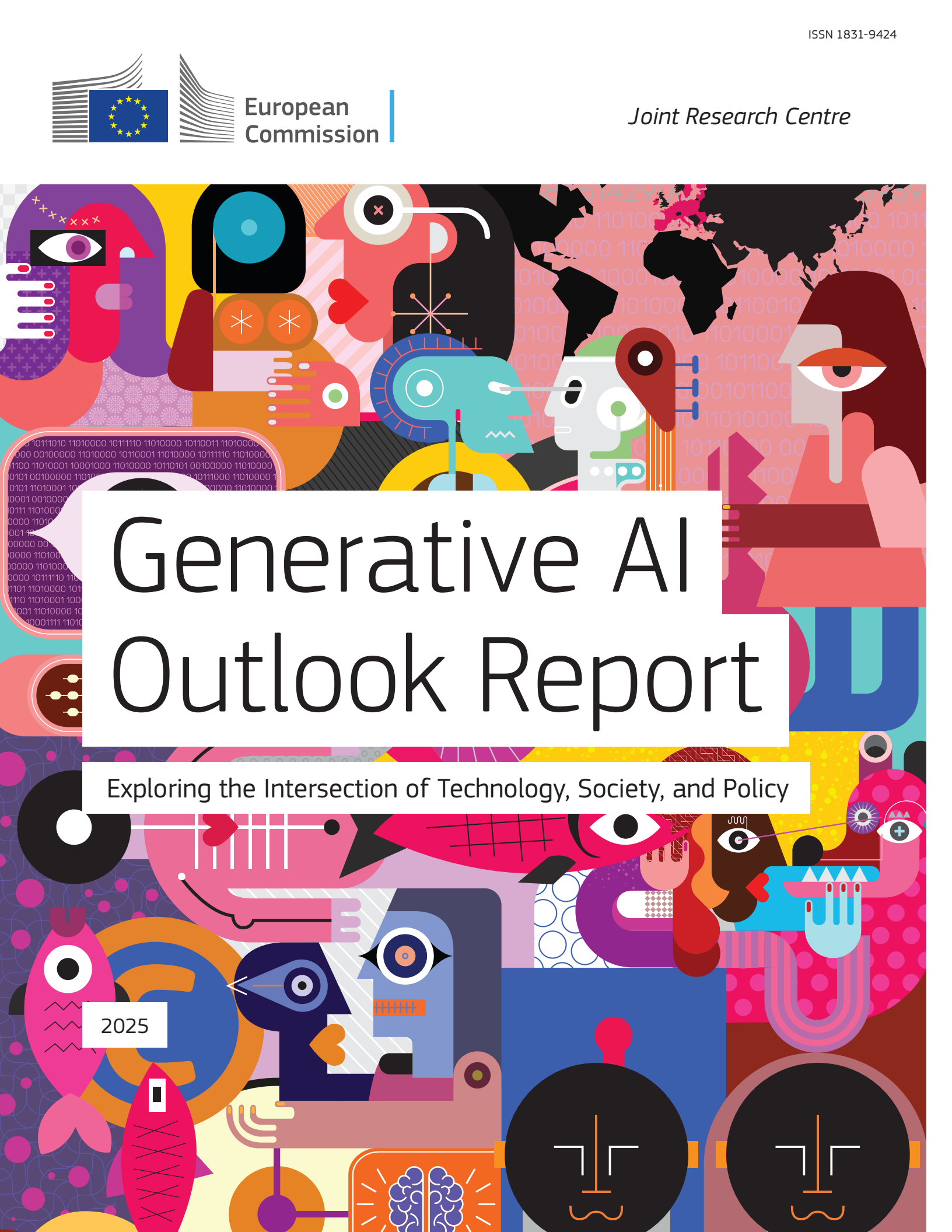




Generative AI Outlook Report

Exploring the Intersection of Technology, Society, and Policy

2025

This document is a publication by the Joint Research Centre (JRC), the European Commission's science and knowledge service. It aims to provide evidence-based scientific support to the European policymaking process. The contents of this publication do not necessarily reflect the position or opinion of the European Commission. Neither the European Commission nor any person acting on behalf of the Commission is responsible for the use that might be made of this publication. For information on the methodology and quality underlying the data used in this publication for which the source is neither Eurostat nor other Commission services, users should contact the referenced source. The designations employed and the presentation of material on the maps do not imply the expression of any opinion whatsoever on the part of the European Union concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries.

EU Science Hub

<https://joint-research-centre.ec.europa.eu>

JRC142598

EUR 40337

Print	ISBN 978-92-68-28247-2	ISSN 1018-5593	doi:10.2760/0991238	KJ-01-25-309-EN-C
PDF	ISBN 978-92-68-28153-6	ISSN 1831-9424	doi:10.2760/1109679	KJ-01-25-309-EN-N

Luxembourg: Publications Office of the European Union, 2025

© European Union, 2025



The reuse policy of the European Commission documents is implemented by the Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39). Unless otherwise noted, the reuse of this document is authorised under the Creative Commons Attribution 4.0 International (CC BY 4.0) licence (<https://creativecommons.org/licenses/by/4.0/>). This means that reuse is allowed provided appropriate credit is given and any changes are indicated.

For any use or reproduction of photos or other material that is not owned by the European Union permission must be sought directly from the copyright holders.

Cover page and decorative elements, © danjazzia / stock.adobe.com

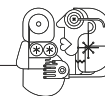
How to cite this report: Abendroth Dias, K., Arias Cabarcos, P., Bacco, F.M., Bassani, E., Bertoletti, A. et al., *Generative AI Outlook Report - Exploring the Intersection of Technology, Society and Policy*, Navajas Cawood, E., Vespe, M., Kotsev, A. and van Bavel, R. (editors), Publications Office of the European Union, Luxembourg, 2025, <https://data.europa.eu/doi/10.2760/1109679>, JRC142598.

Layout: Carmen Capote de la Calle

4

SOCIETAL IMPACTS AND CHALLENGES





This chapter investigates the societal implications of GenAI, emphasising the skills gap and the need for AI literacy among citizens and the workforce. It discusses digital commons and the intersection of AI with environmental concerns, as well as Gen AI in the media and the overall perception, public discourse and narrative around its developments. The chapter also addresses the rights of children and issues of gender bias within AI systems, and the potential to generate false or misleading content. A behavioural approach to GenAI policy analysis is proposed as a means to navigate privacy and data protection challenges. The chapter raises critical questions about how to ensure inclusive and ethical AI applications that align with societal values.

4.1 Skills Gap and AI Literacy for Citizens and the Workforce

KEY MESSAGES

- Addressing the skills gap and enhancing AI literacy in the workforce is a complex challenge that requires coordinated efforts from multiple stakeholders, including businesses, educational institutions, and policymakers.
- By adopting comprehensive strategies that focus on upskilling, reskilling, and fostering AI literacy from an early age through education and continuous through lifelong learning, societies can help ensure that their workforces and citizens are better prepared to harness the potential of GenAI.

INTRODUCTION TO SKILLS GAP AND DIGITAL SKILLS

GenAI has the potential to bring a substantial transformation in the workforce landscape across various industries (see Section 3.4). This shift highlights a potentially significantly evolving skills gap, primarily in digital competencies,

which is becoming increasingly evident as GenAI technologies proliferate. While the propensity for AI and GenAI to replicate some human skills will lead to skills displacement, their widespread diffusion is expected to create a growing demand for digital and data science skills related to the development and maintenance of AI systems, as well as complementary cognitive and transversal skills to enable workers to use and interact with GenAI systems.¹⁴⁹ Such skills, as well as analytical thinking, resilience, flexibility, agility, along with leadership and social influence are the skills required by employers.¹⁵⁰

This skills gap is not merely about understanding how to use GenAI tools but extends to comprehending the broader implications of AI technologies, including their potential to automate tasks and enhance human capabilities, which addresses the economic implications of GenAI. The focus here is on the responses at societal and training levels needed to equip workers, and citizens in general, with the necessary skills.

Digital skills remain high on the EU policy agenda, and existing evidence supports a need for ongoing and sustained efforts. The Digital Decade Policy Programme has a target that at least 80% of persons aged 15–74 should have at least basic digital skills by 2030.¹⁵¹ In 2023, 56% of that population group had basic or above basic digital skills, which is far from the target. The EU also has a target¹⁵² that the share of low-achieving students in computer and information literacy should be less than 15% in 2030. Yet, in 2023,

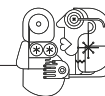
149. https://www.oecd.org/en/publications/oecd-employment-outlook-2023_08785bba-en/full-report/skill-needs-and-policies-in-the-age-of-artificial-intelligence_fe530fbf.html

150. 3. Skills outlook - The Future of Jobs Report 2025 | World Economic Forum <https://www.weforum.org/publications/the-future-of-jobs-report-2025/in-full/3-skills-outlook/#3-1-expected-disruptions-to-skills>

151. https://commission.europa.eu/strategy-and-policy/priorities-2019-2024/europe-fit-digital-age/europes-digital-decade-digital-targets-2030_en

152. As agreed in the Council Resolution on a strategic framework for European cooperation in education and training towards the European Education Area and beyond (2021-2030). <https://op.europa.eu/en/publication-detail/-/publication/b004d247-77d4-11eb-9ac9-01aa75ed71a1>





SOCIETAL IMPACTS AND CHALLENGES

43% of students did not reach the basic level of digital skills.¹⁵³ In 2023, the Council adopted a Recommendation on improving the provision of digital skills and competencies in education and training,¹⁵⁴ recognising the importance to promoting a quality inclusive and consistent approach to the development of digital skills at all levels of education and training.

The 2024–2029 European Commission continues to place a high emphasis on skills (including digital skills) through, for example, its overarching Union of Skills.¹⁵⁵ As part of this policy action, digital skills have been recognised as basic skills in the recently adopted Action Plan on Basic Skills. A forthcoming European Strategy for Vocational Education and Training (VET).¹⁵⁶ The Future of European Competitiveness Report¹⁵⁷ highlights that addressing skills gaps at all stages is crucial to Europe's long-term success and that there is a need for increased efforts in education, training and lifelong learning. It warns that while technology, including AI, is pivotal to preserving Europe's social model, it could exacerbate inequalities if not accompanied by strong investments in education and skills development. The EU Competitiveness compass¹⁵⁸ addresses this need for the domain of AI through the Apply AI Strategy, while the recently published AI Continent Action Plan underlines the need to reinforce AI skills, including basic AI literacy through further developing excellence in AI education, training and research.

- Digital legislation including the AI Act,¹⁵⁹ the Cybersecurity Act¹⁶⁰ and the Digital Services Act¹⁶¹ has implications for citizens in terms of awareness of their rights and responsibilities and understanding of ethical and social implications of digital technologies. Article 4 of the AI Act requires providers and deployers of AI systems to ensure a sufficient level of AI literacy of their staff and other persons dealing with AI systems on their behalf. To do so, providers and deployers should consider staff technical knowledge, experience, education and training and the context in which the AI systems are to be used, including the persons targeted by such AI systems.¹⁶²
- AI literacy is increasingly recognised as a critical component of workforce development and one of the increasing skills demanded by employers.¹⁶³ It involves not only the ability to use AI tools but also an understanding of AI's fundamental concepts, ethical considerations, and societal impacts. As AI systems become more integrated into workplace processes, developing AI literacy is essential for empowering workers to make informed decisions and use AI responsibly and effectively.

The importance of AI literacy needs to be treated not as a standalone skill but as embedded within the broader context of digital competence (see next paragraph on DigComp 3.0). By fostering a workforce that is knowledgeable about AI, organisations can enhance productivity and innovation while mitigating risks associated with AI adoption.

153. International Computer and Information Literacy Study, ICILS <https://op.europa.eu/en/publication-detail/-/publication/59721dc6-a0aa-11ef-85f0-01aa75ed71a1/language-en>

154. <https://eur-lex.europa.eu/eli/C/2024/1030/oj/eng>

155. https://commission.europa.eu/topics/eu-competitiveness/union-skills_en

156. https://commission.europa.eu/document/e6cd4328-673c-4e7a-8683-f63ffb2cf648_en

157. https://commission.europa.eu/topics/eu-competitiveness/draghi-report_en

158. Competitiveness compass – European Commission

159. <https://eur-lex.europa.eu/eli/reg/2024/1689/oj>

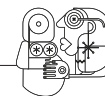
160. <https://eur-lex.europa.eu/EN/legal-content/summary/the-eu-cybersecurity-act.html>

161. https://eur-lex.europa.eu/legal-content/EN/TXT/?toc=OJ%3A2022%3A277%3ATOC&uri=uriserv%3AOJ.L_.2022.277.01.0001.01.ENG

162. See also Q&A on AI literacy: <https://digital-strategy.ec.europa.eu/en/faqs/ai-literacy-questions-answers>

163. The Future of Jobs Report 2025 | World Economic Forum <https://www.weforum.org/publications/the-future-of-jobs-report-2025/digest/>





DIGITAL COMPETENCE FRAMEWORK 3.0

The European Digital Competence Framework for Citizens, commonly known as DigComp,¹⁶⁴ aims to address the digital skills gap by providing a structured approach to enhancing digital competences among citizens. The forthcoming DigComp 3.0 iteration, due to be published by the end of 2025, will further integrate AI-related competences, reflecting the need for citizens to understand and interact with AI systems ethically and effectively. This framework is central in supporting the development of critical digital skills that are necessary for navigating and benefitting from the evolving digital landscape. A systematic and transversal integration of AI, including GenAI, will be included in DigComp 3.0 across all five competence areas of the framework with a focus on conceptual understanding, ethical, and societal implications. The emphasis is on encouraging a critical, reflective, and balanced use of AI systems and their outputs. This integration will align with other key initiatives such as the AI Act and the forthcoming AI Literacy Framework for primary and secondary schools co-developed by the European Commission and OECD ([see Section 6.2](#)).¹⁶⁵ Specifically, GenAI competences will be considered as follows:

- **Citizen cybersecurity:** There is an increased focus on citizen cybersecurity, including the role of AI systems in both cyberattacks and cybersecurity. This is reflected in three out of the 21 competencies: managing digital identity, protecting devices, and protecting personal data and privacy.
- **Rights, choice, and responsibility:** The competence engaging in digital citizenship

will be expanded to include more explicit references to consumer rights, choice, active participation, and influencing. This includes awareness and assertion of rights in relation to recent digital legislation such as the Digital Services Act and AI Act.

- **Misinformation, disinformation, and threats to democracy:** The competence evaluating digital content focuses on dealing with misinformation and disinformation, recognising the role of GenAI and the potential speed at which digital information can spread. This includes identifying biased sources, fact-checking, flagging, and reporting misinformation and disinformation.
- **Twin transition:** The competence protecting the environment includes a consideration of how recent technologies and trends, such as GenAI and social media, are resource-intensive and the role of individual practices in mitigating this impact.

Overall, the updated DigComp 3.0 framework emphasises the importance of digital literacy, critical thinking, and responsible behaviour in the digital age, with a focus on the societal implications of emerging technologies, such as AI and GenAI.

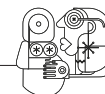
THE ACADEMIC OFFER ON AI

The EU has emphasised the importance of developing a digitally skilled population and workforce to remain competitive in the digital economy. As mentioned above, several policy initiatives, including the Digital Decade Policy Programme and the AI Continent Action plan, aim to strengthen AI capabilities in the workforce. The AI Continent Action plan proposes supporting an increase in EU bachelor's and master's degrees and PhD programmes in key technologies to enlarge the pool of AI specialists. Higher educational institutions play a critical role in shaping the future workforce, and their academic offerings provide valuable insight into

164. Digital Competence Framework for Citizens (DigComp) - European Commission https://joint-research-centre.ec.europa.eu/projects-and-activities/education-and-training/digital-transformation-education/digital-competence-framework-citizens-digcomp_en

165. Empowering learners for the age of AI: draft AI literacy framework launch - European Education Area <https://education.ec.europa.eu/event/empowering-learners-for-the-age-of-ai-draft-ai-literacy-framework-launch>

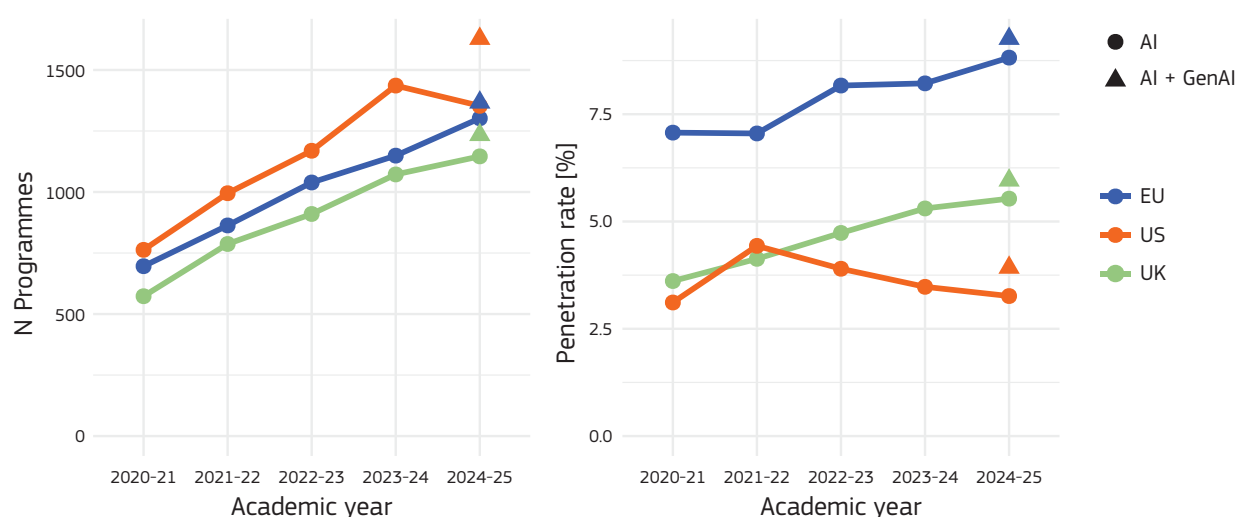




SOCIETAL IMPACTS AND CHALLENGES

the skills that will define tomorrow's human capital. Monitoring trends in academic programmes is therefore essential for understanding how education systems are responding to emerging technological developments. Through the use of text mining, capturing the inclusion of advanced digital technologies in the programmes' syllabus, and drawing on data from StudyPortals, the JRC monitors the availability of English-taught master's programmes, bachelor's programmes, and short professional courses, and studies their characteristics.^{166 167 168} The analysis below focuses on master's degrees, especially relevant for anticipating skills development as they represent the final academic stage before entering the workforce.

Figure 13. AI (and GenAI) related master's degrees by geographic area and academic year, 2020–25.



The penetration rate (right) is defined as the share of AI (and GenAI) master's degrees relative to the total number of masters offered in the geographic area.

Source: JRC elaboration.

Figure 13 shows the trends in the number of programmes by geographical area. For the period 2020–2024, the focus is on AI-related programmes. In 2024–25, the scope is expanded to include programmes explicitly addressing GenAI, as such data were not available in earlier years. The latest set of results confirms a consistent increase in the number of AI master's degrees offered in the EU and the UK, complemented with an additional number of degrees including GenAI. The US shows a marked decline in 2024–25, which is only offset by

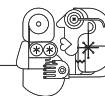
considering GenAI, suggesting a strategic shift towards replacing traditional AI programmes with GenAI offerings. Figure 13 (right) shows the penetration rates – the proportion of AI (and GenAI) master's degrees relative to the total number of master's offered in the geographic area –, reflecting the relative importance of AI across all academic offerings. This plot highlights the EU's strong and sustained leadership in AI adoption in academic offerings between 2020–21 and 2024–25. The EU consistently records significantly higher penetration rates

166. López Cobo M., et al., Academic offer and demand for advanced profiles in the EU. Artificial Intelligence, High Performance Computing and Cybersecurity, EUR 29629 EN, Publications Office of the European Union, Luxembourg, 2019.

167. Righi, R., et al. Academic offer of advanced digital skills in 2019–20. International comparison. Focus on Artificial Intelligence, High Performance Computing, Cybersecurity and Data Science, EUR 30351 EN, Publications Office of the European Union, Luxembourg, 2020.

168. https://joint-research-centre.ec.europa.eu/predict/academic-offer-advanced-digital-technologies_en





SOCIETAL IMPACTS AND CHALLENGES

compared to the UK and the US. Interestingly, the inclusion of GenAI-related master's degrees leads to an increase of 0.4 percentage points in the penetration rate, compared to traditional AI programmes, highlighting the increased interest generated by this emerging field.

STRATEGIES FOR UPSKILLING AND TRAINING

→ **Workplace training programmes:**

Organisations are increasingly implementing in-house training programmes focused on GenAI and digital skills. These programmes often include workshops, online courses, and hands-on projects that allow employees to gain practical experience with AI tools. Such initiatives are crucial for keeping the existing workforce competitive and capable of adapting to technological changes, and to understand remaining capacity building needs (see AI-empowered JRC box). In addition, the Commission has created an AI literacy repository collecting practices from organisations that provide and deploy AI systems¹⁶⁹ in order to support the implementation of Article 4 of the AI Act.

AI-empowered JRC

As European Commission's in-house science service, the JRC has launched an initiative to explore the potential of AI in the workplace. With 75% of colleagues using AI tools like GPT@JRC (in-house secure platform for staff to access and experiment with Large Language Models), the organisation aims at upskilling staff to harness the full potential of AI. The "AI-Empowered JRC" project aims to support staff in using AI tools effectively, addressing concerns around limitations and risks, and developing the skills needed to work with AI systems. As part of this initiative, the JRC is focusing on upskilling and training, particularly for managers, to ensure that AI is integrated

into daily tasks in a way that enhances productivity and efficiency. As an example of specific application at the Science and Policy interface, experiments to deploy "AI Research Assistants" powered by GenAI and agentic AI are ongoing. The project's approach emphasises experimentation, collaboration, and mutual learning, and may offer useful insights and best practices for other organisations looking to adopt AI in their own workplaces. The purposeful use of Generative AI is given particular attention.

→ **Partnerships with educational institutions:**

Collaborations between businesses, public sector organisations and educational institutions can facilitate the development of curricula that align with industry needs. This will also be important in the phase of highlighting what training needs businesses and organisations have. Higher educational institutions and technical schools can offer specialised courses and certifications in AI literacy, ensuring that graduates enter the workforce with the skills required to thrive in a GenAI-driven environment.

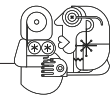
→ **Policy-backed initiatives:** Policymakers have a vital role in supporting the workforce through public initiatives aimed at enhancing AI literacy. Policy programmes can provide funding for training initiatives, create incentives for companies to invest in employee development, and establish national standards for digital literacy.

ROLE OF EDUCATIONAL INSTITUTIONS AND POLICYMAKERS

Educational institutions are at the forefront of bridging the skills gap. By integrating AI literacy into their curricula, education and training institutions at all levels can prepare students for a future where AI is ubiquitous. This includes not only teaching technical skills but also fostering critical thinking and problem-solving abilities, which are essential in an AI-driven world.

169. See <https://digital-strategy.ec.europa.eu/en/library/living-repository-foster-learning-and-exchange-ai-literacy>





SOCIETAL IMPACTS AND CHALLENGES

Policymakers, on the other hand, must create an enabling environment that supports lifelong learning and continuous skill development. This involves crafting policies that encourage educational innovation, support digital infrastructure development, and promote equitable access to learning opportunities for all segments of the population.

ADDRESSING THE GAP

To address the skills gap, a multifaceted approach is necessary, involving formal education and training, and also upskilling and reskilling initiatives. Upskilling refers to enhancing existing skills to meet the demands of new technologies, while reskilling involves training workers in entirely new skills that are relevant to the current job market. A study conducted with 2,307 GenAI decision-makers show that two out of 3 respondents acknowledge that their employees do not have the skills to work with GenAI. About half are planning employee education and training to increase GenAI adoption.¹⁷⁰

4.2 GenAI and Information Manipulation

KEY MESSAGES

- Generative AI models are revolutionising content creation by making it remarkably easy to produce highly convincing manipulative content rapidly and at scale. This capability can be exploited to dominate social media discussions, mimic authoritative news outlets, and create realistic images, videos, and audio. The potential for mis/disinformation is vast, as AI-generated content can be used to mislead the public, erode trust in media, and distort the overall information landscape.

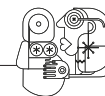
170. NTT Data (2025). Global GenAI Report How organizations are mastering their GenAI destiny in 2025. Global GenAI Report | NTT DATA <https://services.global.ntt/en-us/campaigns/global-genai-report>

Such manipulation can have profound effects, including altering the outcomes of elections, influencing political processes, and shifting public support on critical issues like climate change. By flooding the digital space with misleading information, generative AI can significantly skew public perception and decision-making.

- The rapid generation of mis/disinformation often outpaces the ability to rebut it effectively, as crafting accurate and verifiable responses takes time. This discrepancy means that even well-intentioned use of AI for communication can struggle against the tide of falsehoods. To maintain a commitment to fact-based trustworthy communication to uphold the integrity of information is often an uphill battle in the face of AI-generated mis/disinformation.
- While technical solutions like watermarking AI-generated content and verifying information with trusted sources are valuable, they are not sufficient on their own. The most effective defence against mis/disinformation lies in promoting media and AI literacy among citizens, including policy makers. By equipping individuals with the skills to critically engage with AI-generated content, they can better discern false or inaccurate information. Educating the public to question and analyse the content they encounter empowers them to make informed decisions and resist manipulation. This foundational approach, combined with technical measures, is vital to maintaining a well-informed society resilient to the challenges posed by generative AI.

Various facets of mis/disinformation appear in many sections in this report, but the topic also deserves dedicated consideration when examining the implications of GenAI and how it can be used to create or to fight false or misleading content.





SOCIETAL IMPACTS AND CHALLENGES

Unintentional bias already introduces the risk of misleading content when AI is used to generate or even summarise content ([see also Section 4.8](#)). For instance, using generative AI to obtain information, answers can be heavily biased based on the dataset on which the AI was trained. When AI is used intentionally for manipulating information, the use of GenAI leads to unprecedented challenges by enabling the creation of deep fakes of both audio and visual content at large scale and speed, distorting media integrity.

Independently of the content creation, GenAI can also be used for systemic data poisoning at scale, for example, to pollute free and open knowledge repositories ([see also Section 2.2](#)) or to amplify the dissemination in online news and social media. As an example, bias in AI models can significantly impact attitudes towards climate change. A worldwide change of attitude towards climate change may have a compound effect on global emissions, amplifying or reducing them significantly. Recommender AI systems can spread GenAI-generated disinformation and sensationalist content, exacerbating polarisation on climate change ^{171 172 173} ([see also Section 4.5](#)).

GenAI-based influence operations directed at electoral processes have been monitored closely. Studies reveal a significant increase in artificially generated content with the rise of GenAI capabilities, unveiling Foreign Information Manipulation and Interference (FIMI) attempts.¹⁷⁴

^{175 176 177} One notable interference applying AI based impersonating techniques was the ‘Doppelganger’ campaign that consists of the use of fake clones of legitimate websites – both from media organisations and public institutions.¹⁷⁸

Especially when it comes to social media, GenAI-powered bots boost the spreading of misleading narratives and manipulated information, frequently adding negative sentiment to disrupt public discourse, provoke targeted communities and foster radicalisation ([see also Section 4.3](#)).

Addressing issues caused by manipulated information requires AI literacy among policy-makers and citizens to ensure they can critically engage with AI-generated content and identify false or manipulative information, including deep fakes ([see also Section 4.1](#)). Furthermore, the Regulation on transparency and targeting of political advertising has a specific obligation to ensure transparency of the use of AI to target or deliver political advertisings, and a dedicated Code of Conduct on Disinformation has been adopted by the European Commission and European Board for Digital Services in the context of the DSA along with Guidelines on elections-related risks ([see also Section 5.2](#)). It represents an EU-wide effort to address the issues raised above.

At the same time, it should be underlined that GenAI can also help in combatting information manipulation and its dissemination ([see also Section 6.4](#) for the deep dive on GenAI and cybersecurity). Opportunities include the use of GenAI for clear and targeted communication, fact checking, identifying AI-generated content, and more.

171. Meyerson, E. (2012). YouTube Now: Why We Focus on Watch Time. YouTube Official Blog.

172. Chaslot, G. (2017). How YouTube’s A.I. boosts alternative facts. *Medium*.

173. Falkenberg, M., Galeazzi, A., Torricelli, M. et al. Growing polarization around climate change on social media. *Nat. Clim. Chang.* 12, 1114–1121 (2022). <https://doi.org/10.1038/s41558-022-01527-x>

174. https://edmo.eu/wp-content/uploads/2024/03/EDMO_TFEU2024-Narratives_Report-National_Elections-2nd-edition-1.pdf

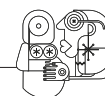
175. https://www.eeas.europa.eu/eeas/1st-eeas-report-foreign-information-manipulation-and-interference-threats_en

176. https://www.eeas.europa.eu/eeas/2nd-eeas-report-foreign-information-manipulation-and-interference-threats_en

177. https://www.eeas.europa.eu/eeas/3rd-eeas-report-foreign-information-manipulation-and-interference-threats-0_en

178. https://euvsdisinfo.eu/uploads/2024/06/EEAS-TechnicalReport-DoppelgangerEE24_June2024.pdf





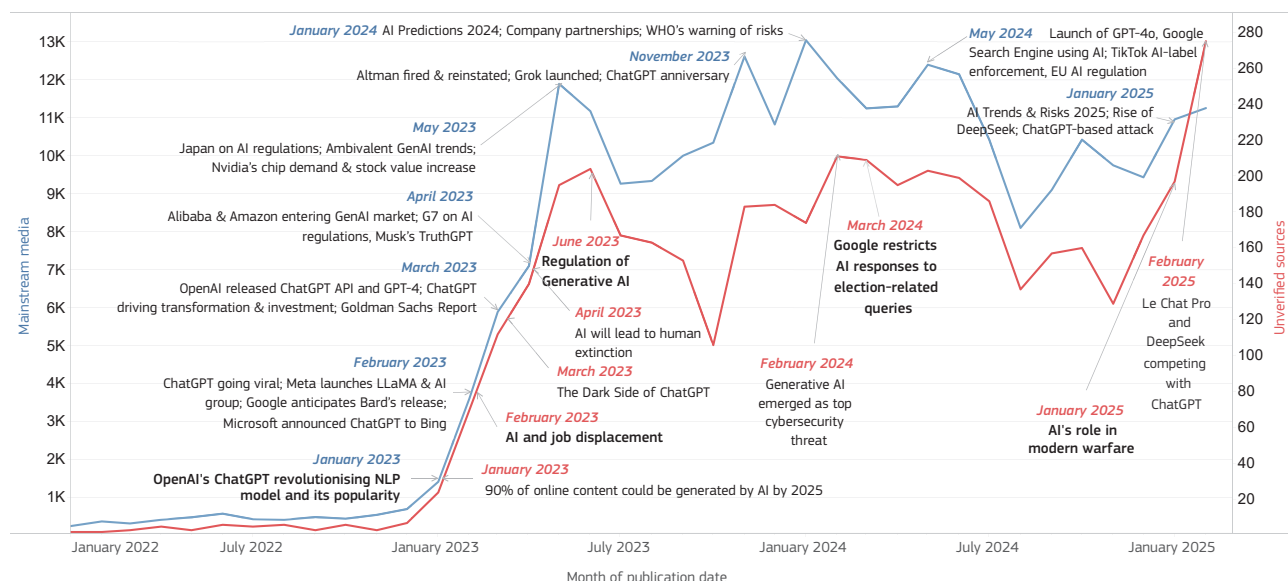
4.3 Generative AI portrayal in the Media

KEY MESSAGES

- The media portrayal of GenAI is often polarised, with utopian visions highlighting its transformative potential and dystopian concerns warning about ethical implications, job displacement, and privacy issues. This dual narrative shapes public discourse and might influence policy perceptions globally.
- Media coverage of GenAI has seen a significant increase since late 2022, particularly after the public release of ChatGPT. Subsequently, reporting peaks followed key events such as the release of new GenAI models, their integration into commercial products, and political regulatory actions from numerous stakeholders, including the active participation of the European Commission in the discourse.
- An analysis of news articles shows that the media convey the high relevance of international governance to manage GenAI and mitigate associated risks.

Understanding the media portrayal of GenAI is crucial to apprehend its perceived role and impact on society and the economy. The perception of GenAI in online media news articles is often framed within a spectrum of narratives, mostly portraying extreme scenarios. On the one hand, utopian visions highlight the transformative potential of GenAI, suggesting how it could revolutionise industries, enhance creativity, and solve complex global challenges, leading to unprecedented innovation and prosperity. On the other hand, dystopian concerns are widespread, with articles warning about the ethical implications, job displacement, and the erosion of privacy and security that could accompany AI's rise. Speculations about strong GenAI fear characterise the debate, with some envisioning a future where machines surpass human intelligence, raising existential questions about control, autonomy, and even potential leading to human extinction. Meanwhile, the take-up of GenAI as an integral part of modern everyday life is increasingly evident, as it becomes embedded in routine activities, prompting discussions about dependency and the reshaping of human capabilities.

Figure 14. Evolution of reporting volume on Generative AI in mainstream media and unverified sources.



Monthly timeline distribution of news articles on Generative AI with annotations. Source: EMM full index. Period: 01-01-2022 to 28-02-2025

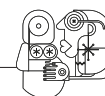


Figure 14 shows how both mainstream and unverified sources, i.e. sources that were indicated by independent fact-checkers as often spreading mis/disinformation, have covered GenAI, based on the reporting trends and peaks, as well as the narratives related to GenAI that influence policy and public perception across the globe.¹⁷⁹ The analysis shows a similar distribution in the reporting trends both for the mainstream media and unverified sources around main GenAI-related events. Nevertheless, unverified sources often distort information by misinterpreting real events and statements. The main reported topics related to GenAI were identified by assessing the key clusters that were computed for each month based on the article sentences' semantic similarity topics, covering technological and corporate developments, regulatory and ethical considerations, economic and industrial impact, global competition and collaboration, and cultural and societal impact. These themes persist in both examined datasets of articles from the mainstream media and from unverified sources.

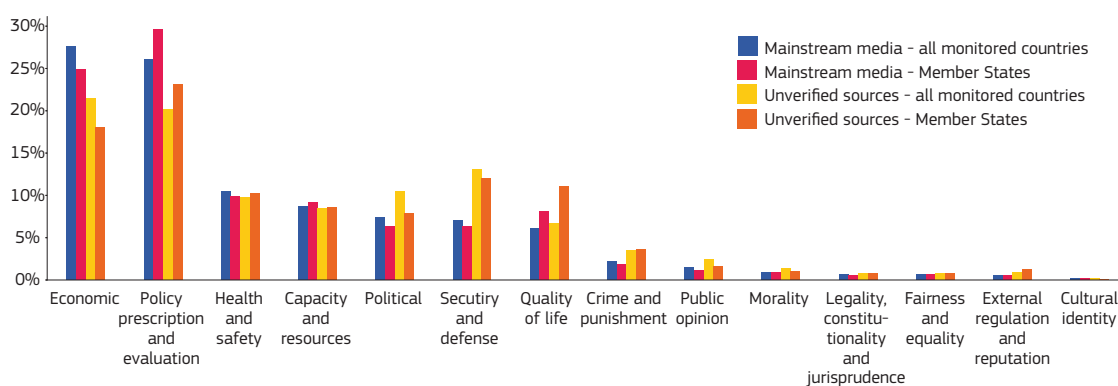
179. The analysis is based on 265 306 online news articles collected via the Europe Media Monitor from both mainstream media and unverified sources^[4] between 1 January 2022 and 28 February 2025, selected by searching for a set of relevant keywords or keyword combinations within the text. These keywords are translated into all 24 EU languages, as well as Russian, Chinese (traditional and simplified), Arabic, Hindi, Turkish, Norwegian, Georgian, and Japanese. The keyword combinations^[5] have been selected carefully to include articles of interest but to avoid irrelevant 'noise' in the news collection. The data-driven analysis used automated techniques including AI, machine learning and large language models (LLMs) to discover trends, patterns and shifts in reporting around GenAI, identify the main stories of mis/disinformation, compare specific trends, discover the associated sentiment and the most common framing dimensions.

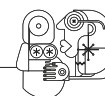
However, the articles in the latter, focus readers' attention on a different perspective within each theme.

FRAMING DIMENSIONS IN NEWS ARTICLES RELATED TO GENERATIVE AI

The examination of media coverage related to GenAI offers additional insights through the analysis of framings detected in the articles. These framing dimensions correspond to specific aspects mentioned in the context of the main topic, thereby revealing the perspective from which issues or news pieces are presented. Figure 15 compares framing dimensions detected in news articles from the mainstream media and unverified sources, colour-coded by source type, including *economic, policy prescription and evaluation, health and safety, capacity and resources, political issues, security and defence, crime and punishment, public opinion, morality, legislation, fairness and equality, regulation, and cultural identity*. The figure reveals differences between mainstream and unverified sources, as well as between articles coming from EU Member State sources and global reporting. We quantitatively analysed framing dimensions and their distribution between the mainstream media and unverified sources. Framing dimensions are assigned to over 75% of the articles from both mainstream and unverified sources across the observed period.

Figure 15. Framings of GenAI news by media type and target.





SOCIETAL IMPACTS AND CHALLENGES

Although *Economic* and *Policy prescription and evaluation* framings were prevalent in both mainstream and unverified sources, the mainstream media exhibited a higher share of these framings. The *Economic* framing stood out due to the substantial number of news articles highlighting the transformative impact of GenAI. This is not only revolutionising the tech sector but also reshaping various other industries, spurring major investment and intensifying competition. The *Policy prescription and evaluation* framing was particularly stressed in the EU mainstream media, addressing the need for and calls for policies to regulate the development and deployment of AI and GenAI technologies, ensuring ethical standards and mitigating potential risks. Both media types frequently used *Health and safety* as well as *Capacity and resources* framings. The news items provide instances on how technological advancements can have a positive impact on health care but also raise concerns about increased risks for the public such as job security. Moreover, the rapid development

requires a significant amount of financial resources and capacities to run AI systems.

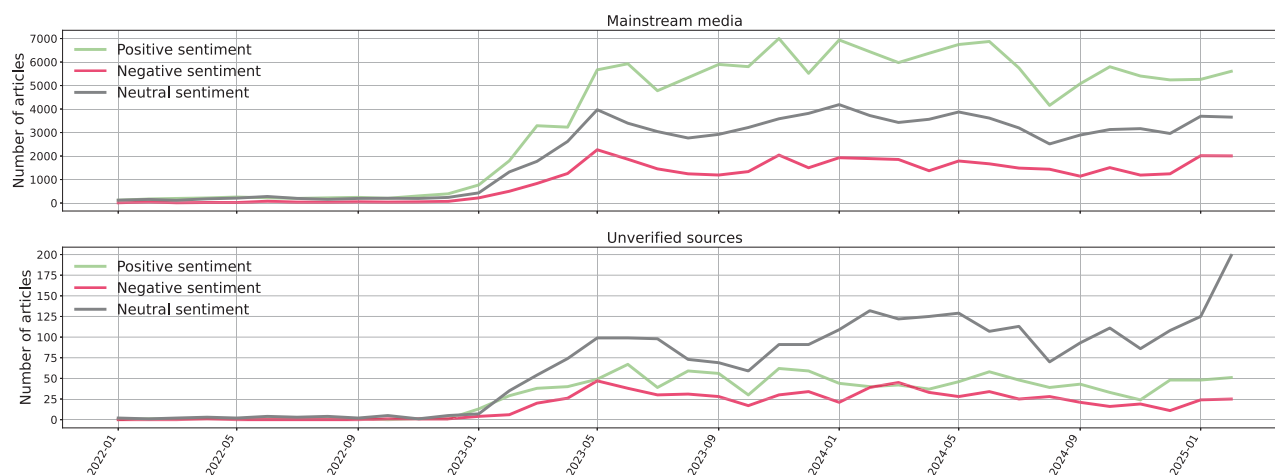
SENTIMENT ANALYSIS OF NEWS ARTICLES RELATED TO GENERATIVE AI

In addition, a pre-trained machine-learning model was applied to the news articles for sentiment analysis to identify the sentiment expressed in these articles.¹⁸⁰ The model assigns positive, neutral or negative sentiment to each article.

Figure 16 presents the evolution of news reporting by sentiment (positive, neutral, negative) divided between mainstream media and unverified sources.

180. Di Nuovo, E., Cartier, E., De Longueville, B. (2024). Meet XLM-RLnews-8: Not Just Another Sentiment Analysis Model. In Natural Language Processing and Information Systems, 28th International Conference on Applications of Natural Language to Information Systems, NLDB 2024, Turin, Italy, June 25–27, 2024, Proceedings (pp. 1). Springer Science and Business Media Deutschland GmbH.

Figure 16. Sentiment evolution related to news on GenAI.



The timeline distribution of GenAI-related articles divided by source type. Source: EMM full and Disinfo index. Period: 01-01-2022 to 28-02-2025.

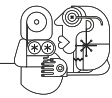
Source: JRC elaboration.

The predominant sentiment of the articles from the mainstream media related to GenAI is positive, prevailing throughout the total observation period. The next prominent sentiment is neutral. Roughly 30% of the total news articles from the mainstream media are portrayed in a

negative light, addressing the concerns about the risks that AI technologies may pose.

Most of the positive headlines refer to economic growth and associated opportunities for multiple industries, featuring the significant advancement





SOCIETAL IMPACTS AND CHALLENGES

and major investments in start-ups focusing on key activities associated with progress, such as technology, healthcare and sustainability. Media also shed light positively on start-ups gaining awards and recognition at international competitions.

In contrast, the negatively connoted headlines refer to start-ups and tech companies facing significant challenges across various industries, including insolvencies, and a lack of investment in start-ups applying GenAI in various domains, lobbying for favourable regulations and more support from governments to address these issues. These negative headlines also capture the threats and risks of the new technologies connoted with a **negative** sentiment.

Sentiment analysis conducted on articles from **unverified sources** reveals the prevalence of the neutral sentiment, closely followed by the positive and the negative sentiments alike. Unverified sources were prone to spreading misleading narratives and disinformation often presenting GenAI with a sensationalist and alarmist sentiment. On the one hand, the headlines with positive sentiment exaggerate the capabilities of GenAI, portraying it as a near-magical technology able to solve all societal problems while, on the other hand, negative headlines precipitate a catastrophic dystopia. Such articles exploit fears of the unknown, suggesting that GenAI could imminently replace human decision-making or even replace humans themselves. Unverified sources often downplay the existing challenges and ethical considerations, oversimplifying complex debates to fit more dramatic narratives. By focusing on extreme scenarios and often ignoring expert opinions, they contribute to a polarised discourse, where the potential benefits of GenAI are either overstated or surpassed by exaggerated risks.

Overall, the analysis shows that the **media conveys the high relevance of international governance to manage Generative AI** and mitigate associated risks. This section is relevant for policymakers as it highlights the need for

action. The results show that applying various text mining techniques on media articles related to this topic provides important insights into the discourse on GenAI in general and can highlight trends in the reporting tonality. This is of particular interest to media analysts, communication experts and policymakers.

The approach described can be used to **inform outgoing communication strategies**, based on the understanding of the way in which issues have been framed in different media. They can also be used to assess the impact of outgoing communication from the EU institutions, by observing whether there is an influence on the discourse in the media around specific issues related to GenAI.

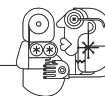
4.4 Digital Commons

KEY MESSAGES

- The digital commons are a crucial element in AI training datasets, and their fate is closely tied to the development of AI technologies.
- As AI continues to evolve, it is essential to protect and support the digital commons to ensure the long-term health and diversity of both since protecting the digital commons is essential for fair and advanced AI.

The digital commons refer to online content and resources – such as code, software, images, texts, and other forms of knowledge – that are shared with free and open licences and are therefore accessible to anyone with an internet connection. They are the online equivalent of common public goods and spaces such as public parks or libraries, and, for example, include various online encyclopaedias and wikis, open source software repositories, digitised cultural heritage archives, openly licensed art and media collections, and online forums where people document and seek answers to questions. The digital commons





comprise a wealth of code, knowledge, and cultural content that is unique because it is diverse, collaboratively made, and free of charge. This has not escaped AI developers and, today, content belonging to the digital commons constitutes a crucial element in practically all AI training datasets. It is therefore not a question if the fate of the digital commons and AI technologies are closely intertwined, but rather *how and with what effects*. While the digital commons are generally considered non-rival – meaning their value and usability do not decrease due to overuse – they are known to be threatened by issues such as pollution, undersupply, and a lack of findability.¹⁸¹ These threats are all reinforced by the introduction of GenAI technologies, opening up new areas of possibility and concern.

OPPORTUNITIES

On the positive side, GenAI may provide new abilities to search and navigate open databases, help fact-check and improve metadata in shared knowledge repositories, and assist in the translation of knowledge and information, thereby improving the findability, accessibility, and quality of content belonging to the digital commons. It is important to emphasise that such uses of AI are not new to the field of the digital commons, which has long relied on machine learning to manage and develop open datasets. Building on such practices, knowledge related to the digital commons could also feed back into and play an important role in improving current AI practices – especially regarding the collection and curation of high-quality AI training datasets. Based on consent and full transparency, projects such as Public Diffusion are, for example, re-envisioning current AI training practices by using content from the commons and the public domain to develop AI models that are fully based on principles of openness and approval,¹⁸² showcasing a radical alternative to commercial AI training practices that are often

based on closed, non-consensual datasets that may violate the rights of copyright holders. With decades of experience in curating text and image databases to address issues concerning bias, discrimination, and cultural gaps, experts from fields relating to the digital commons (including archives, museums, and libraries) could also play an important role in making AI training datasets, models, and systems fairer and more diverse, collaborative, and inclusive.¹⁸³

RISKS AND CHALLENGES

The use of GenAI may have negative effects on the digital commons, including:

- Enclosure and privatisation of free and open knowledge: Restrictive measures to prevent data scraping may inadvertently limit access to online data for public good archives.¹⁸⁴
- Decreased voluntary contributions: As people rely on closed chatbot services, they may be less likely to contribute to shared knowledge databases.¹⁸⁵
- Pollution of free and open knowledge repositories: AI-generated data with errors, hallucinations, and disinformation may enter databases like Wikipedia,¹⁸⁶ requiring costly fact-checking and correction.
- Financial strain on digital commons organisations: Providing open data

181. Dulong De Rosnay, Mélanie, and Felix Stalder. "Digital Commons." Internet Policy Review 9, no. 4 (December 17, 2020). <https://doi.org/10.14763/2020.4.1530>.

182. See Public Diffusion beta, <https://source.plus/public-diffusion-private-beta>

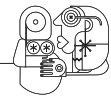
183. Bunz, Mercedes. "The Role of Culture in the Intelligence of AI." In AI in Museums, edited by Sonja Thiel and Johannes C. Bernhardt. transcript Verlag, 2023

184. Longpre, Shayne, and et.al. "Consent in Crisis: The Rapid Decline of the AI Data Commons," ArXiv, 2024. <https://doi.org/10.48550/arXiv.2407.14933>

185. Del Rio-Chanona, R Maria, Nadzeya Laurentsyeve, and Johannes Wachs. "Large Language Models Reduce Public Knowledge Sharing on Online Q&A Platforms." PNAS Nexus 3, no. 9 (September 2, 2024): 1-12. <https://doi.org/10.1093/pnasnexus/pgae400>.

186. Brooks, Creston, Samuel Eggert, and Denis Peskoff. "The Rise of AI-Generated Content in Wikipedia." arXiv, October 10, 2024. <https://doi.org/10.48550/arXiv.2410.08044>.





SOCIETAL IMPACTS AND CHALLENGES

to AI web crawlers incurs significant infrastructural costs (more than half of website traffic is due to automated crawlers compared to human users), with little return, and may lead to financial burdens on organisations hosting open-source content.¹⁸⁷

These risks may ultimately restrict access to and expansion of shared knowledge and place a significant economic burden on organisations that host content belonging to the digital commons.

FUTURE PERSPECTIVES

The future of the digital commons is uncertain and depends on the level of support it receives. There are two possible scenarios:

Best-case scenario:

- The digital commons continue to thrive with financial and infrastructural support from public institutions and commercial AI developers.
- The digital commons remain a diverse, high-quality source of information, promoting culturally balanced AI models and global access to knowledge, information, and cultural history.
- Generative AI technologies provide creative support to communities around the world and thereby facilitate the production of new content that will make the commons grow. They are also successfully adopted by organizations hosting content belonging to the commons, making information management more efficient and accessible.

Worst-case scenario:

- The digital commons deteriorate due to pollution and lack of new data, becoming

an untrustworthy source of culture and information.

- They wither away, reducing the diversity of governance models for creating and managing knowledge.
- This would threaten global access to free and open culture, history, and knowledge, and damage AI development by relying on low-quality or outdated data.

While the GenAI has a potential to enrich cultural practices and support free and open knowledge and information repositories, there is a need to reflect on the potential negative consequences it may bring. Importantly, protecting the digital commons in the age of GenAI is not just a matter of securing open, diverse, and decentralised access to high-quality information for citizens across the globe, but a matter of securing the development of AI technologies that are fairer, more diverse, and therefore also more advanced and useful.

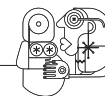
4.5 Environmental Implications of Generative AI

KEY MESSAGES

- The direct impact of AI on the environment is considerable and this can pose problems in particular if data centres are deployed in water poor regions, and recycling of components is an issue.
- Specific measures such as mandatory tracking of energy consumption, better planning of data centre in water - rich areas, and recycling programmes could help reduce the impact of AI on the environment.
- The EU has been pioneering efforts against climate change and integrating environmental protection into policies. Notably, the EU introduced multiple measures to address the impact of data

187. <https://arstechnica.com/ai/2025/03/devs-say-ai-crawlers-dominate-traffic-forcing-blocks-on-entire-countries/>





centres, including the Energy Efficiency Directive and its sustainability rating as well as the Energy Labelling and Ecodesign Directives, and is preparing an upcoming Cloud and AI Development Act focusing on sustainable data centres as well. The EU AI Act includes environmental provisions on AI models and systems transparency, adhesion to codes of practice and information disclosure.

- Emerging technologies in AI, such as energy-efficient transistors, neuromorphic chips, and specialised edge AI chips, are enabling on-device intelligence and reducing energy consumption, but challenges must be addressed to fully realise their potential for sustainable, real-time applications.

DIRECT ENVIRONMENTAL IMPACT OF AI

The increasing use of AI models, particularly GenAI and LLMs, has significant environmental implications and this is expected to grow. These models require massive computational resources, resulting in high energy consumption, water usage, and mineral extraction. The estimated 5,000 data centres in the US and 2000 in Europe¹⁸⁸ are expected to increase their power demand by 50% by 2027 and between 103% and 165% by 2030 according to different estimates.^{189 190}

Based on conservative estimates, data centres' energy demand increased from 194 TWh in 2010 to 204 TWh in 2018.¹⁹¹ The global electricity

188. Statista. Leading countries by number of data centers as of March 2024. Blog post: <https://www.statista.com/statistics/1228433/data-centers-worldwide-by-country/>

189. IEA (2025), Energy and AI, IEA, Paris <https://www.iea.org/reports/energy-and-ai>

190. Goldman Sachs: AI to drive 165% increase in data center power demand by 2030

191. Masanet, Eric, Arman Shehabi, Nuoa Lei, Sarah Smith, and Jonathan Koomey. (2020). "Recalibrating Global Data Center Energy-Use Estimates." *Science* 367, no. 6481: 984–986.

consumption of data centres increased to 460 TWh in 2022. According to recent JRC estimates Data centres in the EU used an estimated 45–65 TWh of electricity in 2022 (1.8–2.6% of total EU electricity use), while telecommunication networks used an estimated 25–30 TWh of electricity (1–1.2% of total EU electricity use)¹⁹² IEA estimated that globally, data centres consumed around 1.5% of electricity consumption in 2024.¹⁹³

While it is challenging to identify the exact share of energy consumption attributable to AI, it is estimated that it currently accounts for 14% of data centre energy consumption, projected to rise to 27% by 2027.¹⁹⁴

The environmental impacts of AI infrastructure include:

- **Energy consumption:** High electricity demand for training and running models contribute to greenhouse gas (GHG) emissions, especially if the energy mix has a low share of renewable sources.¹⁹⁵ For example, different estimates rate that global energy demand for AI infrastructure will reach between 1% and 1.5% of total global energy consumption by 2027.¹⁹⁶ This has led AI companies to invest in nuclear energy and gas power generation.¹⁹⁷

192. Kamiya, G. and Bertoldi, P., Energy Consumption in Data Centres and Broadband Communication Networks in the EU, Publications Office of the European Union, Luxembourg, 2024, doi:10.2760/706491, JRC135926.

193. IEA (2025), Energy and AI, IEA, Paris <https://www.iea.org/reports/energy-and-ai>, Licence: CC BY 4.0

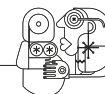
194. Goldman Sachs: AI to drive 165% increase in data center power demand by 2030 <https://www.goldmansachs.com/insights/articles/ai-to-drive-165-increase-in-data-center-power-demand-by-2030>

195. Ramachandran, K., Stewart, D., Hardin, K., & Crossan, G. (2024). As generative AI asks for more power, data centers seek more reliable, cleaner energy solutions. Deloitte Center for Technology Media & Telecommunications.

196. Morgan Stanley (2024). Powering the AI Revolution. Blog post: <https://www.morganstanley.com/ideas/ai-energy-demand-infrastructure>

197. CNBC Why Big Tech is turning to nuclear to power its energy-intensive AI ambitions <https://www.cnbc.com/2024/10/15/big-tech-turns-to-nuclear-energy-to-fuel-power-intensive-ai-ambitions.html>





SOCIETAL IMPACTS AND CHALLENGES

It should be noted that there is high uncertainty surrounding the numbers provided above, for several reasons. On the one hand, these estimates may not be taking into account the possible widespread adoption of reasoning models such as OpenAI's o1, which yield better capabilities at the expense of an increased demand in computational resources. On the other hand, these estimates were made before the release of DeepSeek-R1, which showed that it was possible to compete with OpenAI for a fraction of the energy consumption. At the same time, the observed pace of energy consumption growth can also be under-reflective of the actual demand, given limitations and bottlenecks, for instance on the availability of AI chips, multi-year-long lead times for equipment, and power availability constraints.¹⁹⁸

Obligations to release the full consumption of each data centre used for AI could help evaluate and reduce the total footprint.¹⁹⁹

- **Water footprint:** Projected to account for 4.2–6.6 billion cubic meters of water withdrawal in 2027, this exceeds the total annual water withdrawal of half of the United Kingdom.²⁰¹ This is a significant concern for building data centres in countries with water shortages. Adding to the challenge, other available cooling mechanisms are less efficient, producing

waste heat and potentially utilising harmful cooling agents like fluorinated-gases.

- **Mineral resources:** Extraction of raw materials like silicon, aluminium, copper, tin, tantalum, lithium, gallium, germanium, palladium, cobalt, and tungsten for manufacturing chips has significant environmental costs.²⁰² Hundreds of tonnes of ore are required to be excavated and processed for just one ton of material.²⁰³
- **E-waste:** Data centre hardware has a short lifespan of around 3.5 years,²⁰⁴ resulting in a potential accumulation of 1.2–5.0 million tons of e-waste during 2020–2030.²⁰⁵ Implementing a circular economy could increase the longevity of data centre hardware and reduce e-waste by 16–86% according to the same sources.

However, AI models can also help mitigate climate change by supporting applications like pollution tracking, weather monitoring, and energy optimisation.²⁰⁶ The impact of these applications is still to be quantified and their potential may be hindered by a number of factors, such as interoperability concerns, critical shortage of skills, and limitations in the digital infrastructure. It is worth noticing that data centres tend to be highly concentrated in spatial terms and, given their substantial power and water draw, this poses significant challenges to local reservoirs and to the power transmission

198. Bashir, Noman, Priya Danti, James Cuff, Sydney Sroka, Marija Ilic, Vivienne Sze, Christina Delimitrou, and Elsa Olivetti. (2024). The Climate and Sustainability Implications of Generative AI. *An MIT Exploration of Generative AI*, March. <https://doi.org/10.21428/e4baedd9.9070dfe7>.

199. Crawford, K. (2024). Generative AI's Environmental Costs Are Soaring—and Mostly Secret. *Nature*, 20 February, 2024.

200. Luccioni, A. S., Strubell, E., & Crawford, K. (2025). From Efficiency Gains to Rebound Effects: The Problem of Jevons' Paradox in AI's Polarized Environmental Debate. *arXiv preprint arXiv:2501.16548*.

201. Pengfei Li et al, Making AI Less "Thirsty": Uncovering and Addressing the Secret Water Footprint of AI Models <https://arxiv.org/pdf/2304.03271>

202. Brito, Griffin and Koski (2022), "Nvidia GPU — Design Life-Cycle" <https://www.designlife-cycle.com/nvidia-gpu>

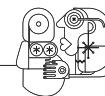
203. Mills (2020), " Mines, Minerals, and "Green" Energy: A Reality Check" <https://manhattan.institute/article/mines-minerals-and-green-energy-a-reality-check>

204. Veau et al, Navigating E-waste for datacenters <https://sustainability-ai.de/static/6d0446a876b6ccb79b58bbaeffcf8b62/Project-Report.pdf>

205. Wang, P., Zhang, LY., Tzachor, A. et al. (2024). E-waste challenges of generative artificial intelligence. *Nat Comput Sci* 4, 818–823.

206. Vinuesa, R., Azizpour, H., Leite, I. et al. (2020). The role of artificial intelligence in achieving the Sustainable Development Goals. *Nat Commun* 11, 233. <https://doi.org/10.1038/s41467-019-14108-y>





SOCIETAL IMPACTS AND CHALLENGES

grid. In Ireland, for example, data centres consume around 20% of the metered electricity supply. Localised environmental impacts of AI are likely to affect arable regions, exacerbating environmental inequalities.²⁰⁷

VOLUNTARY PROGRAMS AND REGULATION FOR DATA CENTRES

The negative environmental consequences of data centres' energy consumption and environmental footprint have been successfully addressed by voluntary programs, such as the EU Code of Conduct for Energy Efficiency in Data Centres.²⁰⁸ Through the commitment on the adoption of agreed and yearly updated best practices more than 400 data centre operators realised significant efficiency gains. The data collected within the program shows that the average power usage effectiveness (PUE) of participating companies fell from 1.8 to less than 1.3 over the last 15 years. Other similar initiatives demonstrate that the industry is actively working on the mitigation of their environmental sustainability with ambitious goals of being carbon neutral or water-positive. The AI community is also calling for more informed and sustainable use of AI.²⁰⁹

Regulators have also recently included sustainability of data centres among their priorities. In March 2024 the European Commission adopted the Delegated Act on a common rating scheme for data centres in the European Union. The Delegated Act implements the recast Energy Efficiency Directive (EED) and details the energy key performance indicators (KPI) that large data centre operators (with a power demand of the installed information technology of at least 500kW) must submit on

regular basis to the European database on data centres.

The EU Taxonomy Regulation²¹⁰ and, in particular, its Climate Delegated Act, also deal with data centres: Section 8.1 of the EU Taxonomy Climate Delegated Act indeed addresses the economic activity "Data processing, hosting and related activities" and sets out the technical screening criteria to identify whether a data centre contributes substantially to, respectively, climate change mitigation and to climate change adaptation, while doing no significant harm to other environmental objectives.²¹¹

EMERGING TECHNOLOGIES AND BREAKTHROUGH INNOVATIONS

Emerging technologies like energy-efficient transistors made from materials such as molybdenum disulfide promise to allow AI processing in small devices, such as smartwatches, enhancing on-device intelligence while reducing reliance on cloud computing.²¹² Similarly, but on a different scale, nanomagnetic computing offers ultra-low-energy computation by leveraging magnetic properties at the nanoscale, which could dramatically reduce energy consumption across AI-powered devices.²¹³

The development of spike neural networks is another promising innovation in the pursuit of energy-efficient AI. By mimicking the brain's

207. Rein S. and Wierman A. (2024), The Uneven Distribution of AI's Environmental Impacts Harvard Business Review <https://hbr.org/2024/07/the-uneven-distribution-of-ais-environmental-impacts>

208. <https://e3p.jrc.ec.europa.eu/en/groups/data-centres-code-conduct>

209. See tools to compare the energy consumption of AI models, e.g. <https://huggingface.co/blog/sasha/announcing-ai-energy-score>

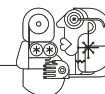
210. Regulation (EU) 2020/852 of the European Parliament and of the Council of 18 June 2020 on the establishment of a framework to facilitate sustainable investment, and amending Regulation (EU) 2019/2088 (OJ L 198 22.06.2020)

211. Bertoldi P., Assessment Framework for Data Centres in the Context of Activity 8.1 in the Taxonomy Climate Delegated Act, European Commission, Ispra, 2023, JRC131733

212. Hsu, Jeremy. "Energy-Efficient Transistor Could Allow Smartwatches to Use AI." New Scientist, 12 Oct. 2023, www.newscientist.com/article/2397235-energy-efficient-transistor-could-allow-smartwatches-to-use-ai/. Accessed 7 May 2025.

213. Imperial College London. "'Nanomagnetic' Computing Can Provide Low-Energy AI." ScienceDaily, 5 May 2022, www.sciencedaily.com/releases/2022/05/220505114646.htm





SOCIETAL IMPACTS AND CHALLENGES

signalling process, SNNs provide event-driven processing, drastically reducing idle energy consumption. These advancements are crucial for edge computing, IoT devices, and mobile AI, where energy constraints are critical.²¹⁴

²¹⁵ ²¹⁶ Furthermore, neuromorphic chips and optoelectronic neurons are optimising energy efficiency by emulating brain-like processing capabilities, which is particularly beneficial for applications in autonomous systems, healthcare monitoring, and wearables.

New specialised chips for edge AI are demonstrating enhanced inference speed and energy efficiency on local devices, minimising the need for data centre dependency. These chips, using technologies such as Binary Neural Networks (BNNs) and memristive crossbar arrays, allow for AI computations directly on devices, reducing latency and bandwidth requirements in smart cities, wearables, and industrial IoT applications.²¹⁷ ²¹⁸

As the use of GenAI increases, semiconductor innovations such as those exemplified will be key to enabling its sustainable and long-term development and uptake. Policy makers must therefore consider measures to foster RDI, as well as industrial scalability, hardware standardisation and supply chain security, ensuring a true twin transition where energy-efficient technologies meet growing adoption and performance demand.

214. Ward-Foxton, Sally. (4 Dec. 2023). "What Is Holding Back Neuromorphic Computing?" EE Times. www.eetimes.com/what-is-holding-back-neuromorphic-computing/.

215. Zhu, Rui-Jie, et al. (27 Feb. 2023). SpikeGPT: Generative Pre-Trained Language Model with Spiking Neural Networks. <https://doi.org/10.48550/arxiv.2302.13939>.

216. Cerf, Emily. 7 Mar. (2023). "SpikeGPT: Researcher Releases Code for Largest-Ever Spiking Neural Network for Language Generation." News. news.ucsc.edu/2023/03/eshraghian-spikegpt/. Accessed 7 May 2025.

217. Tang, Baoshan, et al. (2022). "Wafer-scale solution-processed 2D material analog resistive memory array for memory-based computing." Nature Communications, vol. 13, no. 1. Article number: 3037. <https://www.nature.com/articles/s41467-022-30519-w>.

218. Chen, Yu-Hsin, et al. (Jan. 2017). "Eyeriss: An Energy-Efficient Reconfigurable Accelerator for Deep Convolutional Neural Networks." IEEE Journal of Solid-State Circuits, vol. 52, no. 1, pp. 127–138. IEEE Xplore, <https://ieeexplore.ieee.org/document/7738524>

ENVIRONMENTAL IMPACT OF AI VIA ITS SOCIETAL IMPACT

As mentioned in [Section 4.2](#), biased AI models can influence attitudes, also on climate change. GenAI can also be a source of critical disinformation on climate, with models potentially biased due to hallucinations, training data issues, or reinforcement learning phase biases.²¹⁹ ²²⁰ ²²¹

As described in [Section 5.1](#), the AI Act includes environmental provisions on AI models and systems transparency, adhesion to codes of practice and information disclosure. The environmental requirements extend to any provider, including those based outside of Europe placing GenAI on the European market. While this constraint is a natural consequence of European legislation, it limits the capability of addressing environmental considerations at a global scale. Some US-based providers of AI have threatened not to serve Europe if the regulation is against their interest.²²² ²²³ In parallel, models developed in China are becoming competitive, but they may also fall short on compliance with the European AI rules that prevent the spread of biased information about climate or address direct environmental impacts of models.²²⁴ ²²⁵

219. Gartner website: Experts Answer the Top Generative AI Questions for Your Enterprise <https://www.gartner.com/en/topics/generative-ai>

220. [Turing.com](https://www.turing.com/): Top Generative AI Industry Applications: An In-Depth Look <https://www.turing.com/resources/generative-ai-applications>

221. CNN Analysis: DeepSeek's AI is giving the world a window into Chinese censorship and information control | CNN

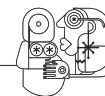
222. DW: Europe's AI bosses sound warning on soaring compliance costs <https://www.dw.com/en/europes-ai-bosses-sound-warning-on-soaring-compliance-costs/a-70243489>

223. Brookings: Are tariffs Big Tech's new tool against EU regulation? <https://www.brookings.edu/articles/are-tariffs-big-techs-new-tool-against-eu-regulation/>

224. Euronews: 'Harmful and toxic output': DeepSeek has 'major security and safety gaps,' study warns <https://www.euronews.com/next/2025/01/31/harmful-and-toxic-output-deepseek-has-major-security-and-safety-gaps-study-warns>

225. Wired: Here's How DeepSeek Censorship Actually Works—and How to Get Around It <https://www.wired.com/story/deepseek-censorship/>





SOCIETAL IMPACTS AND CHALLENGES

To address these challenges, the EU has been pioneering efforts against climate change and integrating environmental protection into policies like the EU Charter of Fundamental Rights [Article 37]. The EU is an ideal region to develop AI technology with strong environmental guarantees.

Climate action in the age of AI requires:

- **Energy-efficient computation:** Advancing innovation in semiconductors and promoting edge computing to reduce overall reliance on energy-intensive data centres (see examples in previous sub-section).
- **Renewable energy for data centres:** Reducing GHG emissions and promoting water-rich regions for data centre placement.
- **Efficient data centres:** Minimising energy and heat consumption as well as e-waste, while considering potential rebound effects.
- **Transparent and trustworthy AI models:** Encouraging citizens to choose models that provide reliable information on climate change.
- **International cooperation:** Establishing global standards to monitor and minimise AI's environmental impacts, considering the indirect effects of policies from main AI powerhouses like the USA and China.
- **Sustainable material use in semiconductors:** Ensuring responsible sourcing of raw materials and fostering innovation in semiconductor design to reduce material demand and enhance circularity through reuse and recycling.

Promoting **Green, EU-centric AI models and systems** and setting up international standards can help ensure that AI development aligns with environmental protection and climate action goals. Having state-of-the-art AIs developed in the EU as “trustworthy by design” or “compliant

by design” is a way to ensure that they comply with EU regulations. Potential strategies on AI and climate in the long term include the promotion of Green, EU-centric AI models and systems and the setup of international standards to monitor and minimise the environmental impacts of AI.

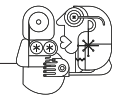
4.6 Generative AI and Children’s Rights

KEY MESSAGES

- GenAI presents both opportunities and challenges for children’s rights. While it has the potential to enhance education, creativity, communication, and information access, it also poses significant risks, such as deceptive manipulation, AI-based harmful content, and lack of privacy and safety.
- To ensure that GenAI is developed and used in a way that respects children’s rights and promotes their well-being, it is essential to design trustworthy GenAI systems, prioritise transparency and accountability, and invest in education and training programmes that promote critical thinking and media literacy skills.
- Child safeguards need to be embedded in GenAI design and must go beyond general privacy and ethics frameworks in order to address specific vulnerabilities of children. In this context, a rights-based, age-appropriate, and inclusive approach is essential to ensure safe, empowering, and equitable AI experiences for all children.

Children’s rights are a fundamental aspect of ensuring the well-being and development of young individuals, and recent advancements in GenAI have significant implications for this area, with children as disproportionately early adopters of technology. This Section reflects on the opportunities and challenges presented by GenAI in relation to children’s rights, with a focus on some relevant aspects, including education, creativity





SOCIETAL IMPACTS AND CHALLENGES

and mental health. Additional relevant input is also reported in [Sections 5.2](#) and [6.2](#).

OPPORTUNITIES

GenAI presents several opportunities for enhancing the lives of children. One key area is education, where personalised and adaptive learning experiences can be tailored to the needs and specificities of individual learners as mentioned in [Section 6.2](#). This can be achieved by using AI-based systems that provide real-time feedback and support, allowing teachers to shift their role to that of scaffolders, as introduced by Vygotsky's concept of a zone of proximal development (ZPD) that represents the gap between what a learner masters and is capable of doing with support from a teacher or a peer with more knowledge or expertise.²²⁶ For instance, AI-powered adaptive learning systems can adjust the difficulty level of educational content based on a child's performance and interests, providing a more effective and engaging learning experience.

GenAI also has the potential to support creativity in children, enabling them to develop new forms of expression using new tools. For example, children can use GenAI to create video games, animations, or other forms of digital content based on their ideas. This can facilitate access to technology and promote the development of creative and problem-solving skills, which are essential for success in the digital age. Moreover, GenAI can support communication by allowing children to efficiently document and sustain their communication, ensuring that their voices are heard. This can be particularly beneficial for children with disabilities or language barriers, who may face challenges in expressing themselves.

In healthcare, generative AI tools can support early detection of health and developmental issue or provide insights into medical data.²²⁷

226. Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

227. Generative AI: Risks and opportunities for children | Innocenti Global Office of Research and Foresight <https://www.unicef.org/innocenti/generative-ai-risks-and-opportunities-children>

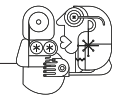
Finally, GenAI can support information access by translating complex content into age-adapted text, images, and other formats, making it more accessible to children. This can also be done in non-dominant languages, promoting inclusivity and diversity.

BARRIERS AND CHALLENGES

Despite the opportunities presented by GenAI, there are several barriers and challenges that need to be addressed. One significant risk is deceptive manipulation which is a form of influence where someone is led to believe something false or is persuaded to act in a way that is not in their best interest. Such deception or misleading tactics are prohibited under the AI Act. Because children's cognitive capacities are still in development, children are especially susceptible to the dangers of misinformation and disinformation. Additionally, AI-based harmful content, such as child sexual abuse material, poses a significant threat to children's safety and well-being. Another source of risk is the use of GenAI in the form of AI companions, which can be particularly detrimental for children, with risks of impersonation or deception in addition to tampering with children's emotional development ([see also Section 4.7](#)) and acquisition of interpersonal skills. The Article 28 of the DSA emphasises the need for minors' protection, highlighting the importance for providers of online platforms of ensuring high levels of privacy, safety and security for children ([see Section 5.2](#)) particularly regarding non-consensual image generation, which can have severe consequences for children's mental and physical health.

The risks of biases in training datasets, constructed and analysed by dominant actors, can also lead to unfair discrimination and to a loss of diversity and richness in GenAI outcomes. This can further result in a form of globalisation of GenAI outcomes, where the perspectives, experiences and needs of marginalised groups are excluded. Furthermore, like many emerging technologies, GenAI can worsen existing inequalities, particularly for





SOCIETAL IMPACTS AND CHALLENGES

children from marginalized communities who may face greater risks and have less access to its benefits. To be effective, generative AI must be inclusive, equitable, and responsible, catering to the diverse needs of all children.

The risks of hallucinations and the increased need for critical thinking skills are also significant concerns. With multiplied capacities compared to search engines and social media, GenAI has even more potential to shape users' perception of the world, social interactions and experiences. As GenAI becomes more prevalent, it is essential to develop and implement effective educational programmes that promote critical thinking and media and GenAI literacy skills (see Section 6.2), enabling children to analyse and evaluate the information they encounter online. Moreover, the potential risks of social discrimination, and negative impacts on children's mental and physical health must be carefully considered and addressed. Starting in this direction, an umbrella review and an expert report on the impacts of social media use (some embedding GenAI tools such as Chatbots) on adolescents' mental health and well-being flag that platforms' adoption of responsible design principles could improve risk mitigation.^{228 229}

WAYS FORWARD

To ensure that GenAI is developed and used in a way that respects children's rights and promotes their well-being, it is essential to design trustworthy GenAI systems. An example of this

is the AI system for collaborative storytelling,²³⁰ which is in line with the EU ethics guidelines for trustworthy AI.²³¹ The adaptation of ethical guidelines has identified relevant aspects for this population, such as stakeholder involvement, risk management, diversity and inclusion, children's rights and capacities, role of parents or carers, and the implementation of age-appropriate behaviours.

Transparency is a core transversal requirement for GenAI systems, particularly when it comes to children. Children-centric GenAI transparency requires informing children about the system's nature using age-appropriate language during child-AI interaction. These measures need to be adapted to age, linguistic, and cultural context, ensuring that children from diverse backgrounds can understand and engage with GenAI systems.

Evaluating the impact of GenAI on children's mental development is essential for better anticipatory governance response, and longitudinal studies can provide valuable insights into the effects of GenAI on children's cognitive processes and brain development.

4.7 Generative AI and Mental Health

KEY MESSAGES

- The use of AI chatbots and companion apps can lead to various mental health issues. These include addiction-like behaviours, validation-seeking tendencies, and in some cases,

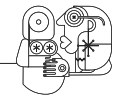
228. Arianna Sala, Lorenzo Porcaro, Emilia Gómez, Social Media Use and adolescents' mental health and well-being: An umbrella review, *Computers in Human Behavior Reports*, Volume 14, 2024, 100404, ISSN 2451-9588, <https://doi.org/10.1016/j.chbr.2024.100404>.

229. Beullens, K., Bozzola, E., Cataldo, I., Hale, L., Kent, M., Montag, C., Nivins, S., O'reilly, M., Rubæk, L., Schiøtz Thorud, H.-M., Sterpenich, V. And Vandenbosch, L., Minors' health and social media: an interdisciplinary scientific perspective, Manolios, S., Sala, A., Sundorph, E., Chaudron, S. And Gomez, E. editor(s), Publications Office of the European Union, Luxembourg, 2025, <https://data.europa.eu/doi/10.2760/3795891>, JRC141090.

230. Escobar-Planas, M. et al. (2025). Implementing and Evaluating Trustworthy Conversational Agents for Children. In: Plácido da Silva, H., Cipresso, P. (eds) *Computer-Human Interaction Research and Applications*. CHIRA 2024. Communications in Computer and Information Science, vol 2370. Springer, Cham. https://doi.org/10.1007/978-3-031-82633-7_29

231. European Commission: Directorate-General for Communications Networks, Content and Technology and Grupa ekspertów wysokiego szczebla ds. sztucznej inteligencji, Ethics guidelines for trustworthy AI, Publications Office, 2019, <https://data.europa.eu/doi/10.2759/346720>





SOCIETAL IMPACTS AND CHALLENGES

encouragement of harmful actions such as self-harm and disordered eating.

- Additionally, the potential for digital impersonation adds another layer of emotional risk, potentially causing distress to individuals affected by such practices.
- The rise of deep fakes and manipulated media can pose significant risks, also in the form of non-consensual explicit content and emerging forms of online harassment.

GenAI is reshaping how individuals interact with technology, introducing new risks concerning mental health. In the coming years, we can expect GenAI to become even more integrated into digital services, and GenAI content to become more pervasive online. This may be accompanied by the emergence of new types of risks.

Recent research has highlighted the potential risks that AI chatbots pose to a person's mental and physical well-being.²³² Specifically, some GenAI characteristics such as the chatbot's perceived sentience, its anthropomorphism or human-likeness, or its surprising ability to tell users what they want to hear (i.e. "sycophancy") may contribute to problematic or addiction-like use of GenAI systems,²³³ with emerging research aiming to understand the complex phenomena at play.

Recent reports have highlighted cases where AI companion apps encouraged users to engage in self-harm, eating disorder behaviour and violence.²³⁴

232. Robert Mahari and Pat Pataranutaporn, 'We Need to Prepare for Addictive Intelligence' (MIT Technology Review, 5 August 2024) <https://www.technologyreview.com/2024/08/05/1095600/we-need-to-prepare-for-addictive-intelligence/>

233. Pat Pataranutaporn and others, 'Influencing Human-AI Interaction by Priming Beliefs about AI Can Increase Perceived Trustworthiness, Empathy and Effectiveness' (2023) 5 Nature Machine Intelligence 1076.

234. Too human and not human enough: A grounded theory analysis of mental health harms from emotional dependence on the social chatbot Replika - Linnea Laestadius, Andrea Bishop, Michael Gonzalez, Diana Illenčik, Celeste Campos-Castillo, 2024 <https://journals.sagepub.com/doi/abs/10.1177/14614448231142007>

Since then, new incidents of self-harm involving teenage users of GenAI have been widely reported, as well as cases involving the digital impersonation of real people, including deceased minors.²³⁵

Moreover, the rise of deep fakes enabled by GenAI, where it becomes difficult to distinguish between real and fake content, poses significant risks to mental health, increasing the risks of cyberbullying. This can take the form of creation and diffusion of non-consensual explicit content, especially among adolescents who are more prone to be vulnerable during this developmental stage. Deep fakes introduce a new dimension of harm through the creation of convincing falsified videos that not only damage reputation but also cause psychological trauma.²³⁶ ²³⁷ The risk is not new, considering that the first photorealistic videos appeared on online platforms as early as 2017,²³⁸ but the capacities of GenAI have exacerbated it tremendously. A recent study carried out by Oxford Research Institute shows that deep fake generators have been downloaded almost 15 million times since late 2022, and 96% of the deep fake models primarily targeted women.²³⁹ A case in point is the Almendralejo case that took place in 2023, where a number of teenage girls received photos of themselves naked that were AI-generated by their peers,

235. Open letter to UK online service providers regarding Generative AI and chatbots - Ofcom <https://www.ofcom.org.uk/online-safety/illegal-and-harmful-content/open-letter-to-uk-online-service-providers-regarding-generative-ai-and-chatbots/>

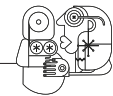
236. Alexander, S. (2025). Deep fake Cyberbullying: The Psychological Toll on Students and Institutional Challenges of AI-Driven Harassment. The Clearing House: A Journal of Educational Strategies, Issues and Ideas, 98(2), 36–50. <https://doi-org.ejournals.um.edu.mt/10.1080/00098655.2025.2488777>

237. Vaccari, C., and A. Chadwick. 2020. Deep fakes and disinformation: Exploring the impact of synthetic political video on deception, uncertainty, and trust in news. Social Media + Society 6 (1):1–13. doi: 10.1177/2056305120903408.

238. Are Deep fakes Concerning? Analyzing Conversations of Deep fakes on Reddit and Exploring Societal Implications <https://www.vice.com/en/article/deepfake-porn-origins-sexism-reddit-v25n2/>

239. OII | Dramatic rise in publicly downloadable deep fake image generators <https://www.oii.ox.ac.uk/news-events/dramatic-rise-in-publicly-downloadable-deepfake-image-generators/>





SOCIETAL IMPACTS AND CHALLENGES

causing them a high degree of distress.²⁴⁰ Researchers are starting to document and study the impact of this phenomenon on its victims²⁴¹ and what legal actions might be taken in such cases. For example, while many schools have general cyberbullying policies, they often do not take into account AI-generated content's unique attributes.²⁴² Services beyond social media and pornographic platforms (see the analysis in the context of DSA in [Section 5.2](#)), such as app stores and search engines, are also exposed to similar risks, for example, the emergence of “nudify” apps that modify images of women to depict them naked has been reported.²⁴³ Moreover, while human participants remain essential in detecting cyberbullying, the use of multiple algorithms could be trained to recognise cultural colloquialisms and slang terminology to facilitate detection in the future.²⁴⁴

240. The Almendralejo case: When AI deepfakes are used to undress teenagers | Euronews Tech Talks Podcast <https://www.euronews.com/next/2023/11/15/the-case-of-almendralejo-when-deepfakes-are-used-to-undress-teenagers-euronews-tech-talks->

241. Are Deepfakes Concerning? Analyzing Conversations of Deepfakes on Reddit and Exploring Societal Implications <https://www.vice.com/en/article/deepfake-porn-origins-sexism-reddit-v25n2/>

242. Alexander, S. (2025). Deep fake Cyberbullying: The Psychological Toll on Students and Institutional Challenges of AI-Driven Harassment. *The Clearing House: A Journal of Educational Strategies, Issues and Ideas*, 98(2), 36–50. <https://doi-org.ejournals.um.edu.mt/10.1080/00098655.2025.2488777>

243. Deepfake Defences – Mitigating the Harms of Deceptive Deepfakes – Ofcom Discussion Paper <https://www.ofcom.org.uk/siteassets/resources/documents/consultations/discussion-papers/deepfake-defences/deepfake-defences.pdf?v=370754>

244. Gomez, C.E., Sztainberg, M.O. & Trana, R.E. Curating Cyberbullying Datasets: a Human-AI Collaborative Approach. *Int Journal of Bullying Prevention* 4, 35–46 (2022). <https://doi.org/10.1007/s42380-021-00114-6>

4.8 Gender – as a Specific Case of Bias and AI Social Implications

KEY MESSAGES

- GenAI systems can perpetuate existing biases and stereotypes, particularly if they are trained on data sets that reflect historical and systemic inequalities, highlighting the need for human-rights-based and ethical AI use.
- The case of credit risk assessment and the potential risks of GenAI in perpetuating gender bias highlights the importance of establishing frameworks and guidelines for fair and transparent AI decision-making, with a focus on mitigating unfair biases and ensuring accountability.

Diversity, fairness, and non-discrimination are fundamental to building Trustworthy AI.²⁴⁵

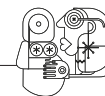
Eliminating unfair bias in AI is crucial to preventing negative outcomes, such as the marginalisation of vulnerable groups and the reinforcement of prejudice and discrimination. AI fairness is assessed by using a set of protected attributes, including g racial or ethnic origin, religion or belief, class, disability, age, gender or sexual orientation, among others, gender, national origin, and age, in line with the EU Charter for Fundamental Rights. In this respect, AI systems should be evaluated considering different protected groups, especially in high-risk scenarios ([see Section 5.1](#)). Furthermore, AI actors should strive to minimise and prevent discriminatory or biased outcomes throughout the AI system's life cycle to ensure its fairness²⁴⁶ and comply with the EU non-discrimination legal framework.²⁴⁷

245. HLEG. (2019). *Ethics guidelines for Trustworthy AI*. European Commission. <https://digital-strategy.ec.europa.eu/en/library/ethics-guidelines-trustworthy-ai>

246. UNESCO. (2022). *Recommendation on the Ethics of Artificial Intelligence* (No. HS/BIO/PI/2021/1). UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000381137>

247. https://commission.europa.eu/aid-development-cooperation-fundamental-rights/your-fundamental-rights-eu/know-your-rights/equality/non-discrimination_en





BIAS, STEREOTYPES, AND FAIRNESS

The growing integration of AI across various sectors has heightened concerns about biases in LLMs, including those related to gender, religion, race, profession, nationality, age, physical appearance, and socio-economic status.²⁴⁸ While AI holds the promise of enhancing efficiency and decision-making in areas like healthcare, education, and business, its widespread use and the high level of public trust it enjoys could also amplify societal prejudices, leading to systematic disadvantages, particularly for women.²⁴⁹

GenAI systems can perpetuate and even amplify existing biases, particularly when trained on data reflecting historical inequalities. Research has highlighted persistent social biases in modern language models, despite efforts to mitigate them. For instance, in gendered word association tasks, recent models still associate female names with traditional roles like “home” and “family,” while linking male names with “business” and “career.” Moreover, in text generation tasks, these models produce sexist and misogynistic content approximately 20% of the time.²⁵⁰

Similarly, an analysis of occupational portraits generated by three popular text-to-image AI generators revealed significant gender and racial biases. Women and black individuals were notably underrepresented, especially in roles requiring high levels of preparation. Women were often portrayed as younger and with submissive gestures, while men appeared older and more authoritative. Alarming, these biases surpassed

real-world disparities, indicating that the issues extend beyond merely biased training data.²⁵¹

Research has demonstrated that AI can exhibit gender bias when used for decision-making processes. For example, in **recruitment**, AI algorithms have been observed to favour male candidates over equally qualified female candidates.²⁵² Moreover, in 2025, the Netherlands Institute for Human Rights found a violation of Dutch and EU anti-discrimination legislation in Meta’s job vacancy advertising algorithm.²⁵³ For example, in violation of the principles of equal treatment and non-discrimination, in 2023, the algorithm in the Netherlands displayed vacancies for receptionist positions to female users in 97% of cases. Similarly, it showed vacancies for mechanics to male users 96% of the time. In healthcare, AI systems have shown gender bias in diagnostic reasoning, leading to treatment recommendations that may misdiagnose or inadequately address conditions in women, often due to training on predominantly male-centric data.²⁵⁴ In **education**, while AI has the potential to create personalized learning paths tailored to individual competencies and needs, it may also unfairly predict higher dropout rates for female students, particularly in male-dominated fields like science, technology, engineering and mathematics (STEM), thereby limiting their access to advanced education programmes.²⁵⁵ Given

251. Zhou, M., et al. (2024). *Bias in Generative AI* (Version 1). arXiv. <https://doi.org/10.48550/ARXIV.2403.02726>

252. Dustin, J. (2022). Amazon scraps secret AI recruiting tool that showed bias against women. In *Ethics of data and analytics* (pp. 296-299). Auerbach Publications

253. The Netherlands Institute for Human Rights rules that Meta’s algorithm engages in prohibited gender discrimination <https://www.prakkendoliveira.nl/en/news/2025/the-netherlands-institute-for-human-rights-rules-that-metas-algorithm-engages-in-prohibited-indirect-gender-discrimination>

254. Zack, T., et al. (2024). Assessing the potential of GPT-4 to perpetuate racial and gender biases in health care: A model evaluation study. *The Lancet Digital Health*, 6(1), e12–e22. [https://doi.org/10.1016/S2589-7500\(23\)00225-X](https://doi.org/10.1016/S2589-7500(23)00225-X)

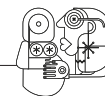
255. Gardner, J., Brooks, C., & Baker, R. (2019). Evaluating the fairness of predictive student models through slicing analysis. In *Proceedings of the 9th international conference on learning analytics & knowledge* (pp. 225–234). <https://doi.org/10.1145/3303772.3303791>

248. Gallegos, I. O., Rossi, R. A., Barrow, J., Tanjim, M., Kim, S., Dernoncourt, F., Yu, T., Zhang, R., & Ahmed, N. K. (2024). Bias and Fairness in Large Language Models: A Survey. *Computational Linguistics*, 50(3). <https://doi.org/10.1162/coli.a.00524>

249. Hall, P., & Ellis, D. (2023). A systematic review of socio-technical gender bias in AI algorithms. *Online Information Review*, 47(7), 1264–1279. <https://doi.org/10.1108/OIR-08-2021-0452>

250. UNESCO & International Research Centre on Artificial Intelligence. (2024). Challenging systematic prejudices: An investigation into bias against women and girls in large language models. UNESCO.





that biased algorithms can lead to discrimination against vulnerable groups, particularly those experiencing the intersection of multiple forms of discrimination, such as racial minorities or individuals with disabilities.

RISK MITIGATION STRATEGIES

Addressing gender biases in AI systems is crucial for developing equitable technologies. This requires multiple strategies, such as enhancing the diversity and representativeness of training datasets, incorporating fairness-focused algorithmic approaches, and increasing diversity among AI developers.²⁵⁶ Since bias mitigation cannot be achieved through one-time interventions, it is essential to conduct regular audits and monitoring using diverse benchmark datasets and methodologies to identify and address biases throughout the AI life cycle.²⁵⁷

Establishing robust policy frameworks at both corporate and governmental levels can guide the ethical development and deployment of AI systems. By implementing comprehensive risk mitigation strategies and fostering inclusive development practices, AI can be harnessed to promote societal equity and prevent the reinforcement of existing inequalities. Under the EU DSA, very large online platforms (VLOPs) and search engines (VLOSEs) must perform annual risk assessments on systemic risks stemming from their services and their algorithmic systems, including risks to fundamental rights, such as non-discrimination, as well as risks of gender-based violence (see Section 5.2).

256. Ho, J., et al. (2025). Gender biases within Artificial Intelligence and ChatGPT: Evidence, Sources of Biases and Solutions. *Computers in Human Behavior: Artificial Humans*, 4, 100145. <https://doi.org/10.1016/j.chbah.2025.100145>

257. UNESCO & International Research Centre on Artificial Intelligence. (2024). *Challenging systematic prejudices: An investigation into bias against women and girls in large language models*. UNESCO. <https://unesdoc.unesco.org/ark:/48223/pf0000388971>

A CASE STUDY: GENDER BIASES IN CREDIT RISK ASSESSMENT

The adoption of AI is also increasing across financial institutions, offering the potential to affect activities such as credit scoring, customer support or regulatory compliance. However, this adoption comes with its own set of challenges, including the mitigation of biases inherent in AI systems.²⁵⁸ Among the many applications, credit risk assessment is one of the key areas where banks are actively exploring the use of AI and which is considered high risk under the AI Act (see Section 5.1). Unlike traditional approaches, which are typically based on statistical modelling, the arrival of AI allows financial institutions to potentially base their credit ratings on characteristics of the loan applicant, make decisions autonomously and generate customer risk profiles much faster, descriptive statistics on the rate of credit approval in the US as of 2022 confirm that there is bias in traditional credit scoring.²⁵⁹ Notably, these report a gender bias of approximately 4%. As an alternative to this traditional approach, the JRC conducted an experiment to assess whether GenAI tools exhibit similar or higher biases, evaluating the validity and reliability of LLMs in supporting financial decision-making.

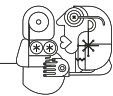
To do so, the exercise relies on a dataset containing over 400,000 short biographies with labelled profession and gender.²⁶⁰ For each profession with enough available observations, 1,000 male and 1,000 female biographies are randomly paired, resulting in a final dataset of 17 professions and 34,000 individuals. Finally, the usage of an LLM as a credit risk scoring algorithm is evaluated by prompting it to choose precisely one person from each pair to provide credit to. To control for

258. "Intelligent financial system: how AI is transforming finance" by Iñaki Aldasoro, Working Papers No 1194 by Leonardo Gambacorta, Anton Korinek, Vatsala Shreeti and Merlin Stein

259. Song, Z., Rehman, S. U., PingNg, C., Zhou, Y., Washington, P., & Verschueren, R. (2024). Do FinTech algorithms reduce gender inequality in banks loans? A quantitative study from the USA. *Journal of Applied Economics*, 27(1).

260. e-Arteaga, Maria, et al. "Bias in bios: A case study of semantic representation bias in a high-stakes setting." *Proceedings of the Conference on Fairness, Accountability, and Transparency*. 2019.





potentially skewed quality of candidates in favour of a specific gender, the results are compared to a baseline alternative in which all gender-identifiable information (names, pronouns) is removed from the biographies and replaced with neutral references (they/them pronouns, no names).

This experiment shows that:

1. The LLM exhibits a similar level of bias as is currently observed in traditional credit provisioning, with a gap of around 4% in favour of men.
2. The situation is heterogeneous across professions, with women in some traditionally female-dominated professions (nurses, teachers) exhibiting a higher likelihood of obtaining a loan than their male counterparts.

4.9 The Contribution of a Behavioural Approach to AI Policy Analysis

KEY MESSAGES

- Integrating behavioural insights helps regulate GenAI to operate transparently and equitably, aligning its developments with societal values and increasing public trust.
- Agentic AI can potentially reduce human errors in decision-making. Nevertheless, it might also be biased, thanks to the human-produced data on which it feeds.
- Behavioural insights can help determine when AI should take precedence over human judgment to enhance social welfare.

In recent years, there has been a significant shift in policymaking, often referred to as the “behavioural turn”. This approach leverages insights from psychology and behavioural economics to

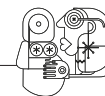
inform policymaking. It does this in a number of ways. For one, it helps to influence policies so that they are better aligned with the natural ways in which people think and make decisions. Rather than assuming individuals always act rationally, behavioural insights help tailor interventions to subtly encourage better choices without coercion. But also, a behavioural approach helps identify those cases where market players take advantage of citizens and consumers’ imperfect thinking to steer their behaviour in a particular way (to polarise them or persuade them to consume). In other words, a behavioural approach in policymaking can help promote the use of behavioural insights for good, while limiting its use for bad.

There is significant potential for such an approach to inform the policy analysis of GenAI. It gives us a broad toolkit to work with. For one, it can help ensure GenAI interacts with humans in ways that protect the user and promote trust in AI. Regulations can help safeguard individuals, particularly the vulnerable populations who may struggle with complex interfaces. By integrating behavioural insights into policy, regulators can provide better guidelines that ensure AI tools operate transparently and are accessible to all users. The aim is to create a framework where AI systems are not only effective but also equitable, ensuring the benefits are widely distributed across the EU’s diverse populace.

We can also leverage behavioural insights to ensure GenAI (and future developments, like Agentic AI, [see Section 2.3](#)) aligns with the broader ethical frameworks that govern societal values and fundamental rights. There are things that we, as humans, consider fair, moral, and/or ethical. Sometimes this can be readily stated, other times it needs to be revealed by systematically observing subjects’ behaviour, e.g. through experimentation. The resulting insights can guide the EU in regulating the development of GenAI so that it respects these core values.

A behavioural approach can also help address the issue of bias in GenAI. This not only involves addressing technical and data-driven aspects,





SOCIETAL IMPACTS AND CHALLENGES

ensuring that the datasets used to train AI systems are diverse and representative, but also requires a deeper understanding of cognitive biases inherent in human decision-making. A behavioural approach here can provide us with a framework to recognise and address these biases within AI systems. Cognitive biases, such as confirmation bias or anchoring, can also inadvertently be mirrored in GenAI if not carefully managed.

Looking towards the future, the further developments of AI will only increase the role for a behavioural approach. Agentic AI, for one, represents an important evolution within the field of Artificial Intelligence, building upon the capabilities of GenAI to function as autonomous entities. While GenAI focuses on creating content – such as text, images, or music – agentic AI extends this functionality, enabling systems to act independently within decision-making processes. These systems are capable of perceiving their environment, making complex decisions, and interacting with both humans and other machines to achieve specified goals. This advancement enhances GenAI's utility but also presents new challenges that must be addressed. Here, a behavioural approach will be particularly valuable.

Agentic AI will build on its ability to learn about user behaviour and preferences in fine detail. While this capability will offer powerful personalisation, it will also present the risk of exploiting users' cognitive biases and lack of information, potentially leading people to make choices that are not in their best interest. Behavioural research in this area can help identify if agentic AI is exploiting such biases and inform policies to counteract such objectionable practices.

However, in addition to risks there are opportunities. Agentic AI's decision-making can also be designed to overcome the cognitive biases that can skew human judgment. Judges and doctors, for example, are subject to biases themselves, and can (and do) make mistakes. Does this mean agentic AI will always be superior? Behavioural research is necessary to determine when it might be better for humans to make decisions, and when there is

scope for agentic AI to take over and improve these decisions for the good of social welfare.

A final consideration: both GenAI and Agentic AI make assumptions of users' behaviour and preferences by building on observed behaviour – such as time spent on social media – which may not accurately reflect individuals' true interests or contribute to their well-being. This flaw in perception can lead to outcomes like increased polarisation or sedentary behaviour. Understanding this dynamic calls for the development of more modern and robust behavioural models, and a more fine-grained oversight, to ensure the behaviours being promoted are conducive to overall well-being.

4.10 Privacy and Data Protection – a Societal Standpoint

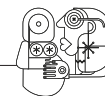
KEY MESSAGES

- GenAI can have very significant societal impacts and raise serious challenges for individuals' privacy and personal data protection. This is because of GenAI's capacity for data analysis and relationship.
- While such challenges should not hamper the development of GenAI, and the positive societal aspects it has to offer, such impact and challenges cannot be disregarded, and a clear technical and legal understanding of the issues raised by GenAI is required across societal sectors, including users, regulators and policymakers, in order to ensure that such issues are adequately – and timely – taken into consideration, discussed and addressed.

THE NOTIONS OF DATA PROTECTION AND PRIVACY

In the EU legal framework, the right to personal data protection and the right to privacy are two fundamental rights, enshrined in key legal texts.





SOCIETAL IMPACTS AND CHALLENGES

They can often be used interchangeably. In the case of the US, for example, the concept of privacy often incorporates that of personal data protection; that being said, they differ considerably.

The right to a private and family life appears in fundamental texts such as the Universal Declaration of Human Rights (Article 12), the European Convention of Human Rights (Article 8) and the EU Charter of Fundamental Rights (Article 7), as well as in the EU Treaties. The right under Article 8 of the European convention of Human Rights and the right under Article 7 of the EU Charter of Fundamental Rights covers the confidentiality of communications, often referred to in the EU context as the right to privacy. Article 8 of the European Convention of Human Rights also covers the protection of personal data, in the absence of a more specific right in that Convention. As stated by the European Court of Human Rights under the abovementioned Article 8, a person's private life "encompasses a wide range of interests", including "a person's identity and personal development, the right to establish and develop relationships with other human beings" and "[a]ctivities of a professional or business nature" – as a result, it is a "notion not susceptible to exhaustive definition".²⁶¹

The right to personal data protection concerns the protection of one's personal data, as stipulated under Article 8 of the EU Charter of Fundamental Rights and under Article 16 of the Treaty on the Functioning of the European Union (TFEU). Such data need to be processed fairly, for specific purposes and subject to a clear legal basis, with individuals being granted specific rights as to the processing of their data, which is overseen by an independent authority.

"Personal data" is a broad concept, which makes any discussions involving it more complex as it applies to all types of personal data, including publicly available personal data.²⁶² The first law

261. See Council of Europe, Privacy and Data Protection, Explanatory Memorandum, par. 2, available at: <https://www.coe.int/en/web/freedom-expression/privacy-and-data-protection-explanatory-memo>

262. Hamburg Data Protection Authority, Discussion Paper: Large Language Models and Personal data, 2024, p. 4.

in Europe to specifically address the protection of personal data dates back to 1970s Germany,²⁶³ and the current General Data Protection Regulation (Regulation (EU) 2016/679), applicable in the EU since 2018, is considered by several as the international data protection standard.²⁶⁴

It is easy to see how personal data and privacy, as rights and as concepts, can be intertwined. The notion of personal data consists of information that relates to an individual (a "natural person") that can render them directly or indirectly identified, or identifiable. As a result, it will very often be the case that the violation of an individual's personal data will result in the violation of their private life.

Within the GenAI life cycle, personal data can potentially be found at all stages, from the training of the language models to the outputs produced by an AI agent. With its use becoming widespread, it is very important to assess the impact that GenAI tools have on the processing of our personal data and on our privacy. Additionally, it is crucial to identify the challenges raised by its use in relation to the processing of our personal data. This section focusses on the societal impact and challenges, leaving out any legal discussions on the topic, which are addressed in [Section 5](#).

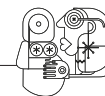
CHALLENGES AND IMPACT FOR DATA PROTECTION AND PRIVACY

Challenges to data protection and privacy stemming from the development and use of GenAI tools exist at all stages of the GenAI life cycle.

263. European Parliament, Understanding EU data protection policy, Briefing, January 2025, available at: [https://www.europarl.europa.eu/RegData/etudes/BRIE/2022/698898/EPRS_BRI\(2022\)698898_EN.pdf](https://www.europarl.europa.eu/RegData/etudes/BRIE/2022/698898/EPRS_BRI(2022)698898_EN.pdf), p. 2

264. "It is arguably the most significant data protection legislation in the world today", King, J., Meinhardt, C., Rethinking Privacy in the AI Era – Policy provocations for a data-centric world, White Paper, February 2024, Stanford University, p. 10; see also Zanfir-Fortuna, G., Why Data Protection Legislation offers a powerful tool to regulating AI February 2025, pdpEcho, available at: <https://pdpecho.com/2025/02/26/why-data-protection-legislation-offers-a-powerful-tool-for-regulating-ai/>





SOCIETAL IMPACTS AND CHALLENGES

A simpler way of thinking about this concept is by dividing it into two phases:

1. “development” phase, which concerns all aspects of the GenAI model/system creation, such as code development, collection of (personal and non-personal) training data, and the training itself, and
2. “deployment” phase, which covers all stages relating to the use of an AI model (including as part of an AI system), as well as any activities in the post-development phase, like fine-tuning.²⁶⁵

The amount of data required to develop an AI model (before it is part of an AI system) is often very significant, and is steadily increasing, together with our ability to collect it in many instances of our daily lives. In the words of King and Meinhardt, “AI’s appetite for data currently knows few bounds”.²⁶⁶ This “insatiable appetite for data” is bound to increase,²⁶⁷ especially in light of the evidence that larger datasets can improve an AI system’s capabilities.²⁶⁸

Deeply linked with quantity is the matter of quality: large datasets may often contain data that are inaccurate, or biased, leading to inaccurate or harmful outputs in relation to individuals, in particular those in less represented communities.²⁶⁹

There is also the question of inference: GenAI systems may allow those using them to infer information about an individual that may go well beyond what that individual intended to disclose in the first place.²⁷⁰ That may result in individuals being subject to misguided decisions, or possibly even discrimination.²⁷¹ In an often-quoted example involving a large retail corporation, an algorithm was developed to identify pregnant women based on their shopping patterns, leading to a father discovering about his teenage daughter’s pregnancy after receiving coupons for baby-related products at home. There, the main concern, as raised by Solove, is that the algorithm could infer sensitive personal data about an individual from data that are quite common to retrieve.²⁷²

The way in which data are collected matters, especially since one of the most common techniques is web-scraping: simply put, it is a technique that “enables the automated collection and extraction of certain information from different publicly available sources on the Internet (such as websites)”, which can then be used, in this context, for the training of AI models.²⁷³ Web-scraping as a technique, however, raises important questions about the impact it has on those whose data are collected, since individuals “may lose control of their personal information when this is collected without their knowledge, against their

265. This is also the approach followed by the European Data Protection Board in its Opinion 28/2024 on certain data protection aspects related to the processing of personal data in the context of AI models, adopted on 17 December 2024. See, in particular, par. 18, p.11.

266. King, J., Meinhardt, C., Rethinking Privacy in the AI Era – Policy provocations for a data-centric world, White Paper, February 2024, Stanford University, p. 17

267. Solove, D., Artificial Intelligence and Privacy, Florida Law Review, Vol. 77, 2025, p. 18

268. King, J., Meinhardt, C., Rethinking Privacy in the AI Era – Policy provocations for a data-centric world, White Paper, February 2024, Stanford University, p. 37. King and Meinhardt also discuss – and challenge – this notion (see p. 38).

269. “An AI system built using a dataset collected from a city will only have a small percentage of certain minority groups, say 5%. If the dataset is used as-is, then the outputs of this AI system will be biased against this minority group because they only make up 5% of the dataset and the AI system has relatively less data to learn from

about them.”, European Data Protection Board Support Pool of Experts, Shrishak, K., Bias Evaluation, AI-Complex Algorithms and Effective Data Protection Supervision, March 2024, p. 6.

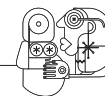
270. Solove, D., Artificial Intelligence and Privacy, Florida Law Review, Vol. 77, 2025, pp. 36-37

271. “The common factor among these risks is the fear of privacy pervasive data collection that allows sensitive inferences to be drawn. These inferences can lead to discrimination, especially when shared with third parties such as insurance companies, financial institutions, or employers.”, Wachter, S., Data Protection in the Age of Big Data, Nature Electronics, Vol. 2, 6–7, January 2019, p. 2

272. Solove, D., Artificial Intelligence and Privacy, Florida Law Review, 2025, p. 37

273. European Data Protection Board, Report on the work undertaken by the ChatGPT Taskforce, 23 May 2024, par. 15.1





SOCIETAL IMPACTS AND CHALLENGES

expectations and for purposes that are different from those of the original publication”.²⁷⁴

Purpose is also an issue: what is the model being trained for? What use will it have? Which type of AI system will it integrate, with what functionalities? Can it be repurposed at some point? This is particularly the case of General Purpose Artificial Intelligence (or GPAI) models/systems. This multi-modality can have a very significant impact on individuals, as systems can be gravely misused (e.g. for fraud, deep fakes as discussed in [Section 5.2](#), and identity theft).

Finally, there is the literacy challenge: are we equipped, as individuals, to understand why and how our personal data are collected, and what they are used for? Do we, as individuals, and as a society, have a clear understanding of how these systems work, the challenges they raise, and the impacts on our lives? This is especially relevant since it is questionable whether even those deploying such systems are able to explain the decisions and outcomes stemming from such systems (the so-called “black box effect”).²⁷⁵

This is essential to determine whether individuals have the capacity – or are given the tools – to understand and act upon the challenges raised by this technology to their personal data and privacy, including through legal mechanisms. It is also relevant to understand whether their actions

produce any effects, be they at the core level of data collection and training, or further down the AI life cycle, when, for example, inferences, relying on GenAI-based tools, are being made about them. Finally, it is also crucial for regulators to understand what measures to put in place, in addition to existing ones, to grant individuals the ability to understand and control, or at least mitigate, the impact this technology may have on people’s data and privacy. ■

274. European Data Protection Supervisor, Opinion 41/2023 on the Proposal for a Regulation on European Union labour market statistics on businesses, 25 September 2023, par. 17. See also par. 18. One well known example is that of Clearview. In the words of the Dutch Data Protection Authority, which imposed a fine on Clearview AI of circa 30 million EUR: “Clearview is a commercial business that offers facial recognition services to intelligence and investigative services. [...] For this purpose, Clearview has a database with more than 30 billion photos of people. Clearview scrapes these photos automatically from the Internet. And then converts them into a unique biometric code per face. Without these people knowing this and without them having given consent for this”, available at: <https://www.autoriteitpersoonsgegevens.nl/en/current/dutch-dpa-imposes-a-fine-on-clearview-because-of-illegal-data-collection-for-facial-recognition>

275. European Data Protection Supervisor, TechDispatch 2/2023, Explainable Artificial Intelligence, November 2023, p. 3

