Attitudes Toward AI in Academic Assessment: A Survey of Croatian Higher Education Teachers

Jurica Grzunov, Ljiljana Zekanović-Korona
University of Zadar
Department of Applied Communication Studies
Dr. Franje Tuđmana 24i, Zadar, Croatia
{jgrzunov, ljkorona}@unizd.hr

Zdravko Kunić
Algebra Bernays University
Department of Information Systems and Business Analytics
Gradišćanska 24, Zagreb, Croatia
zdravko.kunic@algebra.hr

Abstract. *This paper examines how Croatian university teachers perceive the use of AI tools in academic assessment and feedback. A national survey (N = 111) shows that AI is mainly used for technical tasks such as plagiarism detection and language improvement, while its use in more complex or creative assessment remains limited. Teachers see some benefits in efficiency but express concerns about fairness, accuracy, and the quality of AI-generated feedback. Many feel unprepared and lack institutional support. The findings highlight the need for clearer guidelines and targeted training to support responsible and meaningful use of AI in higher education assessment.*

Keywords. artificial intelligence, academic assessment, higher education, teacher attitudes, large language models

1 Introduction

The growing use of artificial intelligence (AI) in higher education is changing how learning is planned, delivered, and assessed. One of the most important developments is the rise of generative AI, especially large language models (LLMs) such as ChatGPT, Claude, and Gemini. These tools can generate and improve written content in real time, which makes them useful for academic assessment and feedback. Their availability has increased expectations about efficiency, personalisation, and managing large volumes of student work. At the same time, they raise concerns about fairness, authorship, academic honesty, and the reliability of automated feedback.

As AI becomes more common in education, universities are trying to define its role in existing teaching and assessment systems. Some focus on using AI to support formative assessment and help students learn better. Others put more emphasis on detecting cheating or plagiarism, which can lead to strict or cautious use. In both cases, the teacher remains essential, not only for creating assessments, but also for interpreting AI results, giving feedback, and ensuring quality.

Research shows that teacher attitudes strongly influence how technologies like AI are used. Confidence, digital skills, institutional support, and the specific subject area all affect how teachers approach AI tools in their work. Even though international research is growing, there is still little evidence about how university teachers in specific countries view and use AI, especially in practice.

This paper addresses that gap by presenting the results of a national survey conducted among university teachers in Croatia. It explores how AI tools are used in assessment and feedback, what benefits and risks teachers see, and what personal and institutional factors shape their experiences. The study also examines the connection between AI use and factors such as digital literacy, self-assessed AI competence, and the presence of institutional guidelines. The aim is to offer useful insights for universities, policy makers, and researchers who are working to integrate AI in higher education in a responsible and effective way.

2 Challenges and perspectives of assessment in higher education in the context of AI

The emergence of ChatGPT and similar tools has sparked debate around academic integrity, while also highlighting the educational potential of such technologies. Media discourse analyses in countries like Australia, the United States, and the United Kingdom reveal that public discussions tend to focus on integrity issues, with limited attention given to students' perspectives (Sullivan, Kelly & McLaughlan, 2023). This imbalance suggests a need for a more inclusive understanding of how AI is shaping academic practices.

Although various recommendations encourage more open and pedagogically sound approaches, many institutions continue to prioritize detection and surveillance. Some scholars argue that this approach

can be counterproductive. Rather than relying on a "catch and punish" model, they advocate for fostering critical understanding of AI and using its potential to support transformative learning. This shift also includes promoting student autonomy and ethical responsibility (Kramm & McKenna, 2023).

A useful contribution to this discussion is provided by Forde-Leaves, Walton and Tann (2023), who propose a framework for studying assessment. Their model helps identify strategic and operational roles within the educational system and provides a comprehensive understanding of the complexity of assessment processes, including the role of AI tools. Similarly, Xia et al. (2024) emphasize the importance of transforming assessment practices to support self-regulated learning and uphold academic integrity. They stress the need for professional development, interdisciplinary cooperation, and educational policies aligned with technological change.

On a broader scale, more universities and international organizations are engaging with the ethical implications of AI in education. Key principles, such as transparency, accountability, inclusiveness, and human oversight, are increasingly highlighted. Research suggests that the most effective institutional responses are grounded in normative frameworks that are applied flexibly in teaching practice and clearly communicate the role of AI in curricula and assessment (Dabis & Csáki, 2024).

At the same time, traditional assessment models, especially summative ones, are becoming increasingly difficult to apply in digital and online environments. In response, a growing number of scholars advocate for "assessment for learning" approaches, which use AI and learning analytics to support collaboration, flexibility, and personalized learning (Williams, 2023). In this context, formative assessment becomes more relevant, as it offers immediate feedback and reduces student stress (Hooda et al., 2020).

Studies focusing on the application of AI in specific educational contexts further illustrate the complexity of the topic. For example, Amin (2023) demonstrates that AI can enhance foreign language instruction by enabling personalized learning, instant feedback, and automated assessment. However, the study also warns of the importance of maintaining ethical standards, equal access, and the human dimension of education. Likewise, Dong (2023) reports that using AI tools in academic writing classes can help improve student skills and increase engagement, opening new possibilities for pedagogical innovation.

Nonetheless, concerns persist, particularly regarding the detection of AI-generated content. Alexander, Savvidou and Alexander (2023) note that both educators and detection tools struggle to reliably identify such content, especially in cases involving partially AI-generated texts. This finding highlights the need for new evaluation criteria and ongoing teacher training.

Generative AI tools also raise questions about educational principles, especially given their ability to produce high-quality content in very short periods. This prompts the need to rethink curriculum design and assessment strategies (Bennett & Abusalem, 2024). In line with this, there is increasing support for learning approaches that develop higher-order thinking and encourage active student participation. Kurtz et al. (2024) propose a model for adopting GenAI technologies based on four stages, which include AI literacy, teacher empowerment, and student collaboration, while also addressing risks such as bias, inaccuracy, and dependency.

These developments have significantly influenced higher education pedagogy. Although concerns about plagiarism remain valid, some authors argue that generative AI can be integrated into assessment design to support deeper learning and metacognitive skills (Eager & Brunton, 2023). In practice, AI tools are already influencing how instructors evaluate student work. Farazouli et al. (2024) found that educators tend to be more cautious and stricter when they suspect AI-generated content, which underlines the need to update quality criteria.

As AI becomes more embedded in educational settings, it is crucial to examine how different types of assessment are affected. While students often express concerns about creativity and authenticity, educators are generally more open to using AI when it promotes critical thinking. These contrasting views highlight the need for inclusive policies and clear guidelines that involve both groups (Smolansky et al., 2023).

Despite advances in AI detection technology, its effectiveness remains limited. Studies show that neither educators nor software tools can consistently identify AI-generated content, reinforcing the argument for rethinking assessment strategies and providing training for all stakeholders (Perkins et al., 2024).

Reviews of institutional policies at leading universities show that many still lack clear strategies for addressing the challenges of GenAI. Nevertheless, some positive examples of integrating AI into assessment design are emerging (Moorhouse, Yeo & Wan, 2023).

Parker et al. (2024) argue that uncertainty in using AI tools is especially visible among early-career academics and teaching assistants, who often lack institutional guidance and clarity regarding professional responsibility. They also indicate that those with prior experience using ChatGPT are more skilled at identifying AI-generated texts, which suggests that experience and training are essential for developing digital literacy and maintaining academic integrity.

Research confirms that generative AI can support problem-solving and critical thinking but also carries the risk of misuse. Ogunleye et al. (2024) stress that task design plays a key role in preventing negative outcomes and ensuring responsible use. To guide this

process, Schön et al. (2023) propose a conceptual model that includes three dimensions: people, learning, and organization. This model emphasizes the need for new competencies, ethical guidelines, and institutional adaptations, as well as a broader awareness of the values shaping education today.

One major challenge in this evolving context is trust. When students are unsure about what is allowed and teachers do not clearly explain their expectations, this can create a sense of surveillance and distrust. Luo (2024) argues that open communication and transparent assessment policies are essential for restoring confidence between students and educators.

Even though the European Union is leading in the development of ethical AI frameworks, applying these principles in educational systems remains difficult. Tubella, Mora-Cantallops and Nieves (2023) emphasize the need for educational strategies and accessible resources to ensure the responsible use of AI in teaching and learning.

The impact of ChatGPT on essay writing demonstrates that it is possible to produce high-quality texts, but also that problems such as citation consistency and reproducibility remain. This supports the need for assessment models that focus not only on knowledge but also on competences and practical skills, aligned with lifelong learning (Kolade, Owoseni & Egbetokun, 2024).

Despite risks of misuse, many scholars advocate for the ethical use of generative AI in education. With responsible teaching and character development, authentic assessment practices can support the constructive use of AI rather than excluding it altogether (Crawford, Cowling & Allen, 2023).

Educators' views on AI remain diverse, which reflects the complexity of its integration into higher education. Research from an Australian university found that many instructors already use AI in teaching, particularly to adapt assessment tasks and stay current with technological developments (Lee et al., 2024). These findings point to the need for ongoing professional development and stronger institutional support.

This trend is also visible in primary and secondary education. Chiu (2024) shows that GenAI tools are increasingly used at earlier education levels, although public and expert debates mostly focus on higher education. Interviews with teachers and school leaders highlight the need to revise educational policies, introduce new professional standards, and support interdisciplinary teaching about AI.

Institutional policies play a central role in shaping these developments. Khlaif et al. (2024), in their study of educators in the Middle East, found that while teachers recognize the benefits of AI, such as increased engagement and reduced workload, they also express concerns about academic honesty and students' essential skills. The authors recommend stronger ethical frameworks and targeted professional development.

From a broader pedagogical perspective, Escotet (2024) cautions that AI can only enhance education when used to support personalization and feedback, not as a replacement for pedagogical judgment. While AI can help in content creation, data analysis, and administrative tasks, its use must remain rooted in human values and professional responsibility.

In the context of language education, Dong (2023) found that AI writing tools support student skill development, offer timely and personalized feedback, and improve engagement. These findings support the transformation of teaching methods, especially in academic writing in English as a foreign language.

Riapina (2024) emphasizes the benefits of personalized learning, adaptive teaching strategies, and interdisciplinary collaboration. At the same time, she highlights the importance of ethical guidelines to ensure fairness and transparency in the use of AI in education.

To sum up, the role of generative AI in higher education is complex and evolving. These technologies are not just technical tools, but also bring important questions about values in education, ethical responsibility, power structures and the goals of teaching and learning. To respond to these challenges, it is important to develop clear theoretical foundations, supportive institutional policies and teaching practices that enable responsible and meaningful use of AI. Balancing the potential of technology with human-centred values will be essential for maintaining academic integrity, encouraging deeper learning and preparing students for the challenges of a rapidly changing world.

3 Methodology

The empirical part of this study was based solely on a quantitative research design, with a structured questionnaire serving as the primary instrument for data collection. The aim was to explore the attitudes and experiences of university teachers regarding the use of AI tools for automated assessment and feedback in higher education.

Participation was voluntary and anonymous, and it was clearly stated at the beginning of the questionnaire that the research was conducted exclusively for academic purposes. The questionnaire consisted of three main sections. The first section gathered demographic data (such as age, gender, academic position, employment status, years of teaching experience, and field of study), along with self-assessments of digital and AI literacy. The second section was addressed only to respondents who had previously used AI tools, asking about the types of tools used, the teaching activities in which they were applied, and the nature of assignments where AI was involved in the assessment process. The third section presented a series of statements on the perceived advantages, limitations, and challenges related to AI-

supported assessment, rated using a 5-point Likert scale.

The questionnaire was distributed online during May 2025, by sending a call for participation to the official contact email addresses (126 in total) of all public and private higher education institutions in the Republic of Croatia, as listed on their official websites. These contacts typically represented institutional coordinators or administrative offices and were asked to further distribute the survey internally among teaching staff. In this way, at least theoretically, all academic staff in the Croatian higher education system had the opportunity to participate, provided the invitation was shared internally by their institution.

The collected data were analysed using both descriptive and inferential statistical methods, and the most relevant findings are presented and discussed in the following chapters.

4 Descriptive analysis of survey results

This chapter presents the descriptive analysis of the survey conducted among university teachers. The analysis is structured according to the thematic sections of the questionnaire and includes an overview of participant socio-demographic characteristics, prior experiences with AI tools, specific educational activities supported by these tools, and respondents' attitudes, perceived challenges, and institutional context. A total of 111 valid responses were analysed, offering an empirical foundation for understanding the current state of AI use in assessment practices within Croatian higher education. In line with the principle of anonymity guaranteed to participants, the survey did not include a question about institutional affiliation, and respondents were informed that answering questions related to their academic rank and disciplinary field was optional. Despite this, most participants voluntarily provided information about their current position and field of work, which allowed for a general overview of the sample without compromising confidentiality. This approach was taken to encourage honest responses on a potentially sensitive topic, while respecting participants' privacy.

4.1 Socio-demographic characteristics

The sample of participants in this study consisted predominantly of women (73.87%), which may reflect the gender composition of certain academic fields, as well as the generally higher participation of women in educational research. The age categories were defined to reflect common generational groupings used in social research, and to facilitate comparative analysis of responses in relation to digital familiarity and adoption tendencies. The age distribution shows that the largest group of respondents were between 40 and 49 years old (35.14%), followed by those aged 50 to 59 (28.83%) and 30 to 39 (26.13%). Younger respondents (under 30, 3.60%) and older ones (60+, 6.31%) were significantly less represented. The academic ranks of respondents varied, with associate professors being the most common (29.09%), followed by assistant professors (18.18%), teaching assistants (16.36%), full professors (10.91%), and tenured full professors (7.27%). Senior lecturers (7.27%), lecturers (4.55%), and postdoctoral researchers (5.45%) were also included. Almost all respondents (93.58%) were permanently employed at public higher education institutions, suggesting that their experiences and attitudes are shaped primarily within the public education sector.

When it comes to teaching experience, the largest share of participants reported between 11 and 20 years of experience (35.14%), followed by those with 21 to 30 years (30.63%). Less represented were those with 0–5 years (16.22%), 6–10 years (12.61%), and more than 30 years (5.41%) of experience. Most respondents came from the social sciences (31.53%) and humanities (16.22%), followed by technical sciences (14.41%), biotechnical sciences (14.41%), and biomedical and health sciences (10.81%). Other fields, such as natural sciences (4.50%), the arts (1.80%), and interdisciplinary sciences (6.31%), were represented to a lesser extent.

Most respondents rated their general digital literacy as high, selecting 4 (52.25%) or 5 (25.23%) on a 5-point scale (mean = 4.02; SD = 0.71; N = 111). However, self-assessed literacy in using AI tools was somewhat lower (mean = 3.20; SD = 1.02; N = 111), with the majority selecting 3 (39.64%) or 4 (29.73%). Only a small number of respondents rated their AI literacy as very high (5 – 9.01%), while 6.31% gave the lowest rating (1). These findings suggest that while the sample is generally digitally literate, there remains a lack of confidence and skills in using specific AI tools, which provides important context for interpreting the rest of the study.

The final question in the first section asked, "Have you ever used artificial intelligence (AI) tools for assessment or generating feedback?" This question played a key role in determining the path through the questionnaire. Of the 111 respondents, 61 (54.95%) answered "yes," while 50 (45.05%) answered "no." Based on their answers, respondents were directed to different parts of the survey. Those who answered affirmatively proceeded to the second section, where they were asked to specify which AI tools they used, and for which activities and types of tasks. Those who answered negatively skipped this section and continued directly to the third section, which focused on general attitudes and perceived challenges related to AI tools. This structure allowed for collecting data tailored to each respondent's experience, gaining detailed insights from AI users and perceptions from non-users.

4.2 Experience with AI tools for assessment and feedback

The second section was presented only to respondents who indicated prior use of AI tools for assessment or feedback purposes. A total of 61 respondents answered the first and second questions in this section, while 60 answered the third question. This part of the survey explored the specific tools, teaching activities, and types of tasks in which AI had been applied in their professional practice.

The first set of items asked participants to identify which AI tools they had used. Results show a clear dominance of large language models (LLMs). Of the 61 respondents, 59 (96.72%) reported using tools such as ChatGPT, Claude, or Gemini, indicating near-universal adoption of LLMs in academic evaluation. This high usage rate may be explained by the accessibility, flexibility, and ease of use of these tools across different educational contexts. In contrast, specialized tools such as Gradescope or FeedbackFruits were not used at all (0%), possibly reflecting limited institutional support, poor integration into existing systems, or a lack of awareness. Moderate use was reported for tools like Turnitin Draft Coach/AI Writing Feedback (29.51%), AI features in LMS platforms like Moodle (26.22%), and Microsoft-integrated AI tools (e.g., Copilot) (19.67%). These findings suggest that some educators rely on AI features embedded within familiar platforms. Less frequently used tools included Google Classroom AI (14.75%) and various niche tools such as Otter.ai, Lumi, MagicSchool AI, Curipod, and Brisk (1.64% to 4.92%). Overall, these results indicate that LLMs are currently the most widely used and accessible form of AI support in assessment, while more advanced or domain-specific tools are still in early stages of adoption. Although the survey did not explicitly ask whether respondents used free or subscription-based versions of AI tools, some comments in the open-ended responses indicated that access to paid versions (such as ChatGPT Plus) was limited. This distinction may influence the range of features used and overall user experience, which represents a relevant variable for future research.

The second group of items focused on the types of teaching and evaluative activities in which AI tools had been employed. The most frequent response was plagiarism detection (65.57%), suggesting that educators see clear value in AI for technical, automated tasks. Other common uses included stylistic or language analysis of student writing (36.07%), generating feedback suggestions (32.79%), and recommending additional materials based on student performance (32.79%). These activities suggest a gradual expansion of AI into formative teaching practices. Fewer respondents reported using AI for customizing feedback (31.15%), developing rubrics (19.67%), and analysing quiz/test responses (16.39%). Even fewer used AI for automated essay grading (11.48%) or short-answer quiz evaluation (14.75%), which may reflect hesitancy to relinquish control over interpretative assessment tasks to AI systems. Rare uses included audio/video content evaluation (8.20%), generating performance reports (14.75%), and creating tailored learning materials (18.03%). These lower rates likely reflect technical limitations, lack of tool integration, or insufficient training. In summary, AI tools are most commonly used for technical and supportive roles in assessment, particularly plagiarism detection, language analysis, and feedback generation. More complex uses, such as direct grading or creative task evaluation, remain limited, reflecting cautious trust in AI as a standalone assessment agent. In most cases, AI is seen as a supplementary tool, not a replacement for professional judgment. However, the study did not collect specific data on how teachers technically interacted with these systems. For example, whether they uploaded entire student assignments or worked with shorter excerpts using prompts. This limits deeper insight into practical workflows and ethical considerations, and will be addressed in future research.

The third set of items explored the instructional contexts in which these tools were used. The results indicate that AI tools are most frequently used for text-based tasks involving analysis, interpretation, and expression. The most common context was language and communication tasks, such as summarization, translation, or editing, as reported by 66.67% of respondents (40 out of 60). This suggests that educators recognize the strength of generative AI in processing and evaluating language-based content. AI is most often used by respondents for research papers and analytical writing (40%), including tasks like information structuring, citation checks, and stylistic improvement. Reflective writing (33.33%) and creative tasks (30%) were also commonly supported, indicating that AI is valued in open-ended tasks that emphasize expression and originality. In contrast, use in multiple-choice quizzes (28.33%) and open-ended quiz questions (10%) was lower, as was use in essay grading (28.33%). Interestingly, AI was least used in multi-phase projects (10%), oral presentations (11.67%), and portfolio assessments (6.67%). A modest share of respondents (20%) also reported using AI for administrative tasks. These findings suggest that AI is more applicable to structured, text-based assessments than to long-term or unstructured tasks, which are still difficult to automate in higher education. Overall, AI tools are most often used for tasks involving textual processing and analysis, while more complex or multidimensional tasks remain less explored.

4.3 Attitudes toward AI tools for assessment and feedback

The third and final section of the questionnaire examined university teachers' attitudes, perceived

challenges, and institutional context related to the use of AI tools for assessment and feedback.

The first set of items assessed to what extent participants agreed with various statements regarding the efficiency, fairness, objectivity, and educational value of AI tools in student evaluation processes. Respondents rated their level of agreement with each statement using a five-point Likert scale, ranging from 1 (strongly disagree) to 5 (strongly agree). A total of 103 respondents completed this question, offering valuable insights into current perceptions of AI tools in the context of higher education assessment. The most positively rated statement was "AI increases the efficiency of student assessment" (mean = 3.09; SD = 0.98), indicating moderate optimism about its usefulness. In contrast, the lowest-rated statement was "AI-generated feedback is more useful to students than teacher feedback" (2.17; SD = 0.99), reflecting significant scepticism about AI's capacity to replace human pedagogical insight. Respondents also expressed low agreement with statements that AI tools are more accurate (2.47; SD = 1.03), fair (2.55; SD = 1.09), or more objective (2.87; SD = 1.07) than teachers. Only the statement about plagiarism detection received a score above the neutral midpoint (3.47; SD = 1.14), showing broader confidence in AI for technical tasks. Statements about the role of AI in the evaluation process, whether as a first step (2.50; SD = 1.15) or last (2.85; SD = 1.20), were also rated relatively low, indicating uncertainty about AI's pedagogical place. Similarly, respondents expressed mostly neutral views regarding whether AI is more suitable for formative than summative assessment (2.88; SD = 0.89).

The second part of this section focused on perceived challenges. A total of 109 respondents completed this question. The greatest concern was that AI tools might provide inaccurate or unfair evaluations (3.98; SD = 0.93), followed by the fear that AI fails to understand context, creativity, or nuance (3.61; SD = 0.97). Lack of clear institutional policies was also a major issue (3.82; SD = 1.16), as was the risk of losing control over the assessment process (3.45; SD = 1.04) and concerns over data privacy (3.36; SD = 1.24). A significant portion of respondents felt insufficiently trained (3.14; SD = 1.25) and lacking access to necessary technical support (2.91; SD = 1.24). The time investment required to learn these tools was also seen as a potential barrier (2.76; SD = 1.09). Some reported colleague or student resistance (3.07; SD = 1.12), and others were uncertain about how to apply AI tools in their current systems (3.08; SD = 1.21) or believed that the tools are not suitable for their type of evaluation (3.12; SD = 1.11). A portion of respondents did not see clear benefits (2.98; SD = 1.18) or viewed AI feedback as generic and unhelpful (3.29; SD = 1.01).

The following question was a single-item statement aimed to capture whether university teachers feel the need for additional training to use AI tools effectively and responsibly in the context of assessment and feedback. Interestingly, this was the highest-rated item in the entire section (3.99; SD = 1.22; N = 111), clearly highlighting a strong demand for structured professional development in this area.

The final question asked whether participants' institutions promote responsible AI use in education. Of the 111 respondents, 57.66% answered no, indicating a general absence of institutional support or initiative. A further 24.32% said that formal policies were in progress, while only 10.81% stated that their institution had a binding official policy in place. Other responses (e.g., "I don't know," "not familiar") were minimal but pointed to broader uncertainty and a lack of communication around institutional expectations and guidelines. This lack of formal institutional support may contribute to hesitation or inconsistent adoption of AI tools in educational assessment. Moreover, it highlights the urgent need for clearer communication, strategic leadership, and policy development to ensure responsible and effective integration of AI in teaching and evaluation practices.

5 Discussion of key findings and emerging patterns

Beyond the descriptive overview, additional statistical analysis was conducted to explore the relationships between demographic factors, digital and AI literacy, institutional support, and perceptions of AI use in assessment and feedback. This section presents the key patterns and correlations that emerged from the data and highlights relevant implications.

5.1 Institutional contexts, privacy concerns, and perceptions of AI fairness

The analysis shows that institutional policies and contexts strongly influence how university teachers use and perceive AI tools in assessment. Respondents from institutions with formal or developing AI policies were more likely to use these tools (60% vs. 52%), though they reported lower average ratings of creativity in assessment, suggesting that such policies may introduce constraints that reduce perceived autonomy.

At the same time, the presence of institutional guidelines was associated with lower concern about student data privacy (mean = 3.13 vs. 3.49), indicating that institutional support can alleviate ethical uncertainties. Perceptions of AI fairness and accuracy varied across academic disciplines and levels of teaching experience, greater trust was expressed by those in fields with well-defined evaluation criteria, such as biomedicine (Table 1), and by less experienced teachers.

Table 1. Average teacher ratings of AI evaluation across scientific fields (source: authors' work)

Scientific Field	Fairer	More accurate	More consistent	More objective
Arts	3.0	3.0	3.0	3.0
Biomedicine and Health Sciences	2.9	3.0	3.4	3.3
Biotechnical Sciences	2.8	2.6	3.4	3.2
Social Sciences	2.7	2.5	2.8	2.8
Technical Sciences	2.6	2.6	3.1	3.2
Interdisciplinary Sciences	2.5	2.2	3.2	2.8
Humanities	2.0	2.0	2.1	2.3
Natural Sciences	1.8	1.8	2.0	2.6
Overall (mean)	**2.54**	**2.46**	**2.88**	**2.9**

Among non-users of AI tools, the most frequently cited barriers were fears of unfair or inaccurate evaluation, lack of institutional guidance, and concerns about student privacy. A clear relationship was also found between privacy concerns and perceptions of fairness: the more concerned teachers were about data privacy, the less fair they considered AI evaluation to be. These findings underscore the need for comprehensive institutional policies and transparent communication strategies to foster trust and promote the responsible use of AI in educational assessment.

5.2 Correlational insights

To better understand the relationships between different attitudes toward AI tools in assessment, a Spearman correlation matrix was calculated based on all variables rated by the respondents. From the full matrix, several statistically notable and theoretically relevant correlations were identified and are discussed below.

A moderate positive correlation ($\rho = 0.50$) was found between respondents' general digital literacy and their self-assessed literacy in using AI tools. This suggests that those who feel more confident in their general digital skills are also more likely to feel capable of using AI in their teaching practice. This finding supports Parker et al. (2024), who argue that previous experience with AI tools like ChatGPT contributes to a greater sense of competence. It also highlights the importance of continuous training that includes both technical and pedagogical dimensions of AI use, as recommended by Xia et al. (2024).

Another significant finding is the correlation between the belief that AI increases efficiency in assessment and the belief that it can recognize context, creativity, and language nuances in student work ($\rho = 0.61$). Respondents who see AI as efficient are also more likely to believe that it can handle complex aspects of student responses. This shows a high level of trust in AI tools and aligns with findings by Dong (2023) and Eager & Brunton (2023), who suggest that AI can support deeper forms of learning and assessment when it is properly integrated into course design.

A particularly strong correlation ($\rho = 0.88$) was found between the belief that AI evaluates more fairly than teachers and the belief that it evaluates more accurately. This suggests a strong belief that AI can outperform human assessors in terms of objectivity. While this confidence may come from negative experiences with human inconsistency or bias, it also contradicts studies warning about the limitations of AI in detecting plagiarism or understanding context (Perkins et al., 2024; Alexander, Savvidou & Alexander, 2023). Still, it fits within the arguments by Riapina (2024) and Kolade et al. (2024), who support the use of AI as a complement rather than a replacement for teacher judgment, especially in more routine or factual tasks.

Another interesting correlation ($\rho = 0.53$) was found between the belief that AI can recognize subtle aspects of student writing and the belief that it can provide useful feedback. This supports the idea that AI tools may have a meaningful role in formative assessment. The finding is consistent with Hooda et al. (2020), who emphasize the value of immediate and personalized feedback, and with authors like Williams (2023) and Bennett & Abusalem (2024), who argue for "assessment for learning" approaches. This also suggests that teachers appreciate AI not just for speed and automation, but for its perceived contribution to the learning process.

Across the full matrix, several variables related to perceived fairness, objectivity, accuracy, and consistency of AI assessment were highly correlated. This suggests that respondents tend to group these dimensions into one overall positive impression of AI. Such patterns can be interpreted as general enthusiasm, but they may also reflect a lack of clear differentiation between different aspects of assessment quality. As Kurtz et al. (2024) point out, teacher training needs to help educators critically understand both the potential and limitations of AI tools.

One of the most relevant findings relates to the correlation between "I do not have enough knowledge about AI tools" and "I do not know how to include AI in my assessment practices" ($\rho = 0.62$). This confirms that lack of knowledge, not resistance, is a major reason why teachers do not use AI. This supports calls from Moorhouse, Yeo & Wan (2023) and Parker et al. (2024) for greater access to professional development and institutional support. Without such support, digital inequalities between educators may deepen.

Another correlation ($\rho = 0.59$) shows that those who do not know how to use AI in their own assessment systems are also more likely to say that AI tools are not useful. This shows how much perception of usefulness depends on feeling in control and

competent. Tubella et al. (2023) argue that even the best technologies remain unused without clearly defined strategies, guidelines, and support structures.

It is also worth noting the connection between two specific beliefs: that AI is more suitable for open/essay-type questions, and that it is more appropriate for formative rather than summative assessment ($\rho = 0.59$). This suggests that teachers who see AI as helpful in complex or creative student tasks are also more likely to see it as a tool for feedback and learning, not final grading. This supports the shift away from purely summative models toward "assessment for learning," as suggested by Williams (2023) and Eager & Brunton (2023). In this context, AI is seen as a support (not a substitute) for teachers.

In summary, these findings show that teachers' attitudes toward AI in assessment are interrelated and shaped by their digital confidence, familiarity with technology, and level of trust in its capabilities. High correlations between various positively perceived aspects of AI suggest that expectations may be high, but not always critically examined. As noted by Farazouli et al. (2024) and Alexander, Savvidou & Alexander (2023), it is important to balance optimism with caution and to promote the responsible and thoughtful use of AI in higher education. Institutions need to support this process by offering practical training, clear policies, and space for open discussion about the benefits and limits of AI in assessment.

6 Conclusion

The findings of this study show that AI tools are starting to find their place in Croatian higher education, but their use in assessment remains cautious and selective. Most university teachers use AI tools, especially large language models, for relatively technical tasks such as detecting plagiarism or improving student language. However, their use in more interpretative or creative forms of assessment is still not that common. This suggests that while AI is seen as helpful in supporting certain tasks, it is not yet trusted to take a more central role in academic evaluation.

Teachers in the study express moderate confidence in AI's usefulness, especially in terms of saving time and improving efficiency. However, they also express significant concerns about fairness, accuracy, and the quality of AI-generated feedback. Many are unsure whether AI can understand context, creativity, or nuance in student work, which are all elements that are essential to meaningful assessment. Importantly, most respondents report a lack of institutional guidelines and feel they do not have enough training or support to use AI tools effectively.

Correlational data from the survey reveal that teachers who feel more confident in their AI skills are more likely to see its benefits. However, general digital literacy alone does not predict AI adoption, pointing to the need for more targeted training focused specifically on educational and ethical use of AI. The absence of clear policies also contributes to uncertainty, raising concerns about data privacy, professional responsibility, and the risk of relying on AI without fully understanding its limitations.

These findings suggest that without institutional guidance and structured professional development, the integration of AI into assessment will remain fragmented and conservative. Educators require not only technical training but also support in designing pedagogically meaningful assessment strategies that include AI as a complement, not a replacement for human judgment. At the same time, institutions must ensure that AI is used responsibly, transparently, and in ways that align with academic values.

This study offers an important first look at the state of AI-supported assessment in Croatian universities. However, it also has limitations. The sample is self-selected and may not fully represent all academic fields or institutions. The survey design is primarily quantitative and does not capture deeper qualitative insights into teachers' reasoning and experiences. Future research should include interviews or focus groups, examine disciplinary and generational differences, and explore how students experience and respond to AI-based assessment. These aspects will be essential for developing ethical, inclusive, and pedagogically sound uses of AI in higher education. In addition, future studies should aim to collect more detailed data on how AI tools are used in concrete teaching and assessment practices, and distinguish between different types of access and tool functionalities. A more granular demographic and disciplinary breakdown of participants will also be important for assessing patterns of adoption and perception across diverse academic contexts.

References

Alexander, K., Savvidou, C., & Alexander, C. (2023). Who wrote this essay? Detecting AI-generated writing in second language education in higher education. *The Journal of Teaching English with Technology*, 23(2), 25–43. https://doi.org/10.56297/BUKA4060/XHLD5365

Amin, M. Y. M. (2023). AI and Chat GPT in language teaching: Enhancing EFL classroom support and transforming assessment techniques. *International Journal of Higher Education Pedagogies*, 4(4), 1-15.

Bennett, L., & Abusalem, A. (2024). Artificial Intelligence (AI) and Its Potential Impact on the Future of Higher Education. *Athens Journal of Education*, 11(3), 195-212.

Chiu, T. K. (2024). The impact of Generative AI (GenAI) on practices, policies and research

direction in education: A case of ChatGPT and Midjourney. *Interactive Learning Environments*, 32(10), 6187-6203.

Crawford, J., Cowling, M., & Allen, K.-A. (2023). Leadership is needed for ethical ChatGPT: Character, assessment, and learning using artificial intelligence (AI). *Journal of University Teaching & Learning Practice*, 20(3). https://doi.org/10.53761/1.20.3.02

Dabis, A., & Csáki, C. (2024). AI and ethics: Investigating the first policy responses of higher education institutions to the challenge of generative AI. *Humanities and Social Sciences Communications*, 11(1). https://doi.org/10.1057/s41599-024-03526-z

Dong, Y. (2023). Revolutionizing academic English writing through AI-powered pedagogy: Practical exploration of teaching process and assessment. *Journal of Higher Education Research*, 4(2).

Eager, B., & Brunton, R. (2023). Prompting higher education towards AI-augmented teaching and learning practice. *Journal of University Teaching and Learning Practice*, 20(5), 1-19.

Escotet, M. Á. (2024). The optimistic future of Artificial Intelligence in higher education. *Prospects*, 54(3), 531-540.

Farazouli, A., Cerratto-Pargman, T., Bolander-Laksov, K., & McGrath, C. (2024). Hello GPT! Goodbye home examination? An exploratory study of AI chatbots impact on university teachers' assessment practices. *Assessment & Evaluation in Higher Education*, 49(3), 363-375.

Forde-Leaves, N., Walton, J., & Tann, K. (2023). A framework for understanding assessment practice in higher education. *Assessment & Evaluation in Higher Education*, *48*(8), 1076-1091.

Hooda, M., Rana, C., Dahiya, O., Rizwan, A., & Hossain, M. S. (2022). Artificial intelligence for assessment and feedback to enhance student success in higher education. *Mathematical Problems in Engineering*, 2022(1). https://doi.org/10.1155/2022/5215722

Khlaif, Z. N., Ayyoub, A., Hamamra, B., Bensalem, E., Mitwally, M. A., ... & Shadid, F. (2024). University teachers' views on the adoption and integration of generative AI tools for student assessment in higher education. *Education Sciences*, 14(10).

Kolade, O., Owoseni, A., & Egbetokun, A. (2024). Is AI changing learning and assessment as we know it? Evidence from a ChatGPT experiment and a conceptual framework. *Heliyon*, 10(4).

Kramm, N., & McKenna, S. (2023). AI amplifies the tough question: What is higher education really for? *Teaching in Higher Education*, 28(8), 2173-2178.

Kurtz, G., Amzalag, M., Shaked, N., Zaguri, Y., Kohen-Vacs, D., Gal, E., ... & Barak-Medina, E. (2024). Strategies for integrating generative AI into higher education: Navigating challenges and leveraging opportunities. *Education Sciences*, 14(5).

Lee, D., Arnold, M., Srivastava, A., Plastow, K., Strelan, P., Ploeckl, F., ... & Palmer, E. (2024). The impact of generative AI on higher education learning and teaching: A study of educators' perspectives. *Computers and Education: Artificial Intelligence*, 6.

Luo, J. (2024). How does GenAI affect trust in teacher-student relationships? Insights from students' assessment experiences. *Teaching in Higher Education*, 30(4), 991–1006. https://doi.org/10.1080/13562517.2024.2341005

Moorhouse, B. L., Yeo, M. A., & Wan, Y. (2023). Generative AI tools and assessment: Guidelines of the world's top-ranking universities. *Computers and Education Open*, 5.

Ogunleye, B., Zakariyyah, K. I., Ajao, O., Olayinka, O., & Sharma, H. (2024). Higher education assessment practice in the era of generative AI tools. *Journal of Applied Learning & Teaching*, 7(1). https://doi.org/10.37074/jalt.2024.7.1.28

Parker, L., Carter, C., Karakas, A., Loper, A. J., & Sokkar, A. (2024). Graduate instructors navigating the AI frontier: The role of ChatGPT in higher education. *Computers and Education Open*, 6. https://doi.org/10.1016/j.caeo.2024.100166

Perkins, M., Roe, J., Postma, D., McGaughran, J., & Hickerson, D. (2024). Detection of GPT-4 generated text in higher education: Combining academic judgement and software to identify generative AI tool misuse. *Journal of Academic Ethics*, 22(1), 89-113.

Riapina, N. (2024). Teaching AI-enabled business communication in higher education: A practical framework. *Business and Professional Communication Quarterly*, 87(3), 511-521.

Schön, E.-M., Neumann, M., Hofmann-Stölting, C., Baeza-Yates, R., & Rauschenberger, M. (2023). How are AI assistants changing higher education? *Frontiers in Computer Science*, 5. https://doi.org/10.3389/fcomp.2023.1208550

Serrano, M. M., O'Brien, M., Roberts, K., & Whyte, D. (2018). Critical Pedagogy and assessment in higher education: The ideal of 'authenticity'in learning. *Active Learning in Higher Education*, 19(1), 9-21.

Smolansky, A., Cram, A., Raduescu, C., Zeivots, S., Huber, E., & Kizilcec, R. F. (2023). Educator and

student perspectives on the impact of generative AI on assessments in higher education. *Proceedings of the Tenth ACM Conference on Learning @ Scale*. https://doi.org/10.1145/3573051.3596191

Sullivan, M., Kelly, A., & McLaughlan, P. (2023). ChatGPT in higher education: Considerations for academic integrity and student learning. *Journal of Applied Learning & Teaching*, 6(1), 1-10. https://doi.org/10.37074/jalt.2023.6.1.17

Tubella, A. A., Mora-Cantallops, M., & Nieves, J. C. (2023). How to teach responsible AI in higher education: Challenges and opportunities. *Ethics and Information Technology*, 26(3). https://doi.org/10.1007/s10676-023-09733-7

Williams, P. (2023). AI, analytics and a new assessment model for universities. *Education Sciences*, 13(10).

Xia, Q., Weng, X., Ouyang, F., Lin, T. J., & Chiu, T. K. (2024). A scoping review on how generative artificial intelligence transforms assessment in higher education. *International Journal of Educational Technology in Higher Education*, 21(1).