

# ChatGPT in Higher Education: A Risk Management Approach to Academic Integrity, Critical Thinking, and Workforce Readiness

Victor Chang<sup>1</sup>, Yasmin Ansari<sup>2</sup> and Mitra Arami<sup>3</sup>

<sup>1</sup>Department of Operations and Information Management, Aston Business School, Aston University, Birmingham, U.K.

<sup>2</sup>Centre for Innovation and Entrepreneurship Education, Aston Business School, Aston University, Birmingham, U.K.

<sup>3</sup>Faculty of Social Sciences, Northeastern University London, London, U.K.

**Keywords:** Artificial Intelligence, Big Data, Large Language Models, Higher Education, Risk Management.

**Abstract:** This paper critically explores the role of ChatGPT in higher education, with a particular emphasis on preserving the academic integrity of student assessments via a risk management paradigm. A literature analysis was conducted to understand existing strategies for addressing academic misconduct and the necessity of equipping students with skills suitable for an AI-driven workforce. The paper's unique contribution lies in its use of a risk management approach to enable educators to identify potential risks, devise mitigation strategies, and ultimately apply a proposed conceptual framework in educational environments. The paper concludes with identifying practical limitations and proposed future research areas, focusing on the uncertainties that emerge from the evolution of AI LLMs and the integration of comprehensive AI tools that pose new risks and opportunities.

## 1 INTRODUCTION

The rising prominence of Large Language Models (LLMs), such as ChatGPT, in the field of Artificial Intelligence (AI), has sparked concerns about potential academic misconduct in higher education, particularly in students' written assignments and exams. Beyond the academic setting, the broader implications of LLMs in the labor market, where they partially or fully automate certain job roles (including management and analysis of Big Data), have caused anxieties across various sectors. This raises the question of how universities can better equip their students to thrive in a rapidly changing workforce while maintaining academic integrity and critical thinking.

This paper focuses on risk management concerning AI LLMs in the educational context of Higher Education Institutions (HEIs). The primary objective is to explore risks associated with the use of tools like ChatGPT productively while avoiding inappropriate and unethical use that could compromise the value of degree programs. Additionally, we aim to address the risks associated with not embracing AI technologies within the curriculum, potentially hindering the development of

job-ready graduates equipped for a technologically optimized labor market.

Initially, existing literature and research surrounding ChatGPT was reviewed which provided insights on uses so far within Higher Education, risks related to ethics such as bias in the training data, implications on the labor market such as displacement of job roles, and the impact on future skills. While most of the literature discussed challenges and opportunities of adopting ChatGPT into practice, little to no risk management strategies were observed and as a result this paper seeks to address these in a practical manner.

By examining two key risk areas associated with ChatGPT in higher education, we seek to understand:

1. The risks of diminishing critical thinking through unethical use, leading to academic misconduct and compromising the value of degree programs.
2. The risks of not embracing AI technologies in the curriculum, potentially hindering the development of graduates prepared for an AI-driven workforce.

A risk assessment was produced based on identified risk areas including privacy risks, general ethical risks, academic integrity, unknown risks,

people/HR risks, and technical risks. Risk mitigation and management strategies are then discussed along with a conceptual framework that practitioners can apply to educational contexts.

Our analysis will culminate in practical limitations based on knowledge on AI LLMs to date and recommendations for further research.

This paper was produced by conducting a literature review and conducting a risk analysis followed by the recommendation to apply a risk management framework and its implications.

## 2 CONTEXT

### 2.1 Introduction to AI LLMs: Functionality and Limitations

Since its inception in November 2022, ChatGPT, a 'weak' or 'narrow' AI, has surpassed 100 million users by January 2023 (Wu et al., 2023). This type of AI is limited to text generation, relying on machine learning from vast training data sets. It can generate human-like dialogue but lacks the ability to provide complex, effectively contextualized examples and explanations (Ausat et al., 2023; Toews, 2021). The novelty of AI LLMs like ChatGPT limits the existing research on risks, including misuse in higher education.

### 2.2 Applications of AI LLMs in Education

ChatGPT can be used to automate tasks, produce written content, give feedback, analyse and synthesise Big Data for research, and collaborate with learners, personalizing the learning experience and making suggestions for improvements (Cotton et al., 2023). Some identified uses include:

Designing classroom exercises, brainstorming, and customizing materials to academic level (OpenAI, 2023) for teachers.

Supporting quick understanding of main text points and organizing thoughts for writing (Kasneci et al., 2023) for academics and students.

### 2.3 Academic Integrity in Higher Education and the Response of Institutions

The growth of AI in Higher Education has raised concerns regarding academic integrity.

Literature Perspective: Crompton & Burke's (2023) (Crompton & Burke, 2023) systematic review

identified key usages and observed a recurring theme of concern for integrity. It reviewed 138 existing articles globally that were used for AIED (Artificial Intelligence in Education) include assessment, prediction, assistance, intelligent tutoring system, and managing learning (Crompton & Burke, 2023).

### 2.4 Labour Market and Skills Development

The demand for creativity and critical thinking in utilizing and managing AI systems are forecasted to grow (Universities UK, 2023).

Employers' Perspective: Universities UK (Universities UK, 2023) conducted a study that found that UK employers will require up to 11million new graduates by 2035. It was emphasised that there is demand for high levels of creativity and critical thinking skills, specifically in utilising and managing AI systems (Universities UK, 2023). The study forecasts a trend towards division of labour, rather than fully automated roles, therefore requiring high level critical skills from new entrants to the workforce. Employers increasingly also require staff that can use Big Data sets for analysis and have the skills to apply data management software.

## 3 RISK ASSESSMENT AND TREATMENT

Considering the intricate relationship between AI Large Language Models (AI LLMs) such as ChatGPT, the academic landscape, and the evolving needs of the labour market, an extensive risk analysis is required (Aven, 2008). The insights and concerns expressed by literature and the understanding of the state of AI in education and the job market culminate in the identification of key risk areas.

Privacy Risks (operational): Concerns student data and personal information, especially in an online environment using AI tools like ChatGPT. This includes AI access to Big Data and implications.

Ethical Risks (operational, strategic, and hazard/legal): Balancing the innovative uses of AI with ethical concerns such as fair use, bias, and accessibility.

Reputational risks (strategic):

Academic Integrity Risks (operational): Issues related to plagiarism and maintaining the authenticity and value of academic work.

People/HR Risks (operational): The potential impact on staff training, student preparation for a changing job market, and interpersonal relations.

Technical Risks (operational): The possibility of malfunction, security breaches, Big Data leaks, or other technical issues with AI tools.

Unknown Risks (strategic): The potential unforeseen consequences of rapidly evolving AI technology.

### 3.1 Privacy Risks

The widespread use of ChatGPT and other AI LLMs raises legitimate concerns about privacy and data security. These concerns extend to students, faculty, and the wider institution. Robust data protection measures must be implemented to ensure that personal and academic information remains secure.

Risk treatment: Students and staff must be trained in the risks of inputting personal and sensitive information. For example, IT services could expand their cybersecurity training or communicate best practices for online safety with AI LLMs.

### 3.2 Ethical and Academic Integrity

The challenge is twofold. AI detectors like OpenAI's Text Classifier and Turnitin's AI checker seek to identify AI-generated work, while institutions like the University of Nottingham (2023) offer guidance to students about distinguishing AI content from authentic critical writing. However, the sophistication of tools like ChatGPT, especially with its ability to mimic a student's writing style, makes detection challenging. The advanced nature of GPT-4 further blurs the line between AI and human-generated content (Wu et al., 2023). Ethical concerns arise when relying heavily on AI detectors, given potential inaccuracies. Overreliance on AI LLMs could also lead to student bias (Ferrara, 2023) and misinformation, hindering genuine learning.

Risk treatment: Some solutions are, of course, beyond the scope of individual institutions and what can immediately be controlled. For example, digital watermarking (Rouse, 2023) can be used to corroborate and verify the legitimacy of information, as well as track source data, and thereby prevent plagiarism and copying. There is also software like ProctorU and Examity (both US based) that monitor computer-based examinations through screen capture and audio monitoring.

### 3.3 Reputational Risks

Reputational damage is a risk that has both financial and strategic implications. Policies and ethical frameworks are being drafted to mitigate these challenges, acknowledging both the potential for misuse and the risk of lagging behind technological trends.

Risk treatment: An institutional commitment to responsible AI usage, ongoing risk assessment, collaboration with other HEIs, and adherence to national and international ethical standards can position the university favourably.

### 3.4 People/HR

While there's debate about AI replacing human jobs, the consensus leans toward AI complementing human tasks (Zarifhonarvar, 2023). The limited capacity of 'weak AIs' only allows excellence in a single task (Lane & Williams, 2023). Additionally, this risk area can be further broken down:

Deskilling of the Workforce: The automation of complex tasks by AI could lead to a reduction in stimulating work (Lane & Williams, 2023), learning opportunities, and worker autonomy. An intentional focus on incorporating AI within curricula will help prepare students for a future where AI complements human tasks.

Loss of Competitive Edge: As organizations like Coca Cola and Bain ((Marr, 2023) leverage AI for personalized marketing, failing to embed AI in education would be a missed opportunity. More examples and context here could elucidate the importance of staying abreast of technological advancements.

Risk treatment: Educators must be proactive in integrating AI responsibly within curricula, equipping students with skills that align with industry needs. This can involve working with employers to inform classroom activities and assessments to provide the necessary training. It also involves ongoing monitoring and review of labour market developments and trends specifically in relation to the usage of AI LLMs.

### 3.5 Technical Risks

Technical malfunctions or security breaches could disrupt learning or compromise sensitive information. A robust infrastructure and regular monitoring are required to prevent and respond to such incidents.

Risk treatment: A robust infrastructure and regular monitoring are required to prevent and respond to

incidents. Cybersecurity considerations must be considered when using AI LLMs where data policies are tenuous (see OpenAI’s privacy policy) or in development. At the time of writing, OpenAI states it stores prompts and generated text for 30 days for ‘learning purposes’ unless deliberately turned off by the user, but there is no clarity on what happens with the data and if there is any danger of breach. Students and staff must be trained in the risks of inputting information that could be personal and sensitive (Gal, 2023), for example student work, CVs, or any other text that contains personal information.

### 3.6 Unknown Risks

Strategically, the rapid development and adoption of AI technologies mean that there may be unforeseen risks requiring adaptive management strategies. Further risk areas, opportunities and implications of AI technologies are unknown as they continue to develop in application and sophistication.

Risk treatment: Adaptive management strategies must be prepared for unforeseen risks, and HEIs must remain vigilant to the shifting landscape of AI.

## 4 FRAMEWORK AND APPLICATION

A risk management framework was utilised to produce this research and guidance on how to apply it to HEIs is recommended below. The framework, adapted from the ISO 31000 guidelines (2018) (ISO Org, 2018) is a tool that HEIs can apply in managing risk related to AI LLMs.

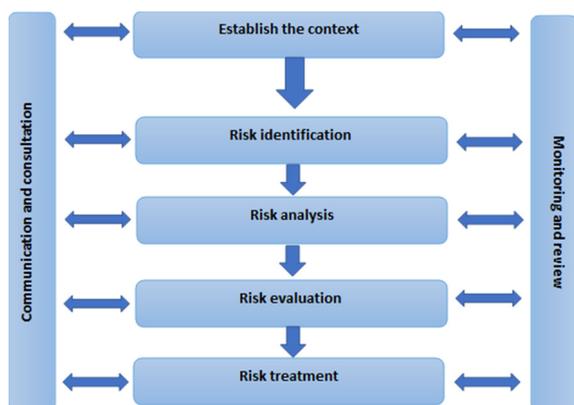


Figure 1: ISO 31000 Risk Management Framework.

### 4.1 Establishing the Context

In this paper, university student assessments and student aptitude in developing skills for post study employment were considered in relation to the growing usage of AI LLMs. This is an issue for HE managers and educators to consider regarding the impact AI has on delivery of their curriculum and assessment.

Practitioners may therefore wish to review what is happening in their institutions, what policies are in development, what advice is being provided for educators and managers.

Government reviews such as the UK’s Department for Education’s paper “Generative Artificial Intelligence in Education” (2023) (UK Government, 2023) and the recent Russel Group’s “New principles on use of AI in education” (2023) (Russel Group, 2023) could direct practitioners to gather best practice and formulate a strategy towards managing risks. These present a starting point, while collaboration with stakeholders and gathering insights into current usage can further determine existing risks and opportunities.

### 4.2 Risk Identification

The risks highlighted in this study were synthesized from a combination of a literature review—with a notable emphasis on ethical concerns like bias—and an overarching examination of AI’s imprint on the HE sector.

If educators can choose to implement ChatGPT as part of assessed work and/or classroom-based exercises for skill development of students, considering risks and their impact will help mitigate some of the issues identified. Involving students and a breadth of colleagues establishes how risks are perceived and their level of impact.

Key areas of exploration might involve understanding perceptions around ChatGPT’s inherent ethical risks, strategizing student education on these issues, and discerning if there are institution-specific risks that might have been overlooked.

For universities seeking to comprehensively identify risks, there are systematic steps to be considered:

- Stakeholder Feedback: Engage with a diverse set of stakeholders within the educational ecosystem. This includes students, faculty, IT personnel, and administrative staff. Their diverse experiences and perspectives can shed light on potential vulnerabilities.

- External Collaborations: Partner with external AI experts or institutions that have already integrated AI technologies. Their experiences can provide valuable insights into potential challenges.
- Continuous Learning Workshops: Organise sessions where the latest findings, research, and anomalies related to AI LLMs like ChatGPT are discussed with staff. This creates a dynamic environment where new risks can be identified in real-time and feed into the ongoing review process outline below.
- Technology Audits: Periodically review the technology's performance and integration within the educational process. Such audits can identify any misalignments or areas of potential concern. This relates to privacy and technology risks identified in our risk assessment (see appendix).

### 4.3 Analysis and Evaluation

To evaluate the impact of risks identified, they need to be fully analysed and understood. The authors recognise further conversations and review are necessary (see ongoing monitoring and review). This means involving students themselves in the risk analysis process, along with educators and assessment methods. For the second risk area, labour market skills, working closely with employers and enabling them to inform curriculum developments in relation to using AI tools can support students for employment.

Risk evaluation involves categorising the impact of risks, this can be done using a simple RAG (Red, Amber, Green) rating, or a risk assessment matrix with numerical scoring. For broader strategic developments, managers may wish to use scenario planning methodologies to build on risk areas and consider outcomes if risks are realised. This can enrich the risk analysis and evaluation process and provide broader mitigation strategies.

### 4.4 Risk Treatment

Risk mitigation and management measures are finally considered to address the identified risks. Following the above RAG rating or risk matrix, designing specific strategies for each risk can include mitigation, transfer, acceptance, or avoidance – this will depend on the severity of the risk. High risk areas require implementation of mitigation strategies to reduce the impact these risks have and can then be monitored. It is of the authors' view that avoidance, or banning AI LLMs outright, is an ineffective strategy considering new tools are on the rise, many of which will be fully

integrated into what we already do. For example, word processing and searching the web already contain GPT, Bard or similar (Yu, 2023).

Building on the above risk identification steps, ongoing stakeholder management (part of communication and consultation below) will leverage a breadth of expertise from students, external collaborators, managers, educators, and technical staff.

Finally, given the dynamic nature of AI and unknown risks related to future developments, ongoing education is imperative. This can build on continuous learning workshops mentioned above, development of working committees (internally and externally) and regularly reviewing policy communications for HE regulators and the government.

Proactive risk mitigation strategies include:

- Assigning a specific task force or working group whose sole responsibility is to track and respond to developments in AI LLMs and their implementation.
- If the institution is already incorporating AI LLMs, like ChatGPT (including integrated versions like Bing AI), then digital service teams should support the function and training of staff and monitor and treat technical faults.
- To ensure students are building the right skills to use ChatGPT, putting on 'how to use ChatGPT' workshops within or outside the curriculum can help minimise students using the tool unethically but encourage appropriate application. Institutions may also want to provide learning materials and resources that students (and staff) can access independently.

### 4.5 Ongoing Communication and Consultation

Communication and consultation must occur at all stages of the risk management framework, to effectively apply risk management strategies. Practitioners can benefit from wider input of colleagues and other stakeholders (like students and employers) to fully understand and manage the risks. It is also likely that this process will highlight areas of concern that are less obvious.

### 4.6 Ongoing Monitoring and Review

Risks can be volatile, in the sense that their risk level and impact can change depending on how variables are affected. For example, as AI LLMs become more sophisticated and more creative, it becomes more

challenging to detect their use in assessed work. Regular reviews and adjustments to strategies are crucial to maintain relevancy.

#### 4.7 Recommendations for Higher Education Institutions Wishing to Adopt AI LLM Tools

Firstly, following institutional guidelines and policies ensures appropriate and ethical integration of any AI tools. With training, both staff and students can use AI LLMs to automate tasks and support their learning journeys. For example

- Students can use ChatGPT to brainstorm for assignments or work through assignment tasks by using iterative prompting. In a reflective practise assignment, a student can prompt "Using the Gibbs reflective model, help me reflect on a teamwork task". ChatGPT will work through the stages of the model and the student can request and respond to feedback and questions. Through this the onus is on the student to generate the content for their work, but they are receiving continuous feedback and support in the process. This approach can also be beneficial to module tutors who have limited time for office hours dedicated to assignment support.
- Educators can use ChatGPT to create classroom-based exercises or make existing exercises more interesting. ChatGPT has the capacity to produce content based on tone and style such as in the form of songs, poetry, recipes, in the 'voice' of a well-known person, as a game or puzzle, and this creates boundless possibilities for creating new tasks.

## 5 CONCLUSION

Limitations of this research include the novelty of the subject area, meaning there are still a lot of unknown and under researched areas in relation to AI LLMs. The quality of existing work is, at times, limited to author opinions and web articles that discuss ChatGPT's implications, which can taint an objective and unbiased view (Neumann et al., 2023). Further research may therefore explore how university policies impact ethical use of AI LLMs and consider developments in pedagogy like assessment methodology. How AI LLMs manage Big Data to produce content and analysis requires ongoing review if it is to be used for educational and academic

research purposes given its impact on privacy, ethical and technical risks.

Practical limitations to applying risk management strategies include resource availability such as setting up working groups to monitor developments and collaborate across stakeholders for ongoing monitoring and review of the state of AI development.

To conclude, the state of AI in Higher Education is a rapidly growing area of concern and cause for research and exploration. Managing risks can help deal with uncertainties and enable HEIs to be prepared as tools grow in sophistication and use.

## ACKNOWLEDGMENT

This work is partly supported by VC Research (VCR 0000230) for Prof Chang.

## REFERENCES

- Wu, T, et al, (2023). A Brief Overview of ChatGPT: The History, Status Quo and Potential Future Development. *Journal of Automatica Sinica*, Vol 10, No 5. <https://doi.org/10.1109/JAS.2023.123618>
- Ausat et al (2023) Can ChatGPT replace the role of the teacher in the classroom: a fundamental analysis. *Journal on Education*, Vol 05, No4, pp 16100-16106 <https://doi.org/10.31004/joe.v5i4.2745>
- Toews, R (2021) What AI still can't do. *Forbes*. Available online: <https://www.forbes.com/sites/robtoews/2021/06/01/what-artificial-intelligence-still-cant-do/?sh=5845f8a166f6> [Accessed June 2023]
- Cotton et al (2023) Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International*. <https://doi.org/10.1080/14703297.2023.2190148>
- Kasneji, E et al, (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, Vol 103. <https://doi.org/10.1016/j.lindif.2023.102274>
- Crompton & Burke (2023). Artificial Intelligence in Higher Education: The State of the Field. *International Journal of Educational Technology in Higher Education*. Vol 20, pp1-22. <https://doi.org/10.1186/s41239-023-00392-8>
- Mearian, L (2023) Schools look to ban ChatGPT, students use it anyway. *Computerworld*. Available online: <https://www.computerworld.com/article/3694195/schools-look-to-ban-chatgpt-students-use-it-anyway.html#:~:text=Several%20leading%20universities%20in%20the,is%20a%20form%20of%20cheating.%E2%80%9D> [Accessed 22nd June 2023]
- Universities Uk (2023) Creativity and critical thinking craved as UK businesses need 11 million new graduates. Available online:

- <https://www.universitiesuk.ac.uk/latest/news/creativity-and-critical-thinking-craved> [Accessed August 2023]
- Aven, T (2008). Risk analysis: assessing uncertainties beyond expected values and probabilities. Wiley. Chichester.
- Ferrara, E (2023), Should ChatGPT be biased? Challenges and Risks of Bias in Large Language Models. Machine Learning with Applications. Available online: <https://arxiv.org/abs/2304.03738> [Accessed June 2023]
- Rouse, M (2011) What is digital watermarking? Technopedia. Available online: <https://www.techopedia.com/definition/23373/digital-watermark> [Accessed June 2023]
- Zarifhonarvar, A (2023). Economics of ChatGPT: A Labor Market View on the Occupational Impact of Artificial Intelligence. Indiana University Bloomington. Available at: <https://ssrn.com/abstract=4350925> [Accessed June 2023]
- Lane, M & Williams, M., 2023. Defining and classifying AI in the workplace, OECD Social, Employment and Migration Working Papers, No.290, Available online: [www.oecd.org/els/workingpapers](http://www.oecd.org/els/workingpapers) [Accessed June 2023]
- Marr, B 2023, 10 amazing real world examples of how companies are using ChatGPT. Available online: <https://www.forbes.com/sites/bernardmarr/2023/05/30/10-amazing-real-world-examples-of-how-companies-are-using-chatgpt-in-2023/> [Accessed June 2023]
- Gal, U. (2023) Chat GPT is a privacy nightmare. If you've ever posted online, you ought to be concerned. Available online: <https://theconversation.com/chatgpt-is-a-data-privacy-nightmare-if-youve-ever-posted-online-you-ought-to-be-concerned-199283> [Accessed June 2023]
- ISO Org (2018) ISO31000 Risk Management Guidelines. Available online: <https://www.iso.org/obp/ui/#iso:std:iso:31000:ed-2:v1:en> [Accessed August 2023]
- Department for Education, UK Government (2023). Generative Artificial Intelligence in Education. Available online: <https://www.gov.uk/government/publications/generative-artificial-intelligence-in-education> [Accessed August 2023]
- Russel Group (2023). New principles on use of AI in education. Available online: <https://russellgroup.ac.uk/news/new-principles-on-use-of-ai-in-education/> [Accessed August 2023]
- Yu, H (2023) Reflection on whether ChatGPT should be banned by academia from the perspective of education and teaching. *Frontiers in Psychology*. 14: 1181712 DOI: 10.3389/fpsyg.2023.1181712
- Neumann et al (2023) "We need to talk about ChatGPT": The Future of AI and Higher Education. [Available online: <https://serwiss.bib.hs-hannover.de/frontdoor/index/index/docId/2467> ] Accessed August 2023