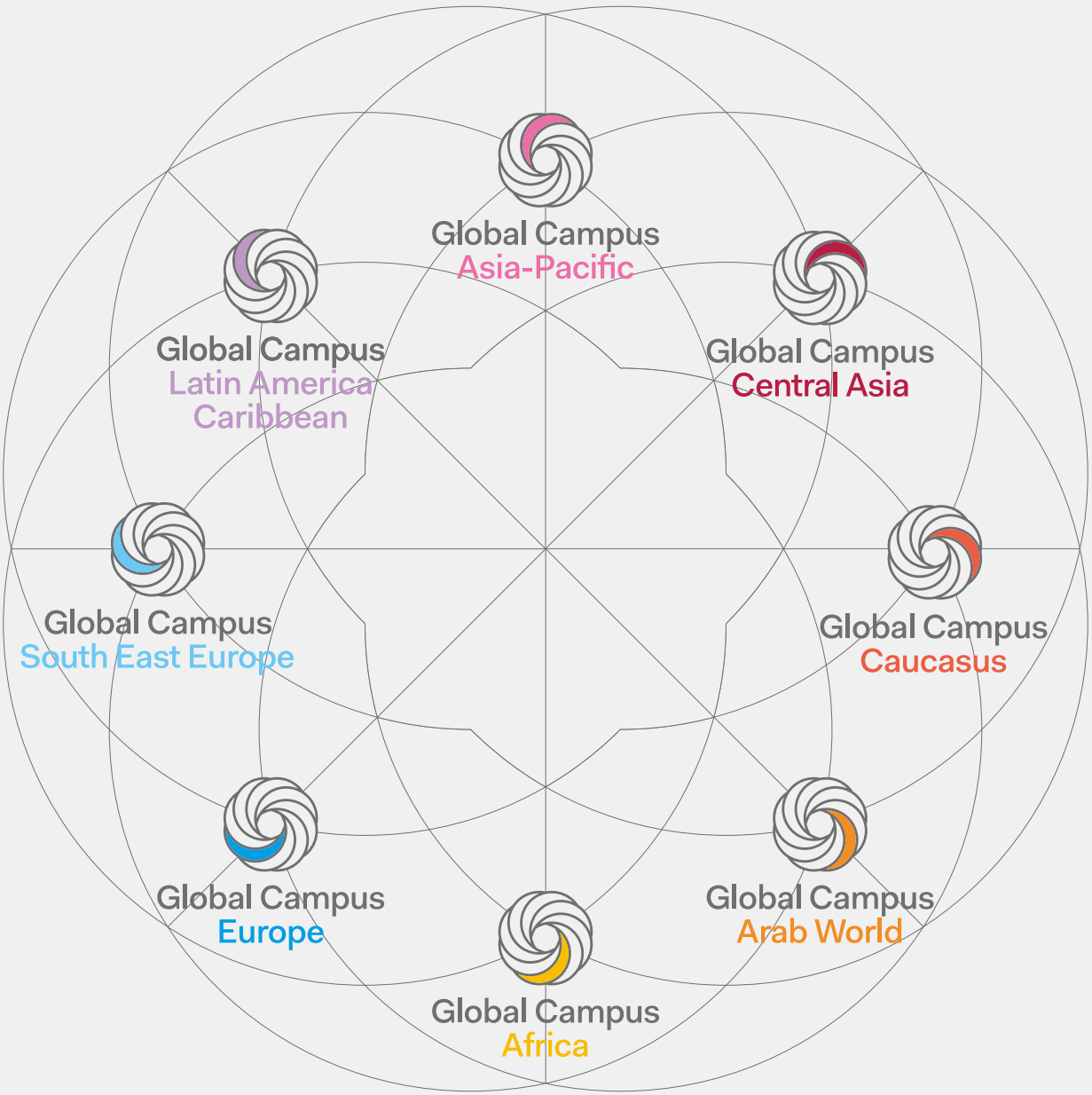
Global Campus Europe

Policy Briefs 2024

Desara Dushi

ChatGPT in Classrooms: A Double-Edged Sword

Supported by


Global Campus
Asia-Pacific
Global Campus
Latin America
Caribbean
Global Campus
Central Asia
Global Campus
South East Europe
Global Campus
Caucasus
Global Campus
Europe
Global Campus
Arab World
Global Campus
Africa

Global Campus of Human Rights

This policy brief is part of the **6th edition of the Global Campus Policy Observatory**, which revolves around the research project on **'The digitalisation of education systems and its impact on human rights, with particular attention to the right to education'**, which was conceptualized and is led by GC Research Manager Dr. Chiara Altafin and which involves a team of seven policy analysts selected among alumni of GC regional programmes, namely Reda Benkhadra (GC Africa), Olga Lucía Camacho Gutierrez (GC Latin America and the Caribbean), Dr. Desara Dushi (GC Europe), Dr. Jean Linis-Dinco (GC Asia-Pacific), Goharik Tigranyan (GC Caucasus), Aida Traidi (GC Arab World), and Dr. Gergana Tzvetkova (GC South East Europe). Research outputs include workshop presentations, policy briefs, advocacy plans, and digital tools (infographics, webinars) developed in cooperation with the GC E-Learning Department.

This policy brief is written by **Desara Dushi**, legal expert with a Double PhD degree in Law, Science and Technology (2019) from the University of Bologna and the University of Luxembourg, and a Master's degree in Democracy and Human Rights in South East Europe (ERMA) (2014) from the University of Sarajevo and the University of Bologna. She is currently a Senior and Postdoctoral Researcher at the Law, Science, Technology & Society (LSTS) Research Group of the Vrije Universiteit Brussel (VUB). Contact: Desara.Dushi@vub.be

This policy brief was finalized in May 2024, and its findings do not account for any legal and technological developments that may have occurred in a subsequent period.

This policy brief is produced with the financial assistance of the European Union and as part of the Global Campus of Human Rights. The contents of this document are the sole responsibility of the authors and can under no circumstances be regarded as reflecting the position of the European Union or of the Global Campus of Human Rights.

This policy brief is realized with the support of the Unit for Analysis, Policy Planning, Statistics and Historical Documentation - Directorate General for Public and Cultural Diplomacy of the Italian Ministry of Foreign Affairs and International Cooperation, in accordance with Article 23 – bis of the Decree of the President of the Italian Republic 18/1967. The views expressed in this policy brief are solely those of the authors and do not necessarily reflect the views of the Ministry of Foreign Affairs and International Cooperation.

Table of Contents

05 **Executive summary**

06 **Introduction**

07 **Problem description**

09 **Rationale for action**

09 **Policy options**

09 **A banning approach: The case of France**

09 **Examples of self-regulatory approach**
The case of The Netherlands
The case of Poland

10 **An example of multiple approaches: The case of Italy**

11 **A liberal approach: The case of Belgium**

11 **Policy recommendations**

13 **Conclusion**

14 **References**

ChatGPT in Classrooms: A Double-Edged Sword

Desara Dushi [1]

Executive summary

This policy brief analyses the promising benefits and negative impact of generative artificial intelligence (GenAI) on higher education, with a focus on ChatGPT. It revolves around questions such as: To what extent should digitalisation become part of education? What are the threats to the right to education and data protection in times of digital technologies? How can digitalisation in education be regulated, and by whom? What should higher education do to address the increasing dependence on technology while safeguarding data autonomy and public values? Can current legislative and regulatory frameworks deliver enough appropriate and desirable guidance?

The policy brief starts with an analysis of the initiatives at the European Union (EU) and Council of Europe (CoE) level. Focusing on the European higher education landscape, initiatives from several EU countries are considered. These analyses underscore the importance of inclusive decision-making processes that involve educators, students and other relevant stakeholders. This provides for a better understanding of the implications of GenAI and ensures that its integration into education systems aligns with human rights principles and standards. To make the tool serve educational purposes, the providers of ChatGPT and other similar technologies must collaborate with educational institutions. The objective should be maximising the potential of GenAI while remaining aware of the possibilities for foreseeable misuse.

1 The author thanks Dr. Chiara Altafin, Research Manager at the Global Campus of Human Rights in Venice, Dr. María López Belloso, Lecturer and Researcher at the Social and Human Sciences Faculty at the University of Deusto, and Prof. Therese Murphy, Professor of Law at Queen's University Belfast and EMA Chairperson, for their valuable and constructive feedback as received in the context of the GC Policy Observatory workshop 'The digitalisation of education systems and its impact on human rights, with particular attention to the right to education', held in Venice on 26 March 2024.

Introduction

GenAI tools like ChatGPT will soon be omnipresent. Since its public release at the end of 2022, ChatGPT – the artificial intelligence (AI) chatbot developed by OpenAI – has experienced rapid growth and widespread adoption. As technology advances to bring us new opportunities, how we educate children and students and prepare teachers must also advance to meet these new opportunities. When education falls behind technological developments, the technology rather than educators defines what counts as educational opportunity.

ChatGPT has revolutionised the educational landscape, redefining how students and educators engage with knowledge. This sophisticated language model offers dynamic learning experiences, personalised assistance and fosters interactive education across various subjects and levels.
In the realm of education, ChatGPT serves as a multifaceted tool, providing real-time support for assignments, clarifying intricate concepts and aiding research efforts. Its natural language processing abilities facilitate a more engaging learning environment, allowing students to ask questions, seek explanations and receive customised feedback. Educators, in turn, utilise ChatGPT to design interactive lessons and create content, tailoring their teaching approaches to diverse learning styles.
However, as with any technological advancement, ChatGPT in education is not without challenges. One significant concern is the potential reinforcement of educational inequalities. Access to this technology might not be uniform across all educational institutions, leading to a digital divide and exacerbating existing disparities.
Moreover, the reliance on AI in education raises ethical questions, particularly concerning privacy and data security. The extensive collection of student data poses risks and maintaining transparency and ethical use of this information is crucial to safeguarding individuals' rights. Additionally, the impersonal nature of AI interaction might impact the development of crucial human skills, such as empathy and emotional intelligence, which are integral to the educational experience.
Furthermore, there are broader implications related to human rights when considering the global implementation of AI in education. Issues such as privacy violations, discrimination and biases within AI algorithms may disproportionately affect marginalised communities, potentially exacerbating social inequalities.
In conclusion, while ChatGPT offers immense potential for educational advancement, a balanced and thoughtful approach is essential to address associated challenges and ensure that its integration aligns with principles of equity, privacy, and human rights.

The text in this box was generated by ChatGPT by using the prompts 'write a 200-word introduction on ChatGPT in education' and then 'make it better including disadvantages of ChatGPT in education and its impact to human rights'.

The launch of ChatGPT marked a significant milestone in the field of AI. The language model garnered widespread attention for its ability to generate human-like text and engage in natural language conversations. The launch generated a spectrum of reactions, from excitement about its transformative potential to thoughtful considerations regarding ethical implications. In education, ChatGPT's potential to provide instant support, clarify complex concepts and offer personalised assistance has resonated with educators and students alike, sparking conversations about its integration into education systems. However, as with any technological advancement, concerns emerged about the responsible use of ChatGPT, addressing issues such as bias, legal and ethical considerations, and potential misuse.

Since 2021, the EU has undertaken steps to explicitly regulate AI, with a proposal for an Artificial Intelligence Act (AI Act), whose final text was adopted by the European Parliament in March 2024. In the final version of the AI Act, AI systems are classified according to their perceived level of risk, from low to unacceptable. The different risk levels mean more or less legal obligations

on the providers and deployers of AI systems.

GenAI was introduced to the public in November 2022, and thus provisions regarding this kind of technology were added later on in the AI Act. A tiered approach was introduced for regulating the so-called 'foundation models' referred in the AI Act as general-purpose AI (GPAI) systems/models, aimed at ensuring a robust protection of fundamental rights, democracy and the rule of law.

Systems falling in the first tier will have to comply with transparency requirements including disclosing that the content was generated by AI, and publishing summaries of content used for training and comply with copyright law and related rights, under Article 53 of the AI Act (European Parliament 2023). More stringent obligations apply for high-impact general-purpose AI models that pose a systemic risk,[2] such as ChatGPT 4. These models will have to undergo thorough evaluations and tests, and any serious incidents will have to be reported to the European Commission, under Article 55 of the AI Act (European Parliament 2023; Hainsdorf et al. 2023).

In addition, the General Data Protection Regulation (GDPR) plays a pivotal role when it comes to technologies handling personal data. Since personal data can be processed in the context of the use of ChatGPT, the requirements of the GDPR apply. Tech companies are obligated to establish a legal basis to collect and use personal data (Article 6), be transparent about their use of people's personal data (Article 7), keep personal data accurate (Article 5), ensure the right to access, correction, erasure (Articles 15-19) and the right to object to the use of their personal data (Article 21). A problematic area of ChatGPT regards compliance with the principle of purpose limitation, as ChatGPT uses any kind of data, including personal data for its training purposes.

From a policy perspective, the Digital Education Action Plan 2021-2027 is an EU policy initiative that focuses on establishing high-quality, inclusive and accessible digital education in Europe, and aims to support the adaptation of the education and training systems of member states to the digital age (European Commission 2021-2027). Among its key priorities are enhancing digital skills and competencies for teachers and educators and fostering the development of a high-performing digital education ecosystem.

EU institutions have issued internal guidelines for their staff on using and interacting with online generative AI models. As such, the European Commission warns staff of the risks and limitations of GenAI models, demanding them to be cautious in their use, and bans their use for critical tasks (Bertuzzi 2023).

Similarly, at the CoE level, the Draft Framework Convention on Artificial Intelligence, Human Rights, Democracy, and the Rule of Law sets forth general common principles that each party shall implement concerning AI systems (CoE 2023). It gives freedom to its member states to decide upon banning any AI systems that they may deem incompatible with human rights, the functioning of democracy and the rule of law.

The CoE has also developed an Education Strategy, with the mission of improving the quality and accessibility of education. One of the three main pillars of the strategy is 'enhancing education through a human-rights based digital transformation' (CoE 2024-2030). In this direction, it has five cross-cutting principles: flexibility in curriculum and programme design; learner autonomy; professional development; inclusion and participation; and democratic and participatory governance of education systems and institutions.

As GenAI is in its first steps, still followed by many uncertainties, it will be interesting to see whether ChatGPT and other similar GenAI tools could be leveraged to revolutionise the learning experience and what the new role of educators could be.

2 Article 3(65) of the AI Act defines 'systemic risk' as 'a risk that is specific to the high-impact capabilities of general-purpose AI models, having a significant impact on the Union market'.

Problem description

GenAI poses many challenges to education systems. Depending on the way of using it, it can replace learning or enhance it. It can cause digital dependency but can also lead to digital savviness. It poses a threat to critical thinking but also offers an opportunity for developing critical thinking.

ChatGPT is a Large Language Model (LLM), a chatbot, which is trained on enormous amounts of online datasets to generate text and quickly provide complex answers to open queries, based on the prompts that users put in. It can 'remember' what the user has said earlier in the conversation, so it can also answer

follow-up queries. As any other LLM, it functions on a next-word prediction basis, which means that it is optimised to generate plausible and pleasing human-like responses, by predicting sequences of words used in their training data, just like a stochastic parrot (Bendler et al. 2021). It cannot reason on its output (Bender & Koller 2020; European Commission 2021-2027; Mitchell 2023), it does not have the ability to understand the meaning of the language it generates (Bender & Koller 2020) and hence it does not know if the information it provides is accurate. Therefore, the answers it provides are not necessarily true or accurate. When it has no information related to the prompt, it will hallucinate (Dahl et al. 2024).

When it comes to education, on the one hand, GenAI, such as ChatGPT, offers several **advantages**. It has the potential to provide instant support both for teaching and learning, including for individuals with special needs, clarify complex concepts and offer personalised assistance. By using it, students and educators alike can save time on simple tasks, such as writing summaries, correcting texts in terms of grammar and programming code. It can reduce teaching workload and help in answering complex questions and brainstorming as well. If intelligently used, it can also help students progress rapidly by serving as a learning assistant or tutor. Moreover, these tools can enable teachers to create content that can make their lectures more dynamic and engaging, to swiftly adapt the course structure to meet students' needs more efficiently than by manual processes, significantly reducing the time required for these adjustments.

On the other hand, the use of ChatGPT also raises several **concerns**. While saving time on more important tasks, educators but also students risk unlearning to do things (Castigli 2023). Too much reliance on GenAI tools risks students losing some important skills such as writing, researching, summarising and even creativity. It could also lead to automating poor pedagogical practices, disempowering teachers and students alike, and undermining the right to quality education (CoE 2024-2030).

Another main concern is its impact on academic integrity (Sullivan, Kelly & McLaughlan 2023), which relates to its use in cheating, creating assignments and essays, or even answering simple homework questions, to plagiarism. It also risks undermining humanistic values like reason, morality and ethics, favouring easily measurable competencies (CoE 2024-2030).

ChatGPT is not based on ethical principles and is unable to distinguish between right and wrong, or true and false (UNESCO 2023). Any cognitive bias existing in the information, from which the tool learns, is transferred into its output. They reflect and amplify the biases of their training data, including racist, sexist, extremist, and other harmful language (Zhang et al. 2022: 119; Bendler et al. 2021). This is an important shortcoming considering that any input provided to a GenAI tool is later used for feeding future generated output. Thus, bias in the input is transferred to bias in the output, leading to wrong, inaccurate or biased answers. It is therefore essential to critically assess and compare its results with other sources (UNESCO 2023).

From a data protection perspective, the processing of all input by users to feed the model poses risks to personal data protection and goes against the principle of purpose limitation of personal data processing as enshrined in the GDPR. Italy was the first country to raise the privacy concerns of ChatGPT by temporarily blocking the tool. Main concerns are the lack of a legal basis for the collection and storage of personal data used to train the tool, and the tool's inability to protect children from being exposed to age-inappropriate responses.

Ultimately, ChatGPT was created by a private profit-driven company. When integrating technologies developed by profit-driven companies, ensuring equitable access to high-quality digital learning environments becomes challenging. There is a need to scrutinise how these endeavours inadvertently support profit-driven motives without necessarily enhancing access to education or its quality. A precautionary approach should be taken towards the involvement of private entities in education, paying particular attention to tools collecting (personal) data for commercial purposes (UNESCO 2023).

The launch of ChatGPT caught higher education unprepared. Initially, only a few educational institutions took steps towards regulating the use of ChatGPT. A few, including Sciences Po University, banned its use completely. While many others have only recently started to react by adopting rules and policies, as analysed in the following sections.

Rationale for action

The introduction of ChatGPT in education (and in any other sector for that matter) is very recent, given its first introduction to the public only in 2022. Consequently, there are presently limited or no policy measures enacted at the national level. While some universities have initiated the development of guidelines for their academic staff, there have been sporadic public reactions from a handful of universities or professors. Hence, conducting a comparison between various policies and identifying best practices is challenging at the time of writing. Nonetheless, in this section, we have pinpointed noteworthy reactions that could potentially serve as examples or reference points for other institutions. With its rapid growth, the conversation about banning ChatGPT or not will soon become irrelevant. While there are a few ban cases such as in Sciences Po University and L'Università della Svizzera italiana (USI), soon the use of ChatGPT in education will become unavoidable. Therefore, the question should not be whether to use it or not, but how to do so safely, effectively and appropriately (Kirk 2023), by mitigating disadvantages and maximising benefits. For instance, ChatGPT might cause educational institutions to go back to traditional methods of assessments, such as in-class and oral assessments, an approach already taken by eight Australian universities (Castigli 2023; UNESCO 2023). While awaiting the traditional top-down approach, which may not be as swift as required for the rapid advancement of technologies significantly impacting education, some examples indicate that self-regulatory methods, where educational institutions (schools, universities or even individual educators) take the initiative, prove to be an effective means of managing such evolution. These approaches should exemplify the necessity for more inclusive methods, wherein practitioners, as well as students or direct users of the system, contribute to the decision-making process of executives.

Policy options

In theory, there are three policy options for dealing with the emerging threads of GenAI: ban, embrace by regulating and/or restricting, or develop it. **Banning** is self-explanatory: the use of technology is forbidden. **Embracing** would need accepting GenAI and regulating it, including by introducing necessary restrictions. Regulation should be flexible enough to ensure the protection of human rights without hindering technological development. The third option promises a middle-ground solution, in cases when an institution or a country would like to explore the benefits of GenAI but they are concerned about the legal and ethical aspects of technology developed by foreign big tech. **Developing** their own language model would allow for the institution or the government to be in control of the legal and ethical aspects of the technology. For instance, as detailed below, this approach was taken by TNO Netherlands, who decided to build their own Dutch language Model, GPT-NL, funded by the Ministry of Economic Affairs and Climate of The Netherlands (Digitale Overheid 2023; TNO 2023), and by the University of Pisa, who developed Sybilla, a ChatGPT made in Italy.

A banning approach: The case of France

Sciences Po University banned the use of ChatGPT to prevent plagiarism. The ban does not apply to cases when the use of ChatGPT is referenced, and when it is used for specific course purposes, under the supervision of a course leader (Sciences Po 2023). The ban is accompanied by a punishment, which may lead to exclusion from the institution, or even from the French higher education system (Sciences Po 2023). It will be interesting to observe whether Sciences Po will rescind its ban at any point in time and what will be its regulatory approach.

Examples of self-regulatory approach

The case of The Netherlands

Six months after ChatGPT was made available to the public,[3] Tilburg University created a working group on GenAI (GenAI WG), and by May 2023 the working group released their first advice for the institution

3 ChatGPT was publicly available in November 2022. The working group was created in March 2023.

(Tilburg University n.d.). Initially, the focus was on assessment, detection and prevention of the use of LLMs in assessment. They also focused on assessing the possibilities of incorporating LLMs in academic education and the need to organise support for lecturers. Interestingly, in November 2023, the WG changed its approach by embracing technological developments and directing towards the need to ensure AI literacy for both its students and staff. They advised the university on the necessity to provide infrastructure and ensure academic integrity and data privacy while using GenAI, in order to implement policies and regulations for GenAI use in education and to build an AI-focused network.

In dealing with the increased use of ChatGPT by students in their assessments, resulting in a high risk of fraud, the GenAI WG advised the university to shift towards more testing on campus rather than home assignments, to move from summative assessments to formative ones, and to change programmes' intended learning outcomes.

A similar approach was followed also by Maastricht University (Maastricht University 2023). Other universities in the Netherlands, while not being so open towards GenAI yet, have taken steps towards embracing the future use of GenAI in education. For example, Universiteit van Amsterdam does not currently allow the use of ChatGPT for teaching purposes but has launched an e-module on GenAI, focusing on the responsible use of AI tools (particularly ChatGPT) in higher education (TLC Science 2023). Utrecht University has created an AI in Education Lab (Utrecht University n.d.).

The case of Poland

Adam Mickiewicz University (UAM) released guidelines focusing on raising awareness among students on the technical limitations of these technologies and the need for a critical view of their outcome. UAM demands students to disclose in a rather detailed way (objectives, method, labels, prompts, etc.) the uses of GenAI in their thesis or other educational output (UAM 2024). Each course instructor is free to set their own rules on the admissibility or not of the use of GenAI by their students (UAM 2024). Supervisors are given the burden of verifying and making sure that their students' use of GenAI does not go against the principles of national law on higher education (UAM 2024).

While this might be a way of urging university professors to carefully assess their students' work against (ab)use of GenAI, it seems that supervisors are given excessive responsibility over their (adult) students' actions, without providing them any specific detection tools, tips or training. The guidelines refer to 'existing detection tools' by simply providing a publicly available link with a list of state-of-the-art tools, some of which necessitate a subscription. Professors are also given the burden of assessing which tools have the least false positive detection rate. UAM does not provide a specific list of the preferred tools, nor does it provide any free subscription or any training to its academic staff, which is the minimum a university should provide when demanding professors to use such tools.

An example of multiple approaches: The case of Italy

Differently from the Netherlands and many other countries, who did not take any steps at the national level, allowing for 'self-regulation' of GenAI by each institution, Italy went one step further by becoming the first authority in the world to put a ban on ChatGPT over privacy concerns (Browne 2023). At the end of March 2023, the Italian National Data Protection Authority (DPA) ordered a temporary ban on ChatGPT over alleged privacy violations (Goujard 2023). This action came as a result of a data breach from ChatGPT that compromised over 100 million users (Zorloni 2023; Lomas 2023). According to the Italian DPA, the company 'lacked a legal basis justifying the mass collection and storage of personal data "to train" the algorithms' of ChatGPT and lacked transparency over this process (Zorloni 2023; Lomas 2023). It also added that the company was processing data inaccurately. Another concern revolved around the lack of age verification filters, to make sure that children under 13, as claimed by OpenAI, cannot use the platform (Zorloni 2023; Lomas 2023). OpenAI seemed to cooperate and at the end of April 2023 ChatGPT became available again in Italy, after fixing the stated issues. However, on 29 January 2024, the Italian DPA went on by notifying OpenAI for breaches of privacy law (Garante per la Protezione dei Dati Personali 2024).

In the education sector diverse opinions have emerged. The prevailing opinion appears to embrace the view that ChatGPT might have the potential to lead to fundamental positive changes in the education system (Maurizio 2023). On 24-25 November 2023, a large event was organised in Bergamo, with various panels focusing on the impact of AI in schools (Stati Generali della Scuola Digitale 2023a). The event involved participants from all relevant stakeholders, including the Italian Ministry of Education and Merit (Ministero dell'Istruzione e del Merito) and students, offering a great opportunity to exchange ideas and concerns, leading to better future policy and regulatory efforts. There was also the launch

of an Experimental Lab on AI at School, which offered student participants the opportunity to solve homework by using ChatGPT accompanied by the 'oversight' of an expert. This experimental lab aims to teach students to use GenAI responsibly, legally, and efficiently (Stati Generali della Scuola Digitale 2023b). This is an interesting approach as an initial step toward students' educational awareness.

The University of Siena (SU) is the first university in Italy to release guidelines on the use of GenAI (Toscana Open Research 2023). In these guidelines, SU announces the launch of courses and tutorials for academic staff and students (Toscana Open Research 2023). SU urges students to reference GenAI use in their thesis and publications and demands a conscious critical use of these tools, while expecting its academic staff to guide their students in this endeavour (Toscana Open Research 2023).

As mentioned, the University of Pisa is the first university to develop its own ChatGPT version, Sybilla, which functions as a real-time assistant able to answer questions in the same language as the posed question, despite having Italian as its basis (Litrico 2023). At its current state, Sybilla seems to function more as a chatbot, than a GenAI tool though: it provides information about roles, guides, schedules, etc.; it does not seem to be generating new content.

A liberal approach: The case of Belgium

While Belgium has shown no intention of banning the use of ChatGPT in education, it has not provided any policy recommendations either. Nevertheless, several educators have embraced ChatGPT and decided to experiment with it. Some university professors demand students to incorporate ChatGPT in their assignments provided that they make a reference to its use and articulate their experience with ChatGPT interaction. The direct involvement of students empowers both educators and students alike to adopt a more cautious approach toward its use and to identify the most beneficial use cases.

In February 2024, Vrije Universiteit Brussel (VUB) released a tool (RAIN) for its employees that provides information on using AI tools in research, such as what AI tools could be used for, what rules academic staff should follow and what AI tools are available for different research purposes. Simultaneously, VUB released some guidelines with 'Tips for using AI-tools and prompt engineering'. [4]

A liberal approach does not mean that anything is allowed. In a recent case, the University of Antwerp is investigating a student over suspicions of generating an academic paper with ChatGPT. The university has released a training on ChatGPT for educators (Universiteit Antwerpen n.d.) and has declared that it will soon develop clear rules on the use of this tool to avoid future fraud (Amies 2023).

4 Both the tool and the guidelines are made available in the VUB internal information system. The author has access as a VUB academic staff member.

Policy recommendations

The challenge for educators is to probe ChatGPT capacities and limitations and make it work for them. The objective should be maximising its potential while remaining aware of the possibilities for foreseeable misuse. Education systems suffer from structural issues: learning systems need to adapt to rapid technological advancements and align with the evolving needs of new generations of learners. A traditional top-down approach would not be enough in this case. Instead, a direct collaboration of educational institutions (not only academic researchers but also lecturers and teachers) with the providers of ChatGPT and other similar technologies, on possible ways of making the tool serve educational purposes would be necessary. In addition, students' input on how they use the tool would assist education institutions and responsible governmental institutions in developing policies on the ethical and legal use of GenAI in education and also the creators of these tools in fine-tuning them.

Overall, collaboration between all these different stakeholders via open workshops and other similar events, like in the case of Italy, is crucial in order to brainstorm and reach concrete guidance on how to positively use GenAI for enhancing students' critical thinking and simplifying their learning capabilities without making them digitally dependent. To reach this scope, students should be made part of collaborative exchanges from the very beginning.

Recommendations for providers of GenAI systems:

- Increase transparency regarding their system training methods (what data are collected, how, and for how long).
- Release disclaimers on potentially foreseen misuse.
- Develop tools to detect AI-generated material (text, images, video) to combat plagiarism. OpenAI is already developing tools to watermark AI-generated images but has not yet developed such a tool for AI-generated text.
- Collaborate with educational entities on possible ways of making the tool serve educational purposes.

Recommendations for educators, course instructors, professors etc:

- Train students how to use the technology wisely and properly. Demand them to document the use of AI, including prompts and output, similarly to how it is done in the introduction section of this paper.
- Consistently educate students about the challenges of AI, by listening to their experiences and concerns, recognising that not all may possess comprehensive knowledge or literacy in AI. Emphasise the importance of verifying AI-generated content for accuracy. Teach them how to test AI-generated output against trustworthy resources.
- Empower students by actively involving them, encouraging them to share their thoughts and experiences regarding their interactions with GenAI tools.
- Exercise caution when employing external fraud detection tools, as they are not infallible and may yield false positive results. This could potentially result in unjust accusations against students. Note that even the widely used plagiarism detection tool Turnitin can get it wrong (Fowler 2023).
- Embrace technological innovation by exploring alternative approaches to assess students' knowledge. Consider options like face-to-face, oral evaluations, traditional pen-and-paper exams (as opposed to open-book exams), in-class group activities and other diversified assessment methods.

Recommendations for education institutions:

- Carefully assess the impact of the integration of new technologies developed by private companies into the institution/education system and adapt policies accordingly.
- Train the trainers (improve the skills of the teaching staff).
- Develop clear plans and guidelines for using AI in education management, teaching, learning and assessment. These guidelines should be kept up to date based on technological developments and incoming regulatory interventions, such as the EU AI Act.
- Develop clear rules for students explaining in which contexts the use of AI is permitted and in which contexts it is not.
- Involve students in the process of policy development.
- Promote transparent communication regarding the institution's fraud and plagiarism policy regarding AI.
- Develop policies to protect academic and educational integrity in light of new technologies.
- Share experiences and findings.
- Launch new courses that focus on GenAI; and/or update existing courses to teach AI literacy.
- Ensure the academic freedom of course instructors while developing guidelines on the use of GenAI in the classroom.
- Assess the possibility of modifying the institution's definition of plagiarism to include for example 'the use of a text written by a generative system as its own'.
- Assess the embrace of GenAI detection tools, such as GPTzero (similar to the use of Turnitin).

Recommendations for national governments:

- Conduct assessments on the necessity of changing learning systems to adapt to rapid technological advancements and identify appropriate methodologies for aligning learning systems with the evolving needs of new generations of learners (i.e. methods employed, strategies for addressing issues such as limited information spans, misinformation, diverse information sources, etc.).
- Ensure no infringement upon academic freedom while attempting to regulate the use of GenAI in academia.
- Hold providers of GenAI tools accountable for any violation of fundamental rights.

Recommendations for international/regional organisations in the context of international cooperation:

- Ensure academic freedom is not violated while attempting to regulate the use of GenAI in academia.
- Enhance cooperation in increasing the willingness of tech companies based in third countries to comply with European regulations and the decisions of supervisory authorities.

Recommendations for Italian institutions:

- The Italian government, like any other government, should adopt a precautionary approach in the adoption and regulation of GenAI in education, balancing between the opportunities presented by these technologies with the risks to human beings and their fundamental rights, and ensuring that such risks are minimised to the greatest extent possible.
- This assessment should include scrutinising the impact of the integration of new technologies developed by private companies into the education system.
- The Ministry of Foreign Affairs and International Cooperation should increase international cooperation to identify the most suitable solutions.
- The government should hold the providers (and deployers) of these services to account, on the basis of assessing their GDPR and AI Act compliance, but also on the basis of further, more specific regulation by national laws.
- The Ministry of Education should include digital skills in all curricula already at secondary school.
- Ensure to not violate academic freedom while attempting to regulate the use of GenAI in academia.

Conclusion

The analysis presented in this policy brief highlights the complexities surrounding the adoption of GenAI in higher education. While these tools offer benefits such as instant support for teaching and learning, or personalised assistance, they also raise concerns about the erosion of essential skills, academic integrity and the reinforcement of biases and stereotypes.

The privacy concerns and the potential commercialisation of user data underscore the importance of regulatory frameworks that prioritise the protection of fundamental human rights and values.

Governments, universities and other relevant stakeholders need to collaborate in developing robust regulatory frameworks that balance the potential benefits of GenAI with the protection of educational values and related human rights. This includes measures to safeguard academic integrity, promote digital literacy and mitigate the risks of bias and discrimination inherent in these technologies.

While GenAI holds promise for transforming higher education, its adoption must be guided by human rights normative standards and the principles of ethics, transparency and inclusivity. By addressing the challenges and opportunities presented by these technologies in a proactive and collaborative manner, we can harness their potential to enhance teaching and learning while upholding the fundamental value of education as a common good.

References

Amies N 'Antwerp University investigates student over suspected Chat GPT-generated paper' (4 February 2023) The Brussels Times, available at https://www.brusselstimes.com/363585/antwerp-university-investigates-student-over-suspected-chat-gpt-generated-paper (last visited 12 March 2024)

Bender E & Koller A 'Climbing towards NLU: On Meaning, Form, and Understanding in the Age of Data' (2020) Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, 5185-5198

Bendler E, Timnit G, McMillan-Major A & Shmitchell S 'On the Dangers of Stochastic Parrots: Can Language Models Be Too Big?' (2021) FAccT '21: Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency, DOI: https://dl.acm.org/doi/10.1145/3442188.3445922 (last visited 12 May 2024)

Bertuzzi L 'EU Commission issues internal guidelines on ChatGPT, generative AI' (31 May 2023) Euroactiv, available at https://www.euractiv.com/section/artificial-intelligence/news/eu-commission-issues-internal-guidelines-on-chatgpt-generative-ai/ (last visited 13 March 2024)

Browne R 'Italy became the first Western country to ban ChatGPT. Here's what other countries are doing' (4 April 2023) CNBC, available at https://www.cnbc.com/2023/04/04/italy-has-banned-chatgpt-heres-what-other-countries-are-doing.html (last visited 12 March 2024)

Castigli M 'ChatGPT all'Università: ecco i progetti educativi con l'AI già attivi' (14 June 2023) Agenda Digitale, available at https://www.agendadigitale.eu/cultura-digitale/chatgpt-alluniversita-ecco-i-progetti-educativi-con-lai-gia-attivi/ (last visited 12 April 2024)

CoE (Council of Europe) Education Strategy (2024-2030) MED-26(2023)08 final rev, available at https://rm.coe.int/education-strategy-2024-2030-26th-session-council-of-europe-standing-c/1680abee81 (last visited 12 March 2024)

CoE Draft Framework Convention (Council of Europe Draft Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law) (2023) CAI(2023)28, available at https://rm.coe.int/cai-2023-28-draft-framework-convention/1680ade043 (last visited 12 March 2024)

Dahl M, Magesh V, Suzgun M & Ho DE 'Large Legal Fictions: Profiling Legal Hallucinations in Large Language Models' (2024), available at https://arxiv.org/html/2401.01301v1 (last visited 12 March 2024)

Digitale Overheid 'Nederland bouwt eigen open taalmodel GPT-NL' (6 November 2023), available at https://www.digitaleoverheid.nl/nieuws/nederland-bouwt-eigen-open-taalmodel-gpt-nl/ (last visited 12 March 2024)

European Commission, Digital Education Action Plan (2021-2027), available at https://education.ec.europa.eu/focus-topics/digital-education/action-plan (last visited 12 March 2024)

European Declaration on Digital Rights and Principles (2022), available at https://digital-strategy.ec.europa.eu/en/policies/digital-principles (last visited 12 March 2024)

European Parliament, 'EU AI Act: first regulation on artificial intelligence' (19 December 2023), available at https://www.europarl.europa.eu/news/en/headlines/society/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence (last visited 12 March 2024)

Fowler GA 'We tested a new ChatGPT-detector for teachers. It flagged an innocent student' (3 April 2023) Washington Post, available at https://www.washingtonpost.com/technology/2023/04/01/chatgpt-cheating-detection-turnitin/ (last visited 19 March 2024)

Garante per la protezione del dati personali 'ChatGPT: Italian DPA notifies breaches of privacy law to OpenAI' (29 January 2024), available at https://www.garanteprivacy.it/home/docweb/-/docweb-display/docweb/9978020#english (last visited 18 March 2024)

Goujard C 'Italian privacy regulator bans ChatGPT' (31 March 2023) Politico, available at https://www.politico.eu/article/italian-privacy-regulator-bans-chatgpt/ (last visited 12 March 2024)

Hainsdorf C, Hickman T, Lorenz S & Rennie J 'Dawn of the EU's AI Act: political agreement reached on world's first comprehensive horizontal AI regulation' (14 December 2023) White & Case, available at https://www.whitecase.com/insight-alert/dawn-eus-ai-act-political-agreement-reached-worlds-first-comprehensive-horizontal-ai (last visited 12 March 2024)

Kirk T 'ChatGPT we need to talk' (5 April 2023) University of Cambridge, available at https://www.cam.ac.uk/stories/ChatGPT-and-education (last visited 13 March 2024)

Litrico A 'Nasce Sibylla, la ChatGPT al servizio delle università' (12 October 2023) Millionaire, available at https://www.millionaire.it/nasce-sibylla-la-chatgpt-al-servizio-delle-universita/ (last visited 12 April 2024)

Lomas N 'Italy orders ChatGPT blocked citing data protection concerns' (31 March 2023) Tech Crunch, available at https://techcrunch.com/2023/03/31/chatgpt-blocked-italy/ (last visited 12 March 2024)

Maastricht University 'ChatGPT Guidelines for Examiners' (7 February 2023), available at https://www.maastrichtuniversity.nl/file/22052-boe-fasos-chatgpt-guidelines-examiners-v-7-2-2023pdf (last visited 22 April 2024)

Maurizio C 'ChatGPT irrompe nelle università: il parere di docenti e studenti, le contromisure' (6 February 2023) Agenda Digitale, available at https://www.agendadigitale.eu/scuola-digitale/chatgpt-irrompe-nelle-universita-il-parere-di-docenti-e-studenti-le-contromisure/ (last visited 8 May 2024)

Maurizio C 'L'impatto di ChatGPT sull'istruzione: cosa ne pensano docenti e studenti' (12 May 2024) Agenda Digitale, available at https://www.agendadigitale.eu/scuola-digitale/chatgpt-limpatto-sullistruzione/ (last visited 12 March 2024)

Mitchell M 'Can Large Language Models Reason?' (11 December 2023) AI: A Guide for Thinking Humans, available at https://aiguide.substack.com/p/can-large-language-models-reason?utm_campaign=email-post&r=2hw8hp&utm_source=substack&utm_medium=email (last visited 12 March 2024)

Proposal for a Regulation of the European Parliament and of the Council Laying Down Harmonised Rules on Artificial Intelligence (Artificial Intelligence Act) and Amending Certain Union Legislative Acts COM/2021/206 final

Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation) [2016] OJ L119/1, DOI: http://data.europa.eu/eli/reg/2016/679/oj (last visited 29 April 2024)

Sciences Po 'ChatGPT: Sciences Po Implements Rules and Opens up Discussion About AI in Higher Education' (9 February 2023), available at https://www.sciencespo.fr/en/news/sciences-po-implements-strict-rules-about-the-use-of-chatgpt-by-students/#:~:text=Non%2Dcompliance%20with%20mentioning%20the,informs%20all%20students%20and%20teachers (last visited 12 March 2024)

Stati Generali della Scuola Digitale (2023a), available at https://www.statigeneraliscuoladigitale.it/# (last visited 12 March 2024)

Stati Generali della Scuola Digitale (2023b) 'Facciamo i Compiti con L'intelligenza Artificiale!', available at https://www.statigeneraliscuoladigitale.it/facciamo-i-compiti-con-lintelligenza-artificiale/ (last visited 18 March 2024)

Sullivan M, Kelly A & McLaughlan P 'ChatGPT in higher education: Considerations for academic integrity and student learning' (2023) 6(1) Journal of Applied Learning and Teaching 31, DOI: https://doi.org/10.37074/jalt.2023.6.1.17 (last visited 12 May 2024)

Tilburg University 'Chat GPT' (n.d.), available at https://www.tilburguniversity.edu/research/institutes-and-research-groups/chatgpt (last visited 13 March 2024)

TLC Science 'Responsible use of Generative Artificial Intelligence (GenAI) in higher education' (n.d.), available at https://rise.articulate.com/share/MyfLgG-cXE1a7XBuctQhndpJB-BgpYny#/ (last visited 12 April 2024)

TNO ‘Nederland start bouw GPT-NL als eigen AI-taalmodel’ (2 November 2023), available at https://www.tno.nl/nl/newsroom/2023/11/nederland-start-bouw-gpt-nl-eigen-ai/ (last visited 14 March 2024)

Toscana Open Research ‘ChatGpt, l’Università di Siena è la prima in Italia a normare l’uso dell’IA’ (6 September 2023), available at https://www.toscanaopenresearch.it/chatgpt-luniversita-di-siena-e-la-prima-in-italia-a-normare-luso-dellia/ (last visited 12 April 2024)

UAM ‘Guidelines on Use of Artificial Intelligence-based content generation systems in teaching and thesis preparation’

UNESCO ‘ChatGPT and Artificial Intelligence in Higher Education: A quick Start Guide’ (2023), available at https://etico.iiep.unesco.org/en/chatgpt-and-artificial-intelligence-higher-education-quick-start-guide (last visited 14 March 2024)

Universiteit Antwerpen ‘ChatGPT: een (vergiftigd?) geschenk voor leraar en leerling? (herhaling 3) (n.d.), available at https://cno.uantwerpen.be/nl/opleiding/chatgpt-een-vergiftigd-geschenk-voor-leraar-en-leerling-herhaling-3-79747 (last visited 18 March 2024)

Utrecht University ‘About the AI in Education Lab’ (n.d.), available at https://www.uu.nl/en/research/ai-labs/ai-in-education-lab/about (last visited 8 April 2024)

Yin L, Alba D & Nicoletti L ‘Open AI’s GPT is a recruiters dream tool. Test shows there’s racial bias’ (8 March 2024) Bloomberg, available at https://www.bloomberg.com/graphics/2024-openai-gpt-hiring-racial-discrimination/ (last visited 16 March 2024)

Zhang D, Maslej N, Brynjolfsson E et al. ‘The AI Index 2022 Annual Report’ (2022) AI Index Steering Committee, Stanford Institute for Human-Centered AI, Stanford University

Zorloni L ‘Il Garante della privacy blocca ChatGPT in Italia’ (31 March 2023) Wired, available at https://www.wired.it/article/chatgpt-blocco-italia-garante-privacy/ (last visited 12 April 2024)

Global Campus of Human Rights

Europe
South East Europe
Latin America-Caribbean
Asia-Pacific
Central Asia
Caucasus
Arab World
Africa

The Global Campus of Human Rights is a unique network of more than one hundred participating universities around the world, seeking to advance human rights and democracy through regional and global cooperation for education and research. This global network is promoted through eight Regional Programmes which are based in Venice (GC Europe), Sarajevo/Bologna (GC South East Europe), Pretoria (GC Africa), Bangkok (GC Asia-Pacific), Yerevan (GC Caucasus), Buenos Aires (GC Latin America and the Caribbean), Beirut (GC Arab World), and Bishkek (GC Central Asia).

The Global Campus Policy Observatory aims to enhance the role of the Global Campus and its regional members in undertaking coordinated research initiatives and providing guidance and expert opinions in response to urgent human rights issues to a broad primary and secondary audience. It provides a virtual hub with the participation of a team of researchers who are alumni from GC regional programmes, for the production of a set of complementary policy analyses on selected topics.

GC Headquarters
Monastery of San Nicolò,
Riviera San Nicolò, 26
I-30126 Venice Lido (Italy)

www.gchumanrights.org

Supported by