The AI for Higher Education

The state of play report

Coordinated by

Pablo de Olavide University

FIM UHK

OSTİM TEKNİK ÜNİVERSİTESİ ANKARA

UNIVERSITY OF SZCZECIN

SGroup UNIVERSITIES IN EUROPE

UNIVERSIDAD PABLO DE OLAVIDE SEVILLA

DECEMBER 2025

AI4Uni project

ai4uni.usz.edu.pl

The AI for Higher Education – the state of play report **is a result of the activities conducted within 2024-1-PL01-KA220-HED-000250324 project.**

Authors:

Elif Akagün, OSTIM Technical University

Hasibe Aysan, OSTIM Technical University

Elena de la Cova, Pablo de Olavide University

Kinga Flaga – Gieruszyńska, University of Szczecin

Saim Karabulut, OSTIM Technical University

Aleksandra Klich, University of Szczecin

Blanka Klímová, University of Hradec Králové

Alicia Troncoso Lora, Pablo de Olavide University

Marcel Pikhart, University of Hradec Králové

Angela Troncoso, Pablo de Olavide University

Co-funded by the European Union. Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or Foundation for the Development of the Education System (FRSE). Neither the European Union nor Foundation for the Development of the Education System (FRSE) can be held responsible for them.

Table of Contents

A. Legal report

1. Legal, practical and ethical implications of Artificial Intelligence usage 5

2. Legislation of AI 6

2.1. Main legal challenges 6

2.2. European Union 8

2.3. Other legislations 12

2.4. Countries from the project consortium 13

3. Ethics of AI 21

3.1. Overview of collective and individual implications 21

3.2. Main proposals 23

4. Practical implications for Higher Education 26

4.1. Impact of AI on advanced skills demand 26

4.2. Impact of AI on learning 26

4.3. Impact of AI on teaching 27

B. Interviews report

1. Interviews about AI usage 29

1.1. Purpose of the interviews 29

1.2. Analysis approach 29

2. Results and Key Insights 31

2.1. Findings by cluster 31

2.2. Mapping intersections across themes and cluster 36

3. State of the Play: Conclusions 47

References 49

Table of Figures

Figure 1. Codification of the interviews..... 30
Figure 2. AI Usage for students. 31
Figure 3. AI Usage for academic staff. 33
Figure 4. AI Usage for administrative staff. 35
Figure 5. Overall usage of AI. 37
Figure 6. Frequency of use. 38
Figure 7. AI Knowledge. 40
Figure 8. Trustworthiness by cluster. 40
Figure 9. Reported usefulness of AI..... 42
Figure 10. Perceived usefulness of AI. 44
Figure 11. Concerns of AI use. 45
Figure 12. Concerns by age group I. 46

AI4Uni

Erasmus+ Cooperation partnerships in higher education

2024-1-PL01-KA220-HED-000250324

The general objective of the project is to increase the organisational capacity of the universities to responsibly use AI solutions. This aim will be achieved by the following specific objectives: 1. raising awareness about responsible AI usage, 2. developing necessary competences and skills, and 3. implementing tailor-made AI solutions for the participating universities.

"Education does not compete with artificial intelligence in remembering more, but in shaping minds capable of understanding, interpreting, and making better decisions"

A. Legal report

1. Legal, practical and ethical implications of Artificial Intelligence usage

Alicia Troncoso

Artificial Intelligence (AI) is having an increasing influence on people's daily lives and plays a key role in many sectors of society, including higher education, thanks to its ability to automate and facilitate decision-making. The benefits of intelligent systems are extraordinary, but the risks they bring with them can be even greater. As Ignatieff (2023) points out, we should consider that we have created AI, and therefore we must be able to understand what we have created. If we can understand it, we will be able to keep it under control.

In recent years, much emphasis has been placed on the need for “trustworthy” (“responsible”) AI that respects legal and ethical issues and enables the construction and operation of robust intelligent systems. Currently, the UN (2024) indicates that there is a global governance deficit with respect to AI, and most organizations do not have adequate governance and management for intelligent systems. In fact, the National Institute of Standards and Technology (NIST) states that much more research is still needed, as well as standardization in various areas of AI (Schwartz, 2022).

This report aims to introduce the regulatory and ethical challenges of AI and to take a brief tour of the current AI legislation in the European Union as well as in the countries that are members of the project consortium: Spain, Poland, Türkiye and Czechia. In addition, it provides an overview of the ethical aspects of AI that have been collected in publications from the European Union as well as from other important organizations. Finally, the report concludes by addressing the implications and impact that the use of AI has on higher education institutions.

2. Legislation of AI

Alicia Troncoso

In order to control the possible risks resulting from the implementation of AI systems, and to ensure compliance with ethical principles, most countries have drawn up laws that aim to regulate AI, which we will summarize in this section.

2.1. Main legal challenges

This section summarizes the main legal challenges raised by widespread use of AI.
One of the most critical issues is how to ensure that when AI algorithms are in operation, they fully respect the fundamental values and rights established in the European Union, as well as its ethical principles (art. 2 of the Treaty on European Union). The key areas of concern are the possible forms of discrimination by algorithms and the need to develop not only ethical guidelines but also design principles and tools to include ethical aspects from the early design stages of AI. Other lines of action include the obligation to assess their impact on human rights, to assess discrimination when algorithms affect people or are used in public administration, or to require auditing mechanisms for algorithms. In this context, it is necessary to reflect on whether and when automated decision-making should be allowed. Naturally, some users and sectors such as justice or health are more sensitive than others in this respect.

In terms of personal data protection, the General Data Protection Regulation (GDPR) of the European Union (EC, 2016a) contains provisions that address some of the concerns mentioned above and that could serve as inspiration for other areas. Specifically, the GDPR takes into account: i) data protection by design and by default, ii) privacy impact assessments, iii) the (relative) right to explanation regarding the logic involved in automated decision-making, and iv) the right not to be the subject of a decision based solely on automated processing (although the art. 22.2 in GDPR recognises exceptions to this right, the data subject will always keep the right to obtain human intervention on the part of the controller, to express his or her point of view and to contest the decision). As the GDPR came into force in May 2018, it is necessary to monitor its application in the context of AI. Some authors question the suitability of the GDPR for dealing with AI (see Poullet 2018; Delforge and Gerard 2017; Watcher 2017) due to the difficulty of complying with the principles of purpose limitation (since AI systems are, by nature, general-purpose systems) and proportionality, as well as the requirement to obtain legitimate consent. Criticism also refers to the perceived limitations regarding the so-called right to explanation.

In this context, the transparency and explainability of algorithms are considered essential

for understanding how decisions are made and providing citizens with the possibility of questioning them and thinking critically. Although it is not yet clear how these requirements should be addressed, some references can be found in public documents on related aspects such as:

i. the processes to be implemented
ii. the interference of AI with the protection of intellectual property
iii. who are the people responsible for examining the decision-making carried out by an algorithm
iv. the development of guidelines to help explain AI.

The European Commission commissioned a study in order to analyse these and related aspects. This study was published in 2018: https://actuary.eu/wp-content/uploads/2019/02/AlgoAware-State-of-the-Art-Report.pdf

Responsibility is another issue that frequently arises in legal literature and policy debates and can be considered another aspect of AI to be legislated both in the European Union and in the national legislation of individual countries. It is an essential aspect to guarantee that potential victims of AI-based applications have legal certainty and also have a system of compensation for the damage caused. The complexity of AI systems and the variety of actors involved in the value chain, together with the learning capabilities of AI, can make it very difficult to assign responsibilities.

Therefore, the main question is: does the current framework provide adequate mechanisms to address the liability and safety of AI products and services? If not, what adjustments are necessary? To analyse these aspects, the European Commission created the Expert Group on Liability and New Technologies (LNT) in 2018 with the mission of providing the Commission with expertise on the applicability of the Product Liability Directive (Council Directive 85/374/EEC of 25 July 1985) to traditional products, new technologies and new societal challenges and assist the Commission in developing principles that can serve as guidelines for possible adaptations of applicable laws at European Union and national level relating to new technologies. The group was divided into two subgroups. The first subgroup focused on the Product Liability Directive and examined which provisions of the Directive are adequate to resolve liability issues in relation to traditional products, new technologies and new social challenges. The other subgroup focused on technologies and assessed whether existing liability regimes can be adapted to new realities in the development of new technologies, such as AI. This group of experts has helped the Commission develop principles, which can serve as guidelines for possible adaptations of laws at European and national level with regard to new technologies.

On November 21, 2019, the group published its Report on Liability for Artificial Intelligence

and other emerging technologies (LNT, 2019), which determines when EU accountability frameworks will continue to work effectively in emerging digital technologies such as AI or the Internet of Things. This report presents the group's conclusions and recommendations.

AI products or services, like all others, cannot be guaranteed to be 100% safe. One possible way to manage safety is to define acceptable levels of safety for the identified risks that are characterized by the probability of occurrence and impact criteria. This requires assigning quality levels for the software and processes that are applied from the design of the product or service, according to the criticality of the software modules and the corresponding datasets that contribute to the identified risks. In addition, it may be necessary to implement additional mitigation measures of a technical or operational nature when the residual risks are considered unacceptable.

In addition to the legal challenges outlined above, cybersecurity-related challenges and the need to provide testing platforms, including regulatory testing environments as well as other legal areas such as consumer protection and competition law, are also considered important. Both legal experts and the media have recently warned of possible anti-competitive behaviour that could take place in this context. Such concerns relate to collusion (including, for example, tacit collusion in price fixing), price discrimination, data concentration, and the role of privacy in the assessment of competition. National public authorities and stakeholders have reiterated the need to remain vigilant in this regard (UK House of Lords, 2018). The growing power of certain platforms also raises concerns about their responsibility as quasi-public actors in the digital society (EPSC, 2018). The regulatory regime for machine-generated data, such as data from sensors, has been a recurring issue in the political debate on AI, and in general on Industry 4.0.

2.2. European Union

The European Union, with the aim of becoming a leader in AI, approved a coordinated plan on Artificial Intelligence (EU, 2021), which it updated in April 2021, to promote AI, through the Digital Europe and Horizon Europe programs, and through the European funds of the Recovery and Resilience Facility, of which 20% has been allocated to the digital transition of companies. In this plan, the European Union advocates for people-centered AI with the aim of creating the right conditions for its development and implementation, promoting its excellence, ensuring that it serves people and is a force for social good, and promoting leadership in high-impact strategic sectors such as sustainable production, health, the public sector, mobility, and agriculture, among others.

In addition, the European Union has developed a regulation establishing harmonized standards on AI (EC, 2021a) and a regulation for machines (EC, 2021b), which were also approved in 2021, and an ethical guide (EC, 2020), published in July 2020 and developed

by a group of high-level experts based on a consultation to give people the confidence to adopt these technologies, while encouraging companies to develop them. These regulations and guidelines propose new rules and directives to ensure that AI systems used in the European Union are safe, transparent, ethical, impartial and under human control. *Machines* are understood to include a wide range of consumer and professional products, from robots to lawnmowers, 3D printers, construction machines and industrial production lines.

The European Union, in the context of its digital strategy, has published different regulations related to data and digital services:

- GDPR, which aims to protect individuals with regard to the processing of their personal data and the free movement of such data in the European Union and the European Economic Area (EEA).
- Digital Services Act, which aims to modernize and harmonize the European Union legal framework for dealing with illegal or potentially harmful online content, the liability of online intermediaries for third-party content, the protection of users' fundamental rights online, and the reduction of information asymmetries between online intermediaries and users. Available at: https://digital-strategy.ec.europa.eu/en/policies/digital-services-act-package
- Digital Markets Act, which aims to resolve imbalances in European Union digital markets resulting from the dominance of large online platforms, by establishing harmonized rules defining and prohibiting certain unfair practices by these platforms. Available at: https://digital-markets-act.ec.europa.eu/index_en
- Data Act, which aims to facilitate access and use of non-personal data, including business-to-business, business-to-government and business-to-customer data, and to revise the rules on the legal protection of databases. Available at: https://digital-strategy.ec.europa.eu/en/policies/data-act
- Data Governance Act, which aims to establish robust mechanisms to facilitate the reuse of certain categories of protected public sector data, increase trust in data intermediation services and encourage “data altruism” in the European Union. Available at: https://digital-strategy.ec.europa.eu/en/policies/data-governance-act

It has also been publishing several regulations as part of its cybersecurity strategy, such as the directive on measures for a high common level of cybersecurity across the European Union, which aims to establish a standardized level of protection in all European Union Member States. Available at: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32022L2555&from=ES

In this context, the AI Act (https://artificialintelligenceact.eu/), known as the “AI Law”, has been published, with the aim of ensuring that Europeans can trust what AI has to offer, addressing the risks of certain AI systems to avoid undesirable results. Available at: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai

On the other hand, the European Commission released in 2022 the *Artificial Intelligence Liability Directive* (AILD), which deals with claims for damage caused by AI systems, or the use of AI, adapting the rules of non-contractual civil liability to AI. This directive complements the AI Law by introducing a new liability regime that aims to ensure legal certainty, increase consumer confidence in AI and assist consumers in their liability claims for damage caused by AI-based products and services. However, the European Commission revealed in February 2025 that it will withdraw the proposal for a new directive on the liability of AI. All information can be found at: https://www.ai-liability-directive.com

Below is an overview of the AI law passed in the European Union in June 2024. The objective of the AI Act is to: improve the functioning of the internal market by establishing a uniform legal framework, in particular for the development, market introduction, commissioning and use of AI systems in the European Union, in accordance with Union values, in order to promote the adoption of human-centric and trustworthy AI, while ensuring a high level of protection of health, the safety and fundamental rights, including democracy, the rule of law and the protection of the environment, protect against the harmful effects of AI systems, as well as support innovation in the European Union.

The law adopts a risk-based approach, in which four levels are defined: unacceptable, high, limited and minimum risk.

The unacceptable risk is the risk presented by a very limited set of AI systems that may pose a threat to people's safety and rights and are therefore prohibited. For example: subliminal techniques that modify behaviour in a harmful way, exploitation of vulnerabilities, biometric categorization and social classification, remote biometric identification in real time in public spaces, databases through non-selective extraction of facial images from the internet, inference of emotions, analysis of recorded video in public spaces that use biometric identification, etc.

The high risk refers to AI systems that may have an adverse impact on people's safety or fundamental rights. For example: AI systems used in critical infrastructure, educational or professional training, product safety components, employment and worker management, essential public and private services, migration management and border control, administration of justice and democratic processes, etc. In this case, the AI Law imposes important responsibilities on suppliers and those who deploy high-risk AI systems on the market: adequate risk assessment and mitigation systems, high quality of the datasets that

feed the system to minimize risks and discriminatory results, recording of activity to ensure traceability of results, detailed documentation providing all the necessary information about the system and its purpose for the authorities to assess its conformity, clear and adequate information to the implementer, adequate human oversight measures to minimize risk, high level of robustness, security and accuracy, etc.

Limited risk is associated with the lack of transparency in the use of AI. For example, when using foundational deep learning models, chatbots, generative AI systems, etc. In this case, the AI Act introduces requirements for transparency, informed choice and safety; that is, users should be aware that they are interacting with a machine.

The minimum risk is related to most AI systems currently in use, for example, in many video games or *spam* filters. In this case, they have no additional legal obligations, but the providers of these systems could voluntarily apply the requirements for trustworthy AI. This risk-oriented approach is also reflected in the Law, as article 27 states that: “before deploying one of the high-risk AI systems, those responsible for the deployment shall carry out an assessment of the impact that the use of such systems may have on fundamental rights.”

Also, always from the perspective of risks, the AI Law determines in its article 57, that “controlled testing spaces for AI” be created, understood as a controlled framework established by a competent authority that offers suppliers and potential suppliers of AI systems the possibility of developing, training, validating and testing, in real conditions where appropriate, an innovative AI system, in accordance with a controlled testing space plan and for a limited period of time, under regulatory supervision. With regard to sanctions, the Law states that “Member States shall establish the system of sanctions and other enforcement measures, such as warnings and non-pecuniary measures, applicable to infringements of this Regulation”.

Furthermore, in 2023 the European Commission created the European Center for Algorithmic Transparency (ECAT), establishing its headquarters at the Joint Research Center (JRC) in Seville (Spain). The center's objective is to provide technical assistance and practical guidance for transparent and trustworthy algorithmic systems. Its work is carried out by combining methodologies from different disciplines to integrate technical, ethical, economic, legal and environmental perspectives.

2.3. Other legislations

As a consequence of its AI strategy, Canada launched in 2017 the Digital Charter Implementation Act (C-27 law), which includes privacy reform, the creation of a Personal Information and Data Protection Tribunal, and the introduction of a comprehensive Artificial Intelligence and Data Act (known as AIDA) (IAPPAI, 2024). The AIDA is expected to come into force in 2025, and it proposes the factors to determine which AI systems would be considered high impact. Available at: https://ised-isde.canada.ca/site/innovation-better-canada/en/artificial-intelligence-and-data-act-aida-companion-document

In the United States (US) AI strategy updated in 2023, one of the objectives set is to promote responsible, safe and secure AI systems. In October 2022, the draft "AI Bill of Rights: Making Automated Systems Work for the American People" was published by the White House Office of Science and Technology Policy, which articulates five principles to guide the design and deployment of automated systems: safe and effective systems; protections against algorithmic discrimination; data privacy; notification and explanation; and human alternatives, consideration, and fallback. Although it does not establish formal standards for AI systems, the draft bill has served as a basis for discussions on US AI policy at the federal level. The next executive order on AI (Executive Order 14110) was published in October 2023, on the safe and reliable development and use of AI, requiring the adoption of more than 150 measures by dozens of federal agencies. This decree has a major impact as it applies directly to most organizations in the AI value chain that do business with the federal government (IAPPAI, 2024).

United Kingdom presents a more cautious approach when it comes to "pro-innovation" legislation, with a flexible and sectoral approach that empowers regulators to create measures adapted to the diverse needs and risks posed by different sectors. It should be noted that the Guide on AI and data protection has been updated, and that a report on foundational models has been published by the Competition and Markets Authority (CMA, 2023).

Finally, in AccessNow (2024) you can find a report on public policies and AI laws in Ibero-America. Most of these countries are in the process of debating and approving laws like the European AI Law, in a real race for regulation (León Coronado, 2023).

2.4. Countries from the project consortium

2.4.1. Spain (UPO)

Alicia Troncoso

In line with European policies, the Spanish government approved the Artificial Intelligence Spanish Strategy (AISS, 2020), which sets out major objectives to place Spain in a leadership position alongside other European Union countries, which are at the forefront. This strategy is cross-cutting and encompasses numerous disciplines in the technological, scientific, social, legal and humanistic fields. Joint and interdisciplinary work across all these disciplines will be more necessary than ever to face the challenges ahead and achieve AI in the service of Humanity.

Among these objectives, the aim is to promote scientific excellence and innovation in Artificial Intelligence; to develop AI tools, technologies and applications for the projection and use of the Spanish language; to promote the creation of qualified employment, promoting training and education, by stimulating national talent and attracting global talent; to incorporate AI as a factor for improving the productivity of Spanish companies, of efficiency in the Public Administration and as an engine of sustainable and inclusive economic growth; generate an environment of trust in relation to AI, both at the technological level and in terms of regulation and social impact; promote the debate on the technological development of humanist values, focused on ensuring the well-being of society, and promote an inclusive and sustainable AI to face the great challenges of our society, specifically reducing the gender gap and the digital divide, supporting the ecological transition and territorial structuring.

In 2023, the Spanish government created the Spanish Agency for the Supervision of Artificial Intelligence with the aim of carrying out tasks of supervision, advice, awareness-raising and training for public and private companies for the adequate implementation of regulations regarding the appropriate use of AI and specifically of algorithms.

The Spanish Artificial Intelligence Strategy 2024 has recently been approved, giving continuity to the initiatives deployed by the Spanish Government in the field of AI to date, adapting them to the notable changes experienced in this technology in recent years. This strategy reinforces Spain's commitment to cutting-edge technology, reaffirming its position as a leader in the development and application of AI solutions. It is an ambitious plan, designed to consolidate and expand the use of AI throughout the economy and in public administration.

2.4.2. Poland (USZ)

Kinga Flaga-Gieruszyńska, Aleksandra Klich

Poland is quite actively implementing the provisions of the EU AI Act to ensure the safe and ethical use of AI technology. To this end, the Ministry of Digital Affairs has drafted a law on AI systems, which was published on 16 October 2024. In the meantime, the Ministry of Digital Affairs conducted public consultation on the draft law on AI systems, taking into account the comments of stakeholders. On 11 February 2025, a new version of the draft law was published and forwarded for further legislative work. In addition, a conference entitled 'AI Act – Revolution in Artificial Intelligence Regulations in the European Union' was held on 14–15 March 2025, during which the progress of the Act's implementation was discussed and experiences from other member states were shared. The Ministry of Digital Affairs continues to work on adapting Polish law to the provisions of the AI Act to ensure compliance with EU regulations and to support the development of AI technology in Poland

The draft bill will enable the introduction of mechanisms into the Polish legal system that will allow the obligations specified in chapters I, II, III section 4, V, VII, IX and XI of the AI Act to be fulfilled within the legally required time. The Act covers, among other things, issues related to the designation of a national supervisory authority, the procedure before this authority, control, the notifying authority, certification, the protection of citizens' rights in court proceedings, as well as the procedure for initiating proceedings and imposing administrative financial penalties for violating the provisions of Article 5 of the AI Act. The scope of application of Regulation 2024/1689 and the Act does not include matters related to national defense, which have been excluded on the basis of a reference to the relevant provisions of the Act on government administration departments. Furthermore, according to the regulation, the scope of these provisions also excludes issues concerning the internal security of the state and the constitutional order, as defined in the provisions on the Internal Security Agency and the Intelligence Agency, as well as on the Military Counterintelligence Service and the Military Intelligence Service. These regulations also do not apply to basic research, applied research and development work that does not include real-world testing. AI systems made available under free open-source licences are also excluded from the scope of the Act, unless these systems are placed on the market or put into service as high-risk AI systems or as AI systems covered by Article 5 or 50 of the AI Act. The regulations also do not apply to systems put into use exclusively for purposes in the area of basic research, application research and development work before the AI Act came into force.

The draft law on AI systems does not fully implement all the obligations arising from the AI Act. The content of the draft, which corresponds to the requirements of Polish and European law, is the result of the work and substantive analyses of the Ministry of Digital

Affairs, consultations with representatives of the European Artificial Intelligence Office and dialogue with other public institutions and social stakeholders. As the AI Act gives member states a certain degree of flexibility, including in terms of designating the appropriate authorities and creating administrative and judicial procedures, in the case of Poland this means the need to prioritise activities related to the establishment of a market surveillance system for AI systems, to enable the implementation of regulations that will come into force by 2 August 2025, while conducting communication and administrative activities. With regard to the remaining issues covered by the AI Act, the Ministry of Digitalisation is planning further legislative measures in cooperation with the relevant authorities, which will be in force in 2026-2027, including those concerning high-risk AI systems and regulatory sandboxes.

The proposed law defines the organisation and method of exercising national supervision over the market for AI systems and general-purpose AI models. The establishment of a new authority is also justified from the point of view of the public finance sector, because the current legal status does not provide an entity that would meet the requirements set out in Article 70(3), i.e. simultaneously guaranteeing an adequate number of specialists in the field of AI technology, data, personal data protection, cybersecurity, fundamental rights and legal standards.

The proposed law defines the organisation and method of exercising national supervision over the market for AI systems and general-purpose AI models. The establishment of a new authority is also justified from the point of view of the public finance sector, because the current legal status does not provide an entity that would meet the requirements set out in Article 70(3), i.e. simultaneously guaranteeing an adequate number of specialists in the field of AI technology, data, protection, etc. The new regulations provide for a system of sanctions in case of non-compliance with the requirements of the AI Act. The regulations will be enforced by an appointed supervisory body, which in Poland will be the Commission for the Development and Security of Artificial Intelligence. The Commission will have the power to impose penalties, conduct audits and issue recommendations on improving the security and transparency of AI systems. In addition, in the event of a gross violation of the regulations, the Commission will be able to apply to the courts for orders to suspend the operation of the AI system. The proposed regulations ensure the Commission's independence in the supervision of AI systems, enabling it to autonomously perform all tasks and exercise market supervision. They also ensure equality of all members of the Commission, who will have the right to participate in meetings, access available information, vote and make decisions. The President of the Commission will be appointed by the Sejm of the Republic of Poland with the consent of the Senate of the Republic of Poland. He or she may be dismissed in the same manner. The proceedings before the Commission in cases of violation of the regulations on AI will be governed by the provisions of the Code of Administrative Procedure.

The bill requires companies and institutions using AI systems to ensure full transparency in how AI makes decisions and to allow supervisory authorities easy access to documentation on AI systems. The aim is to increase accountability and trust in AI technology in Poland. The draft law also includes educational programmes and support for companies, including training and legal and technological advice. In addition, the introduction of regulatory sandboxes will enable companies to test new AI solutions in a controlled and safe environment, which will help to bring about innovation more quickly and effectively.

The AI Act regulations in Poland will work closely with personal data protection regulations, especially with the General Data Protection Regulation (GDPR). All AI systems that process personal data will have to meet additional requirements related to privacy, transparency and minimising the risk of personal data breaches. The regulations are also aimed at supporting the development of the AI sector in Poland by promoting innovation and attracting investment. In turn, the regulations on responsibility and ethics are aimed at creating an environment in which entrepreneurs and investors will be able to operate in a responsible manner and with respect for human rights. Chapter 6 of the draft law concerns the reporting of serious incidents related to the use of AI systems. The provider of an AI system is obliged to report to the Commission any serious incident that occurs in connection with the use of AI systems.

As expected, the implementation of the AI Act in Poland will have a major impact on public administration. The new regulations will enable the implementation of AI systems in decision-making processes, such as the granting of social benefits, the assessment of tax applications or processes related to public safety, if they comply with regulations on transparency, fairness and accountability. Regulations in Polish law also require the issue of administrative fines imposed by the supervisory authority.

The implementation of the above solutions is aimed at creating a stable and predictable environment for the development of AI in Poland, while ensuring the protection of citizens ‘and consumers’ rights.

2.4.3. Türkiye (OSTIM TU)

Elif Akagün, Hasibe Aysan, Saim Karabulut

As of beginning 2025, Turkey does not yet have a comprehensive, direct legal framework for AI. However, significant developments indicate a growing national commitment to AI governance, ethics, and strategic planning. These efforts align with global regulatory trends, particularly the European Union’s AI Act.
Turkey’s AI governance efforts began with the National Artificial Intelligence Strategy 2021-2025, which serves as a roadmap for AI development. The Ministry of Industry and

Technology oversees AI initiatives through the Artificial Intelligence and Digital Transformation Office. In addition to public-sector initiatives, civil society and academia play a key role in shaping AI ethics and awareness. Organizations such as the Turkey Artificial Intelligence Initiative (TRAI) have been instrumental in fostering AI dialogue and innovation.

Turkish legal and regulatory initiatives gained pace with the proposition of AI law on June 24, 2024. On this date, the Turkish Grand National Assembly (TBMM) received the country's first AI Law proposal, marking a milestone in AI regulation. This proposal follows the European Union's adoption of the AI Act on May 21, 2024, reinforcing Turkey's intent to align with international legal standards. The proposed AI Law (will be mentioned as PAIL from now on) aims to ensure safe, ethical, and fair AI usage by protecting personal data and privacy rights. In order to achieve these aims, the PAIL intends to establish a regulatory framework for AI development and deployment. One of the key characteristics of PAIL is the definition of AI. For the first time in Turkey, PAIL explicitly defines AI as: "A computer-based system capable of performing human-like cognitive functions, with abilities in learning, reasoning, problem-solving, perception, and language understanding."

The PAIL mandates AI systems to adhere to core ethical principles such as safety which ensures AI does not pose risks to individuals or society and transparency which intends to provide users with clear information about AI decision-making processes. Other ethical principles aimed in the PAIL involve fairness and prevention of discriminatory outcomes in AI applications, accountability through establishment of liability mechanisms for AI-related harm and data protection. The regulatory approach induced by PAIL is risk-based AI regulation, differentiating between low-risk and high-risk AI applications. Therefore, high-risk AI systems require special evaluation and compliance measures. PAIL involves some penalties for non-compliance which include:

- Up to 35 million TL (or 7% of annual turnover) for prohibited AI applications.
- Up to 15 million TL (or 3% of annual turnover) for compliance violations.
- Up to 7.5 million TL (or 1.5% of annual turnover) for providing false information.

Notably, PAIL does not explicitly clarify whether foreign-developed AI systems used in Turkey fall under its jurisdiction, which may raise concerns regarding cross-border AI governance.

Turkey's AI regulations currently interact with existing laws. Some of the existing regulatory touchpoints involve The Personal Data Protection Law (KVKK) that indirectly regulates AI applications handling personal data and data protection. Besides, Turkey closely monitors the EU AI Act, suggesting future alignment with European AI governance frameworks. Other proposed amendments to the Turkish Civil Code and Criminal Code explore AI liability issues, including expansion of strict liability rules in Turkish Contract Law

to include robotic entities and extension of criminal liability under Turkish Penal Code for AI-related harm, similar to negligence laws concerning dangerous animals.

There are some institutional and academic responses to these regulatory efforts in Turkey as well. For example, The Council of Higher Education (YÖK) issued the "Ethical Guide for Generative AI in Scientific Research and Publications" (May 2024), setting national guidelines for AI use in academia. Some higher education institutions published an AI Ethics Guide requiring students to disclose AI usage and cite AI-generated content. Some others developed internal AI usage guidelines, particularly for assignments and projects and updated their academic integrity policy to address AI usage explicitly. Finally, there are AI engineering departments recently established in some universities.

It is possible to observe Civil Society Engagement in Turkey. For example, The Turkey Artificial Intelligence Initiative (TRAI), established in 2017, plays a crucial role in raising AI Awareness, Strengthening the AI Ecosystem and Organizing AI Seminars, Summits, and Networking Events. TRAI's and similar organizations' and initiatives' efforts have facilitated key discussions on AI regulation, innovation, and industry collaboration.

The PAIL across currently under parliamentary review. Across the near future, after PAIL's enactment, it is expected to be finalized within 45 days following standard legislative procedures. Moreover, further legislative efforts may be required to address liability gaps and cross-border AI governance. Turkey's approach to AI regulation is expected to evolve alongside international best practices, particularly those established by the EU AI Act. Turkey's regulatory landscape for AI is rapidly evolving, balancing innovation promotion with ethical safeguards. The proposed AI Law, coupled with institutional frameworks and academic initiatives, signals a comprehensive, forward-looking approach to AI governance.

2.4.4. Czechia (UHK)

Blanka Klímová, Marcel Pikhart

The Czech Republic has been actively working on establishing a comprehensive framework for the regulation and development of AI. Here is an overview of the current rules, official regulations, and university policies regarding AI in the Czech Republic.

National AI Strategy

The Czech Republic's approach to AI is primarily guided by the National Artificial Intelligence Strategy of the Czech Republic 2030 (NAIS), which was updated in July 2024 and is a very detailed report on the use of AI (https://mpo.gov.cz/assets/cz/rozcestnik/pro-media/tiskove-zpravy/2019/6/NAIS_eng_korektura_06-19_web.pdf).

This strategy aims to leverage AI for the benefit of the Czech economy and society, positioning the country as a leader in AI innovation. The strategy outlines several key areas:

- Research, Development, and Innovation: Emphasizing the importance of AI research and development to maintain technological progress and competitiveness.
- Education and Skills Development: Focusing on improving digital skills and integrating AI education into the curriculum at various educational levels.
- Ethical and Security Aspects: Ensuring that AI development adheres to ethical standards and addresses security concerns.
- Support for Businesses: Providing subsidies, manuals, and retraining courses to help businesses adopt AI technologies.
- Public Administration: Streamlining public services through the implementation of AI solutions

The Czech Republic is also influenced by the EU Artificial Intelligence Act (AI Act), which provides a harmonized framework for AI regulation across the European Union. The AI Act categorizes AI systems into four risk levels: unacceptable risk, high risk, limited risk, and minimal risk. It sets out specific requirements for high-risk AI systems, including transparency, accountability, and human oversight.

Prohibited AI Practices

Under the AI Act, certain AI practices are prohibited, such as systems that manipulate human behaviour, exploit vulnerabilities, or use real-time biometric identification in public spaces for law enforcement purposes. These prohibitions are designed to protect fundamental rights and ensure the ethical use of AI.

University Policies

Universities in the Czech Republic are aligning their policies with the national strategy and the EU AI Act. Many institutions are incorporating AI ethics and governance into their curricula and research programs. For example, universities are developing new AI-focused degree programs and promoting interdisciplinary research to address the societal impacts of AI.

The implementation of AI regulations in the Czech Republic involves multiple stakeholders, including government bodies, academic institutions, and private sector organizations. The Ministry of Industry and Trade (MIT) plays a central role in coordinating these efforts, ensuring that the AI strategy is effectively executed.

The Czech Republic is committed to fostering a responsible and innovative AI ecosystem. Through the National AI Strategy and alignment with the EU AI Act, the country aims to harness the potential of AI while addressing ethical, legal, and societal challenges. Universities are also playing a crucial role in this ecosystem by advancing AI research and education.

3. Ethics of AI

Alicia Troncoso

Although there is often talk of “ethical AI”, in reality no AI can be ethical, since ethical behaviour requires a conscience. It is more correct to speak of “AI ethics” as part of ethics that addresses the profound ethical dilemmas arising from the potential for AI-based systems to reproduce biases, contribute to climate degradation and threaten human rights, among others. Thus, the ethics of AI is a subfield of applied ethics that studies the ethical issues raised by the development, deployment and use of AI. Its fundamental concern is to identify how AI can improve or raise concerns for people’s lives, whether in terms of quality of life or the human autonomy and freedom necessary for a democratic society.

3.1. Overview of collective and individual implications

3.1.1. At individual level

It is important to think about how AI could involve new challenges in relation to individual human beings. In this context, it is crucial to consider how the concepts of autonomy, identity and dignity of individuals, as well as issues of security and privacy, could change under the influence of AI.
Regarding autonomy, it represents the idea that everyone has the capacity to know what is good or bad for them, to live according to their standards and that they have power over their thoughts and actions, thus promoting individual choice, rights and freedoms. AI might contribute to human development by extending human capabilities or could stimulate productivity and prosperity and lead to active work until a later age.

According to Fearon (1999), identity is considered a social category with its own rules of belonging or “a set of attributes, beliefs, desires or principles of action that a person believes distinguish them in a socially relevant way and of which they are especially proud”. As we interact with AI systems or robots, we need to better understand the potential transformations of how we perceive them and how they shape our conception of the world. For example, profiling, targeted advertising or other AI-based techniques can profoundly affect our identities (Floridi, 2015). The way in which human intelligence and cognitive abilities are affected by interaction with machines and AI is an important area of research today.

Human dignity is the right of a person to be respected and valued and treated in an ethical manner, as mentioned in article 1 of the Charter of Fundamental Rights of the European Union. Individual rights and responsibilities could be lost as a result of the increasing

interaction between humans and machines (EDPS, 2018). At the moment, intelligent devices have no moral responsibility (EGE, 2018). However, although the European Parliament asked the European Commission to consider a specific legal status for robots, at present liability is ultimately related to human liability (EESC, 2016).

In the world of devices, the crucial questions that arise are related to privacy, security and data protection. Meters, toys, refrigerators and smartphones, to mention just a few, have integrated AI systems that send our data to the manufacturers of these devices, often without our knowledge. AI applications in the field of healthcare are even more sensitive, for example, when they suggest diagnoses or treatments (AI Now, 2017) or when insurance companies or technology companies misuse data (Powles and Hodson, 2017). Therefore, the right to the protection of personal information and privacy is increasingly recognized as crucial in European societies (EGE, 2018).

3.1.2. At societal level

AI has broader consequences for society and the well-being of communities around the world. These changes have a direct impact on restructuring of power relations and the transformation of the social contract. It is important to examine these changes in order to mitigate risks and dangers. Several challenges at societal level can be summarized, among others, as follows: fairness and equity, responsibility, accountability and transparency, datafication, democracy and trust and cumulative knowledge.

Regarding fairness and equity, we can distinguish two levels: social inequalities generated by the application of AI in the workplace, and individual level inequalities as a result of unfair AI-driven decision-making. Research has shown that the use of algorithms could discriminate against particular groups or individuals, for instance in the criminal justice system (Chouldechova, 2017) and in recruitment processes (O'Neill, 2016).
Accountability is connected to responsibility, and it is extremely important because it deals with biases and discriminations caused by data mining and profiling. Accountability is seen as a necessary condition for the social acceptability of AI (Mission Villani, 2018). Transparency means not only transparency of algorithms (often referred to as 'black boxes'), but also of data and automated decision making. The importance of making decisions is particularly emphasized when it is used by public agencies, such as predictive policing or risk assessment in the criminal justice system (AI Now, 2018). The decisions made by algorithms may not be understood or explained and it is not clear who is responsible for them.

Researchers in digital politics and media argue that we live in a culture of mass surveillance, characterized by a computerization of people's lives from which it is difficult to escape (Van Dijck, 2014; Rathenau Institute, 2017; Korff et al., 2017). The world's

biggest digital companies control a large part of the online public sphere and hold most personal data.

While dealing with ethical and societal implications of using algorithms, it is crucial to tackle their impact on democracy and the trust of citizens. AI has been used as a political tool to influence voter behaviour, and to manipulate the public by influencing public opinion. The use of political bots to spread false information is now easier than ever. Finally, in the long term, it is also necessary to reflect at all levels of society whether the digital data that constitutes the accumulated knowledge about society should be considered a national asset, similarly to cultural assets and landscapes.

3.2. Main proposals

Many organizations are developing AI codes of ethics with the aim of establishing a sound and coherent regulatory framework that promotes AI technology in the service of stakeholders. In this section we summarize the main proposals on principles and other ethical issues for AI systems.

3.2.1. Rome appeal for IA ethics

The "Rome Call for AI Ethics" is a document signed on February 28, 2020, in Rome, to promote an ethical approach to AI. The underlying idea is to promote a sense of shared responsibility among international organizations, governments, institutions and technology companies in an effort to create a future in which digital innovation and technological progress give humans their centrality. This document is available at: https://www.romecall.org/the-call/

This call addresses three areas of impact: ethics, education and rights, and defines six principles such as transparency, inclusiveness, accountability, fairness, reliability, security and privacy.

3.2.2. European Union

The High-Level Expert Group on Artificial Intelligence set up by the European Commission in June 2018 published "Ethical Guidelines for Trustworthy AI" in 2019 (EU, 2019). These guidelines are available at: https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai .

According to these guidelines, AI reliability relies on three components that must be satisfied throughout the system's life cycle: a) AI must be lawful, i.e., comply with all applicable laws and regulations; b) it has to be ethical, so as to ensure respect for ethical principles and values; and c) it must be robust, both technically and socially, since AI systems, even if the intentions are good, can cause accidental harm.

3.2.3. The Council of Europe

The “2023 Council of Europe Framework Convention on Artificial Intelligence, Human Rights, Democracy and the Rule of Law” established principles and obligations aimed at ensuring that the design, development, use and decommissioning of AI systems are fully consistent with respect for human dignity and individual autonomy, human rights and fundamental freedoms, the functioning of democracy and the observance of the rule of law. These principles were: transparency and oversight (article 7), accountability and responsibility (article 8), equality and non-discrimination (article 9), privacy and protection of personal data (article 10), safety, security and robustness (article 11) and safe innovation (article 12).

3.2.4. The United Nations Educational, Scientific and Cultural Organization (UNESCO)

UNESCO developed in 2021 the first global standard on AI ethics: the “Recommendation on the Ethics of Artificial Intelligence”, adopted by all 193 Member States. This recommendation states that values and principles should be respected by all actors during the life cycle of AI systems in the first place and, where necessary and appropriate, should be encouraged through the modification of existing laws, regulations and business guidelines and the development of new ones. All of this should be in accordance with international law, in particular the United Nations Charter and human rights obligations of Member States, and be in line with internationally agreed social, political, political, environmental, educational, scientific and economic sustainability goals, such as the United Nations Sustainable Development Goals (SDGs). These recommendations are available at:
https://unesdoc.unesco.org/ark:/48223/pf0000381137
In addition, UNESCO in 2023 published an assessment of the ethical impact of AI, which is available at:
https://unesdoc.unesco.org/ark:/48223/pf0000386276

3.2.5. Inter-American Development Bank

In 2021, the Inter-American Development Bank launched an initiative entitled “fAIr LAC”, which consists of five tools for the application of ethical principles in the design, development and necessary audits of AI-based solutions. These tools are an ethical self-assessment tool for the public sector, one ethical self-assessment tool for entrepreneurs, two tools for responsible use of AI for public policy, and a tool for auditing AI-based decision support systems.

3.2.6. Other proposals

In recent years a multitude of proposals about the ethics of AI have emerged. Australia has defined eight ethical principles to ensure that AI is safe, safe to operate and reliable. It is available at: https://www.industry.gov.au/publications/australias-artificial-intelligence-ethics-framework/australias-ai-ethics-principles

The US White House proposed "The Blueprint for an AI Bill of Rights", which is a guide for incorporating a set of principles into policy and practice, including detailed steps to actualize these principles in the technology design process. In 2023, the G7 in Hiroshima proposed a set of international guiding principles for organizations developing advanced AI systems. The Spanish Observatory of the Social and Ethical Impact of Artificial Intelligence (OdiseIA) has published the GuIA framework, one of the most comprehensive that includes several ethical principles applicable to AI. In 2023, IndesIA, the Industrial Spanish Association for the Promotion of the Data Economy and Artificial Intelligence, in collaboration with OdiseIA, has prepared a document that presents the state of the art of AI ethics applied to the industrial sector. In addition, standards such as ISO/IEC 24368, an ethical certification by the IEEE, the IEEE 7000 standard and even maturity models organized according to 5 levels have been proposed in order to evaluate how ethical an IA system is (Krijger, 2023).

4. Practical implications for Higher Education

Alicia Troncoso

4.1. Impact of AI on advanced skills demand

The development of new AI models requires high-level skills in different areas. The basic skills needed to perform cutting-edge work in this field require advanced levels of scientific, mathematical and technical knowledge that are not easy to acquire. In particular, the development of AI methods requires knowledge of statistics, linear algebra, as well as computer architecture and infrastructures and programming languages. The skill set required is scattered and the actual number of people with AI skills is quite small.

It is expected that the high visibility of AI and the current demand will relatively quickly attract talent to the field. For example, since its launch in May 2018, some 90000 students from more than 80 countries have enrolled in the free online six-week course called "Elements of AI" (https://www.elementsofai.com/), which is part of the AI Education program organized by the Finnish Centre of AI (https://fcai.fi/education). Thus, current students of statistics, mathematics, physics and computer science could rethink their career paths and find new opportunities as AI experts. In addition, high-level AI skills may also emerge from unexpected places, for example, through the open software and hardware communities.

As a consequence of this situation, high-level AI talent could be offered as a service, similar to infrastructure or software as a service. This may mean that high-level AI skills will not be needed on a massive scale. Another important issue is also the likely impact that AI may have on the demand for skills for individuals in general which will be addressed in the next section.

4.2. Impact of AI on learning

How interactions with machines and AI affect human intelligence and cognitive abilities is an important area of research. While recent scientific literature has focused on interactions between AI systems and adults, there are important differences when it is children interacting with artificial systems that need to be deeply investigated.

In general, AI can be used in three distinct ways that may have different implications for the development of human cognitive abilities in both children and adults.

First, AI can support already existing competencies. When competencies are understood as combinations of domain-specific expertise, AI can reduce the need for domain-specific knowledge and thus make cross-cutting and generic competencies independent of the application domain more important.
Second, AI can accelerate cognitive development and create cognitive capabilities that would not be possible without technology. Work automation has made things possible that would be impossible without technology. For example, it would be impossible to design a modern microprocessor or neural chip without computer-aided design tools.
Third, AI may reduce the importance of some human cognitive abilities or make them obsolete. For example, because AI can convert speech to text and vice versa, and do mathematical calculations, dyslexia or dyscalculia may become socially less important than they have been in the past. This has obvious advantages for individuals, but it is difficult to predict the impact that AI has in relation to skill duplication. From a pedagogical point of view, it may be more beneficial to use AI to help individuals develop skills to overcome difficulties in reading and counting, rather than using AI to make skills that support important cognitive abilities redundant.

It is also often assumed that AI systems enable new levels of personalization and diversity for information systems. However, much of this personalization is the result of precise categorization that classifies users into predefined classes. While these systems can effectively simulate personalization, they do not necessarily allow for deeper levels of diversity. AI systems can be excellent predictive machines, but this strength can be a major weakness in domains where learning and development are important.

4.3. Impact of AI on teaching

There are some clear opportunities for AI in teaching, such as assessing students in different ways and personalized tutoring systems. AI models can help describe a learner's state of knowledge and learning.

An expert system based on a pedagogical model can manage the introduction of learning materials to the learner through an adaptive and interactive user interface. Student behaviour and learning can also be monitored in great detail in learning management systems (LMS). In addition, intelligent tutoring systems have also been an important source of data for learning research (Porayska-Pomsta 2015).
Because AI supervised learning algorithms are based on historical data, they can only see the world as a replay of the past. This has profound ethical implications. When, for example, students and their achievements are assessed by these AI systems, the assessment is necessarily based on criteria that reflect cultural biases and historically relevant measures of success. Supervised learning algorithms create unavoidable biases,

which are currently the subject of extensive debate.

Rapid advances in natural language processing (NLP) and AI-based human-machine interfaces will also generate new pedagogical possibilities. For example, learning by teaching machines shows clear potential, while real-time machine translation also opens up new possibilities in language learning.

While it is possible to imagine many exciting possibilities for AI in education, without clear policies regarding the emerging technical possibilities in the broader context of transforming education and the future of learning, educational AI is likely to be offered primarily as a solution to existing problems. Thus, instead of using AI to renew the system and orient it to the needs of a post-industrial economy and knowledge society, AI will only be used to automate and reinvent obsolete pedagogical practices making them increasingly difficult to change. It is therefore crucial to look beyond current practices and ask fundamental questions about what competencies and skills are needed in a digitally transformed society and how we should teach them.

B. Interviews report

1. Interviews about AI usage

María Elena de La Cova Morillo-Velarde, Carlos Salvador Silva Perea

As part of the objectives of the AI4UNI project, interviews were conducted with representatives from universities that are not participants in the project but are interested in sharing and/or acquiring knowledge about AI implementation in higher education institutions (HEIs). To this end, the SGroup conducted interviews with ten universities, each including a representative from one of the project's three target groups or clusters (academic staff, administrative staff, and students). In total, 30 interviews were carried out.

The participating universities were: Aarhus University (Denmark), Justus Liebig University Giessen (Germany), Kaunas University of Technology (Lithuania), University of Los Andes (Colombia), University of Minho (Portugal), University of Lille (France), Stellenbosch University (South Africa), University of Westminster (United Kingdom), University of Catania (Italy), and Ghent University (Belgium).

1.1. Purpose of the interviews

The interviews aimed to explore how members of the three target groups (clusters) use and perceive AI, to identify their main concerns and potential training gaps, and ultimately to inform the design of future training activities within the AI4UNI project. In particular, the main topics covered in the interviews were the following:

- General knowledge about AI.
- AI use.
- Usefulness of the AI use.
- Concerns about the use of AI.
- Awareness about potential institutional guidelines about AI use.

Questions varied depending on the university cluster to which the interviewee belonged.

1.2. Analysis approach

To analyse the content of the interviews, the qualitative data analysis software ATLAS.ti was used. This tool enables researchers to systematically categorize (code) textual data, organize themes, and identify patterns and relationships within the responses, facilitating the extraction of rigorous, evidence-based conclusions.

Based on the topics covered in the interviews and on an initial review of the interview transcripts, tentative code categories were developed inductively, which evolved considerably as the analysis progressed. The final set of code categories used to analyse the interview transcripts is as follows:

Codes	Description
Scenarios	Situations or contexts in which AI tools can be used.
Frequency of use	Frequency of AI use reported by an interviewee.
Trustworthiness	Reported levels of confidence in AI-provided results
AI knowledge	Demonstrated AI knowledge and training needs
AI guidelines & permission	Awareness about institutional AI use policies.
Reported usefulness and Advantages	Perceived usefulness of AI tools or lack thereof
Concerns about the use of AI	Concerns expressed regarding the use of AI
Reported AI tools	Explicit AI tool identification

Figure 1. Codification of the interviews.

2. Results and Key Insights

María Elena de La Cova Morillo-Velarde, Carlos Salvador Silva Perea

This section presents the interview findings, with particular attention to the general patterns observed within each of the three university groups (Findings by cluster) and the differences identified across clusters and thematic codes (Mapping intersections across themes and clusters).

2.1. Findings by cluster

2.1.1. Students

Students reported having comparatively lower knowledge of AI and, consequently, expressed a greater need for training compared to the other clusters, which they identified as a concern. Although their trust in AI tools is limited, they still consider them useful, primarily for studying (content assimilation, exam preparation), brainstorming support (idea generation), and writing assistance (academic assignments, grammar corrections) (see AI uses in Figure 2).

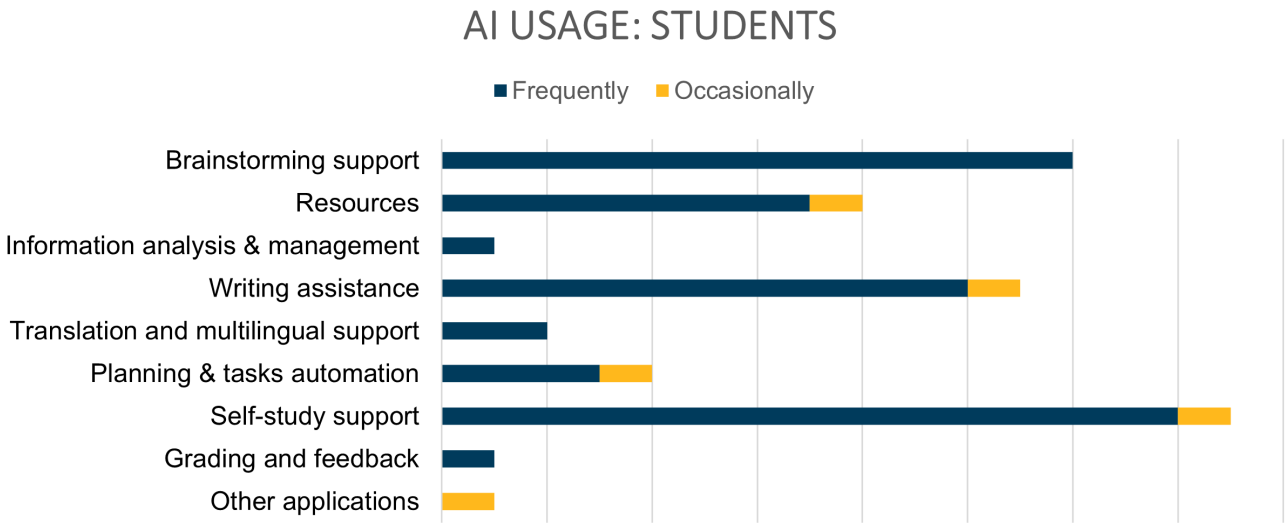

Figure 2. AI Usage for students.

Their main concern regarding the use of AI is that it may produce unreliable output (hallucinations, unverified content, biases). Regarding this, one student stated in an interview:

> "it can't, it doesn't always get the accurate, like the key details, especially in law. So if you use AI in law, it pulls out like things that sound good, but misses the core

> themes or the core point. So then it actually doesn't help you if you just rely on it.”.

Similarly, another student noted:

> (...) “I think it's when you are looking for literature, some of the references that AI would generate (especially I saw that with ChatGPT) are not… they don't exist. They just quote very random books or authors. And then when you Google them up, you… well, they're not available, those texts are not available. That person doesn't even exist.”

This last concern was echoed across clusters.

Students also raised concerns about AI’s potential negative impact on learning processes, particularly cognitive offloading. Students reported experiences such as: “sometimes I have the feeling that my skills and my knowledge decreased since I use it, because I think, okay, I will ask AI, so I don't have to think by myself”, “it can make you kind of more lazy, where if you always get AI to do the things, then maybe you lose your own thought of the process”, or the following:

> “it doesn't engage your brain. So it's too easy to offload or there's this cognitive offloading. So it's too easy to just offload your work to AI and have it do it for you. And you haven't developed any of the skills that you're meant to in that task”.

The main AI tools reported by students are: Cecile AI, ChatGPT, Copilot, DeepSeek, DeepL, Grammarly, HyperWrite, NotebookLM, PDF AI. These tools indicate that their use of AI is closely tied to study support, writing assistance, and managing course materials. General-purpose models such as ChatGPT, Copilot, DeepSeek, and Cecile AI are widely used for clarifying concepts, generating explanations, and assisting with problem-solving across subjects. Tools like Grammarly and HyperWrite show that students may rely on AI for improving written assignments, refining grammar and style, and drafting essays or reports. The inclusion of DeepL reflects frequent use for translating course readings or communicating in academic contexts, while PDF AI and NotebookLM suggest that students probably employ AI to summarise readings, extract key points from documents, or organise study materials.

2.1.2. Academic Staff

Professors demonstrated the highest level of knowledge about AI among the clusters, including an understanding of what trustworthiness entails, and expressed the lowest need for training—an expected finding given the previously mentioned factors.

They generally perceive AI as helpful, particularly for improving workflow efficiency and reducing time and costs. However, it is noteworthy that many also commented on AI's limitations, describing it as sometimes ineffective or unreliable. Academic staff reported using AI tools primarily for finding resources, providing writing assistance, supporting translation and multilingual tasks, and generating course-related content.

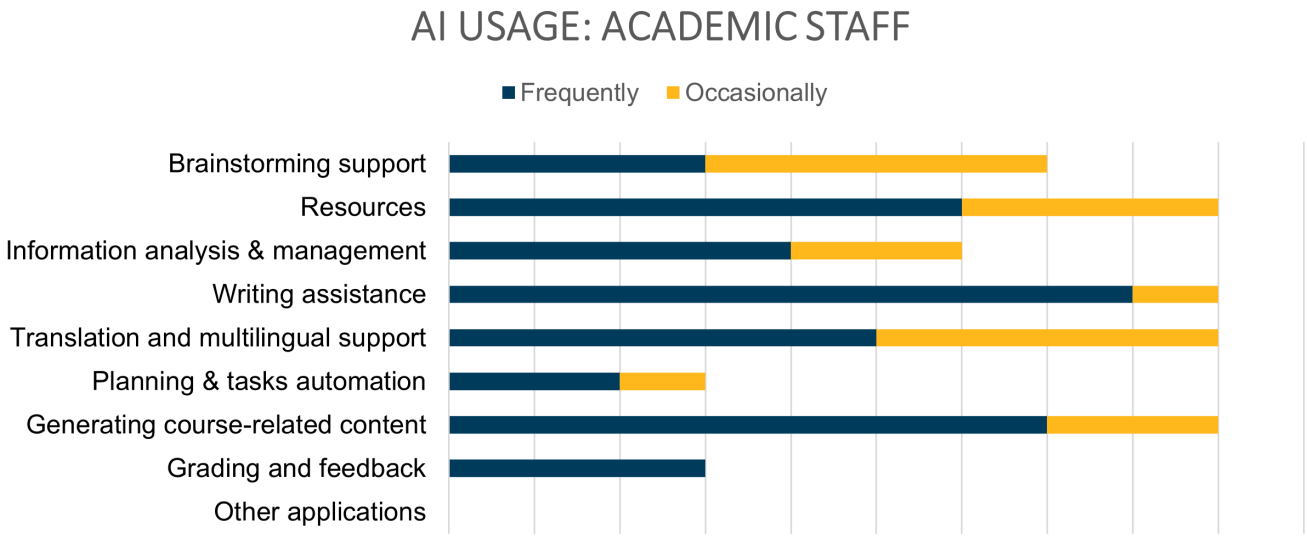

Figure 3. AI Usage for academic staff.

Regarding their main concerns about AI, the most common issue by far was the unreliability of AI output. Many participants mentioned hallucinations and stressed the need to verify any academic references provided by AI chatbots—for example, by checking them in Google Scholar. A second major concern was the impact of AI on learning processes. Several interviewees emphasized that when people rely on AI to perform various tasks—especially during training or skill-development stages—they miss important learning opportunities. As a result, their ability to understand and evaluate whether an AI-generated output is correct or accurate may be diminished. One professor noted the following:

> "I see it as a problem mostly for students because I teach in computer engineering courses where they do programming (...) and I noticed that there is an increasing trend in having the AI write the code for them instead of writing that themselves.

> And that reflects on their ability to write the code, on their ability to understand the code, and in general on their ability to actually learn the new concepts. They start relying on the AI a lot without actually learning things."

Data privacy also emerged as a significant concern for academic staff. In this respect, several participants mentioned their caution when handling personal data (mainly students' data) in AI tools, for example when evaluating or grading student work. Interestingly, some academics also pointed to uncertainty about "what is sensitive and what is not in a given context," including in relation to the use of customized chatbots. Such uncertainty suggests a clear need for further training and a more robust understanding of institutional AI regulations.

The main AI tools used by academic staff according to their responses are: ChatGPT, Claude, Consensus, DeepSeek, DeepL, Elicit, Gemini, Google Translate, Grok, Mistral, OpenAI, Scopus AI, Writefull. This list suggests that professors are integrating AI across both teaching and research in a pragmatic and task-oriented way. General-purpose language models such as ChatGPT, Claude, Gemini, Grok, DeepSeek, and Mistral indicate widespread use of AI for drafting materials, generating explanations, summarizing texts, and supporting reasoning-intensive tasks. At the same time, the inclusion of research-specific tools like Elicit, Consensus, Scopus AI, and Writefull shows that faculty are increasingly relying on AI to streamline literature discovery, analyse academic sources, and enhance the clarity and quality of scholarly writing. The presence of translation tools such as DeepL and Google Translate further suggests that AI is being used to support multilingual communication and access to international research.

2.1.3. Administrative Staff

Administrative staff, together with students, showed a lower level of understanding of what trustworthy AI entails, although they reported feeling confident about what AI is. They nevertheless expressed a high need for training.

Given the nature of administrative work, it is not surprising that participants emphasized AI's usefulness for reducing time and costs—this being the highest-rated benefit among the clusters. It is also expected that their main uses of AI tools involve *planning and task automation*, such as managing appointments and emails or generating automated transcripts of meeting notes. Another common use is *information analysis and management*, particularly for processing data or creating statistics. AI tools also appear to be useful in enrolment processes for "certain basic tasks like initial screenings of

applications, initial screenings of data, etc." One participant explained their use of AI as follows:

> I used to interview like so many students for scholarships. So, I used to process all the information for the scholarship. So, in the moment of interviewing students, I'll be able to organize everything. But then, I used to record some interviews. I send the recording to TurboScribe. Then, I upload the script of the interviews to ChatGPT. So, it helps me just like organizing all the information of my students. (...) Let's say students from Europe, from the States, from South America. ChatGPT helps me just like making an equivalent between the letters they send me and also the transcripts they have, like really everything. So, I just send everything to the chatbot I have already prompted.

The third main use of AI tools was for *translation and multilingual support*.

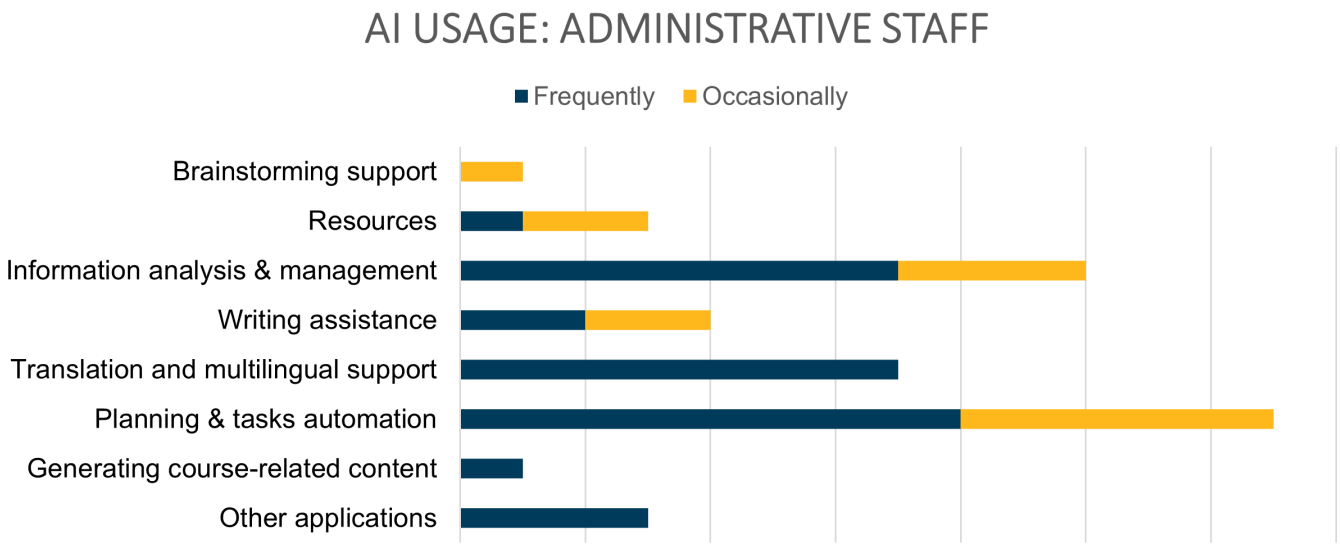

Figure 4. AI Usage for administrative staff.

Due to the fact that administrative staff handle, for example, student applications and staff personal information, their main concern when using AI tools relates to data privacy, as illustrated in the following quote:

> (...) we manipulate a lot of sensitive data, either sensitive because there are personal data about our staff members, our students, or sensitive because research sensitive data should be secret at that moment and open afterwards. But before being open, the data are closed and should stay closed. So using a system where you

> know that your data is then sent to the provider of the system and used for training other AI systems afterwards is a big problem for sensitive data. So yeah, we warn people about that, and we forbid the use of AI systems for sensitive data.

Almost as important is their concern about the unreliability of AI outputs. Many participants from the administrative cluster stated that they need to double-check the results produced by these tools to ensure that no errors are made. For example, one participant noted:

> it happened to me once (...) that I asked for some information about some universities and it gave me names and departments that didn't exist. Definitely, yes. I was checking the information at that time, luckily. But yeah, these kinds of hallucinations were the ones that I didn't trust.

Third, interviewees expressed concern about their need for further training, which reinforces the initial training needs they had already identified.

The main AI tools used by administrative staff are: Adobe Firefly, ChatGPT, Claude, Consensus, Copilot, DeepL, Descript, Firefly AI, Gemini, O3 Pro, O4, OpenAI, Speechly, Synthesia, Transcriptor AI, TurboScribe. The list indicates that their adoption of AI is strongly oriented toward practical, task-specific support functions. While general-purpose language models such as ChatGPT, Claude, Gemini, and OpenAI's O-series are used, a large portion of the tools reflect the operational demands of administrative work. Applications like Descript, Speechly, TurboScribe, Transcriptor AI, and Synthesia indicate a significant reliance on AI for transcription, note-taking, meeting documentation, and the creation of audiovisual materials. Tools such as Adobe Firefly and Firefly AI point to the use of generative AI for producing visual assets, presentations, and design materials needed in administrative communication. Meanwhile, translation and language-related tools like DeepL and Consensus support multilingual correspondence and information processing.

2.2. Mapping intersections across themes and cluster

2.2.1. AI use across clusters

In order to analyse the interviews regarding the types of use that participants made of AI tools, the following sub-categories were used to code their answers: Resources, Writing Assistance, Brainstorming support, Information analysis and management, Translation and

Multilingual Support, Planning/Tasks Automation, Self-study Support, Generating course-related content, Grading and Feedback, Other applications.

As shown in the following pie chart (figure 5), the main use of AI across clusters is *writing assistance*, which is also among the top three uses for both students and academic staff, as reported in the cluster-specific results above. There appears to be a tie between three other AI uses: *brainstorming support*, *resources*, and *planning and task automation*. The differences in usage percentages compared with other uses, such as *information analysis and management* or *translation and multilingual support*, do not appear to be significant.

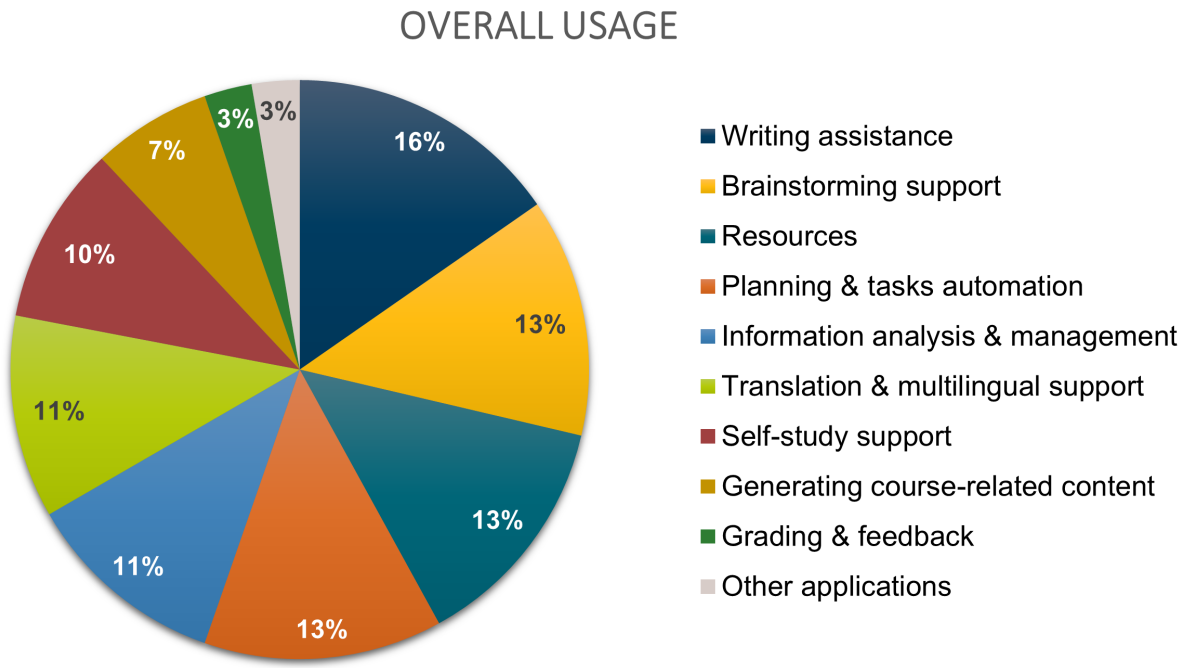

Figure 5. Overall usage of AI.

2.2.2. Frequency of use

As shown in the following graph (Figure 6), *writing assistance* emerges as one of the most frequent AI uses across clusters followed by *brainstorming*, *resources*, and *self-study support*. These patterns are consistent with earlier findings: students and academic staff tend to rely on AI primarily for learning-related tasks, idea generation, and content development, reflecting their focus on improving academic performance and enhancing the quality of their outputs. Regarding less frequent uses, participants mentioned AI tools being used occasionally for *planning and task automation*, *resources*, and *information analysis and management*. Interestingly, although *planning and task automation* was the least frequent AI use across clusters overall, it was one of the main uses of AI tools

reported by administrative staff. This aligns with their role-specific priorities, as administrative tasks often involve scheduling, workflow coordination, and repetitive processes that can be streamlined through automation.

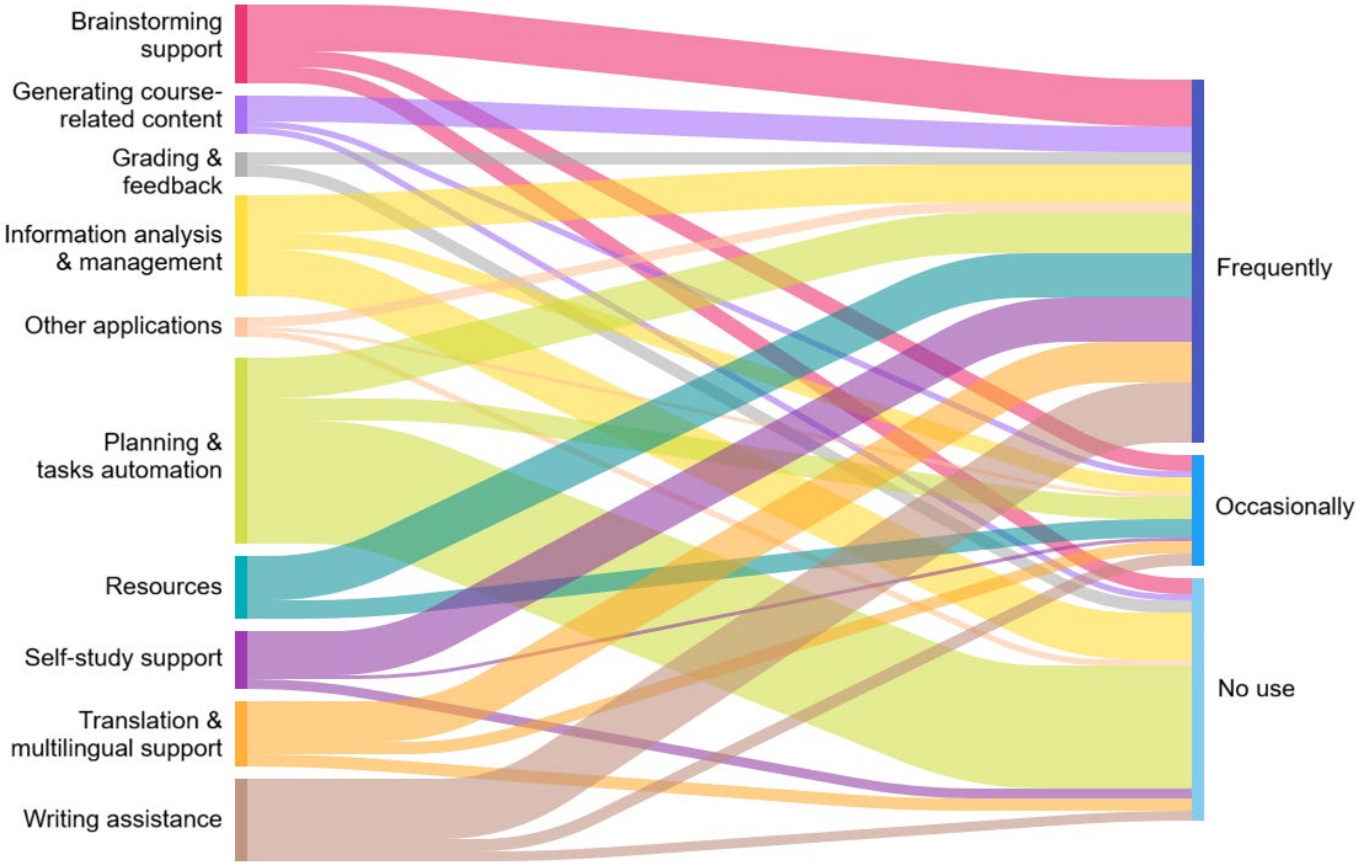

Figure 6. Frequency of use.

2.2.3. Knowledge about AI

To assess participants' knowledge about AI, three sub-categories were used when coding their responses: Basic AI Knowledge, Knowledge about Trustworthiness, Training need identified.

Both academic and administrative staff demonstrate a comparable baseline understanding of AI, as illustrated in the graph below. In terms of training needs, administrative staff and students express similar concerns, as reflected in the comparable number of quotations coded under this category.

However, when it comes to knowledge about trustworthy AI, academic staff appear to possess a more developed grasp of the concept than the other two groups. According to the DigComp 3.0 European Digital Competence Framework, trustworthy AI has three components:

(1) it should be lawful, ensuring compliance with all applicable laws and regulations

> (2) it should be ethical, demonstrating respect for, and ensure adherence to, ethical principles and values and (3) it should be robust, both from a technical and social perspective, since, even with good intentions, AI systems can cause unintentional harm. Trustworthy AI concerns not only the trustworthiness of the AI system itself but also comprises the trustworthiness of all processes and actors that are part of the system's life cycle.

Evidence from participants' statements illustrates these differences. Academic staff often referred to issues such as bias and data limitations, which are central to robustness and ethical considerations:

> So it is about understanding the limitations, which in my opinion is often a limitation of the data. Also understanding if the data may eventually be biased. (...) So if there's only a certain perspective in the data that the AI will summarize, there will be that exact perspective in the produced report. If that is fine, then it's fine. But if not, and this is so preparing to work, understanding the data that you supply, understanding the shortfalls or the limitations or the bias that you introduce yourself by supplying that information to the AI, and then taking that into consideration for the outcome.

Administrative staff, by contrast, emphasized transparency and data handling practices, showing awareness but in a more operational sense:

> I look into the settings and see where they got their data… what they do with my data? What will they do with my data?

Students' responses were generally narrower, focusing on security and source reliability:

> So maybe if it's data secure and if the information that you get from this AI, you can rely on it and maybe it shows you which source it used to answer you.

These patterns suggest that while all groups recognize aspects of trustworthiness, academic staff demonstrate a deeper and more nuanced understanding of the concept compared to administrative staff and students.

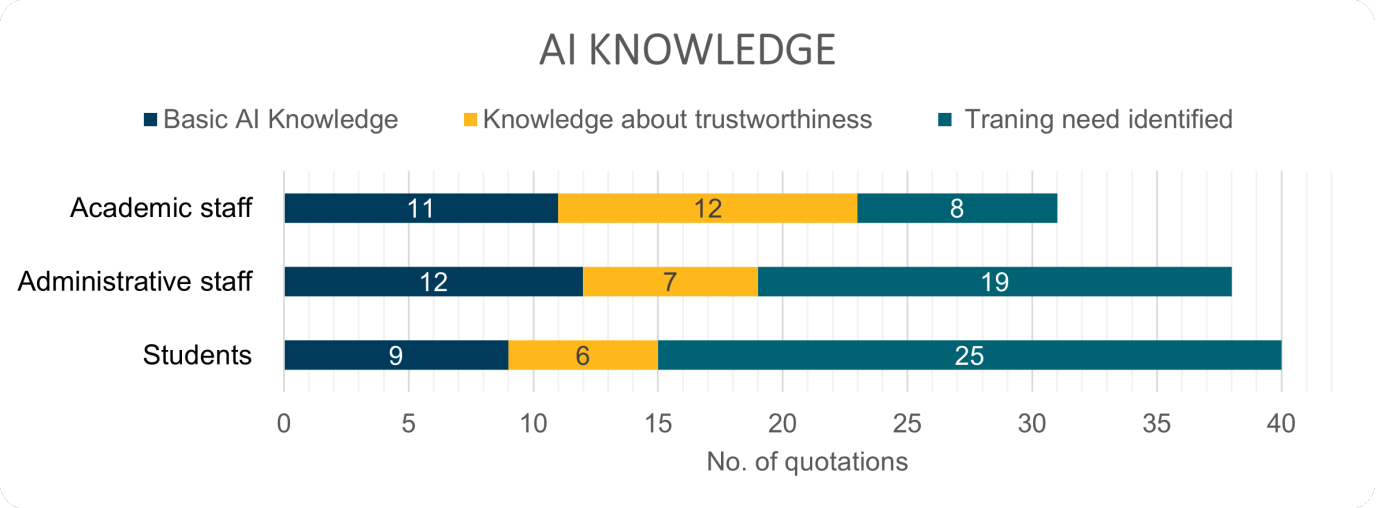

Figure 7. AI Knowledge.

2.2.4. Trustworthiness

To evaluate the trustworthiness attributed to AI by the different clusters, the following sub-categories were used: blind trust, conditioned-trust and mistrust.

Across all clusters, *conditioned trust* accounted for the largest share of trust-related quotations — representing 51.2% among academic staff, 59.5% among administrative staff, and 65.7% among students.

Instances of *blind trust* were comparatively rare, particularly among students (2.9%), while *mistrust* appeared with moderate frequency across clusters (31–37%).

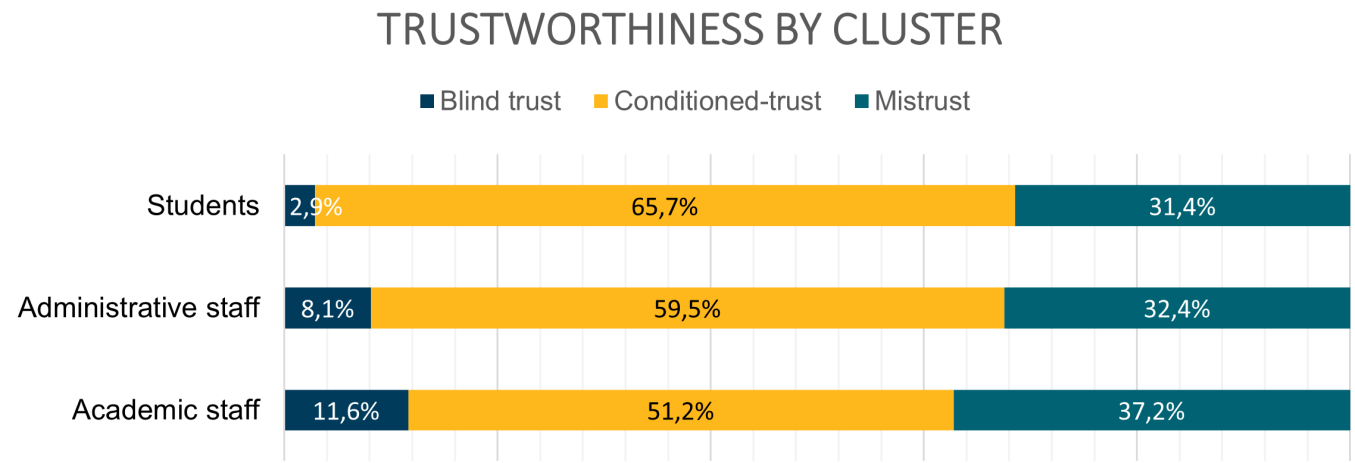

Figure 8. Trustworthiness by cluster.

2.2.5. AI Guidelines and permissions

A large number of participants appeared uncertain or unaware of the institutional guidelines on AI use within their universities. Even when they believed that some form of guidance existed, they were often unsure about its scope or implementation—an issue

that emerged even among members of the same institution.

Key points:

- Status of institutional guidelines on AI use:
  - Eight of the ten interviewed universities appear to have implemented some form of AI-related guidelines, either general or specifically targeting student use (the latter being more common).
  - Two university representatives indicated that their institutions are currently developing guidelines.
  - One institution reported having no formal guidelines in place.
- Students generally reported that they are permitted to use AI as a supplementary tool—comparable to the use of search engines—but are prohibited from submitting AI-generated work as their own. They are also required to clearly acknowledge any use of AI in academic assignments. As one student explained:

  > So the main approach is that you must declare the use of AI, that you cannot pass it off as your own. So it constitutes plagiarism. If you try and pass off AI work on its own. I think, yeah, those are the general guidelines that you basically have to cite that you've used that will be transparent in your use and you cannot pass it off as your own work.
- Some institutions explicitly promote the use of certain AI tools, such as Grammarly, to improve academic writing quality.
- Overall, the responses underscored the importance of ethical considerations in the academic use of AI. These include the requirement to credit AI-generated content to avoid plagiarism, and concerns around data privacy—particularly the expectation that personal data should not be shared with AI systems hosted outside Europe.

2.2.6. Reported usefulness and advantages

To assess how the university community perceives the usefulness of AI tools, the following sub-categories were used to code participants' responses: Lack of usefulness, Generic usefulness, Enhanced quality in human production, Cost and time reductions, Workflow improvement and enrichment, Other advantages.

The data on perceived usefulness reveal distinct priorities across clusters:

- Academic staff value AI primarily for *workflow improvement and enrichment*, indicating that they see these technologies as tools to enhance the quality and efficiency of their professional tasks. Interestingly, the second most reported category—*lack of usefulness*—suggests that a portion of academic staff remain

skeptical or encounter limitations when applying AI in academic contexts.

- For administrative staff, the highest perceived usefulness is linked to *cost and time reductions*, which reflects the efficiency-driven and process-oriented nature of their work. *Workflow improvement and enrichment* also ranks highly for this group, while *lack of usefulness* appears less frequently but remains present.
- Among students, the main perceived benefit of AI also lies in *cost and time reductions*, followed by *generic usefulness*, which points to a broader, less task-specific appreciation of AI tools. *Workflow improvement and enrichment* is the third most frequent category, suggesting that students recognize AI's potential to support learning processes, though its perceived value is overshadowed by its efficiency-related benefits.

Overall, these patterns indicate that perceptions of AI usefulness closely align with the distinct roles and responsibilities of each cluster: academic staff prioritize quality enhancement, administrative staff focus on efficiency, and students emphasize general support and time-saving advantages.

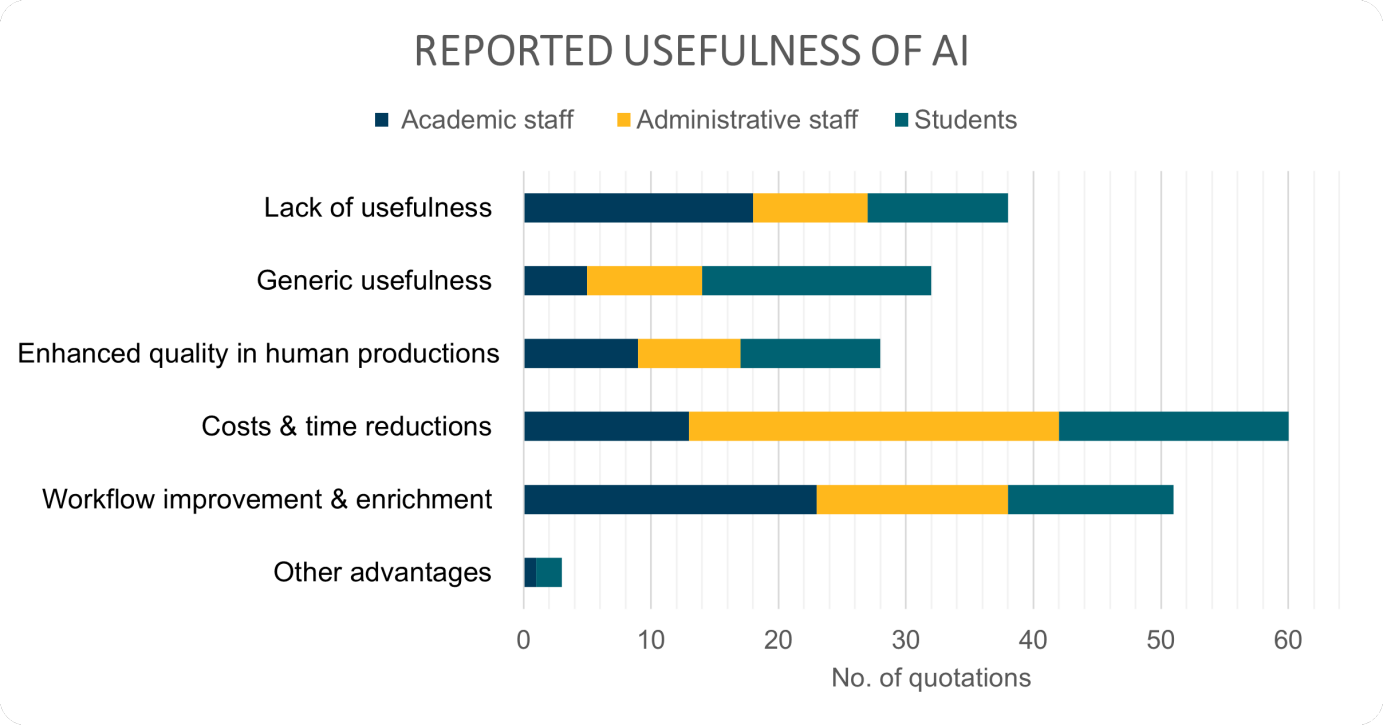

Figure 9. Reported usefulness of AI.

2.2.6.1. Cross-Analysis of Perceived Usefulness and AI Uses

As shown in Figure 10, the analysis of interview responses reveals differing assumptions about how participants attribute usefulness to the various activities they carry out using AI tools:

- Using AI for brainstorming support is strongly associated with *workflow improvement and enrichment*, suggesting that participants view AI-generated ideas as helpful for expanding or structuring their work.
- Writing assistance is connected with several advantages simultaneously—including *enhanced quality in human production*, *cost and time reductions*, and *workflow improvement and enrichment*—indicating that this use case is perceived as broadly beneficial across clusters.
- The use of AI for planning and task automation is closely linked to *cost and time reductions*, confirming that participants largely view automation as a means of increasing operational efficiency.
- Translation and multilingual support is primarily associated with *enhanced quality in human production*, something students specifically highlighted. Interestingly, this use case also shows parallel links to *workflow improvement and enrichment* and *lack of usefulness*, suggesting inconsistent experiences or expectations among participants.
- The use of AI tools for obtaining or consulting resources shows the highest rate of *perceived lack of usefulness*. This may indicate a knowledge or experience gap between participants who perceive clear advantages in these tools and those who consider them of limited value.

2.2.7. Concerns about using AI

To thoroughly evaluate the concerns, the following sub-codes were used: Unreliable output, AI Impact on Learning, AI Impact on environment, Unethical Use of AI, Data Privacy, Training needs, Human Replacement, Other Concerns.

The following graph (Figure 10) expands on the information provided above for each cluster. As shown in the figure, *unreliable output* emerges as the main concern across all groups, indicating a widespread lack of trust in the accuracy and validity of AI-generated information. This aligns with the qualitative findings discussed earlier, where participants frequently mentioned hallucinations, fabricated references, and the need for constant verification of AI outputs.

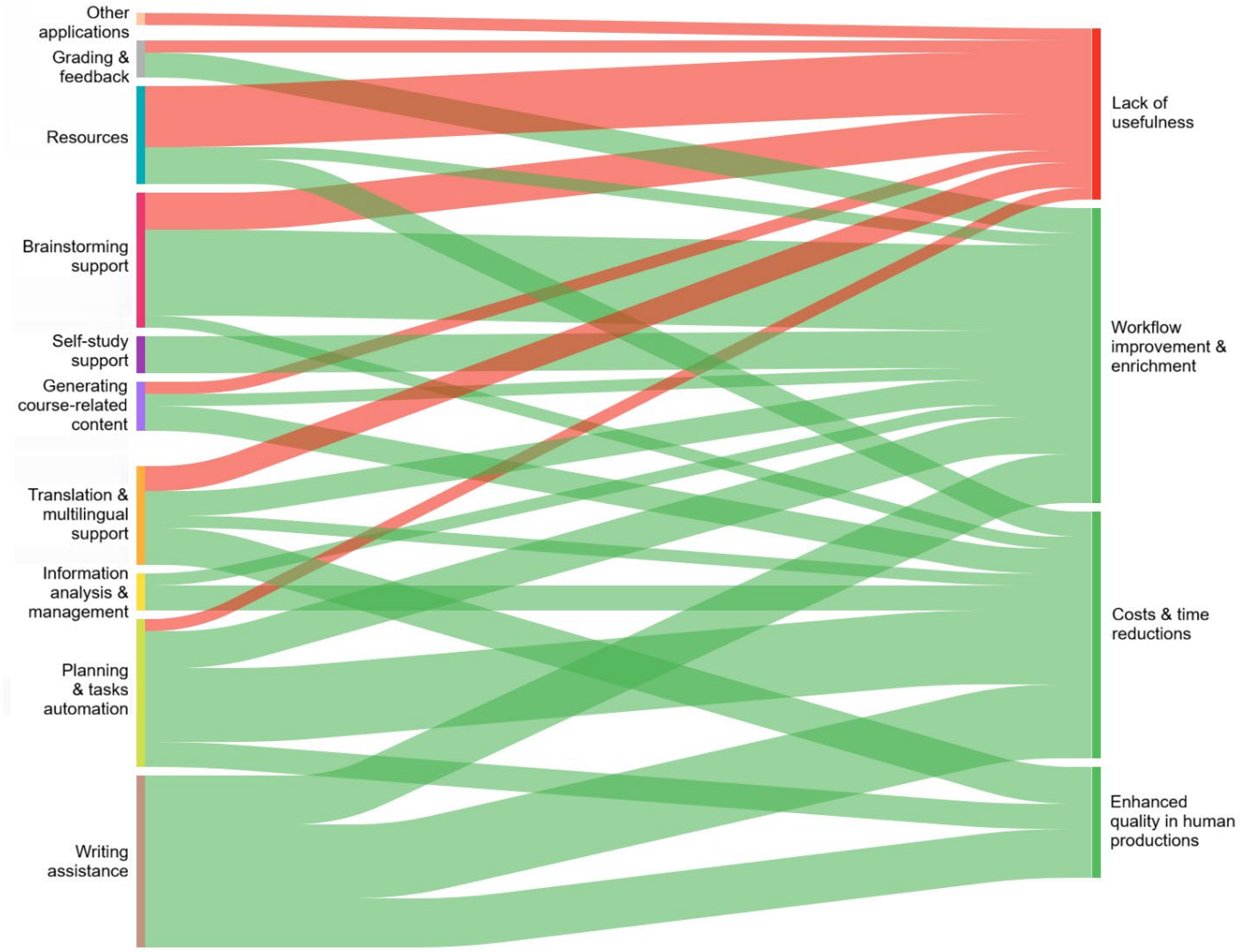

Figure 10. Perceived usefulness of AI.

The second most prominent concern is *data privacy*, which reflects the university community's heightened awareness of the risks associated with handling personal, sensitive, or institutional data through external AI tools.

Training needs also appear as a relevant concern. This suggests that, despite growing exposure to AI, members of the university, in particular students and administrative staff, still feel insufficiently prepared to use these tools effectively and responsibly. The identification of this need is consistent with the variability in AI competence observed in the interview responses.

It is also significant that the impact of AI on learning emerges as the fourth major concern across clusters. Its presence as a shared concern—almost equally raised by academic staff, administrative staff, and students—highlights an underlying tension: while AI tools offer efficiency and support, they may also undermine critical thinking, skill development, and learning autonomy. This illustrates an important point about the university community's

core mission: the generation, transmission, and safeguarding of knowledge. Concerns about the erosion of learning processes and cognitive offloading signal an awareness that AI adoption must be balanced with pedagogical integrity and long-term educational goals.

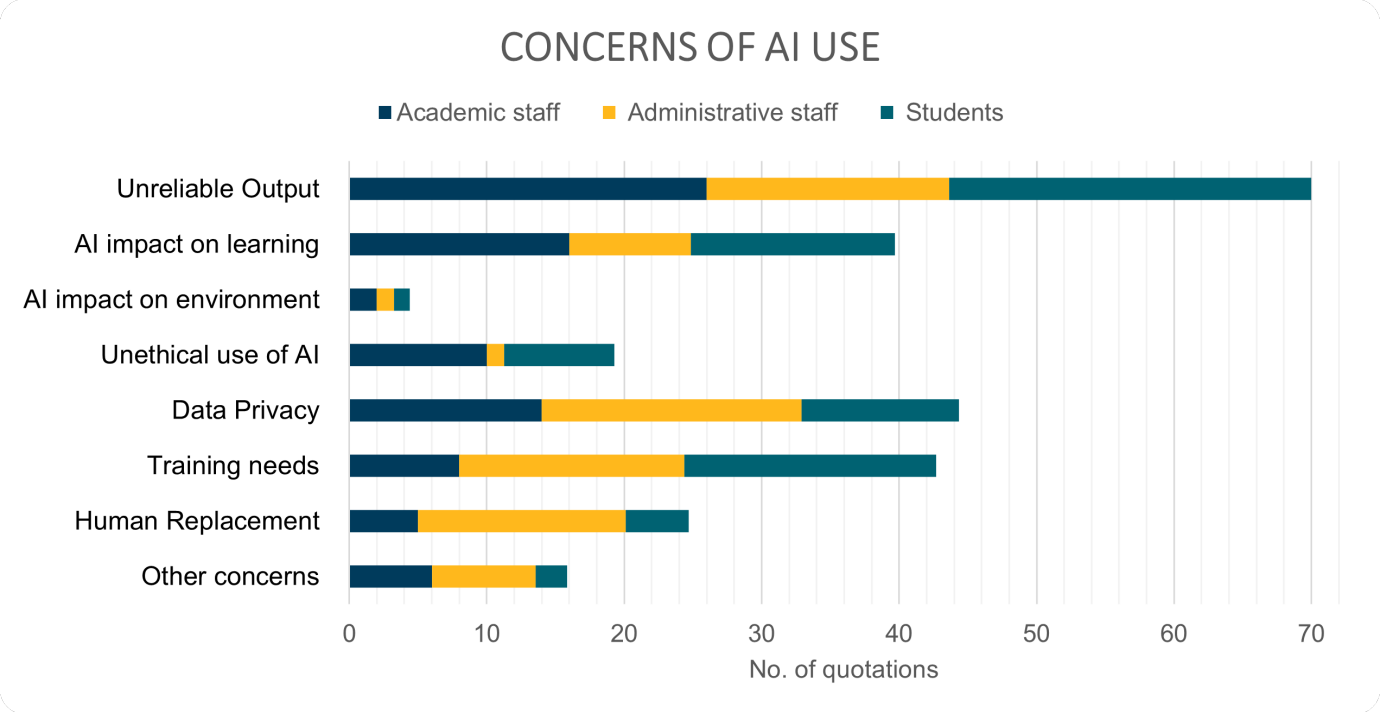

Figure 11. Concerns of AI use.

2.2.7.1. Concerns by age-group

Thus, section provides a nuanced cross-analysis that complements the broader findings detailed in the general concerns about using AI where *unreliable output* and *data privacy* emerged as the most critical issues. As illustrated in the figure 12, these worries often align with the specific roles and developmental stages of the three clusters; for instance, students (typically representing the younger demographic) express significant anxiety regarding cognitive offloading and their own training needs, while academic staff focus more deeply on the impact of AI on learning and the erosion of critical thinking skills. By relating these age-specific patterns to other epigraphs such as ”Knowledge about AI” and ”Reported Usefulness and Advantages”, the report highlights that while the perceived usefulness of AI is high across all generations, the specific nature of the "conditioned trust" identified in the trustworthiness analysis is heavily influenced by the professional and academic responsibilities inherent to each age group. Ultimately, this integrated view allows the AI4UNI project to identify specific training gaps and tailor future educational activities to the distinct concerns of the entire university community.

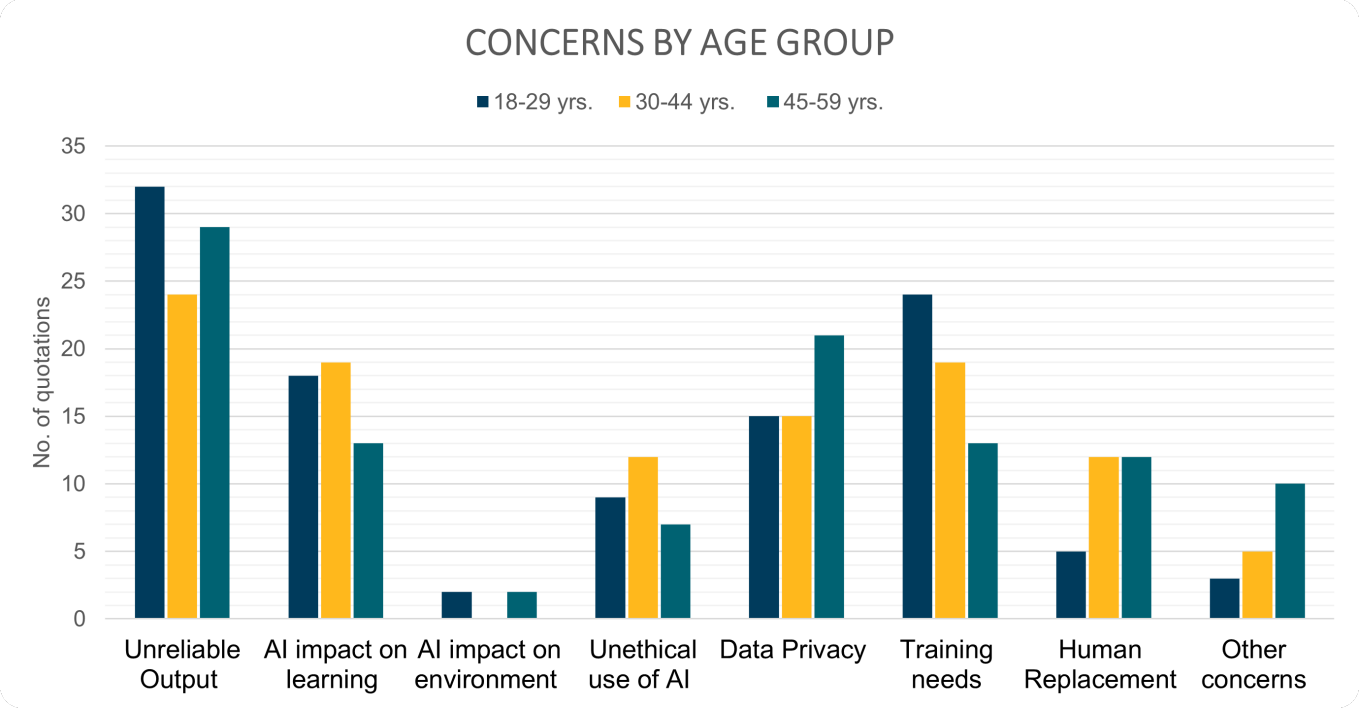

Figure 12. Concerns by age group I.

3. State of the Play: Conclusions

María Elena de La Cova Morillo-Velarde, Carlos Salvador Silva Perea

This report confirms that AI is rapidly increasing its influence across all sectors, including higher education, bringing both extraordinary benefits and significant risks. Internationally, there is a recognized global governance deficit concerning AI, underscoring the necessity for responsible and trustworthy intelligent systems.

In terms of Legislation and Ethics, the European Union (EU) is leading the way by implementing the AI Act, which establishes a uniform legal framework based on a risk-based approach (categorizing systems into unacceptable, high, limited, and minimum risk). The project consortium countries—Spain, Poland, Türkiye, and Czechia—are actively working to develop national AI strategies and align their proposed regulations with the EU's requirements. Key ethical proposals from organizations like the EU and UNESCO emphasize that trustworthy AI must be lawful, ethical, and robust.

The analysis of the 30 interviews conducted by SGroup as part of the ERASMUS+ AI4UNI project (specifically addressed in WP2 of the project) reveals several critical findings. The interviews aimed to explore how members of the three target groups (academic staff, administrative staff, and students) use and perceive AI, identifying concerns and training gaps to inform future AI4UNI training activities.

First, regarding usage and concerns, the most common use of AI across all university groups is writing assistance. However, the primary concern is the unreliability of AI output—such as hallu

cinations and fabricated references—which has resulted in a widespread sentiment of conditioned trust rather than blind acceptance. Data privacy and the negative impact of AI on learning (cognitive offloading) are also significant concerns.

Second, knowledge and training: While academic staff demonstrated the highest level of AI knowledge, both students and administrative staff expressed a high need for further training. Additionally, many participants were uncertain or unaware of their institution’s specific AI guidelines, revealing gaps in policy communication and awareness of implementation.

Finally, perceived usefulness: Perceptions of AI usefulness align closely with professional roles. Administrative staff value AI most for cost and time reductions, academic staff prioritize workflow improvement and enrichment, and students focus on general support and time-saving advantages.

References

Ignatieff, M. (2023). “Hope for troubled times”. Program Tech & Society, Telefónica Foundation.

UN (2024). Governing AI for Humanity (AI Advisory Board). United Nations, September

Schwartz, R., Vassilev, A., Greene, K., Perine, L., Burt, A., Hall, P. (2022). “Towards a Standard for Identifying and Managing Bias in Artificial Intelligence”
NIST Special Publication 1270. https://doi.org/10.6028/NIST.SP.1270

EC (2016a). European Commission (EC), 2016a. 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), OJ L 119.

Poullet, Y., (2018). “Is the general data protection regulation the solution?”, Computer Low & Security Review. The International Journal of Technology Low and Practice. Vol. 34, no. 4, pp. 773-778.

Delforge, A., Gerard, L. (2017), “Notre vie privée est-elle réellement mise en danger par les robots?: Étude des risques et analyse des solutions apportées par le GDPR”, L'intelligence artificielle et le droit, pp. 143-188, Larcier, Brussels.

Wachter, S., Mittelstadt, B., Floridi, L. (2017). “Why a Right to Explanation of Automated Decision-Making Does Not Exist in the General Data Protection Regulation?”, International Data Privacy Law. http://dx.doi.org/10.2139/ssrn.2903469

LNT, (2019). Report on Liability for Artificial Intelligence and other emerging technologies. https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=63199

EPSC (2018). European Political Strategic Centre (EPSC). Election Interference in the Digital Age: Building Resilience to Cyber-Enabled Threats. https://digital-strategy.ec.europa.eu/en/events/election-interference-digital-age-building-resilience-cyber-enabled-threats

EU, (2021). Coordinated Plan on Artificial Intelligence 2021 Review. https://digital-strategy.ec.europa.eu/en/library/coordinated-plan-artificial-intelligence-2021-review

EC, (2021a). 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.
https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence

EC, (2021b). Proposal for a Regulation of the European Parliament and of the council on machinery products.
https://ec.europa.eu/commission/presscorner/detail/es/ip_21_1682

EC, (2020). Ethics guidelines for trustworthy AI.
https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai

IAPPAI, (2024). Global AI Governance Law and Policy. International Association of Privacy Professionals (IAPP).

CMA, (2023). Competition & Markets Authority (CMA). AI Foundation Models Initial Report.
https://assets.publishing.service.gov.uk/media/650449e86771b90014fdab4c/Full_Non-Confidential_Report_PDFA.pdf

AccessNow, (2024). Radiografía normativa: ¿dónde, qué y cómo se está regulando la inteligencia artificial en América Latina? Access Now. https://www.accessnow.org/wp-content/uploads/2024/02/LAC-Reporte-regional-de-politicas-de-regulacion-a-la-IA.pdf

Leon Coronado, C., (2023). La carrera por la regulación de la inteligencia artificial. Revista Latinoamericana de

Economía y Sociedad Digital, 4.

AISS, (2020). Artificial Intelligence Spanish Strategy. https://portal.mineco.gob.es/es-es/ministerio/areas-prioritarias/Paginas/inteligencia-artificial.aspx

Fearon, J. (1999). What is identity (as we now use the word)? https://web.stanford.edu/group/fearon- research/cgi-bin/wordpress/wp-content/uploads/2013/10/What-is-Identity-as-we-now-use-the-word-.pdf

Floridi, L. (ed.) (2015). The Onlife Manifesto. Being Human in a Hyperconnected Era. London: Springer Open.

EDPS, (2018). European Data Protection Supervisor (EDPS). "Towards o digital ethics". Report from EOP5 Ethics Advisory Group. https:!!edps.europo.eu!sites!edp!files!publication!18-01-25_eog_report_en.pdf

EGE (2018). European Group on Ethics (EGE). "Statement on Artificial lntelligence, Robotics and Autonomous Systems". European Group on Ethics in Science and New Technologies. Brussels: European Commission. https://ec.europa.eu/ research/ege/pdf/ege_ai_statement_2018.pdf

EESC, (2016). European Economic and Social Committee (EESC). "Opinion on Artificial lntelligence". https://www.eesc.europa.eu/en/our-work/opinions-information-reports/opinions/artificial-intelligence/opinions

Al Now, (2017) Report. https://ainowinstitute.org/AI_Now_2017_Report.pdf

Powles, J. and Hodson, H. (2017). "Google DeepMind and healthcare in an age of algorithms". Health and Technology, Vol. 7, no. 4, pp. 351-367.

Chouldechova, A. (2017). "Fair prediction with disparate impact: A study of bias in recidivism prediction instruments". Big Data, Vol. 5, no. 2, pp. 153-163.

O'Neil, C. (2016). "Weapons of math destruction. How big dota increases inequality and threatens democracy". New York: Crown Publishers.

Mission Villani, (2018). "For a meaningful artificial intelligence towards a French and European Strategy". https://www.aiforhumanity.fr/pdfs/MissionVillani_Report_ENG-VF.pdf

Al Now, (2018). "Algorithmic lmpact Assessments: A Practical Framework for Public Agency Accountability". https://ainowinstitute.org/aiareport2018.pdf

Van Dijck, J. (2014). "Datafication, dataism and dataveillance: Big data between scientific paradigm and ideology". Surveillance & Society, Vol. 12, no. 2. pp. 197-208.

Rathenau lnstitute, (2017). "Human rights in the robot age". Council of Europe Report. https://www.rathenau.nl/sites/default/files/2018-02/Human%20Rights%20in%20the%20Robot%20Age-Rathenau%20Instituut-2017.pdf

Korff, D. Wagner, B. and Powles, J. et al. (2017). "Boundaries of Low: Exploring Transparency, Accountability, and Oversight of Government Surveillance Regimes". University of Cambridge Faculty of Law Research Paper no. 16/2017. https://ssrn.com/abstract=2894490

Krijger, J., Thuis, T., de Ruiter, M. et al. (2023). "The AI ethics maturity model: a holistic approach to advancing ethical data science in organizations". AI Ethics, Vol. 3, pp. 355–367. https://doi.org/10.1007/s43681-022-00228-7

Porayska-Pomsta, K. (2015). "Al in Education as a Methodology for Enabling Educational Evidence-Based Practice". ln 17th lnternational Conference on Artificial lntelligence in Education (AlEO 2015), pp. 52-61. Madrid.