

<https://doi.org/10.53656/ped2026-6s.05>

*Artificial Intelligence
and the Future of Education
Изкуствен интелект
и бъдещето на образованието*

AXIOLOGICAL FOUNDATIONS FOR THE IMPLEMENTATION OF AI IN EDUCATION: ON THE ISSUE OF NEUTRALITY

Elena Lavrentsova

Trakia University – Stara Zagora (Bulgaria)

Abstract. The article is devoted to the issues of the development of AI in modern socio-cultural conditions and its active implementation in the field of education. A comprehensive analysis of the processes of use and integration of AI technologies in the education system has been carried out, which have led to a profound transformation of pedagogical practices thanks to the possibilities of scaling in the processing of educational content, automation of routine aspects of teaching activities, personalization of training, etc. The interaction of AI with the education system is considered not only as an expression of technological modernization, but as the overall transformation of the institutional nature of this key social institute, affecting various aspects of its functioning, and above all its axiological foundations. Special emphasis is placed on understanding the ethical and social dimensions of the expansion of AI in the field of education through the prism of post-non-classical scientific rationality and in bringing to the forefront the category of “subjectivity” and socio-humanitarian criteria for the development of the educational system in the conditions of the crisis of technogenic civilization.

Keywords: artificial intelligence; axiology; education system; subjectivity; academic integrity; digital inequality

Introduction this the research problem

In recent years, artificial intelligence has increasingly penetrated various areas of life, rapidly transforming not only the form and structure of individual human activities, but also the fundamental foundations of the institutional nature of society. This phenomenon has become one of the key objects of scientific knowledge and is actively attracting the attention of both representatives of technical sciences and researchers from the humanitarian sphere. With the active implementation and expansion of the influence of this correlate of public life, the ways and forms of social interaction and interpersonal communication undergo profound changes, traditions are being reconsidered, and various dimensions of culture are being transformed.

The processes of emergence of new material and spiritual realities acquire an irreversible character, new principles and meaning-forming foundations are being constituted, which accompany not only people's everyday life, but also affect the existential dimensions of their being.

The rapid development of artificial intelligence (AI) is increasingly impacting the processes of production, employment, consumption, innovation and economic growth, which often leads to its being declared an instrument of the new industrial revolution (Barratt, 2019). This is not just a matter of progress in information technology, it is essentially a real transformation of the production, economic and social structure of society. The increased expansion of AI is changing the position of man in the economic, labor and educational spheres (Sullivan et al., 2023; Rudolph et al. 2023). In view of this, human interaction with digital space and AI is becoming one of the central problems in both the economy and employment, as well as in education. According to consolidated international assessments, the education system as an area of intellectual activity is among the sectors that are most susceptible to transformation due to the introduction of AI (Shestakova and Djanelidze, 2025).

Many positive changes can be pointed out that education is undergoing as a result of the active implementation of AI. But It is here that it is important to highlight a fundamental question: does the impact of AI on the very nature of man, his inherent immanent nature, remain neutral, especially from an ethical, value-based perspective, or does it hide such potential risks that, in the absence of adequate regulations, including in the field of education, could generate powerful destructive processes capable of deeply eroding the authenticity of man, his subjectivity. The purpose of this theoretical and philosophical study is to explore and analyze the possibilities for applying AI in the educational process, its benefits and potential risks, and to emphasize the need and importance of responsible and ethically supported use of AI.

Methodologically, the study is based on the use of a rapid review approach and a philosophical-analytical approach. This allows for a review of recently published articles and their key ideas with subsequent synthesis and systematization. Accordingly, this review, combined with analytical reasoning, could provide valuable information, allowing scientists, educators and politicians to respond in a timely manner to the impact of artificial intelligence in all its ambiguity and contradictions on the field of education. Several key electronic databases were used: Academic Search Ultimate, ACM Digital Library, ERIC, Scopus, and Web of Science. The publication period was mainly defined from 2022 to the present. To be included in this rapid review, publications had to discuss the use of AI in education, without further restrictions on a specific educational context. A notable limitation is the predominant attraction of publications in English.

1. Using AI in education: application areas

In science, AI is often considered as such systems, essentially neural networks, which are designed to solve various intellectual tasks and capable of machine learning. The concept of “intelligence” in turn can be defined as the ability to solve intellectual tasks and problems in the absence of a known algorithm. Another characteristic of intelligence is the ability to make rational choices in the absence of information. In addition, when clarifying the semantic and conceptual nature of AI, the property of intellectual systems to perform creative functions that are traditionally attributed to humans is brought to the fore. An intellectual system in turn is a technical or program system that is able to solve tasks belonging to a specific subject area, the knowledge of which is stored in the memory of such a system. Artificial intelligence tools are often seen as including all technologies capable of mimicking human cognitive functions (learning, creative task and problem solving, decision-making).

In other words, AI can be defined as computing systems that are capable of participating in human processes, such as learning, adapting, synthesizing, self-correcting, and using data for complex processing tasks” (Popenici, 2023). This phenomenon describes such a type of technological solutions that can perform tasks usually associated with human abilities. At the same time, a point of ambiguity and even confusion emerges here: while it is true that AI is able to perform tasks that are usually the product of human conscious efforts, this does not mean that AI is capable of absolutely accurately reproducing the human act of thinking and has the ability to be self-aware.

The impact of AI is very clearly evident in the field of education and teaching, reflecting on various aspects of the educational process, including those related to academic integrity. In recent years, AI has transformed from a niche innovation into an integral part of educational practice, which completely changes the way of communication and forms of social interaction between the main subjects - teachers and students. The integration of artificial intelligence in modern education is moving from a purely technological function to the role of a powerful catalyst, which requires a complete rethinking and reassessment of pedagogical approaches, the design and architecture of curricula and the fundamental goals of education. Its rapid adoption (especially after the emergence of generative AI like ChatGPT, and later other forms) and the unprecedented, avalanche-like growth in the number of users in a very short time came as a “shock to education, similar to the COVID-19 pandemic” (Mills et al., 2023, p. 16), carrying within itself both powerful creative possibilities and hidden destructive potential.

The research literature has widely described various forms of AI’s impact on education and its implementation in learning, however, it should be noted that most studies in this area, especially those related to the use of generative AI (GAI – representing a set of trained AI models and techniques that use statistical methods to

create content based on its probabilities), are characterized by their hypothetical nature and are based on rather limited, local practical experience. As the leading lines of influence, the following trends are most often noted: the use of gamification based on artificial intelligence; the spread of adaptive learning technologies; the implementation of intelligent robotic solutions in educational processes; the inclusion of questions and studies related to artificial intelligence in educational programs; the use of artificial intelligence in micro- learning in the field of corporate education; the spread of artificial intelligence on massive open platforms for online courses and training; the use of generative artificial intelligence in education.

One of the widely discussed topics concerns the possibilities of attracting AI as an extremely convenient tool for working in planning and preparing educational courses (for example, such commercial tools as MagicSchool, Eduaide). Taking examples from the field of foreign language teaching, a number of studies have shown that AI, and in particular GAI, can be used even at the stage of preliminary planning of the educational course - as a kind of partner for generating ideas, thereby assisting in the development of both the overall content of the course (which is especially useful if the teacher has not yet led a course with a similar thematic focus) and individual thematic blocks, accompanied by the selection of appropriate tasks and offering various ideas for improving its structure (Alqahtani et al., 2023; Bao et Li, 2023).

In addition, AI can be actively used in the development of educational materials with an educational and methodological (lectures, tests, tools, textbooks, methodological aids, collections of practical tasks) and even research focus (books, articles) (Lo, 2023). It can be attracted as a valuable resource for creating illustrative materials for presentations and learning cases with the help of image generators. Thanks to its inherent ability to train permanently with open data and to generate a practically infinite number of different variants of learning tasks - tests, examples, tasks, questions to individual text fragments, etc., AI is capable of significantly increasing *variability* within the learning process both when planning a new course and when updating and enriching already existing ones, thus influencing the increase in the effectiveness of training.

At the same time, this does not eliminate the issue related to legitimization and validation of the generated proposals and results by the teacher, since AI does not demonstrate sufficient “sensitivity” to the pedagogical parameters of the educational process, “does not understand” to a large extent the instructions affecting the pedagogical component and the ways of implementing it in the proposed options, and generally does not capture the ideological and conceptual, including axiological, foundations of the pedagogical impact, determining various aspects of teaching activity.

The potential positive impact of AI on the state of the education system is also associated with additional opportunities for *personalization/individualization* of

education. AI is able to take into account the academic success of students, to analyze its nature and dynamics, as well as the influence of various factors - individual-personal, institutional, environmental, on its course. It is able to be maximally relevant to the interests and preferences of each student, his specific pace and style of learning, by offering learning materials and assignments, individual disciplines and training courses, and more broadly, complex educational trajectories, adapted to individual needs.

The research literature describes a wide range of possibilities for the application of AI in this regard: from flexible modification of tests and exam tasks with the help of AI, which can be complicated or vice versa depending on the level of knowledge and characteristics of the student (Southworth et al., 2023) , to the use of voice assistants, performing a facilitating function through direct participation in discussions and debates on the given topic in real time, giving instructions and guiding the course of the discussion, with subsequent summarization and providing feedback. In an even more expanded learning format, a new, richer educational environment such as Metaverse can be constructed for example, with an extremely high level of interactivity and opportunities for conducting scientific experiments, simulations, virtual laboratories and excursions, increasing the effectiveness and attractiveness of learning.

More routine teaching activities, such as checking homework and exam assignments, can also undergo a change towards further individualization by attracting AI chatbots, which allow not only express assessment, but also provide precise and personalized feedback according to each student's profile on common errors and difficulties, which can serve to further optimize teaching. Here, another point is emerging, related to increasing *analytics* and *prognostics*, since thanks to AI's unique ability to operate with big data on student success, there is an opportunity to track certain trends in a wider time frame and predict potential learning difficulties or, conversely, the most optimal heuristic triggers for the child's development.

Adaptation with the help of AI to the characteristics of children with special educational needs and disabilities in the context of inclusive education, including task reworking and preparation of specialized learning materials, seems particularly promising (Zhai, 2023). Personalized and immediate feedback and individualized learning paths lead this increased motivation and engagement, better outcomes, and learners' satisfaction and are essential for those with special educational needs as well (Adiguzel and Kaya, 2023). The latter highlights the potential of AI in *expanding access to education* by creating accessible educational resources for people with limited opportunities, low socio-economic status, migrants and refugees who have difficulty learning the language of the host country, as well as those living in remote areas. Special attention is paid to international students, who often face language difficulties in their studies. For them, AI tools can assist not only in language learning, but also in understanding and preparing materials in another language,

reducing the need for additional consultations and increasing confidence and satisfaction in the learning process (Farrelly and Baker 2023).

As numerous studies have shown, for teachers, AL can serve as a very useful virtual tool to support teaching process. Table 1 categorizes its main functions into three key activities, namely teaching preparation, assessment and individualization of training.

Table 1. AI main functions to support teaching process

Activity	Functions
Teaching preparation	Generating course materials, new productive ideas, texts, explanation of concepts
	Generating images, tables, diagrams, charts, presentations
	Providing suggestions, based on data analysis
	Discovering recurring, persistent, causal relationship between phenomena
	Performing language translation
	Systematization, generalization, outlining of trends for future development
	Analysis of concrete, specific documents
	Recognition and analysis of plant and animal species, geographical regions
	Use of Virtual Labs, Simulations
Assessment	Generating assessment tasks, case studies
	Evaluating student performance
	Assessing student progress
Individualization of training	Creating inclusive AI models
	Generating personalized instructional and training materials, teaching approaches
	Generating specific learning materials and learning content designed for students from vulnerable communities
	Generating teaching materials with a focus on equality, tolerance, positive acceptance of difference, social-emotional learning

Overall, AI technologies help to increase the level of inclusiveness of the education system by overcoming physical, economic, territorial, linguistic and institutional barriers, thus expanding the accessibility of quality education. In addition, AI can be used to analyze new social and labor market trends, helping to develop up-to-date educational programs and thereby closing the traditional gap between the needs of business and the offers of the education system.

Perhaps one of the most discussed issues in the field of scientific discourse is the issue related to the use of AI in the performance of educational tasks by students. At the same time, it is characterized by a high dose of ambivalence depending on the intentions, interests and expectations of the main participants in the educational process – teachers and students. Its ambivalence is manifested in the fact that students strive to accelerate the pace and increase the efficiency of completing educational tasks, and teachers try to monitor compliance with the relevant regulations and restrictions when using AI for such purposes.

Undoubtedly, AI tools enables learners to develop new digital skills that are very important in today's technological society. Learners need to formulate the correct questions and prompts, operationalize the tasks, and give precise instructions to get satisfactory answers (Ivanov and Soliman, 2023). They assist students in completing written assignments such as reports, abstracts, essays or extended answers to questions on the course they are studying in order to prepare for the exam, to generate new ideas and develop the topics of their independent research - course projects, bachelor's theses or master's theses. Students can benefit from using AI as a tool to prepare their initial draft, then refine it and correct errors (Aleedy et al., 2022). They can also use technology on AI for testing knowledge and self-esteem, so they can ask the chatbot and evaluate the technology text/completed task.

The research literature notes that for students, AI can serve as a virtual tutor to support their learning. Table 2 categorizes its main functions into three key activities, namely preparation for learning (i.e. answering questions, summarizing information, etc.), assessment (i.e. checking concepts, assisting with project preparation, etc.), and individualization.

Table 2. AI main functions to support learning process

Activity	Functions
Learning preparation	Generating learning materials, ideas, texts, explanation of concepts and tasks
	Generating images, tables, diagrams, charts, presentations
	Answering questions, possibility of modulating questions
	Analysing, structuring, systematizing and summarizing information
	Facilitating collaboration and cooperation, team forms of work
	Increasing the level of interactivity in the learning process
	Performing language translation
	Providing suggestions, based on data analysis
	Systematization, generalization, outlining of trends for future development
	Analysis of concrete, specific documents
	Recognition and analysis of plant and animal species, geographical regions

	Use of Virtual Labs, Simulations
Assessment	Drafting assistance
	Exam preparation
	Concept checking
	Providing feedback in the learning process
	Supporting the functioning of self-assessment mechanisms
Individualization of learning	Using inclusive AI models
	Personalized assistants/teachers in a chosen subject
	Generating personalized instructional and learning materials, educational paths of learning
	Generating specific learning materials and learning content designed for students from vulnerable communities

In general, two responses to the use of AI tools by students are taking shape in pedagogical and academic circles: *prohibition* and *integration*. (Strzelecki, 2024). Proponents of the restrictive regime towards the use of these tools by students, who perceive resorting to AI-assisted approaches to academic misconduct focus on identifying and preventing such cases. However, this is becoming increasingly difficult as tools for detecting text generated by large language models have limited accuracy (Weber-Wulff et al., 2023). In response to such a challenge, the transformation of control materials and procedures is most often proposed in the direction of changing the structural components of the tasks themselves and including more open-ended questions requiring deeper reflection, critical thinking and argumentation skills. Supporters of the integrative vision, who emphasize the increased impregnation of AI in the educational process, in turn, believe that subject-centered training in the use of AI is necessary, an integral part of which should be regular discussion of the problems of the application of AI in education and the implementation of special training aimed at revealing the problematic moments, risks and challenges associated with its use in the context of value principles and dimensions of human life.

2. Axiological foundations and ethical social dimensions of the use of AI in education

The integration of AI into the educational process brings with it impressive prospects, but at the same time it hides significant challenges related to the confidentiality of the data used, the formation of a predisposition to algorithmic reactions and decisions, increasing difficulties in ensuring information security and numerous other risk factors, most of which in one way or another affect the value mode not only of education as one of the key social institutions, but also of human existence as a whole. This outlines the need to fit the issues related to the implementation of

AI into the relevant axiological framework, revealing the ethical and philosophical foundations with the theoretical focus on values – their nature, origin, structure and place in reality.

One of the issues that requires precise attention in terms of the use of AI is *the problem of trust and reliability* of data and results. When interaction with AI chatbots takes place over a longer period of time, one or another distortions, inaccuracies, erroneous or false results become noticeable, revealing the unreliable nature of the generated options. Chatbots can offer different, sometimes diametrically opposed answers to identically formulated questions, ignore the given instructions, mention non-existent facts, events, as well as the names of authors and titles of publications when setting the task of generating an academic text (Duah and McGivern, 2024; Perkins, 2023). The problem arises with the extent to which it is possible to rely on the information received, trust in its authenticity in view of the possible and periodically appearing “hallucinations” inherent in such technologies. The most frequently launched recommendations in this direction are related to additional verification of reliability, objectivity, credibility, logical coherence and verifiability of all received answers and information data.

From here arises another, even more fundamental problem, related to *the risk of losing the sense and perception of objective reality*. In the long term, a person may develop a specific perception of the Self and the world, which will be characterized by a total dominance of the virtual. Starting from the more insignificant aberrations of the real, such as, for example, a generated “false” answer to an inaccurately formulated question, the movement towards post-reality may continue through the multiplication of false statements in the Internet environment (if the answer in question is uploaded and becomes available on the Internet) and ultimately lead to the emergence of not just invented, imaginary, but false-fictitious names, processes, phenomena and even entities, constituting under certain conditions a distorted image of the universe, its false, but with all the attributes of the real, simulacrum.

This in turn raises another important issue related to the active implementation of AI and the digitalization of the social as an expression of the powerful transformations embodied in the digital society. We are talking about an increasingly noticeable decrease in live, face-to-face communication, reflecting the simultaneous *erosion of the value of direct, personally colored communication*, filled with both the actualized meaning through the eyes of the Other, and the flickering in its fleeting givenness and shared emotion. But it is precisely such communication that builds a sense of true belonging, lays the foundations for mutual respect, trust and understanding, contributes to mental health, facilitates self-knowledge and knowledge of the world, and leads to the formation and development of the human personality as a concentrated “condensate” of social energy and an expression of all social relations.

Additional concerns are raised by the fading of such a key human ability as imagination, the formation of the so-called “*clip*” *thinking* and the increase in *social isolation*, which causes the emergence of such a specific phenomenon as “social autism”. The threat of *reducing mental workload*, the orientation towards intellectual convenience and comfort at the expense of intense mental, but at the same time truly developing individual, intellectual activity is also assessed as significant. This in the future may lead to a serious decrease in the intellectual level of students from different age groups and levels due to the shift of their cognitive, heuristic and analytical functions to AI.

In this context, the efficiency and productivity that AI provides is assessed as an axiological benefit only if it frees up cognitive resources for higher thought processes, and not if it leads to an atrophy of basic analytical abilities. Therefore, in the new sociocultural conditions, the role of the teacher is transformed into a mentor and facilitator, who guides and at the same time directs the student towards an ethically supported and rational use of technology, so that it supports the ultimate goal of education - the formation of an intellectually autonomous personality. In practice, learning management can involve designing a decision architecture based on the principle of libertarian paternalism (Elstad, 2008): learning environments must be designed so that students have incentives (cohesion and structure) to overcome the mental pain that effort gives. Paternalism comes into play with the teacher setting the framework for teaching: instructions, expectations and control. The libertarian part comes into play because students will have some autonomy and co-determination. The goal of libertarian paternalism is to influence each student’s decisions in ways that will make them better off through perseverance while preserving choice. Perseverance must be learned; it is not innate.

The focus of researchers’ attention is also on the risk of *the degradation of critical thinking* and *creative skills*. Developing a tendency to receive accessible, ready-made and processed information that can be easily obtained and at the same time often seems well-founded and convincing, a person, especially if he is a child of school age, may succumb to the temptation not to problematize the provided content. The more often such a situation is repeated and the corresponding attitude of “blind”, uncritical acceptance of the information offered is reinforced, and the focus often shifts from the process of making sense to the mechanical receipt of the final product, the more the ability to analyze, meaningfully compare different opinions and facts begins to fade. Another not infrequent feature is superimposed on this situation: unlike true understanding, the answers generated by artificial intelligence, in many cases, lack depth and insight. This further reinforces concerns about the loss of valuable skills such as introspection, reflection, resilience, and the ability to think critically and independently with unrestricted use or even misuse of AI.

The axiological aspect of the problem here is rooted in the value that society attributes to the authentic nature of cognitive effort. Proceeding from this, the educa-

tional value of AI is determined not by its technological capabilities, but by whether it supports or suppresses the intellectual activity and efforts of the individual. If technology seeks to replace the thought process, it leads to the degradation of the learning experience, but if integrated through conscious pedagogical strategies and practices, it can serve as a valuable tool for upgrading. For this reason, the ethical regulation of learning in the age of AI requires students to develop a high degree of self-regulation and the ability to differentiate the role of technology: when and how it enriches them and makes them more capable and when it is intellectually passive. Ultimately, modernity presents education with a fundamental challenge: to build such institutional mechanisms that allow artificial intelligence to function as a complement to human potential, while preserving the central place of personal development and moral integrity in the field of the educational and teaching-educational process.

The existing limitations of AI in terms of recreating and adequately interpreting one or another contextual environment, along with processing non-trivial ideas, judgments and conclusions, as well as non-standard situations, raise many additional questions in the ethical and pedagogical fields, related to the implementation of facilitating, selecting, evaluating, controlling and, more broadly, tutoring functions of AI. This also concerns *the problem of relevance* and adaptability of the generated content to the individual needs and requirements of the learners, which often implies taking into account even the smallest nuances in the attitudes, aspirations, expectations and other personal and behavioral characteristics of the individual. To what extent is an automated and algorithmic assessment able to be not only purely indicative in terms of the various aspects of the learner's cognitive movements, efforts and their effectiveness in the form of one or another academic achievement, but also to be commensurate with the entire fullness and richness of his personality? To what extent is an algorithmic, over-rationalized forecast, not respecting a special "logic of the heart" in the words of the great Blaise Pascal, capable of directing not only the educational, but also the existential trajectory of an individual's development in the right direction? And how should the fundamental question of assuming responsibility by the algorithmic system for any mistakes and decisions made be ultimately resolved? In this sense, the excessive fascination with AI in the direction of its unlimited use and absolutization of its capabilities in the educational process is fraught with a significant weakening of the role of the teacher and the importance of the pedagogical impacts exerted by him in the context of their emotional, educational and socializing underpinnings.

Additional ethical threats from the use of AI in education relate to the inherent bias and *bias of algorithms* based on the initial training matrix used for language models, which is constructed on a limited set of data, distinguished by a certain semantic, political, ideological and cultural orientation. As a result, the biases may appear in relation to certain ideas and views, speech structures and turns, models

and argumentative moves, especially noticeable in the field of social sciences and the humanities, where the action performed by a person and the meaning embedded in it are necessarily taken into account. Such specificity highlights a real lack of neutrality in the functioning of AI technologies. Language models like ChatGPT, Gemini, Llama and others simply have their own political and ideological bias. In some cases, this can lead to the use of discriminatory, sexist, marginalizing, labeling and stigmatizing words and sentences (Dwivedi et al., 2023).

In addition, the research literature describes other additional effects of constructive bias in the use of AI, reflecting the axiological aspects related to selectivity, favoritism, hierarchy and inequality. These include such effects as, for example: selection of sources only from certain resources and databases; the tendency of artificial intelligence to create clickbait headlines and to use the most popular, discussed materials to generate new texts; bias towards statements and opinions of a confirmatory nature, if such a task is initially fixed; focus on points of view that have gained popularity during the training of the language model; bias towards favorable/negative points of view on some issues depending on the training set; authority bias, i.e. emphasis on the views of famous people and companies; and many others (Ray, 2023).

Another very important point affecting the value ideals of social justice concerns the emergence of a number of new social threats and risks associated with the development of modern technogenic civilization and AI: the possibility of *deepening digital inequality* and its transformation into a *digital divide* in access to educational resources (both between different educational organizations and between students, for example, depending on income level or place of residence, etc.), changing skill requirements in the labor market, the impact on socialization and the development of soft skills of students. It is important to understand that the digital divide today is not just a lack of Internet or access to computers. In the context of the development of AI, it also includes differences in the following areas: access to infrastructure (uneven coverage with high-speed Internet, availability of modern devices); digital literacy and skills (mastery of digital tools, ability to effectively use the opportunities provided by AI); access to high-quality content and services (the ability to use new AI solutions that can improve professional productivity, expand its capabilities and reduce costs); participation in the digital economy (the ability to create, monetize and use products and services based on both its presence and the use of AI tools). Furthermore, the implementation of AI requires appropriate resources, significant investments and the availability of technical infrastructure, which is not always available in subsidized regions and educational institutions with insufficient budgets (Shestakova and Djanelidze, 2025).

AI-generated social inequalities are also manifesting themselves in another dimension, affecting professional development and identity. Artificial intelligence doesn't just automate routine tasks—it's dramatically changing the entry point for

young professionals. As tasks previously assigned to junior employees are now handled more quickly and more affordably by ChatGPT, Midjourney, GitHub Copilot and other tools, the junior level as a career step is eroding, and the entry point into the market is becoming a mid-level position. This disrupts the traditional career ladder and creates a new kind of inequality – the *generative-gap* (Brynjolfsson and Mitchell, 2017). A similar trend may in the future affect a number of other professional fields and areas.

The discussion about the value bases of AI and its implementation in education cannot fail to add the issue of *plagiarism, authorship, and academic integrity*. Students can sometimes present text generated by AI as their own, ignoring the normative principles and rules of academic ethics (Duah and McGivern, 2024). The same applies to colloquia, laboratory and any other examination forms intended for assessment. There is also another point regarding the possibility that the AI-generated text may contain incorrect references and citations, possible passages from texts and the works of other researchers.

Eliminating such abuses and manifestations of academic dishonesty requires, on the one hand, additional training in the ethical use of AI, including within existing academic disciplines, and, on the other hand, the development of technologies to identify AI-generated content in student work (Sullivan et al., 2023). In addition, the issue of authorship in the context of determining and deriving one or another criterion for authenticity and originality of course projects and theses submitted for evaluation has caused considerable discussion. In any case, whenever academic integrity is put to the test in one way or another, it requires the introduction of clear, transparent, and consensual principles and rules according to which AI is used only as a complement to natural intelligence.

The systematization of the main risks of using AI can certainly be carried out on various foundations, but still the more significant problems relate to the following aspects, categorized in Table 3 – reliability, accuracy, construct reasoning, plagiarism and academic integrity, human autonomy and equality.

Table 3. Main potential issues associated with the use of artificial intelligence

Aspect	Issue
Reliability	Resorting to biased data in the content offered, Misleading content generation
	Possessing limited up-to-date knowledge and information
	Citing non-existent facts, documents, or events – phenomenon of hallucination
Accuracy	Providing inaccurate information
	Providing outdated information about persons and events
	Artificial intelligence developers often resort to selective information for training data

Construct reasoning	Difficulties in establishing causal relationships
	Errors when performing mathematical calculations
	Failure when operating with unknown or unusual input data
	Misclassifying images when backlighting or context changes
Plagiarism and academic integrity	Student and academic staff plagiarism
	Significant limitations in detecting plagiarism
	Destruction of academic integrity and undermining the principles of academic ethos
Human autonomy and equality	Possibility of negative impact on a person's mental abilities (including memory, critical thinking, decision-making, creativity)
	Opportunity to change the way people cognitively engage with tasks, including skill formation, developing competencies over time.
	Formation of "automation bias" in people, when they rely too much on AI-generated data and results and ignore contradictory information.
	Increasing social isolation and loneliness among frequent users
	Destruction of authenticity in the formation of attitudes and value orientations
	Increasing digital and educational inequality

In the research literature, one can find various proposals for ethical rules regulating the use of AI in education, the most common and key ones being: *the security principle*, which guarantees the confidentiality and security of sensitive information; *the accessibility principle*, emphasizing the priority of human rights; *the principle of inclusion in accessibility*, with an emphasis on non-discrimination and *the principle of human-centeredness*. It is the latter that expresses the deep essence of the axiological foundations outlining both the ethical framework for the implementation of AI in education and the leading paradigm for the development of education in modern conditions.

Modern education must be *human-centered*. The educational system must understand and manifest that the most important thing is not knowledge, but the whole person, an autonomous subject who emerges from it, accumulating a certain experience. Knowledge in itself is not primary, but secondary to the bearer of this knowledge. It is not enough for the individual to be perceived in such a way that he can absorb the maximum volume of knowledge, because today the person already lives in the era of information overload, when all the information is contained in his mobile phone, in Chat GPT. And the problem is how he interacts with this information, what worldview he has, how he can formulate a task and make sense of it, how he is able to react on a psycho-emotional level and how he can connect all this within himself.

It is the system of transmitting knowledge from the teacher to the student not so much through contact with information as through contact with the experience of the

other person, who in a state of interest, insight, acceptance and love, transmits certain models of behavior and attitude towards other people and the world. The fundamental thing here is the interaction with the personality of the teacher: the anthropos must meet the anthropos. The most valuable thing is the encounter with the spiritual principle embodied in the personality of the teacher. In this perspective, it can be recalled how Friedrich Fichte laid the foundations of the system of modern education, the main goal of which was the production of the soldier of the line infantry and the factory worker. Then, within the framework of the specific historical era, the state itself, as a kind of supersubject in a certain integrity due to its survival, reproduction and preservation in history, created such a school system that produced not the autocratic subject, but the production objectified human resource. But today, it is not possible to give a purely technological answer to technological problems. And the only protection from sophisticated information manipulation, which destroys consciousness much more effectively than direct violence, is precisely this autocratic anthropos, woven into the idea of a hierarchy of mutual service and mutual responsibility.

Conclusions and summary

Nowadays, the very understanding of ethics and axiology in the context of their phenomenal nature is radically changing: if earlier ethical relations were understood only as relations between people, now artificial intelligence has entered the orbit of ethics and axiology. AI acquires, along with the object status, also the status of the subject, intertwined in a wide field of value-moral relations. Artificial intelligence is undoubtedly a powerful catalyst for change in education. It has already become deeply embedded in everyday learning, covering various aspects of the learning process. And the question is not whether it should be used or not. With proper and precise design of tasks and courses, AI is capable of becoming a driver for the development of a prosperous and creative personality thanks to the assistance provided in deeper analysis, the search for non-standard solutions, and the capture of new perspectives. It is important, along with its intensive implementation, to take into account the effects of its influence on human autonomy in the context of human-AI interaction. Such studies are currently hampered by the lack of consensus on the conceptual nature of the category of “autonomy”, and by practical difficulties in assessing it, and by the limited empirical data and measurement methods, especially in the context of sustainable and socially complex use of chatbots – for which there are few longitudinal studies assessing cumulative, long-term, and systemic impacts on individual autonomy.

Understanding the systemic, long-term effects of human-AI interactions is particularly important and valuable in the field of education, where children’s interactions with AI systems at a very early age can significantly influence the formation of their key skills and their further development. Ultimately, joint, focused efforts involving educators, researchers and policymakers are needed to

ensure the ethical and responsible use of AI in education and to develop measures to build social resilience, ensuring the ability of societal systems to adapt to possible shocks and potential harms associated with AI. The guiding intention here should be the pursuit of building a more equitable and successful education system that prioritizes the ethical use of AI, provides children with individualized learning and support, cultivates digital literacy and critical thinking, optimally utilizes the advantages of AI technologies, and develops frameworks for their safe application while simultaneously taking into account and minimizing the problems they pose.

The potential of AI can be fully realized only with a balanced, ethically and pedagogically justified approach to its implementation, combined with precise consideration of both its creative potential and the threat hidden in it. In view of this, the integration of artificial intelligence in education should not be perceived as a technological race for its own sake, but as a conscious process of affirming human values in a new, digital reality (Vasileiou, 2025). The ultimate success of this radical transformation depends on the fundamental human ability to overcome utopian attitudes and hopes for technological neutrality and to take active social responsibility for the ethical defects that algorithmic systems can scale within the educational space.

REFERENCES

- Adiguzel, T., Kaya, M. H., Cansu, F. K. (2023). Revolutionizing education with AI: Exploring the transformative potential of ChatGPT. *Contemporary Educational Technology*, 15(3). <https://doi.org/10.30935/cedtech/13152>.
- Aleedy, M., Atwell, E.; Meshoul, S. (2022). Using AI Chatbots in Education: Recent Advances Challenges and Use Case. *Artificial Intelligence and Sustainable Computing: Proceedings of ICSISCET*, 661 – 675. https://doi.org/10.1007/978-981-19-1653-3_50.
- Alqahtani, T., Badreldin, H.A., Alrashed, M., Alshaya, A.I., Alghamdi, S.S., Bin Saleh, K., Alowais, S.A., Alshaya, O.A., Rahman, I., Al Yami, M.S., Albekairy, A.M. (2023). The emergent role of artificial intelligence, natural learning processing, and large language models in higher education and research. *Research in Social and Administrative Pharmacy*, 19(8), 1236 – 1242. <https://doi.org/10.1016/j.sapharm.2023.05.016>.
- Bao, Y. & Li, B. (2023). A Preliminary Study on Graduate Student Instructors' Exploration, Perception, and Use of ChatGPT. *International Journal of Computer-Assisted Language Learning and Teaching (IJCALLT)*, 13(1), 1 – 23. <https://doi.org/10.4018/IJCALLT.332873>.
- Barrat, D. (2019). *The Last Invention of Humanity: Artificial Intelligence And the End of the Homo Sapiens Era*. Moscow: Alpna non-fiction. ISBN 978-5-00139-082-4.

- Brynjolfsson, E. & Mitchell, T. (2017). What Can Machine Learning Do? *Workforce Implications, Science*, 358(6370), 1530 – 1534. <https://doi.org/10.1126/science.aap8062>.
- Duah, J.E. & Mcgivern, P. (2024). How generative artificial intelligence has blurred notions of authorial identity and academic norms in higher education, necessitating clear university usage policies. *The International Journal of Information and Learning Technology*, 41(2), 180 – 193. <https://doi.org/10.1108/IJILT-11-2023-0213>.
- Dwivedi, Y.K., Kshetri, N., Hughes, L., Slade, E.L., Jeyaraj, A., Kar, A.K., Baabdu Lah, A.M., Koochang, A., Raghavan, V., Ahuja, M., Albanna, H., Albashrawi, M.A., Al-Busaidi, A.S., Balakrishnan, J., Barlette, Y., Basu, S., Bose, I., Brooks, L., Buhali, D., Wright, R. (2023). Opinion Paper: „So what if ChatGPT wrote it?“ Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management*, (71), 102642. <https://doi.org/10.1016/j.ijinfomgt.2023.102642>.
- Elstad, E. (2008). Towards a Theory of Mutual Dependency between School Administrators and Teachers: Bargaining Theory as Research Heuristic. *Education Management and Administration Leadership*, 36(3), 393 – 414. <https://doi.org/10.1177/174114320809059>.
- Farrelly, T. & Baker, N. (2023). Generative Artificial Intelligence: Implications and Considerations for Higher Education Practice. *Education Sciences*, 13(11), 1109. <https://doi.org/10.3390/educsci13111109>.
- Ivanov, S. & Soliman, M. (2023). Game of algorithms: ChatGPT implications for the future of tourism education and research. *Journal of Tourism Futures*, 9(2), 214 – 221. <https://doi.org/10.1108/JTF-02-2023-0038>.
- Lo, C.K. (2023). What Is the Impact of ChatGPT on Education? *A Rapid Review of the Literature, Education Sciences*, 13(4), 410. <https://doi.org/10.3390/educsci13040410>.
- Mills, A., Bali, M., Eaton, L. (2023). How do we respond to generative AI in education? Open educational practices give us a framework for an ongoing process. *Journal of Applied Learning and Teaching*, 6(1), 16 – 30. <https://doi.org/10.37074/jalt.2023.6.1.34>.
- Perkins, M. (2023). Academic Integrity considerations of AI Large Language Models in the post-pandemic era: ChatGPT and beyond. *Journal of University Teaching and Learning Practice*, 20(2). <https://doi.org/10.53761>.
- Popenici, S. A. D. (2023). The critique of AI as a foundation for judicious use in higher education. *Journal of Applied Learning and Teaching*, 6(2), 378 – 384. <http://journals.sfu.ca/jalt/index.php/jalt/index>.

- Ray, P.P. (2023). ChatGPT: A comprehensive review on background, applications, key challenges, bias, ethics, limitations and future scope. *Internet of Things and Cyber-Physical Systems*, (3), 121 – 154. <https://doi.org/10.1016/j.iotcps.2023.04.003>.
- Rudolph, J.; Tan, S.; Tan, S. (2023). War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. *Journal of Applied Learning and Teaching*, 6(1), 364 – 389. <https://doi.org/10.37074/jalt.2023.6.1.23>.
- Shestakova, N. N. & Djanelidze, M. G. (2025). The Artificial Intelligence: Impact on the Education – Labor Market System. *Administrative Consulting*. (5), 91 – 105. <https://www.acjournal.ru/jour/article/view/2829>.
- Southworth, J., Migliaccio, K., Glover, J., Glover, J., Reed, D., Mccarty, C., Brendemuhl, J., Thomas, A. (2023). Developing a model for AI Across the curriculum: Transforming the higher education landscape via innovation in AI literacy. *Computers and Education: Artificial Intelligence*, (4), 100127. <https://doi.org/10.1016/j.caeai.2023.100127>.
- Strzelecki, A. (2024). To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. *Interactive Learning Environments*, 32(9), 5142 – 5155. <https://doi.org/10.1080/10494820.2023.2209881>.
- Sullivan, M., Kelly, A., McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning and Teaching*, 6(1), 31 – 40. <https://doi.org/10.37074/jalt.2023.6.1.17>.
- Vasileiou, I. (2025). Introduction to the special issue on “Artificial Intelligence in Education”, *Journal of Pedagogical Research*, 9(2), 1 – 3. <https://doi.org/10.33902/JPR.202514532>.
- Weber-Wulff, D., Anohina-Naumeca, A., Bjelobaba, S., Foltýnek, T., Guer Rero-Dib, J., Popoola, O., Šigut, P., Waddington, L. (2023). Testing of Detection Tools for AI-Generated Text. *International Journal for Educational Integrity*, (19), 26. <http://dx.doi.org/10.1007/s40979-023-00146-z>.
- Zhai, X., (2023). ChatGPT for Next Generation Science Learning, *XRDS: Crossroads. The ACM Magazine for Students*, 29(3), 42 – 46. <https://doi.org/10.1145/3589649>.

✉ **Dr. Elena Lavrentsova, Assoc. Prof.**

ORCID iD: 0000-0002-4920-275X

Faculty of Education

Trakia University – Stara Zagora

Stara Zagora, Bulgaria

E-mail: elena.lavrentsova@trakia-uni.bg