting and Social Change*, 130. — P. 99–113. https://doi.org/10.1016/j.techfore.2017.07.027 (in English).

Chen H., Chiang R.H.L., and Storey V.C. (2012) Business intelligence and analytics: From Big Data to big impact. *MIS Quarterly*, 36(4). — P. 1165–1188. DOI: https://doi.org/10.2307/41703503 (in English).

Gurcan F., Gudek B., Menekse Dalveren G.G., and Derawi M. (2025) 'Future-ready skills across Big Data ecosystems: Insights from machine learning-driven human resource analytics', *Applied Sciences*, 15(11). — 5841 p. https://doi.org/10.3390/app15115841 (in English).

Martins P., Cardoso F., Váz P., Silva J., and Abbasi M. (2025) 'Performance and scalability of data cleaning and preprocessing tools: A benchmark on large real-world datasets', *Data*, 10(5). — 68 p. https://doi.org/10.3390/data10050068 (in English).

Miller S. (2014) 'Collaborative approaches needed to close the Big Data skills gap', *Journal of Organization Design*, 3(1). — P. 26–30 DOI: https://doi.org/10.7146/jod.9823 (in English).

Gosudarstvennaya programma «Tsifrovoy Kazakhstan». Postanovlenie Pravitel'stva Respubliki Kazakhstan ot 12 dekabrya 2017 goda № 827. [State program «Digital Kazakhstan». Resolution of the Government of the Republic of Kazakhstan dated December 12, 2017 No. 827]. [Electronic resource] – URL:https://primeminister.kz/assets/media/gosudarstvennaya-programma-tsifrovoy-kazakhstan-rus.pdf (in Russian)

Nastasa A., et al. (2025) Skills and epistemic cultures in artificial intelligence research: Evidence from academic job postings. *Frontiers in Sociology,* 10. — 1655903 p. DOI: https://doi.org/10.3389/fsoc.2025.1655903 (in English).

Nazyrova A., Miłosz M., Bekmanova G., Omarbekova A., Aimicheva G., and Kadyr Y. (2025) 'The digital transformation of higher education in the context of an AI-driven future', *Sustainability*, 17(22). — 9927 p. https://doi.org/10.3390/su17229927 (in English).

NJIT Learning Initiative (2024). *AI core competency framework for undergraduate programs*. Available at: https://www.njit.edu (in English).

Nurbekova G.F. (2023) *Zhogary oku ornynda bilim alushylardy ulken kəlemdi derekter boiynsha dayarlaudyn teoriyalyk-praktikalyk negizderi* (Doktorlyk dissertatsiya) [Theoretical and practical foundations of training university students in Big Data (Doctoral dissertation)]. L.N. Gumilyov atyndagy Eurasia ulttyk universiteti [L. N. Gumilyov Eurasian National University], 146 p. [Electronic resource] – URL: https://clck.ru/3SxhjW (in Kazakh).

Orynbassar M., Zhumadilova M., and Abdykerimova E. (2024) *Artificial intelligence in Kazakhstan's education system: Analysis and prospects*. *Yessenov Science Journal, 48*(3). — P. 71–76. https://ysj.yu.edu.kz/the-role-of-transformation-centers-in-the-multimodal-freight-transportation-system/ (in English).

Persaud A., and Bayon M.C. (2020) 'Key competencies for Big Data analytics professions: A multi-method study', *Information Technology and People*, 34(1). — P. 178–203 (in English).

Reinsel D., Gantz J., and Rydning J. (2018) *The Digitization of the World: From Edge to Core.* An IDC White Paper No. US44413318, International Data Corporation, Framingham. — 28 p. [Electronic resource] – URL: https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf (in English).

Serik M., Abdrakhmanova A., and Zhanatbekova N. (2023) 'Developing digital and analytical competencies of university students in the context of higher education digitalization in Kazakhstan', *Education and Information Technologies*, 28(5). — P. 6123–6141(in English).

Serik M., Nurbekova G., Mukhambetova M., and Zulpykhar Z. (2022) 'The educational content and methods for Big Data courses including Big Data cluster analysis', *World Transactions on Engineering and Technology Education*, 20(3). — P. 203–208. [Electronic resource] – URL: https://www.scopus.com/inward/record.uri?eid=2-s2.0-85136273484andartnerID=40andmd5=1fb09a972a528d961228cd83335a258e (in English).

Zhukabayeva T., Baumuratova D., Zholshiyeva L., Karabay A., and Abdrakhmanov K. (2025) 'Digital transformation in higher education: Toward a national model of digital university in Kazakhstan', *Sustainability*, 17(24). — 11132 p. https://doi.org/10.3390/su172411132 (in English).

Zhunusbekova A., and Askarkyzy S. (2025) Students' perceptions of artificial intelligence use in higher education and its impact on academic integrity. Pedagogy and Psychology, 61(4). — P. 145–155. https://doi.org/10.51889/2960-1649.2024.61.4.008 (in English).

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 CenterLLP 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
ISSN 2518–1467 (Online),
ISSN 1991–3494 (Print)
http://www.bulletin–science.kz/index.php/en

Managing Editor: A.Shormakova
Editors: D.S. Alenov, M.Konyrbekov
Computer layout: G.D. Zhadyranova

Signed for print: June 30, 2026
Format: 70×90 1/16. 48.0 printed sheets. Order No. 3

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