Iulian ÎNTORSUREANU, PhD (corresponding author)
iulian.intorsureanu@ie.ase.ro
Bucharest University of Economic Studies, Romania

Roxana VOICU-DOROBANȚU, PhD
roxana.voicu@rei.ase.ro
Bucharest University of Economic Studies, Romania

Codrin-Florentin NISIOIU, PhD
codrin.nisioiu@ie.ase.ro
Bucharest University of Economic Studies, Romania

Cătălin PLOAE, PhD
catalin.ploae@rei.ase.ro
Bucharest University of Economic Studies, Romania

Generative Artificial Intelligence and the Academic Integrity of Graduation Works in Economics – Exploring Perceptions of Romanian Academia

Abstract. *Technologies based on generative artificial intelligence (GAI) and large language models (LLM), available on platforms such as ChatGPT, have raised intense concerns about their impact on education. In the context of knowledge assessment, a particular challenge is to ensure academic integrity for bachelor's and master's theses, henceforth, graduation works (GWs).*
This article presents the results of an exploratory study on the perceptions of academia in Romanian economic higher education regarding the academic integrity of GWs in the context of using artificial intelligence (AI) tools based on LLMs, such as ChatGPT.
The survey-based research provides a detailed snapshot of the target group's attitudes and perceptions; the study also identifies significant relationships between constructs and proposes two latent factors – a structural and a consultative dimension – that reflect two facets of LLM use in authoring GWs.

Keywords: *higher education, economics, graduation works, academic integrity, generative artificial intelligence, ChatGPT, large language model (LLM).*

JEL Classification: A22, A23, O33.

1. Introduction

Developed in the past decade, Artificial Intelligence (AI) and Large Language Models (LLMs) are revolutionising education in general and higher education in particular, with its linkage to research. In education, AI, alongside other learning technologies, can revolutionise institutions by providing personalised learning experiences, automating tasks, managing and allocating intelligently teaching

DOI: 10.24818/18423264/58.2.24.08
© 2024 The Authors. Published by Editura ASE. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

resources, improving the utilisation and effectiveness of educational resources, enhancing assessment and smart grading, and making education more accessible.

However, as highly complex models based on the technique of artificial neural networks trained on enormous text volumes, LLMs achieve startling results by predicting the most likely words in a sequence based on given criteria – a threat to academic integrity, especially in the case of paper/essay topics, including graduation works (GWs). These challenges are also linked to academic and research integrity, with concerns regarding the increasing ease of cheating, which may erode public trust in higher education and research and undermine the intrinsic value of the academic degree.

Improvements in natural language processing (NLP) brought the concept of Generative Pre-Trained Transformers (GPTs). Model GPT-3 released in 2020 by OpenAI was a revolutionary model in generating realistic and coherent text, and November 2022 saw OpenAI release ChatGPT, a free, user-friendly application rendering GPT-3.5 (and GPT-4 as of March 2024) technology available to anyone. This launch led to an increase in LLM applications, with other companies such as Meta, Google, or Anthropic releasing big language models as well (LLaMa, Bard carried by PaLM 2 – later Gemini Advanced, Claude, respectively).

The research aims to investigate the perceptions and attitudes of the Romanian economic university environment regarding the challenges posed by the new LLM-based AI technologies at the end of the first academic year of the "ChatGPT era." The presented results may improve the academic response to the challenges posed by the widespread deployment of GAI tools. The discussions and measures taken in this context can be based on current data on perceptions and attitudes, such as those provided by this study. The research focusses on economic higher education in Romania, where plagiarism, particularly at the doctoral level, is considered a notorious topic and a gateway to corruption. No studies on the impact of GAI on higher education have been published nationally, and no studies are dedicated to economics or business education in Romania, although insights into the effects of AI exist (Năstasă, Matei, and Mocanu, 2023).

Considering the risk to academic integrity, the assessment through written work should be rethought, considering the challenges posed by LLM technology (Yan et al., 2023). Similar studies have focused on the impact of GAI/LLM in education on a more general level, both from the perspective of teachers, pupils, and students. Although integrity issues related to report/essay topics are present in this ensemble, they must be dealt with in detail. Given the importance of bachelor's and master's theses as evaluations upon completion of studies, the authors considered that the mentioned aspects deserve detailed research. Of particular interest for this paper is the literature related to investigating the topic in economic research. Authors such as Korinek (2023) discuss that GAI can significantly boost productivity in economic research by automating micro-tasks, with potential long-term implications for cognitive automation.

2. A literature review on GAI and ethical concerns in academia

AI-powered education has rapidly evolved to shape the use of research, learning, and assessment. AI in education (AIED) is a topic of discussion in policymaking and day-to-day academic operations. With LLMs, there is an even more radical transformation of approaches in higher education, altering the mechanisms of formulating and producing ideas while overall mitigating the downsides of qualitative research in phygital (physical + digital) settings (Schmitt, 2023).

The gains and benefits offered in education through various AI tools for teaching, learning, and research will be realised after weighing some ethical considerations. One of the primary concerns remains plagiarism detection and academic integrity. AI in general (Amani et al., 2023), and ChatGPT in particular (Michel-Villarreal et al., 2023) as a tool are themselves the topic for several articles, some listing also the perceptions of university teachers (Firat, 2023). It impacts all elements of higher education, from academic examinations to academic paper writing.

The ethical concerns regarding AIED are a key topic, with more than 3 million hits from the search terms "ethics" and "AI" on Google Scholar in December 2023 and 4.5 million hits from the same search terms in March 2024, with constant needs for systematic literature reviews, which must be updated with a very high frequency. In this context, academic integrity is one of the five major concerns of GAI in education (alongside "impact on learning outcomes and skill development, limitation of capabilities, policy and social concerns, and workforce challenges" as listed in (Li et al., 2023)).

The European Union (EU) and other educational organisations have taken the initiative to mitigate AIED's risk of dilution of academic integrity, although it may enhance academic performance or provide the context for empowering learners (Perkins, 2023). Through the Joint Research Centre, the EU has been issuing a series of reports on this area of vulnerability in the past years, which may lead to a dismantling of values (Tuomi, Cachia and Villar-Onrubia, 2023). The advent of sophisticated LLMs does not diminish the importance of core competencies drawn by the European Framework for the Digital Competence of Educators (*DigCompEdu*), as they hinge on four fundamental, immovable elements: agency, humanity, social fairness, and justified choice.

Academic misconduct (*FFP – fabrication, falsification, plagiarism*) is still a growing concern overall. Still, there is a need for a nuanced understanding of plagiarism among students, particularly in higher education settings, as there appears to be a critical gap in their comprehension of the concept and its implications due to a need for more understanding and awareness. The U.S. Department of Health and Human Services (n.d.) defines this type of academic misconduct, exacerbated by the advent of digital technology, as the unauthorised use of someone else's "ideas, processes, results, or words without providing proper acknowledgment". This unethical practice extends beyond just using another person's words; it encompasses

the uncredited use of ideas, graphics, designs, statistical data, methodologies, and computer programmes and has several forms, including buying papers, taking content from online sources without attribution, neglecting to use quotation marks for direct quotes, and failing to cite sources when paraphrasing. As a challenge to higher education, plagiarism has various forms and underlying causes, from culture to the prevalence of information and information tools, and lack of awareness of academic misconduct. Another reason for this faulty behaviour may be a need for more understanding and unfamiliarity with plagiarism, laziness, and a reluctance to read extensively. Most countries have policies to counteract academic misconduct, and teachers and students are targeted. The emergence of technological tools that put libraries just a click away from students enhances the different types of plagiarism, including paraphrasing.

Although intrinsically a benefit of the Internet, the digitalisation of knowledge fosters a range of unethical behaviours, and the latest buzzword is ChatGPT (Montenegro-Rueda et al., 2023), with GAI being perceived as an actual threat to transparent science. It may be used for academic plagiarism and other forms of intellectual theft, provided the teachers have specific tools to detect AI-generated content, which can negatively affect academic integrity. To avoid this situation, researchers and developers of AI need to optimise self-regulatory mechanisms for technologies such as ChatGPT to improve their safety and applicability. The deep integration of education with artificial intelligence technologies such as ChatGPT requires a reevaluation of the role of GAI in AIED and joint efforts from AI research and development personnel, educators, and students to achieve optimal educational outcomes. This reevaluation may provide alternatives, such as considering GAI an author. Moreover, the fact that GAI is not present in AIED solely through chat tools, but also with voice assistants, comprehension assessors, research assistants, translation tools, or other support tools makes the reevaluation process highly complex.

A first step may be to identify where academic institutions stand on AIED topics through their administration's attitudes, students, and faculty. The students and the faculty have a wide range of attitudes towards using technology in the classroom, specifically AIED. Studies such as (Chan and Hu, 2023) underline this mix of enthusiasm and reluctance, with differentiators in adoption based on gender, age, academic major, prior perception of technology (including in Romania), and prior exposure to technology in education. The faculty's caution regarding GAI appears to be cross-cultural, with studies from various countries listing it, particularly given the risks to academic integrity. Following a research by Moorhouse, Yeo, and Wan (2023) on GAI and assessment based on guidelines from the world's top-ranking universities, Luo (2024) investigates the policies on GAI in 20 universities when it comes to assessments, highlighting the main problem of GAI being considered as an external help, which decreases the value of students' independent work and their originality, thus emphasising the need for a sophisticated approach to originality in the context of technology-based education. Policies in place range from outright

banning the use of GAI in dissertations to a more nuanced approach, acknowledging its benefits.

Students are intrinsically motivated to use GAI (Lai, Cheung and Chan, 2023), expect high value from using AIED (and GAI, specifically) while reducing their own agency and relying more on than learning from GAI (Darvishi et al., 2024). Potential costs are perceived as non-deterrent to usage (Chan and Zhou, 2023). This conclusion may prove the gap between the faculty's cautionary approach and students' enthusiastic adoption, often disregarding academic misconduct.

Ethical Considerations

A series of studies propose solutions to the plagiarism conundrum in higher education. These solutions, apart from the radical one of banning the use of GAI (Yu, 2023), range from a basic anti-plagiarism tool (APTs) to complex honour code systems, with each tool carrying a certain impact in the institution in which it is used and a certain range of effectiveness and suitability. Based on trust and mutual respect, these systems can foster a culture in which students take personal responsibility for maintaining honesty in their academic endeavours.

In this comprehensive approach, needed given the mainstream emergence of LLMs (Yan et al., 2023), the role of the graduation thesis as written work for evaluation (being for undergraduate, dissertation, or doctoral thesis) is emphasised, with some authors even discussing the usefulness of these works in their current format (Kenwright, 2024). There is an international standard for theses (ISO 7144); however, its latest review is from 2019. Higher education branches targeting assessments that are mainly written work (such as humanities or law) are particularly vulnerable to the temptation of using a GAI tool for delivering a paper. Adilov et al. (2023) propose a Course Vulnerability Index meant to "measure the level of susceptibility of courses to cheating by using ChatGPT". Although it is an innovative and quantitative tool for heat-mapping an organisation's curriculum, this index does not consider the wide range of academic misconduct via GAI. Tools such as ChatGPT may prove to generate valueless content, which, in turn, may lead to flooding academia with superficial "word salads".

Ideas such as a more comprehensive educational approach to address plagiarism, particularly for graduation theses, emphasise the need for strict policies and educational programs to foster ethical research practices. There is a growing need for a multi-faceted approach to mitigate the issue, highlighting the students' proactive role in reducing plagiarism through self-motivation and enhanced awareness, and stressing the ineffectiveness of punitive measures. Strategies, lists of principles, and guidelines at the European or national level align with recommendations from multinational organisations, such as the European Network for Academic Integrity: ENAI, on navigating the complexity of ethical concerns in written works, not only given GAI. It is more important than ever to focus on moral and legal issues surrounding artificial intelligence technologies and establish sound regulatory mechanisms through joint efforts to promote their safe, reliable, and sustainable application. Ultimately, a more comprehensive framework is needed,

which should be supported by strong leadership (Crawford, Cowling and Allen, 2023), with the issue of authorship and attribution having to be addressed in a shifting mindset to digitally mediated knowledge production.

3. Research Methodology

Given the novelty of the current situation related to the wide-scale availability of LLMs and their implications for university assessment based on graduation works (GWs), this study has a mainly descriptive and explorative character.

The main objective of this research is to describe in detail the perceptions and attitudes of the Romanian economics academia regarding the challenges raised by the new generative AI technologies (GAI), more specifically LLMs, such as ChatGPT, in the context of undergraduate and dissertation works, GWs.

Furthermore, the study explores the relationships between these attitudes and factors such as familiarity with the LLM technology, coordination experience, and the coordinator's field of economics; also, we look at how the attitudes influence perceptions of academic integrity risk, the acceptability of specific LLM usage scenarios as well as the support for specific measures to mitigate integrity risks.

To this purpose, a cross-sectional survey-based research was carried out, with the following specific research objectives:

- RO1: describe the attitudes towards the use of LLM/ChatGPT in GWs in general;
- RO2: describe the perception of the risk posed by LLM/ChatGPT to the academic integrity of GWs;
- RO3: describe the perceptions of the required risk mitigation measures;
- RO4: describe the perceptions of acceptable use scenarios of LLM/ChatGPT in GWs, at the faculty level – coordinators and evaluators of GWs, from the Romanian economic higher education, as well as:
- RO5: explore how these attitudes and perceptions are related to each other, as well as with factors such as: experience in coordinating and evaluating GWs, degree of familiarity with LLM/ChatGPT, etc.
- RO6: Assess the possibility of identifying latent constructs that capture or explain the perception of LLMs acceptability in GWs authoring.

A series of constructs were identified and detailed, in turn, by variables to operationalise the objectives. These are presented in Table 1.

Table 1. Detailing the constructs and variables

Constructs	Variables
ATT: Attitudes towards using LLM/ChatGPT to write GWs	ATT-OVR: Overall feeling about using LLM/ChatGPT to write GWs ATT-OVRb (binned): negative, neutral, positive ATT-AD: Attitude regarding the adoption or rejection of LLM/ChatGPT in the context of GWs
EXPCT: Expectations w.r.t. future changes caused by LLM/ChatGPT	EXPCT-MAG: Expectation on the magnitude of changes caused by LLM/ChatGPT

Constructs	Variables
	EXPCT-SPD: Expectation on the speed of changes caused by LLM/ChatGPT EXPCT-EFF: Expectation of additional effort required from coordinators
RSK: Perception of risk to academic integrity using LLM/ChatGPT	RSK-BEF: Perception of plagiarism risk, before LLM/ChatGPT RSK-AFT: Perception of plagiarism risk after the emergence of LLM/ChatGPT
ACC: Perception of the degree of acceptability of LLM/ChatGPT usage scenarios, from the perspective of academic integrity	ACC-V1..ACC-V9: Perception of the degree of acceptability of the scenario S1..S9
FAM: Level of familiarity with LLM/ChatGPT technology	FAM-OVR: Level of familiarity with LLM/ChatGPT, in general (self-assessment) FAM-USE: Degree of experience in using LLM/ChatGPT FAM-INF: Information sources used FAM-LIM: Degree of knowledge about LLM/ChatGPT limitations FAM-WRK: Degree of knowledge about LLM working principles
M-INST: Perception of the need for institutional measures in response to the challenges raised by LLM/ChatGPT	M-INST-V1..V6: Perception of the need for institutional measure 1..6
M-PROC: Perception of the need for measures for the GW coordination process in response to the challenges raised by LLM/ChatGPT	M-PROC-V1..V11: Perception of the need for measure 1..11 within GW coordination and evaluation
XPR: Level of experience in coordinating and evaluating GWs	XPR-YRS: Number of years of experience
DOM: Scientific field (subfield of Economics)	DOM-V: Scientific field of respondent
UNIV: University	UNIV-V: University affiliation of respondent

Source: authors' realisation.

The data was gathered via a questionnaire with nineteen questions, specifically designed to investigate the specified constructs and variables. The types of questions used are multiple-choice questions with one or more possible answers, a matrix-type question (acceptability of LLM/ChatGPT usage scenarios), and an open-answer question ('other proposed measures'). Five-response Likert scales were used for the perceptions and expectations questions.

The questionnaire was administered online through the Qualtrics platform in the timeframe Sept. 2023–Feb. 2024. For data collection, a convenience (non-probability) sample was used, and the answers provided voluntarily were collected following the dissemination of the survey in the targeted academic environment – faculty members from the state universities in Romania, belonging to the economics departments. Out of 201 collected responses, 170 were retained after a validity verification.

Based on an own web-based data collection, the statistical population was estimated to consist of 1763 faculty members; the collected sample comprises therefore approximately 10% of the population.

The data analysis was carried out using the IBM SPSS software and the R environment, including the psych package (Revelle, 2023); statistical procedures to measure relationships among variables were used. Further, an exploratory factor analysis was performed, in order to identify latent factors influencing acceptability.

4. Results and Discussion

After processing the collected data, the results are presented further from the point of view of the respondents' characteristics and the aspects that were the subject of the study.

4.1 Respondents' profile

From all the questionnaires collected, regarding the field of economic sciences in which the respondents coordinate GWs, the following structure was recorded: Cybernetics, Statistics and Economic Informatics 17.6%, Finance 16.5%, Business Administration 15.9%, Management 14.7%, Accounting 12.9%, Marketing 11.8%, Economics 7.1%, International Business and Economics 3.5%. The field taxonomy was based on the official listing of scientific domains for bachelor and master studies in Romania.

Regarding coordination and evaluation experience, the distribution of respondents was as follows: 1-5 years 11.8%; 6-10 years 10.6%; 11-15 years 15.9%; 16-20 years 21.2%; over 20 years 40.6%. It can be observed that faculty members with extensive experience are predominant, with respondents with more than 15 years of practice representing 61.8% of the total.

For the degree of familiarity with the LLM/ChatGPT technology, the proportion of responses was: 'not at all' 7.6%, 'to a small extent' 17.1%, 'to some extent '42.9%, 'to a large extent' 24.1%, 'to a very large extent' 8.2%. Respondents who reported a medium-high degree of familiarity are predominant (75.3%) compared to those less familiar (24.7%). This proportion is reflected in the user experience results: 24.7% of respondents say they have not used LLM/ChatGPT directly; 42.9% accessed LLM/ChatGPT just out of curiosity, for testing purposes, while 32.4% used it for teaching and other purposes.

To the question about knowing the limitations of the LLM/ChatGPT technology regarding the generated text, the responses were: 'not at all' 11.2%, 'to a small extent' 23.5%, 'to some extent' 37.1%, 'to a large extent' 22.9%, 'to a very large extent' 5.3%. From the perspective of integrating LLM/ChatGPT in writing GWs, a percentage of 34.7% of coordinators with a low level of awareness of LLM limitations may be a strong signal regarding the need for information/training.

Regarding the familiarisation with the LLM type GAI technology, the answers were: ‘not at all’ 20.0%, ‘I understand the general principle of operation’ 71.2%, ‘I know in detail the principles and techniques on which it is based’ 8.8%.

The most used sources of information about LLM/ChatGPT were discussions with peers (67.6%) and media materials (60%). The data confirm the general interest in the subject of GAI shown in society and reflected in the media. 42.4% of the respondents reported consulting scientific articles, while 22.4% attended related training courses.

4.2 Attitudes and perceptions of LLM/ChatGPT use in GW writing and academic integrity risks

When asked "How do you feel about the possibility of ChatGPT (or similar AI tools) being used for writing bachelor's / master’s theses?", the predominant feeling reported was worry (44.1%). A slight predominance of rather negative feelings (scared, frustrated, bothered, worried) can be observed, amounting to 56.4%, compared to the positive ones. However, the extreme forms do not have a large weight (scared 0.6%, excited 4.1%).

The attitude regarding the integration of LLM technology registers a majority of answers for an adaptive approach (61.8% - "ChatGPT is a reality that cannot be avoided, but measures are needed to limit the negative consequences). 29.4% adopt an attitude favourable to the integration of technology proactively ("ChatGPT is a technology that can be useful in achieving LLD and should be integrated into the educational process and used by students under the coordination of teaching staff"), and 8.8% believe that a total ban on the use of LLM is required.

The perception of the risk posed by LLM/ChatGPT to the academic integrity of GWs reflects a major increase in the level of perceived risk after the advent of ChatGPT, from 36.4% to 70.6% of respondents (cumulative responses high + very high risk).

Regarding the expectation of changes in the way of assessment through GWs, the majority of the respondents expect significant changes (‘to some extent’ 40%, ‘to a large and very large extent’ 34.1%). The expected time horizon for these changes is 2-5 years (67.3% of respondents); 19.5% believe that they will occur in the next year, and 13.2%, in more than 5 years.

4.3 Acceptability of using LLM/ChatGPT and necessary measures

To explore the perception of acceptability of some LLM/ChatGPT usage scenarios for the realisation of GWs, the survey included nine scenarios (use cases), which were rated on the scale of unacceptable - acceptable, under certain conditions - acceptable. The most acceptable scenarios, considering the cumulative percentages for conditional acceptance and acceptance, were "using LLM/ChatGPT for language, spelling, grammar corrections in a paper (without using the tool elsewhere)" – 90.6% and "using LLM/ChatGPT for identifying relevant

references/bibliography for a thesis (without using the tool elsewhere)" – 77.6%. At the opposite end, the least acceptable scenarios were: "using the text generated by LLM/ChatGPT, without changes" – 24.7% and "making a first draft with LLM/ChatGPT, which is then critically reviewed, corrected and completed by the student" – 52.4%. The results of these scenarios are shown in Figure 1.

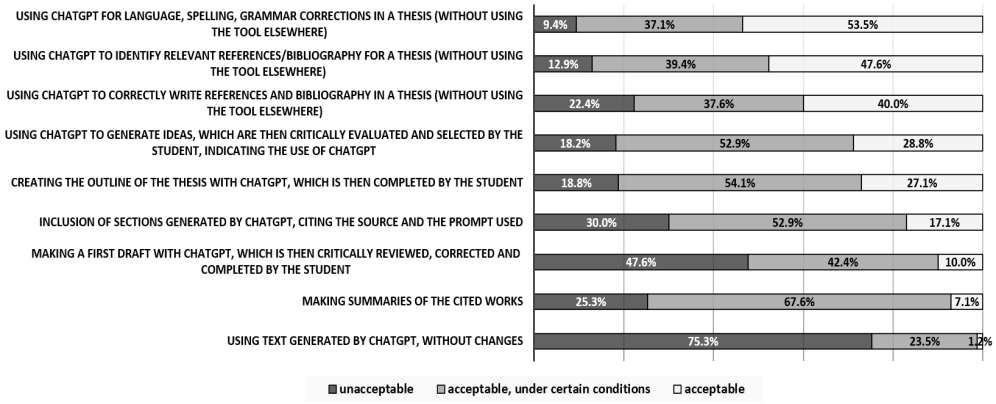

Figure 1. The degree of acceptability of LLM usage scenarios
Source: authors' realisation, diagram based on SPSS results.

The measures whose necessity was assessed through the questionnaire were divided into two categories, with separate questions: institutional measures taken at the level of the educational system and, respectively, measures for coordination and evaluation activities. As measures in the first category, the option "training activities and student awareness regarding academic integrity in general and the rules applicable to the use of LLM/ChatGPT" was rated as necessary to the greatest extent, with 62.4% of the responses. It is noted that, among the measures presented in the survey, most met over 50% of the opinions; the least selected was "eliminating assessment through GWs (substitution with other forms of assessment for the completion of studies)" with 31.2% - see Table 2.

Table 2. Perceptions of the need for institutional measures

No.	Measure	Favourable responses (%)
1	Training and awareness actions on academic integrity in general and on the rules applicable to the use of ChatGPT, for students	62.4
2	Training teachers on the capabilities and limitations of generative AI such as ChatGPT	57.1
3	Extending checks through anti-plagiarism programs, assessing the likelihood that text was generated by tools such as ChatGPT	55.3

No.	Measure	Favourable responses (%)
4	Adapting the university's regulations regarding plagiarism through ChatGPT, with detailed specification of accepted / inacceptable use cases	51.2
5	Providing recommendations for teachers on measures to limit cases of misconduct, in the process of coordinating theses	47.1
6	Eliminating assessment through bachelor and master theses (replacement with other forms of graduation assessment)	31.2

Source: authors' calculations using SPSS.

Regarding the measures related to coordination and evaluation, those considered necessary were "putting a greater emphasis on thesis presentation and debate" – 54.7% of respondents, "formulating creative topics, more difficult to solve through LLM/ChatGPT" – 54.1% and "detailed verification of the student's knowledge regarding the content of the work, for evaluation (defense)" – 51.2%.

Another question concerned the expectation of additional effort in the coordination activity. The overwhelming majority of respondents (85.9%) believe that additional workload will be required (40.0% 'to some extent', 31.8% 'to a large extent', 14.1% 'to a very large extent').

4.4 Relationships between variables

The collected data were evaluated regarding relationships among variables based on the chi-squared test (χ^2) with two variables, appropriate for nominal qualitative data. The hypothesis H_0 - the two variables are independent - was rejected, for the pair of variables AT-OVRb - ATT-AD: $\chi^2(df=4, n=170) = 43.468$, $p < 0.001$, $\varphi_c^2 = 0.256$. The variable AT-OVRb was derived from AT-OVR by binning, obtaining three new categories: negative, balanced, and positive feeling.

The φ_c^2 indicator (squared Cramer's phi coefficient) quantifies the size of the effect, representing the proportion of the variation explained between two qualitative variables. Values of 0.25 and higher are considered to show a large effect (Witte and Witte, 2017). This shows that the overall feeling towards LLMs and GWs has a strong influence on the attitude related to the integration / adoption of such tools.

Among the proposed institutional measures, a detailed analysis of the choice to opt out of GWs as a form of assessment (M-INST-V1) revealed its association with risk perception (RSK-AFT). Respondents who perceived the risk for academic integrity at a "very high" level favoured, to a greater extent (48%) the elimination of GWs compared to the overall proportion (31.2%), the relation being statistically significant for a small-medium effect: $\chi^2(df=4, n=170) = 10.27$, $p < 0.05$, $\varphi_c 2 = 0.06$.

4.5 Exploratory Factor Analysis

An exploratory factor analysis was conducted as a method of dimensionality reduction, in order to identify meaningful latent constructs which capture or explain

the perception of LLMs acceptability in GWs elaboration (variables ACC_V1.. ACC_V9).

The correlation structure of the nine ordinal items was examined through the polychoric correlation coefficients, as an alternative to the rank correlations coefficients, the items being measured on a Likert scale (Watkins, 2020). A similar approach was presented in (Covrig et al., 2023). All correlation coefficients are positive, and a large proportion of about 86% of all paired correlations are greater than 0.3. Furthermore, in order to assess whether the exploratory factor analysis was appropriate to factor the correlation matrix, we applied the Bartlett test of sphericity. The calculated value of the χ^2 test statistic, for 36 degrees of freedom, was 1063.839, with a p-value much lower than the 0.01 significance level; therefore, the null hypothesis that the correlation matrix is an identity matrix, is rejected. As well, the Kaiser-Meyer-Olkin measure of sample adequacy was 0.861, much higher than 0.7, the recommended value. Parallel analysis and Kaiser criterion recommended two factors to be extracted. After applying a principal axis factoring method of estimation, varimax rotation, the standardised loadings of the two factors (the correlation coefficients between each factor and the nine manifest variables) were obtained and are presented in Table 3, where they are displayed in decreasing order for readability.

Table 3. Standardised loadings of factors

Item	Item text	F1	F2
ACC_V8	Using ChatGPT for language, spelling, grammar corrections in a thesis (without using the tool elsewhere)	**0.947**	0.123
ACC_V9	Using ChatGPT to identify relevant references/bibliography for a thesis (without using the tool elsewhere)	**0.932**	0.048
ACC_V7	Using ChatGPT to correctly write references and bibliography in a thesis (without using the tool elsewhere)	**0.720**	0.343
ACC_V6	Using ChatGPT to generate ideas, which are then critically evaluated and selected by the student, indicating the use of ChatGPT	**0.562**	0.475
ACC_V3	Creating the outline of the thesis with ChatGPT, which is then completed by the student	**0.548**	0.508
ACC_V4	Making a first draft with ChatGPT, which is then critically reviewed, corrected and completed by the student	0.195	**0.881**
ACC_V1	Using text generated by ChatGPT, without changes	-0.023	**0.829**
ACC_V2	Making summaries of the cited works	0.431	**0.649**
ACC_V5	Inclusion of sections generated by ChatGPT, citing the source and the prompt used	0.433	**0.516**

Source: authors' calculations using R, psych package.

The two identified factors of the perception of acceptability of LLMs accounted for 67.6% of the total variance of the data. The proportion of the total variance

explained by each factor is balanced: factor F1 accounted for 36.8% of the total variance and 54.4% of the common variance, while factor F2 accounted for 30.8% of the total variance and 45.6 % of the common variance.

The first factor, F1, is highly and positively correlated to items addressing the facilities of ChatGPT / LLMs in elaborating or creating the structure of a GW (grammar, spelling, references – identification and organising the bibliography, outline, ideas generation), so it represents the structural dimension, therefore we propose to be denoted F1 – 'structural dimension'.

The second factor, F2, shows the 'consultative' dimension of using ChatGPT / LLM in GW elaboration, because the items with which it correlates strongly and directly capture aspects of the inquiry and consultation side of LLMs for the creation of content.

Analysing the shape of the scores' distribution for the two factors, we noticed that while the first Factor 1 – the 'structural dimension' of acceptability exhibits a negative skewness, for the second F2 – 'consultative dimension', smaller scores have higher frequencies, indicating a positive skewness.

A more detailed analysis of the factor scores was performed using nonparametric tests, such as Kruskal-Wallis Rank Sum Test or Jonckheere-Terpstra Test, to analyse whether there are significant differences in the medians of Factor 1 and Factor 2 scores, given the grouping variables ATT-OVR and ATT-AD.

The results show that for each factor, the smallest median corresponds to the group of respondents who have a negative feeling about using LLMs in GWs, while the group of respondents with a positive feeling has the greatest median, far above the general median of each factor. A similar finding was obtained for the grouping variable ATT_AD – attitude regarding adoption or rejection of LLM in the context of GWs. Respondents who considered banning the use of ChatGPT when writing GWs have the smallest medians; the opposite is true for those who favoured the adoption and integration of LLMs. This suggests that the way professors perceive the use of LLMs may change over time, as it turns it into an integrated tool within the educational settings.

5. Conclusions

The ethical concerns surrounding GAI must be included in the academic institutions' code of conduct within a larger framework to increase the students and the faculty's resilience to technological disruptions. The OECD is considering this approach and suggests "legislation, ethics frameworks, technical AI standardisation, audits, model release, and access plans" (Lorenz et al., 2023) as answers to the dangers and difficulties GAI may face in the future. Given that the universities depend heavily on changes in various industries and business contexts, as well as future perspectives on skills and the job market that indicate a potentially increased impact of language-based GAI on jobs, with up to 32.8 percent of jobs in the International Standard Classification of Occupations potentially affected on a full scale, it is important to weigh both the advantages and disadvantages of this

technology. The academic top management should use the strategic vision and create inclusive, long-term ethical instruments driven by policy.

The study's results provide a detailed picture of the attitudes and perceptions of university teachers in Romania, with the role of coordinating GWs in economics. They are primarily familiar with LLM technology, but the limitations of the technology need to be more well-known. The predominant feeling of the subjects was concern, and the general attitude regarding GAI was favourable to integrating LLM tools in the educational process. Significant changes are expected in the assessment based on graduation works in 2-5 years. The study also includes perceptions regarding the degree of acceptability of some scenarios for using LLM-based tools and perceptions regarding the need for institutional and procedural measures. Correlations between variables were investigated, which has led to identifying several statistically significant relationships. An exploratory factor analysis has found two latent constructs which explain the perception of LLMs acceptability: a structural dimension and a consultative dimension for employing LLMs to assist with authoring GWs.

A potential line of further research may be linked to the revisiting of the Glendinning Scorecard for Academic Integrity (2017), by adding aspects related to a) the speed and effectiveness of the institutional response to GAI-related challenges, b) the adaptability of academic practices to integrate GAI ethically, c) the effectiveness of existing policies against GAI misuse, and d) the faculty and staff preparedness to handle GAI challenges.

The Romanian National Framework for AI 2023-2027 aims to support education for research, development, and innovation and the training of AI-specific skills in the broader population. In this context, the presented results may contribute to the academic environment's effort to respond effectively to the challenges created by the widespread use of tools based on GAI. Current data should inform discussions and actions taken in this context on perceptions and attitudes, such as that provided by this study. This type of assessment of attitudes builds upon a desiderate for a more informed, data-driven (or evidence-based) decision-making process, particularly in education, in which a long-term approach is needed.

References

[1] Adilov, N., Cline, J., Hanke, H., Kauffman, K., Meneau, L., Resendez, E., Singh, S., Slaubaugh, M., Suntornpithug, N. (2023), *ChatGPT and the Course Vulnerability Index*. *SSRN Electronic Journal*, https://doi.org/10.2139/ssrn.4433334.

[2] Amani, S., White, L., Balart, T., Arora, L., Shryock, K.J., Brumbelow, K., Watson, K.L. (2023), *Generative AI Perceptions: A Survey to Measure the Perceptions of Faculty, Staff, and Students on Generative AI Tools in Academia. arXiv preprint* arXiv:2304.14415.

[3] Chan, C.K.Y., Hu, W. (2023), *Students' Voices on Generative AI: Perceptions, Benefits, and Challenges in Higher Education. arXiv:*2305.00290.

[4] Chan, C.K.Y., Zhou, W. (2023), *Deconstructing Student Perceptions of Generative AI (GenAI) through an Expectancy Value Theory (EVT)-based Instrument. arXiv:*2305.01186.

[5] Covrig, M., Goia (Agoston), S.I., Igreț, R.Ș., Marinaș, C.V., Miron, A.D., Roman, M. (2023), *Students' Engagement and Motivation in Gamified Learning. Amfiteatru Economic*, 25 (Special issue No. 17), 1003-1023.

[6] 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 and Learning Practice*, 20(3), https://doi.org/10.53761/1.20.3.02.

[7] Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., Siemens, G. (2024), *Impact of AI assistance on student agency. Computers & Education*, 210, 104967.

[8] Djamba, K.J. (2022), *Cloud-Based Centralizing system for academic history, plagiarism prevention management in Higher Education Institution IN DRC: Benefit, Challenges. British Journal of Multidisciplinary and Advanced Studies*, 3(2), 142-152, https://doi.org/10.37745/bjmas.2022.0062.

[9] Glendinning, I. (2017), *Scorecard for academic integrity development: Benchmarks and evaluation of institutional strategies. Plagiarism across Europe and Beyond-Conference Proceedings,* 25-34, https://academicintegrity.eu/conference/proceedings/2017/Glendinning_Scorecard.pdf.

[10] Kenwright, B. (2024), *Is it the end of undergraduate dissertations?: Exploring the advantages and challenges of generative ai models in education.* In *Generative AI in teaching and learning*, 46-65, IGI Global.

[11] Korinek, A. (2023), *Generative AI for Economic Research: Use Cases and Implications for Economists. Journal of Economic Literature*. https://doi.org/10.1257/jel.20231736.

[12] Lai, C.Y., Cheung, K.Y., Chan, C.S. (2023), *Exploring the role of intrinsic motivation in ChatGPT adoption to support active learning: An extension of the technology acceptance model. Computers and Education: Artificial Intelligence*, 5, https://doi.org/10.1016/j.caeai.2023.100178.

[13] Li, L. Ma, Z., Fan, L., Lee, S., Yu, H., Hemphill, L. (2023), *ChatGPT in education: A discourse analysis of worries and concerns on social media. arXiv preprint arXiv:*2305.02201.

[14] Lorenz, P., Perset, K., Berryhill, J. (2023), *Initial policy considerations for generative artificial intelligence. OECD Artificial Intelligence Papers*, 1, OECD Publishing, Paris, https://doi.org/10.1787/fae2d1e6-en.

[15] Luo, J. (2024), *A critical review of GenAI policies in higher education assessment: a call to reconsider the "originality" of students' work. Assessment and Evaluation in Higher Education,* https://doi.org/10.1080/02602938.2024.2309963.

[16] Michel-Villarreal, R., Vilalta-Perdomo, E., Salinas-Navarro, D.E., Thierry-Aguilera, R. Gerardou, F.S. (2023), *Challenges and Opportunities of Generative AI for Higher Education as Explained by ChatGPT. Education Sciences*, [online] 13(9), 856, https://doi.org/10.3390/educsci13090856.

[17] Montenegro-Rueda, M., Fernández-Cerero, J., Fernández-Batanero, J.M., López-Meneses, E. (2023), *Impact of the Implementation of ChatGPT in Education: A Systematic Review*. *Computers*, https://doi.org/10.3390/computers12080153.

[18] 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, 100151.

[19] Năstasă, A., Maer Matei, M.M., Mocanu, C. (2023), *Artificial Intelligence: Friend or Foe? Experts' concerns on European AI Act. Economic Computation & Economic Cybernetics Studies & Research*, 57(3), 5-22.

[20] 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/10.53761/1.20.02.07.

[21] Revelle, W. (2023), *psych: Procedures for Psychological, Psychometric, and Personality Research.* [online] Northwestern University, Evanston, Illinois. R package version 2.3.6. Available at: https://CRAN.R-project.org/package=psych.

[22] Schmitt, B. (2023), *Transforming qualitative research in phygital settings: the role of generative AI. Qualitative Market Research: An International Journal*, https://doi.org/10.1108/qmr-08-2023-0107.

[23] Tuomi, I., Cachia, R., Villar-Onrubia, D. (2023), *On the futures of technology in education: Emerging trends and policy implications. Publications Office of the European Union,* https://doi. org/10, 2760, 079734.

[24] US Department of Health and Human Services, T.O. of R.I., n.d. *Definition of research misconduct*. https://ori.hhs.gov/definition-research-misconduct#:~:text=Research%20misconduct%20means%20fabrication%2C%20falsification,and%20recording%20or%20reporting%20them.

[25] Watkins, M. (2020), *A Step-by-Step Guide to Exploratory Factor Analysis with R and RStudio (1st ed.).* Routledge, https://doi.org/10.4324/9781003120001.

[26] Witte, R.S., Witte, J.S. (2017), *Statistics.* John Wiley & Sons.

[27] Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., Gašević, D. (2023), *Practical and Ethical Challenges of Large Language Models in Education: A Systematic Scoping Review,* [online] https://doi.org/10.1111/bjet.13370.