GENERATIVE ARTIFICIAL INTELLIGENCE. OPPORTUNITIES, RISKS AND POLICY CHALLENGES

EPTA report 2023

EPTA

European Parliamentary Technology Assessment

GENERATIVE ARTIFICIAL INTELLIGENCE. OPPORTUNITIES, RISKS AND POLICY CHALLENGES

EPTA report 2023

EPTA
European Parliamentary Technology Assessment

cap cit Consell Assessor del Parlament sobre Ciència i Tecnologia

Contents

Preface

Distinguished reader,

Generative artificial intelligence (AI) is a subset or specialised branch of AI that specifically deals with the generation of content or data that is often creative, human-like, or novel. Generative AI systems are designed to create new content, such as text, images, music, or video that can mimic human creativity or generate content that is contextually relevant. As a result, generative AI technologies have drawn significant attention in recent years due to their ability to create realistic and often convincing content in various creative domains.

AI technologies continue to advance and affect various facets of society, including the economy, privacy, security or its ethical sphere. In this context, Parliaments are called to play a pivotal role in navigating the complex landscape of AI. Democratically elected bodies have to take the lead in regulating this burning issue. In doing so, they serve as the voice of the people and are entrusted with crafting and passing laws that can effectively address AI-related risks.

Moreover, parliaments provide a platform for informed, inclusive, and transparent deliberations, allowing experts, stakeholders, and citizens to contribute to AI policy making. Their involvement ensures that regulations are comprehensive, forward-looking, and reflective of societal values, striking a balance between fostering innovation and protecting against potential harm.

AI has the potential to reshape our societies as a whole. For this reason, the active engagement of parliaments is critical to safeguarding the public interest and ensuring that AI technologies are both developed and deployed responsibly.

You have in your hands the collaborative efforts of the European Parliamentary Technology Assessment (EPTA) network presented to the 2023 Conference on Generative Artificial Intelligence – Opportunities, Risks, and Policy Challenges (A time for technology assessment to parliaments), held in the premises of the Parliament of Catalonia in Barcelona on 9 October 2023.

This collective endeavour brings together the expertise, alternatives, and perspectives of different parliaments in a multi-level governance scheme.

This report underscores the significance of collaboration in tackling complex global issues and, at the same time, serves as a roadmap for policymakers and the public in general to navigate the AI frontier responsibly. Together, we are shaping the future of AI by championing ethical AI principles, promoting transparency, and guiding the development of AI technologies that benefit society as a whole.

I wish you an instructive and enjoyable reading!

Anna Erra i Solà
President of the Parliament of Catalonia and President of CAPCIT
(Advisory Board of the Parliament of Catalonia for Science and Technology)

Barcelona, October 2023

Executive Summary

Why is it relevant to talk about generative artificial intelligence (AI)?

In the last decade, deep learning methods (a subfield of machine learning, in turn a subfield of AI) have seen remarkable improvements in accuracy and generalisation, leading to breakthroughs in healthcare (e.g. cancer diagnosis) or science (e.g. weather prediction or protein modelling).

Some of the most successful developments using a deep learning methodology are large language models (LLMs). These models are aimed at generating sequences of words that are plausible but not necessarily truthful. The purpose of the developers is to achieve good statistical approximations to what humans could possibly write according to the many human-written texts used to train the model.

Several corporations have built applications on top of LLMs like ChatGPT or Bard. These applications serve many purposes like searching for information, summarising texts, or creating content based on a dialogue with a user. A portion of the public have perceived these systems as an example of machines acquiring human-like intelligence and this has led to warnings from the scientific community against that misperception.

There are a number of issues that need to be put on the table to inform a discussion about a correct use of this family of applications.

No trustworthiness. As mentioned above, the output generated by these systems is not aiming at being truthful but plausible. This is a risky property as uninformed members of the public may take decisions based on erroneous information produced by the systems. Appropriate safeguards must be introduced to protect the citizens. The connection with the sources of information that current search engines have is missing in LLMs.

Black box behaviour. These models are the result of a tedious process of optimisation that computes billions of parameters that are then used to generate the results. The capacity of these systems to explain why they generate something is extraordinary limited as the output is the result of millions of arithmetic operations over those parameters.

Rigidity. Systems are trained with a set of data to fix the parameters and then no further training is done and no further changes are introduced. If we limit the training set to documents prior a given date, no information in documents generated after that date will influence the generation of output.

High energy consumption. The cost of training the systems account for thousands of megawatts. In addition, as systems need to be retrained frequently due to the lack of adaptability, the consumption has to be repeated every time.

Proprietary systems. Most of these applications belong to a handful of companies that can afford their high development cost. This legitimate business has to be monitored to protect citizens from potential biases and misuses. For instance, there is opacity on the data used to train the systems and thus it is unclear if these systems respect regulations like GDPR.

Copyright. As texts are generated from the documents used to be trained and no reasonable trace can be made to them, the notion of copyright seems to be at stake here. Some discussion has to be had on how to preserve the rights of human creators.

In summary, although the new developments based on deep learning methods are producing significant advances in several areas of science, their use in the context of LLMs needs to be closely scrutinised to avoid misinformation and cyber-security threats.

How does TA addresses generative AI?

We asked ChatGPT the question "What is technology assessment?", and this is the beginning and the end of its answer: "Technology assessment, often referred to as 'technology assessment (TA)' or 'techno-

logical assessment,' is a systematic and interdisciplinary process used to evaluate and understand the potential impacts, consequences, and implications of new technologies. It involves the careful examination of technological developments, innovations, or proposed technological solutions to assess their social, economic, environmental, ethical, and political aspects. The primary goal of technology assessment is to inform decision-makers, policymakers, and the public about the risks and benefits associated with the adoption and implementation of specific technologies (…) It plays a crucial role in shaping public discourse, policy development, and regulatory decisions related to new and emerging technologies".

As it is a general definition that also contains enough details (please note the open enumeration of the scope, the object, the public and the impact of TA), it is difficult to criticise.

In other EPTA reports, we have stated that technology assessment (TA) explores the relationship between technology and society, and we have indicated that TA searches how current technological developments affect the world we live in and aim to contribute to the formation of public and political opinion. Considering those, it is the duty of the members of the EPTA network to inform and advise parliaments on the interrelationships between STI (science, technology and innovation) and politics and society.

In the 2018 EPTA Report it is said that "artificial intelligence (AI) is emerging as one of the most hotly debated technologies on the horizon". This year's EPTA report builds upon a foundation of previous assessments, delving deeper into the ever-evolving societal implications of this transformative technology.

Drawing on the insights and lessons learned from the EPTA member's reports, this executive summary offers a nuanced examination of AI's multifaceted impacts. It navigates the intricate landscape of AI adoption, taking into account its influences on economic structures, labour markets, healthcare, ethics, democracy and governance. The EPTA members' reports underscore AI's capacity to exacerbate inequalities and ethical dilemmas.

The Austrian report points out the need of TA to react quickly to new AI products, such as AI chatbots, and it states that "TA has to be even faster than before".

In the same vein, the Swiss report writes, "Given the wide gap between the unusual speed of these technologies and the time needed to adopt potential measures against their risks, it seems crucial for TA to be able to react responsibly and proactively". The report also says, "TA has a critical contribution to make for the democratic legitimation of the use and regulation of generative AI: by furnishing independent information that duly takes into account the many facets of its impact on society, TA can help decision-makers (both citizens and political actors) form an opinion and take considered decisions on these technological developments".

As DBT (Denmark) writes, "There is a role for TA in equipping MPs to better understand the technology" and remembers the foresight capacities of TA in order "to make educated predictions about how generative AI can affect our societies".

The German report written by TAB emphasises the fact that "the established procedures of technology assessment need to be put to the test". The German report also points out the potential of generative AI in the field of TA and the political decision making process.

The Portuguese report includes a set of questions for TA, such as, "Should AI developers release or restrict their models?" or "To what extent can platforms and AI developers form meaningful partnerships that can aid in the detection and removal of inauthentic content? ".

The Rathenau Instituut (Netherlands) report serves as a call to action to TA with the following words: "TA can broaden the debate on generative AI and work". The report explains, "TA shows that the potential impact of generative AI on the labour market is not unambiguous, but varied and complex, with both positive and negative possible outcomes".

The Greek report, focused also in the labour implications, draw attention to the fact that "the consequences of new technologies are not determined in advance. They must be examined and controlled".

The Japanese report shows the growing concerns and fears regarding the unchecked proliferation of AI technologies and calls for "new measures and discussions" which "are needed to balance the risks while taking advantage of the potential of generative AI". The Japanese report also claims for actions "to educate people on how to use generative AI and to develop human resources to create them".

The Swedish report expounds, "The need for relevant knowledge of AI must be met through education and training, continuing education and research".

The NBT (Norway) report explains its experience in stakeholder involvement in the field of Generative AI in a "continuous effort to involve citizens, stakeholders and those affected to help formulate future policy advice for parliamentarians in the next three-year period". TA often involves engaging a diverse group of stakeholders, including experts, policymakers, industry representatives, affected communities, and the general public. This ensures that a wide range of perspectives is considered in the assessment process.

The STOA report (European Parliament) centres at showing that TA experts have posed the emphasis at the scope of the legal regulation of the AI. What is the risk of general-purpose AI systems? What kind of rules are to be set up for AI? The report emphasises the need for responsible AI development and thoughtful regulatory frameworks to ensure the equitable distribution of benefits and mitigate unintended consequences.

The UK report explains the UK Government and UK Parliament debate on the way to regulate AI, with positions defending AI regulation through a code of conduct rather than relying on legal norms as a pragmatic and flexible approach to address the rapidly evolving landscape of AI. But also with positions defending that "workers to have the right to consultation and notification where the application of technology in the workplace will result in the surveillance of a worker, or result in a significant change to their work".

The Lithuanian report in the same path states, "It is very important to consider various mechanisms of regulation and governance in the field of generative AI" and "to prepare society for the use of artificial intelligence".

The Catalan report (CAPCIT) follows the same line saying "that regulation should be implemented to avoid the potential dangers of generative AI systems," because, if we look at generative AI technology such as Large Language Models, we find, as it is stressed in this report, that "they are disembodied systems that are not prepared to solve the challenges of robotic systems. In terms of intellectual capacities, these systems show very shallow reasoning capabilities; they are unaware of their world, lack common sense, and have not experienced anything."

What do parliamentarians think about generative AI?

Serving the legislative through TA demands a good knowledge of the context in which the legislator is involved. It is as important to know how to deliver scientific and technological information to parliamentarians, as to know what the knowledge and the legal and political context is in which the legislator finds itself, for all TA must be contextualised to legislative needs if it wants to be useful and contribute to legal debates and legal drafting of parliamentarian activities.

With AI the need to know where do parliamentarians stand is greater, for is a technology that evolves at such rapid speed that society, hence parliaments too, are quickly behind the knowledge of the technology, its uses and its challenges. The rapid evolution of this technology makes it difficult to have a deep knowledge of the first up to date and, thus, a deep knowledge on which policies and regulations should be on the agenda.

What do parliamentarians know about generative AI? In order to have an answer to that we have given the chance to all EPTA parliaments to answer six questions, each being related to the areas that are discussed in this report and in this year's EPTA conference.[1]

Taking into consideration the answers given to the question on the concept of generative AI, it is clearly seen that most members of parliament associate it with the generation of content (text, audio, and video)

---

1. The questions asked to EPTA parliamentarians were:
   i. What is the first thing that comes into your mind, when you hear "generative artificial intelligence"? If you are not sure, what does artificial intelligence suggest to you?
   ii. Do you think that there is a relationship between generative AI and democracy? Could you briefly explain your thought on that?
   iii. If you were to think about policy areas where generative AI will most likely have to be dealt with, which would they be?
   iv. Can you imagine how generative AI can affect the evolution/structure of the job market/employment sector/work?
   v. Are you aware if the Parliament has discussed issues regarding generative AI, or are there plans to do so?
   vi. What kind of knowledge do you need to make decisions about generative AI? And how can your TA institution support you?

by a machine without human supervision. Most members of parliament clearly stressed that AI, and generative AI in particular, is a tool and, like any tool, it can bring opportunities but also risks. This very balanced answer in the first question, though, tends to lose its balance as soon as we have asked them about the relationship between generative AI and democracy, as well as with key policy areas such as health, education and the job market. Here the answers given show that there are more concerns about the risks of the use of generative AI than ideas regarding the benefits that it can bring to societies.

When it comes to generative AI and democracy, most of the answers connect both concepts and they generally do so stressing the perils they foresee. A much commented one is disinformation, together with the fact that generative AI has a great capacity to enhance societal biases (often the need of a "quality control" of the data that feeds generative AI is mentioned in many of the answers provided). Moreover, some parliaments seem already active and so shows the answers of their MPs in the field of LLMs. Here parliamentarians have pointed out the need to evaluate whether there must be a LLM in their own language, in order to feed generative AI with the societal values and data of that given country (e.g. Norway). Quite a number of answers related LLMs with the lack of control of the quality and "mind-set" of data, which can conduce to foster values that clash with those from democracies. Nonetheless, many have stressed that they expect a positive impact too and even to be able to use it in their parliamentary work.

Regarding generative AI and health, parliamentarians have pointed at the relationship between generative AI (and generally, AI) with better diagnoses and treatment. Nonetheless, what has been more recurrent in the deputies' answers has been the fear that personal data might be compromised as well as the fact that it is necessary that final decisions are always in human hands. Some have also stressed the fact that the biases embedded in the healthcare data may contribute to enhance those biases in the healthcare system and health research.

Concerning generative AI and education, it is probably one of the areas where most parliamentarians have shown a clear understanding of the actual challenge that the educational system is already going through due to generative AI. That is, while recognising that it is a tool that may help in educational tasks, it has been stressed the need to train teachers and to design an educational program that takes into consideration this technology and, thus, that is able to teach and evaluate without generative AI stopping those procedure from being achieved.

Finally, if we take into consideration the answers given to the relationship of generative AI and the job market, here parliamentarians have tended to be more balanced between the advantages and the risks of generative AI. While stressing that there will be tasks, rather than entire jobs, substituted by this technology, they stress too that new jobs that will be created and point out that what is crucial in this area is the capacity of a society to provide the skills for those new jobs and to take a special care with the fact that no one is left behind by not having "technology-literacy".

Generative AI and democracy

Generative AI tools can reduce barriers to democratic participation, make public services more efficient and inclusive, and give people better access to knowledge and enhance skills. New text generation tools, and universally designed tools for search, translation and simplification of complex texts, could enhance citizens' access to information, and empower more people to express their views on societal issues. Danish MPs, for example, underscore that AI creates new possibilities for democratic participation through, for instance, public polls for hearing citizens' opinions.

Disinformation, bias and trust in democratic institutions

Generative AI might dramatically increase the pace and volume of mis- and disinformation, deepfakes and fake news online. The digital public sphere could be flooded with automated, targeted, and manipulative content, making it increasingly difficult for people to establish whether information is true or false, human or AI generated.

The new tools can be used to confuse the public, aggravate social tension and increase political polarisation, and thus undermine the integrity of democratic processes and trust in institutions. A study on deepfakes among Swiss MPs finds that many representatives view deepfakes and AI content as a high and concrete risk for Swiss democracy, with a political party using AI to generate a campaign poster in the current federal election as a case in point.

Furthermore, research shows that the predominant large language models have biases in both their data and training, and thus can perpetuate different social stereotypes. The models can also hallucinate and generate false content that appears to be true. As a result, the outputs of the AI systems can inadvertently or intentionally influence and manipulate opinions, posing a threat to individual autonomy, as well as democratic processes.

Bias can also lead to wrongful decisions and even discrimination, if used, for instance, by public sector institutions that are unaware or unable to discover the mistakes produced by the systems. This can undermine the public's trust in democratic institutions.

Democratic control of AI

An ongoing policy debate within many European countries concerns how to strengthen democratic governance of AI, without hampering innovation. Large technology corporations are today behind the major breakthroughs within AI. These companies possess insight, resources, and expertise in AI systems with the potential to impact and transform societies. Yet the models made available to the public are opaque. To strengthen public oversight and control mechanisms, it will be crucial for democracies to establish a robust regulatory framework, governing both the development and deployment of AI.

The EU is currently negotiating the AI Act – a risk-based legislative framework setting standards for transparency, reliability and accountability. AI systems that can significantly impact democracy, rule of law, or individual rights and freedoms will either be prohibited or considered high risk. High-risk systems will be subjected to strict obligations, such as adequate risk-assessments and mitigation systems, high-quality datasets, detailed documentation and logging, and appropriate transparency and human oversight mechanisms.

Regulatory discussions are also taking place within the OECD and the Council of Europe, and, on a national level, regulatory supplements to the international frameworks are currently being considered. Ethical guidelines on responsible use of AI, in addition to legal frameworks, can also help address common challenges such as misinformation, privacy and data protection, bias, intellectual property and language inequality. For example, the Danish National Center for Ethics is developing ethical recommendations on the use of generative AI, and the Norwegian Digitalisation Agency has developed guidelines for responsible use of AI in the public sector.

Building AI infrastructure in Europe?

ChatGPT has been characterised as multilingual, but monocultural.1 In general, the public and authorities have limited insight into how large language models function. When generative AI tools are embedded in everyday digital services, gaining insight into the functions of the models are crucial for safeguarding fundamental principles such as privacy, transparency and reliability. An accessible and safe AI infrastructure could also better support underrepresented languages, represent cultural values, and enable public administration to realise the potential of generative AI.

In countries like Norway, Sweden and the UK, options for building national large language models are currently being explored. Developing such models could be a way to strengthen democratic oversight over the design and deployment of generative AI and reduce dependence on external players. However, many questions remain unsolved, for instance who should build, operate and access such a model. Additionally, building a model from scratch is resource- and cost-intensive. Several contributions to this EPTA report highlight that access to data and sufficient computing power is also essential for countries and regions to benefit from the adoption of generative AI.

Generative AI and health

The health care sectors in many European countries are under pressure due to an aging population, a widespread lack of skilled personnel and the increased costs of new and more specialised treatments.

Thus, for health care, the surge in generative AI technologies could not have come at a much better time, and politicians and technology companies are hailing AI solutions as a key ingredient to solving these challenges. In the G7 Hiroshima Process on Generative AI, improving healthcare is mentioned by six of the seven respondents as one of the key opportunities of generative AI.[2]

As pointed out in a study by the Norwegian Board of Technology, AI holds several different promises for healthcare.[3] Digital health assistants can help health personnel in making a diagnosis, selecting treatment, monitoring the patient and warning of complications. By using machine learning, they can analyse medical literature, interpret images and other data, and plough through thousands of patient records. This may become the key to treatment that is better adapted to each individual, and less dependent on the doctor's experience.

More specifically, so-called "generalist medical AI"[4] can automatically draft radiology reports that take into account the patient's history; assist surgical teams with spoken queries and annotated visualisations; provide bedside clinical decision support with detailed explanations and recommendations for future care; and draft patient notes and discharge reports.

Moreover, computer systems with AI can speak the patient's language and respond quickly and accurately when people make contact. Such "digital first lines" can provide better healthcare across the country or region and enable health personnel to spend less time on the telephone and more time on treatment. Watches and wristbands with sensors can register everything from heart rate to tone of voice. AI interprets the data and provides continual information about the users' physical and mental health. This makes it easier to monitor patients at home, detect illness earlier and start treatment more quickly.

In addition to increasing efficiency and freeing up hands, generative AI has already demonstrated its applicability in health research, both for improving our understanding of our brains and bodies, but also for developing new treatments, diagnostic and predictive methodologies and tools. The Alphabet subsidiary DeepMind announced in 2022 that its AlphaFold tool had mapped all known proteins, a game changer in understanding our biology, the potential uses of which are immense, in terms of diagnostics and treatments. Drug discovery can be accelerated with AI by generating new drug candidates and predicting which protein design will be most effective.

Society faces some significant legal, ethical and technical challenges for AI to be a success in healthcare. In order to succeed with generative AI in health care, health data must be made more easily accessible in a manner that is compatible with good data protection. That will be a challenge when clinical measurements are combined with patient records and genomic data, as well as behavioural data from sensors. Adding to this, large models have a larger risk of exposing sensitive patient data from the training sets.[5]

Historical health data contain biases, such as an overrepresentation of Western men. Machine learning is generally based on historical health data, and this can increase with model scale. When AI health care solutions are introduced, it must be ensured through validation and continuous auditing that they do not contribute to discrimination, for example, based on gender or ethnicity.

In relation to this, there is a need for discussing how well a doctor should understand the automated process. This is both a question on what a doctor needs to know in order to act responsibly and meaningfully on the outputs of the process, but also a question of ensuring transparency and trust with a patient through

---

2. https://read.oecd-ilibrary.org/science-and-technology/g7-hiroshima-process-on-generative-artificial-intelligence-ai_bf3c0c60-en?utm_source=pocket_saves#page13

3. Teknologirådet 2022: *Artificial intelligence in the clinic – Six trends for the health service of the future* https://media.wpd.digital/teknologiradet/uploads/2023/01/Artificial-Intelligence-in-the-Clinic.pdf

4. Moor, M., Banerjee, O., Abad, Z.S.H. et al. Foundation models for generalist medical artificial intelligence. *Nature* 616, 259–265 (2023). https://doi.org/10.1038/s41586-023-05881-4

5. Moor, M., Banerjee, O., Abad, Z. S. H., et al. Foundation models for generalist medical artificial intelligence. *Nature* 616, 259–265 (2023).

being able to understand and explain the background for the output of an automated or partially automated process.

AI systems cannot always explain why they have arrived at a result in a way that humans can understand, such as someone being at high risk of an illness. Therefore, it is critical to develop new methods and standards for how AI models explain their advice or insights. Moreover, the better the digital assistants, the more difficult it will be to know when the system should be overridden by humans. Routines for maintaining the competence of doctors and nurses are therefore necessary. It is important that staff is adequately equipped to utilise the systems correctly and with a clear grasp of what the system is doing. An added element to this is ensuring that implementation of AI does not lead to so much increase in speed of processes that doctors no longer have an actual opportunity to make a human judgement.

New diagnostic and treatment capacities have meant that the budgetary pressure on hospitals and care has increased significantly. The increase in diagnostic and predictive capacity that generative AI is expected to leverage means that more diseases and ailments will be diagnosed, which will increase this pressure.

At the same time, precision and niche medicine is proportionally more expensive than more general medicine, both because the production scale is smaller and because it is often more expensive to produce. Thus, adds onus to the already existing need to prioritise diagnoses and treatments.

Therefore, in conclusion, some fundamental questions remain to be clarified when it comes to utilising generative AI in health care. Some of these are not exclusive to this sector, but in health care, the stakes are often life-or-death, and this sharpens the edge on which responsible and successful implementation balances.

Generative AI and education

The impact of generative AI on education is hard to overstate. The debate about the use of AI in education has been going on for more than 40 years. With the advent of ChatGPT (and other powerful text generators), the various visions of future learning technologies – from learners having their own personal learning companion to AI helping to bypass formal assessments – became realistic and readily available options for the first time. However, despite their ability to convincingly simulate the human use of natural language, generative AI systems still exhibit very shallow reasoning capabilities and lack any experience of the "real" world. Their use in educational contexts can have profound direct and indirect effects and requires a great deal of scrutiny. The proposed EU AI Act, for instance, classifies the use of AI in education as high risk. The question of how the educational system should deal with this technological development therefore takes on a new urgency.

In the debate, and the MPs' responses to the EPTA questions for this year's report, two main themes emerge: the role of education in preparing future generations for a world where humans and intelligent machines interact ("learning about generative AI"), and the role of AI-based systems in improving education ("learning with AI").

The need for learning (and educating) *about AI* is recognised by most MPs who responded to the question on AI and education. This applies to learners at schools, but also to the further training of professionals and the training of teachers as well, whose "usage skills need to be significantly increased", as a German MP puts it. There are also considerations about the age at which such education should start and about the competencies that should be taught – competencies for creating AI, but also for dealing with AI, also in a critical way. "It must be ensured that our future generations understand the risks of AI from a young age, but at the same time can harvest the benefits", says an Austrian MP. An MP from Germany stresses that students still need to learn to solve problems on their own. In addition, it is pointed out that schools also have a role to play in "democratically spreading out" the competencies to many people, as a Danish MP puts it.

The country contributions to this report by individual EPTA members detail a number of political initiatives to develop educational programmes on AI. For example in Sweden, universities are expected to offer programmes and courses related to the new needs of the labour market. In the UK, the House of Lords

Science and Technology Committee has identified several actions that the Government should take to address the existing STEM skills gap, including the provision of courses below degree level and improvements to apprenticeships programmes. The contribution on Denmark points to the need to engage citizens and stakeholders in developing principles for using and learning about generative AI in the education system. Lithuania is taking a highly inclusive approach, involving not only the expert community and business, but also philosophers in the debate on the coexistence of humans and AI. Last but not least, a majority of two thirds of European citizens call for more education and training to develop their digital skills, according to recent survey results cited by the European Parliament (STOA).

When it comes to learning *with AI*, most MPs see opportunities, with only a few considering an outright ban. Most governments seem to agree with the positive view; school authorities in Germany, for example, responded to ChatGPT by allowing its use as an aid in exams, and many educators are open to using AI. Generative AI has the potential to enhance learning, enable greater participation, and reduce inequalities by adapting the learning process to the individual needs, including for learners with special needs (e.g., dyslexia). "There are tremendous possibilities with this as a resource", says a Danish MP. However, in order to realise these opportunities, MPs say that certain regulations need to be in place. The personal data of students and teachers must be protected, as must the special rights of minors. AI systems need quality checks and certifications "to support school boards, principals, and teachers in deciding whether or not to use a product" (German MP). Educators need to be aware that AI-powered tools can "become highly addictive" (comment by Norwegian MPs), and public institutions should avoid becoming dependent on particular vendors. Instead, an Austrian MP suggests promoting open source products. A legal framework is called for, so as not to "leave it to the market", as a Dutch MP puts it. In addition, generative AI, with its vast amount of training data from often undisclosed sources, raises issues of copyright and fair compensation for creators, as highlighted in Japan's contribution to this report.

In addition to learning about AI and learning with AI, a third issue can be identified, namely that generative AI may change education in very fundamental ways ("*AI transforming education*"). One issue is the role of exams and formal educational qualifications: "Many students have started using ChatGPT to complete assignments, possibly cheating in assessment situations", as the Norwegian MPs became aware. Indeed, a German student was found cheating in his final *Abitur* exam in Germany with the help of ChatGPT. As a consequence, "testing of what has been learned must probably be held differently in the future", according to an Austrian MP. "After all, we want to know the skills of the learners, not the AI" (German MP). Another issue is the revision of curricula and teaching methods across all subjects, not just in media education. "AI will also have an impact on the content and competences that we have to teach students" (German MP). "It might not be the best use of time to have students spend a lot of time editing grammar", says a Danish MP. In Catalonia, the Department of Education has already started studying the introduction of AI into primary and high school curricula. In Japan, the Ministry of Education, Culture, Sports, Science and Technology released tentative guidelines regarding the use of generative AI in (primary and secondary) education in July 2023 and urged universities to develop their own guidelines.

While the disruptive character of generative AI on education is widely recognised, it remains to be seen what role it will play vis-à-vis and/or in conjunction with other, more traditional types of AI-systems.

Generative AI and work

Along with imageries of Terminator and SkyNet, the prospect of AI induced mass unemployment is probably one of the most enduring representations of AI in the public imagination, and one of the potential impacts that has most often made the rounds in newsrooms. For that reason, it is unsurprising to see that the impact of generative AI on work more broadly is a topic that features prominently in conversations with politicians presented in the country reports below.

The impact of generative AI on employment is far from certain and is more complex than the binary scenarios of mass unemployment or the creation of new industries. There are currently some professions increasingly experiencing a shortage of qualified workers, and more sectors which are expected to experi-

ence a shortage in the not-too-distant future. Here the hopes and predictions are that generative AI, coupled with for example robotics technology, can alleviate these shortages. When talking about AI's benefits, its potential in radiology is often mentioned. Here is a potential case of generative AI alleviating a workforce shortage, such as in the UK, where the Royal College of Radiologists has estimated that three in four hospitals are lacking sufficient radiologists.

The flip side of this is that the fear of mass unemployment is not wholly unwarranted. There are several sectors in which automation of jobs with generative AI is a very likely scenario, and where some jobs that exist today will disappear. What remains unclear is whether new jobs within these same companies and sectors will arise which those displaced by AI can carry out, whether they will have the competencies and prerequisites to receive upskilling and move into other sectors, or if these people will be displaced permanently from the workforce. For some, the first two possibilities will not be the case.

An additional variable in the equation on AI's impact on employment is the likelihood that a by-product of generative AI will be the emergence of new products and services, the development and utilisation of which will foster new types of jobs in existing companies and new industries.

The impact of generative AI on work spreads far beyond the question of (un)employment. Implementation of generative AI in workplaces will mean a change in the nature of tasks. Beyond full automation, as described above, at least two scenarios can be imagined. In one scenario, the task itself will be automated, and the new job will be to administer the automated process, e.g., from doing actual welding to the digital administration of the welding machine. In another scenario, the generative AI system will assist in carrying out the job, e.g. AI decision support for radiology, where AI can point out areas of interest or concern for the radiologist. In both cases, there will be a fundamental need for upskilling and further training of workers, in order to carry out these jobs responsibly, efficiently, and with high quality.

Generative AI is different from previous automation technologies in that it is likely to impact blue-collar and knowledge workers alike, as is seen in the example of radiologists above. The following is just a few instances of how generative AI already impacts different types of work. Within journalism there are already widespread reports of media organisations utilising generative AI to produce news items. Part of the Hollywood screenwriters' strike centres on the concern that generative AI can be used to do much of the work in developing screenplays. There are vast obvious potentials within accounting, and AI has already for quite a while been utilised within finance. Generative AI can do much legislative research work for lawyers faster and cheaper. Graphic design and illustrations can increasingly be done automatically.

In general, the hope is that generative AI and the systems utilising it will be able to take care of jobs that are dull, dirty and dangerous, freeing up workers to do more meaningful and rewarding jobs, and in many sectors also enable workers to focus on delivery of their core task, rather than peripheral tasks such as documentation or journalisation. Nonetheless, there are also concerns. One of these is that the integration of generative AI will lead to an acceleration of work processes. A by-product of this is that each worker, particularly in white-collar jobs, will be expected to accomplish significantly more tasks and to complete them at a faster pace, which can be problematic in different ways depending on the sector. The quality of work output risks suffering simply because there would not be sufficient time to complete them. It also counts the added risk of implementation in a way that puts so much emphasis on speed that meaningful human oversight of processes will not be possible. There is also a risk to mental well-being associated with this future scenario. The WHO has already dubbed stress the health epidemic of the 21st century, and further accelerated tasks could risk adding to this trend. There is thus a need for a fundamental discussion on how we want to allocate the resources that can be set free by implementing generative AI. This connects to a wider societal discussion of work-life balance, a 4-day workweek, and of economic equity in society, in light of the increasing concentration of wealth and the widening pay gap between regular employees and top-level management.

Related to this disparity is the use of AI as a management tool which has already been observed. The use of AI for management is not in itself problematic or concerning. Generative AI could help managers get a better overview of workers' current and future tasks, help provide better feedback and better, more tailored task allocation. If used correctly, it could be a management assistant, which helps managers improve working conditions and well-being among subordinates. However, the gig economy has already demon-

strated how generative AI can also be used to degrade working conditions. Here, AI has been used to design schedules that are nearly impossible to live up to, create opaque rating mechanisms that punish delivery delays regardless of their cause, and incentive structures designed by algorithms to ensure that bonuses are only achievable for a few employees. In addition to these dehumanising working conditions, data driven management can also lead to questionable practices of employee surveillance.

The final point of importance concerns the working conditions behind the development of generative AI. Producing generative AI rests on large amounts of annotated data. Where this data does not exist in already well-documented databases, it requires substantial human labour to do this annotation. The largest AI companies have turned to outsourcing this task to companies in the global south who exploit workers, provide next to no fair standards of work, and utilise an algorithmic management approach similar to the platforms described above.[6] There is thus a task of making AI companies comply with fair working conditions for the people annotating and training the models.

Both the above discussion and the following country reports underline that generative AI is still in a phase of uncertainty as to the impacts it will have on work. The positive potentials are immense, but harnessing these while curbing the potential negative effects requires careful political consideration, especially on how to offset negative employment impacts while stimulating innovation and emergence of new services, and how to allocate the dividends of increased efficiency.

Outlook

Generative AI has the potential to impact many areas of our society and economy. Technological development is taking place at an ever-faster pace. Whether and how generative AI can be used responsibly and sustainably remains an open question. It is an ongoing challenge for parliamentary TA to keep up with the pace in order to provide MPs with timely and relevant guidance and decision-making knowledge and to contribute to a fruitful societal debate on the topic.

---

6.   https://fair.work/en/fw/publications/fairwork-cloudwork-ratings-2023-work-in-the-planetary-labour-market/

Austria

Author: Walter Peissl

Stimulated by the next generation of generative AI systems, the development of AI and its perception in the public has entered a new stage, also in Austria. AI is already being used in various ways, for example, to detect structures in large amounts of data (pattern recognition), to draw conclusions from them and thus to prepare decisions. With generative AI systems, AI is entering new waters. This new generation of chatbots fosters human-machine interactions based on natural language, easing the processing of requests (also called prompts). This is made possible by a so-called "large language model" (LLM). In response to concrete requests, AI facilitates generating new content, however based on processing existing data. Be it a text, image, video, or sound. As all kinds of content become easier to generate, this results in additional challenges for society, especially for democracy, i.e. with more content also comes more disinformation. ChatGPT, an AI chatbot from the company OpenAI, is the best-known product of this type of AI. Other companies have also entered this market already.[7]

AI in Austria

As in many other European countries, the launch of ChatGPT has attracted a great deal of attention in the media, among the general population, and in politics in Austria. The generative AI systems represent a new development with potentially disruptive societal effects. However, this tremendous public attention and high expectations also obscure the view of realistic assessments of the possibilities and limits of these systems. In Austria, the high level of attention has also led to a public discussion about funding AI research via the media, highlighting possible deficits.[8]

The leading players in the Austrian AI landscape are the Federal Ministry of Finance (BMF) with its State Secretary for Digitalisation, the Federal Ministry of Labour and Economy (BMAW) and the Federal Ministry of Social Affairs, Health, Care and Consumer Protection (BMSGPK). All political parties in Parliament are actively contributing to the debate. AI-Austria[9] is a think-tank and lobby organisation for AI application in Austria, bringing together players from science, business, education and society. The main players from academia are several Universities and other research institutions like the Institute of Science and Technology Austria (ISTA), the Austrian Institute of Technology (AIT), and the Austrian Academy of Sciences (OeAW). And a new specialised university is being founded in Linz, the Institute of Digital Sciences Austria (IDSA).[10]

Austria's R&D sector is not at the leading edge of AI worldwide but has some potential. For instance, besides some university research groups, there are some AI start-ups; however, monitors report that 'Artificial intelligence remains in first place among the innovation and technology trends' in the Austrian Start-up landscape.[11] A recent study detects considerable challenges.[12] Additionally, a fundamental study on the potentials of AI applications in the media sector may be highlighted, namely AI.AT.Media, which was conducted

---

7. https://www.pcguide.com/apps/chat-gpt-alternative/
8. https://science.orf.at/stories/3218860/
9. https://aiaustria.com/
10. https://www.bmbwf.gv.at/Themen/HS-Uni/Aktuelles/idsa.html
11. https://austrianstartupmonitor.at/wp-content/uploads/2023/04/ASM-2022.pdf
12. https://www.ipa.fraunhofer.de/de/Publikationen/studien/ki-austria.html

by the Austrian Press Agency Media Lab in 2021 (Krawarik et al. 2021). At the beginning of 2023, APA presented its future AI strategy "APA Trusted AI for journalism and communication"(APA 2023).

As in many countries, in Austria, AI is seen as a general-purpose technology with corresponding importance for the country's business location and competitiveness (Huang/Peissl 2023). The basis of the Austrian AI policy is the National AI Strategy from 2021 "Artificial Intelligence Mission Austria 2030 (AIM AT 2030)". It focuses on pursuing the following three objectives:

1. A broad deployment of AI oriented towards the common good is targeted, carried out in a responsible manner on the basis of fundamental and human rights, European fundamental values, and the upcoming European legal framework.
2. Austria should position itself as a research and innovation location for AI in key areas and fields of strength, and
3. Through the development and use of AI, the competitiveness of the Austrian technology and business location should be secured (BMK 2021).

The Artificial Intelligence Mission Austria aims to examine the safe use of AI applications in public administrative processes. To promote a modern and efficient public sector, data-based machine learning and AI shall be used to (partially) automate work processes and decision-making in the administration (BMK 2021).

AI in the debates of the Austrian Parliament

The topic of AI has been raised several times in the Austrian Parliament, including questions on AI and child pornography, the use of AI in educational contexts, AI and inner security, impacts on marketing and consumers' behaviour, and many more issues.[13] Most recently, an open forum on "Impact of Artificial Intelligence on Society and Democracy" was held with a specific focus on ChatGPT. By incident, the day after the event, the Parliamentary Committee for Research, Innovation and Digitalisation also held an expert hearing on the issue.

The EPTA questionnaire was sent out to the members of the Parliamentary Advisory Board for Foresight & TA. The sending out overlapped with the beginning of the Parliament's summer recess, so that out of the potentially five, only two answers arrived in time at the ITA, stemming from the Greens and the Social Democrats. The two MPs agree in predicting a potential for AI to impair democracy; one also sees a potential for promoting democracy. In particular, they point to the issue of fake news and deepfakes and their danger to democracy. Accordingly, political communication, the labour market, education policy, and the military use of AI are considered particularly sensitive policy areas. One MP cites data protection for education and health data, ensuring high-quality systems and training data, avoiding dependency on IT corporations, and promoting open-source solutions. The potential gains in quality and efficiency would have to be distributed in a socially equitable manner, and, especially in the case of applications in the medical field, the ultimate responsibility would have to remain with the treating physician. In the Austrian Parliament, there have already been some discussions on AI (see above). Concerning knowledge necessary, both MPs indicated wanting to be informed about the fundamental technical aspects, the implications and possible application areas. The underlying values in the design of such systems were also acknowledged, and TA was given an important role.

Last but not least, the so-called monitoring reports delivered to Parliament semiannually covered many aspects of AI, including generative AI,[14] deepfakes, AI in health, AI risks, and others. These thematic outlines triggered particular interest among the MPs. For instance, the issue of deepfakes was taken up in a parliamentary motion.[15]

---

13. https://www.parlament.gv.at/suche?VTS_01searchType=all&VTS_01searchScope=all&VTS_01category=Parlamentskorrespondenz&VTS_01gp_liste=XXVII&VTS_01search=K%C3%BCnstliche+Intelligenz
14. https://www.parlament.gv.at/dokument/fachinfos/zukunftsthemen/143_generative-ki.PDF
15. https://www.parlament.gv.at/PAKT/VHG/XXVII/E/E_00104/

Societal and political relevance and debate

ChatGPT raised much awareness in the Austrian public. Many news and commentaries in the media paved the way for public and political discussions in Parliament. They led to specific use-case discussions in NGOs, adult education institutions, and other settings. However, in a survey in March 2023,[16] about 18% of respondents said they already used chatGPT, with the figure for those under 28 (Gen Z) being about 34%. Overall, the Austrians surveyed use chatbots to translate and compose foreign-language texts (75%), as a search engine (73%) or to compose invented texts for entertainment purposes (61%). Fundamental scepticism is evident in fears of fraud (78%), fake news (77%), lack of privacy (67%), and environmental impact (61%). 81% call for stricter regulations of ChatGPT & AI to ensure privacy, neutrality, and correctness.[17]

The (previous) federal government established the Austrian Council for Robotics and AI and presented the Whitepaper „ How to Shape Austria's Future with Robotics and AI Positively" in 2018.[18] Moreover, a study on "AI in Austria" has been commissioned.[19] In 2021, the previous federal government finally issued the "Artificial Intelligence Mission Austria 2030" (BMK 2021).

So far, there is no special AI legislation in Austria. Like other EU countries, Austria is involved in developing the AI-Act on the European level. Besides this, there is a proposal from the State Secretary on Digitalisation to set up an Austrian AI Authority with a service character in order to prepare citizens and companies. For example, it would collect all institutions that have obtained AI certification. An Austrian seal of quality for AI is also under consideration. A transparency obligation is also seen as crucial when confronted with AI.[20] From academia, there was a White Paper from 2021 from TU Vienna and JK University Linz in this sense, supposedly launching the first commercial Certificate for AI systems, which was widely discussed across the EU.[21]

The role of TA in the debates

The ITA has been dealing with AI in general and specific issues (on labour, welfare institutions, education, etc.) for years. On ChatGPT specifically, the ITA wrote a short info in May 2023 for the Austrian Parliament as part of the semiannual monitoring activity (see already above). Colleagues from the ITA contributed to the TAB expert hearing to prepare the first parliamentary report on Generative AI for the German Bundestag.

Among the ITA studies on AI in the past five years are:
- Critical AI Literacy (2023–2024)[22]
  (Strauß 2021)
- Automating Welfare (2022–2025)[23]
- Artificial Intelligence: Towards more comprehensibility and transparency (2021–2022)
  (Udrea et al. 2022)[24]
- Technology Assessment in Austrian AI Start-ups (2020–2021)
  (Tanja Sinozic et al. 2022)[25]

---

16. https://science.apa.at/power-search/14071413255035448015

17. https://www.ots.at/presseaussendung/OTS_20230328_OTS0064/pwc-chatgpt-ki-studie-fast-zwei-drittel-der-oesterreicherinnen-sind-fuer-verbot-an-schulen-jeder-dritte-fuerchtet-um-arbeitsplatz

18. https://www.bmk.gv.at/themen/innovation/publikationen/forschungspolitik/ki/whitepaper.html

19. https://www.bmk.gv.at/themen/innovation/publikationen/ikt/ai/zahlen.html

20. https://www.parlament.gv.at/aktuelles/pk/jahr_2023/pk0740

21. https://www.jku.at/fileadmin/gruppen/219/LIT_AI_Lab/News_Seite/White_Paper_-_Trusted_Artificial_Intelligence_-_Towards_Certification_of_Machine_Learning_Applications_web_s.pdf and https://www.elise-ai.eu/events/first-commercial-certificate-for-ai-systems-now-on-the-market

22. https://www.oeaw.ac.at/en/ita/projects/current-projects/cail-critical-ai-literacy

23. https://www.oeaw.ac.at/en/ita/projects/auto-welf

24. https://www.oeaw.ac.at/en/ita/projects/artificial-intelligence

25. https://www.oeaw.ac.at/en/ita/projects/artificial-intelligence-technology-assessment-in-austrian-ai-start-ups

- The AMS [Labour Market Service Autority] algorithm (2019–2020)
  (Allhutter et al. 2020)[26]
- Artificial intelligence and labour (2018–2019)
  (Cas/Krieger-Lamina 2020)[27]

Besides the reports above and numerous invitations to expert panels (e.g. in UNECSO), public lectures, conferences, etc., the ITA was invited to contribute to a session on generative AI of the Parliamentary Commission on Research, Innovation and Digitalisation (the ITA nominated external experts and one ITA staff was giving oral input). Basic functional details of LLMs, the European scene, and ethical reflections on AI, in general, built the basis of the discussions. The topics of transparency, competence building and research were central to many MP's questions. Moreover, under the heading of "AI fitness," the MPs also discussed the acceptance and competence of users with the experts.[28] ITA was also invited by the Austrian Court of Audit to give a keynote on the uses of AI in auditing and auditing public sector AI.

The lessons learned from TA

The well-known phenomenon that logic and timelines of politics and science are different turned out to be true again in this issue. TA always tries to be at hand with information at the right time, but sometimes, TA has to be even faster than before. The public release of ChatGPT and its potentially disruptive power produced a high political demand to know better. So, as long as TA has done some previous work in a similar or more general aspect, it is easier to react quickly to a particular new issue, such as LLMs. Dealing with many other fields of technology may also help to ask the right questions in a specific issue. In short, methodological and procedural experience helps as much as knowledge in a specific area of research.

Literature

Allhutter, D., Mager, A., Cech, F., Fischer, F. and Grill, G., 2020, *Der AMS-Algorithmus – Eine soziotechnische Analyse des Arbeitsmarktchancen-Assistenz-Systems (AMAS)*, November 2020, Wien: ITA.

APA, 2023, *Leitlinie zum Umgang mit Künstlicher Intelligenz*: APA – AUSTRIA PRESSE AGENTUR.

BMK, 2021, *Strategie der Bundesregierung für Künstliche Intelligenz – Artificial Intelligence Mission Austria 2030 (AIM AT 2030)*, https://www.bmk.gv.at/themen/innovation/publikationen/ikt/ai/strategie-bundesregierung.html.

Cas, J. and Krieger-Lamina, J., 2020, KI und Arbeitswelt, in: Christen M.; Mader C.; Cas J.; Abou-Chadi T.; Bernstein A.; BraunBinder N.; Dell'Aglio D.; Fábián L.; George D.; Gohdes, A. (Ed.): *Wenn Algorithmen für uns entscheiden: Chancen und Risiken der künstlichen Intelligenz*, Zürich: vdf, 144-164.

Huang, L. and Peissl, W., 2023, Artificial Intelligence—A New Knowledge and Decision-Making Paradigm?, in: Hennen, L., Ladikas, M., Hahn, J., Peissl, W., Est, R. v. and Lindner, R. (Eds): *Technology Assessment in a Globalized World – Facing the Challenges of Transnational Technology Governance*: Springer, 175-201, https://link.springer.com/content/pdf/10.1007/978-3-031-10617-0.pdf.

Krawarik, V., Schell, K., Ertelthalner, V., Thallinger, G. and Bailer, W. (BMK), 2021, AI.AT.*Media – AI and the Austrian Media Sector: Mapping the Landscape, Setting a Course*.

Strauß, S., 2021, Don't let me be misunderstood - Critical AI literacy for the constructive use of AI technology, *TATuP - Zeitschrift Für Technikfolgenabschätzung In Theorie Und Praxis 30*, 44-49.

Tanja Sinozic, Steffen Bettin and Udrea)., T., 2022, *Cultures of responsibility for AI start-ups*, commissioned by: ITA, Vienna.

Udrea, T., Fuchs, D. and Peissl, W., 2022, *Künstliche Intelligenz – Verstehbarkeit und Transparenz*, Wien: ITA.

---

26. https://www.oeaw.ac.at/en/ita/projects/ams-algorithm
27. https://www.oeaw.ac.at/en/ita/projects/the-social-effects-of-artificial-intelligence
28. https://www.parlament.gv.at/aktuelles/pk/jahr_2023/pk0740

Catalonia

Catalonia - The Advisory Board of the Parliament of Catalonia for Science and Technology (CAPCIT)
Generative AI for Education. The case of Catalonia.

What is it about

The release of ChatGPT has allowed millions to experience AI for the first time. AI is an ancient dream of humanity: to delegate our unpleasant tasks to machines so we can devote our time to enjoying life. ChatGPT is just one of many examples of 'generative AI'. Before the current wave of ChatGPT-like systems and for over forty years, the AI field has tried to emulate intelligence following a *symbolic* approach. Our reasoning and thinking seem to be based on symbols, as words representing concepts and relationships among them. The research program of symbolic AI was to give instructions to computers, written in languages that combined those symbols so that computers could decide on actions. However, symbolic AI has proven to perform very poorly when confronted with tasks for which human perception is needed: seeing, hearing or sensing. We humans do those tasks very well, but they are difficult to express as symbol manipulations.

An alternative way explored in AI for quite some time is to computationally model structures we see in the brain, the so-called neural networks. A neuron looks for patterns in the signals of its incoming connections, and when a pattern is detected, a signal is sent to the outgoing connections. A rather simple individual behaviour that, when scaled to billions of interconnected neurons, produces the complex combined behaviour of a human brain. These neurons are structured in layers of input/output signals. Here is where the concept of 'generative AI' appears. When given certain inputs at the first layer of neurons, these systems propagate the signals to 'generate' a final output signal at the final layer.

Since the 40s, we have produced different computer models of neurons and developed algorithms to determine the parameters that lay at the connections between those neurons. The parameters quantitatively determine the relationships between the incoming and outgoing signals of a single neuron. The access to data and computing power has allowed researchers to develop networks with billions of parameters. ChatGPT has 1.5 billion parameters, and GPT4 has 1.7 trillion parameters. Incredible numbers.

How does it work? In straightforward terms, in the context of text generation, these vast networks learn the likeliest words that continue a particular sequence of words and use algorithmic tricks to select one. The systems iterate the process by adding the just generated word to the previous input sequence in order to generate another word. And so on. As the text over which the networks are trained is human-generated (the web, books, newspapers), the text generated 'seems' human-generated. Similarly, for images, combinations of images and text, or videos.

Generative AI has had a serious social impact on education from the very beginning. As its performance was so similar to human-generated content, educators were scared that they would not be able to distinguish human-created content from content generated by an AI. Some countries opted to ban these systems from the classroom, e.g., some states in the US; systems appeared to try to detect AI-generated content, e.g., DetectGPT for text. Things have changed in the last months, and educators are seeing these systems and AI in general as an opportunity in education.

What is the state of play

In 2019, the Catalan Government defined the Catalan AI strategy, named CATALONIA.AI, around several axes. Among them, we find support for the ecosystem, the impulse to research and innovation, talent attraction, and the promotion of an ethical approach to AI. This strategy identifies four mechanisms: CIDAI to

promote innovation, AIRA to promote research and talent attraction, OEIAC to monitor an ethical approach to AI and DCA-AI to integrate the communities.

Catalonia has a vibrant ecosystem with fifteen research centres working on AI and research groups in all its universities. Since 1994, Catalonia has its own independent AI association, ACIA, which is a member of the European Association of AI, EurAI.

According to ACCI0, Catalonia currently has a strong ICT sector with over 22000 companies representing 12% of the GNP, growing very fast. Although we lack more recent data, already in 2018, the AI sector represented above 1B€ and employed more than 8,000 people. With its research, industrial and computing infrastructure, Catalonia is an attractive hub for AI companies.

Since the first works on deep learning -the technological basis of generative AI- appeared, several Catalan research groups have made significant contributions. The Catalan researcher with the most impact in generative AI is undoubtedly Oriol Vinyals, educated at UPC and currently working for Google. His work has influenced the subsequent research at UPC and many other Catalan research centres. Catalan researchers have been doing fundamental research and applying generative models to create content like art material, music, or images.

A very significant project to produce LLMs for Catalan is the project Aina, coordinated by the Barcelona Super Computing Center (BSC). The Aina project is reasonably funded, and its models will be open to everybody, including companies interested in developing applications.

In terms of education, the education sector, including the Department of Education, are reacting to the AI and generative AI wave positively, seeing opportunities in the application of AI to education. A very positive sign is that last May, the Department of Education has opened a research call for non-university levels for the first time and it includes Artificial Intelligence as one of its research objectives.

Stakeholders in Catalonia

There are numerous actors in the generative AI landscape in Catalonia. First, there are two main Government Departments with responsibilities. The Department of Research and Universities via the General Director of Research and the Department of Business and Work via the Secretary of Digital Policies are responsible for implementing the Catalan strategy on AI.

At the academic level, several university departments are researching AI and applying AI to different areas of knowledge. In the Catalan strategy on AI, as mentioned before, the industrial stakeholders are organised around CIDAI and the academic stakeholders around AIRA. ACIA, the civil association on AI, has gathered researchers in the field since 1994.

Since March 2023, the city of Barcelona has been a node of the European Network ELLIS. 'Ellis Barcelona' is called to be the main interlocutor on generative AI. The node integrates researchers from the main research centres and universities with a strong international recognition in AI, especially in Machine Learning.

Many companies have established research departments on AI in Catalonia or have shown a strong interest in using AI in their business, including Amazon, Huawei, Microsoft, NTT and SAP.

In the area of education, it is worth mentioning that in March 2023, the Department of Education and the Department of Research and Universities signed an agreement to facilitate research activities in compulsory Education. AI can be found among the priorities set in that agreement.

Importance to Catalonia

As already mentioned, the ICT sector is very dynamic and strategic for Catalonia. There is a rich fabric of important software developers and small start-up companies with a strong capacity to generate new and innovative solutions. Barcelona is among Europe's top five favourite cities to establish a company. AI is creating and will continue creating a significant amount of qualified jobs. The wave of generative AI will certainly fuel the creation of new companies and developments. The Catalan Strategy on AI has clearly identified the importance of AI.

The application of AI to education is also very important for Catalonia. UNESCO, as early as 2019, in a document signed by more than 100 countries warned about the responsibility of the governments to introduce AI in the curricula. As the pace of development accelerates, new reflections are needed on how to adapt the contents to generative AI concepts. Catalonia has a long tradition of innovative education methods, and the introduction of AI in both primary and high school curricula has caught the attention of educators and is already under study by some working groups in the Department of Education.

Ongoing debate

A large part of the population perceives Generative AI systems as the onset of AI, fulfilling the promise of AI. This perception is generating some debate within the AI community.

As in previous occasions in AI history, a particular success with a specific technology makes people believe 'this is it,' we have achieved 'true AI.' If we analyse this in a bit of detail, we can see that this is not quite the case (yet). The human intelligence capacities AI wants to automate can be divided into physical and intellectual. A true AI should be doing the same as humans do in both. Let's look into Large Language Models (aka LLMs, e.g. ChatGPT) or image generation systems (like MidJourney). They are disembodied systems that are not prepared to solve the challenges of robotic systems. In terms of intellectual capacities, these systems show very shallow reasoning capabilities; they are unaware of their world, lack common sense, and have not experienced anything. They manipulate natural language in awe-inspiring ways but are not the whole path to true AI; they may be a component but not the entire solution.

Given the expectations that ChatGPT generated in the general public, a letter issued in March 2023 was signed by many relevant researchers and public figures asking for a moratorium on the public release of generative AI systems until proper regulation is in place. This letter generated a wave of interest in the media and reactions by major companies arguing that they would self-regulate better. These reactions have been seen as a manoeuvre to avoid regulation in the USA.

Another significant aspect of the public debate has been the positioning of the three 'fathers' of Deep Learning that received the Turing Award (a kind of Nobel prize in computing), Bengio, Hinton and Lecun. Deep Learning is the technology behind all generative AI systems. Bengio signed the letter mentioned above, Hinton resigned from Google and then warned about the dangers of misinformation that LLM could bring about. LeCun is proposing to go back to older proposals in AI based on cognitive architectures to overcome neural network limitations.

The debate in Catalonia is held similarly, and researchers believe that regulation should be implemented to avoid the potential dangers of generative AI systems.

AI in Parliament

The Catalan Parliament has not yet legislated in this field directly. This being said, along the past five years Ai has increasingly appeared in parliamentary debates, in particular in connection to the so-called "motions" and "resolutions" that are used to foster a particular action from Government. Since 2019, the Parliament asked the Catalan Government to establish a strategy on AI following the coordination plan for AI of the European Commission (Moció 66/XII).[29] Later, during the pandemic, the Parliament asked the Government to use AI technology to conduct epidemiologic surveillance (Resolució 940/XII).[30] In different terms, AI was also envisaged as an opportunity to promote the use of the Catalan Language and to ensure that citizens could interact with chatbots with their mother tongue (Resolució 911/XII).[31] Finally, two more parliamentary initiatives are to be selected taking into account the approach they have to AI. One the one hand, Parliament has stressed the need to work for an ethic use of data, as well as to ensure the respect of human rights within AI systems (Resolució 482/XIV). On the other, Parliament has pointed at the relationship

---

29. Publicació de l'aprovació en el Ple
30. Publicació de l'adopció en el Ple
31. Publicació de l'adopció en el RPle

between AI and the touristic sector (Moció 247/XIV), being this the latest parliamentary initiative, which dates from July 2023.[32]

The role of TA in the Catalan parliament

The current awareness that policymakers show in Catalonia about generative AI shows, on the one hand, that the term is not yet part of the consolidated parliamentary vocabulary and that the term is often confused with other terms like AI, machine learning, or misinformation. There is a shared understanding that the Parliament has to address the issue and that CAPCIT should play a central role in informing the politicians. The negative side of AI is often stressed, especially concerning the dangers to democracy. The opportunities it represents are evident to most policymakers, as well as the need for its regulation. Identified opportunities are numerous and include improving health, education, and public administration. In the concrete case of education, there is agreement that its use is unavoidable, but there is no consensus on how it should be introduced in the education system. The protection of data privacy and some ethical concerns are raised in the application of AI to health. The impact on the labour market is mainly seen as improvement opportunities and increased efficiency. The possibility of addressing challenges like climate change has been signalled as well.

It is thus perceived that AI in general and generative AI in particular, are technologies that need to be explained to the decision-makers as well as to the Catalan society for the impact they already have and expect to have in the near future. CAPCIT is thus called to play a pivotal role providing scientific inputs to guide political decision-making in this field.

---

32. Publicació de l'adopció en el Ple

Denmark

Generative AI and Democracy; Danish MPs and a Country Perspective

This paper will introduce and investigate different perspectives on the topics of generative AI and democracy. Denmark is an interesting case for investigating these topics as the country is praised for its high level of digitalization as well as being ranked high on numerous democracy indexes. With a key enabling technology as generative AI, significant changes will happen and its impact on society – and democracy – is therefore highly relevant. This paper will begin with an overview of themes connecting generative AI and democracy, as they emerged in six interviews with Danish Members of Parliament (MPs). Subsequently, we will unfold the themes further by referencing reactions from other societal stakeholders. We will end this paper by identifying the role that technology assessment (TA) can play in equipping politicians to grapple with generative AI.

Problem(s) at Stake

The digital and democracy position that Denmark is in, set the country in a unique situation whenever new, key enabling technologies emerges. The emergence of generative AI is no exception. It is therefore natural to investigate how politicians imagine the impact of generative AI on democracy in Denmark. With our inherent relationship with digitalization, it is no surprise that Danish MPs in general have a positive and welcoming attitude towards generative AI. While not much relation between generative AI and democracy was directly mentioned, the topics of geopolitics, free speech and public education was mentioned as concerns for how generative AI could have a negative impact on society.

Positive attitudes on generative AI and democracy

There are visions for how generative AI can influence democracy positively. It relates to the accessibility and efficiency where the aspirations are that generative AI can function as an assistant to the individual. Starting with the politicians themselves, they can and are already using generative AI in their everyday work. The positive benefit of generative AI is not only in the MPs own work but also for the public. Generative AI can for example empower people with disabilities. For instance, generative AI empowers people with dyslexia to express themselves in writing with much more confidence. On a broader societal level, the Danish MPs see that generative AI can open new possibilities of hearing citizens' opinions on different topics. The examples mentioned during the interviews are that generative AI can enable a new form of public polls for hearing citizens opinions, and that AI tools will be cheaper and more accessible, which allows for larger part of the population to participate in the democratic conversation. Furthermore, it was mentioned that generative AI tools could help citizens get more knowledge and thereby skills, which can empower the individual citizens and enhance their participation in democracy.

Concern on generative AI's impact on democracy

Two topics of concern were recurrent across the conversations with MPs: geopolitics and free speech. While it was not a direct concern, another persistent theme was a need for changes in the primary education system if society is to keep up with the technological development.

Geopolitics are no longer solely a matter of military power or geographical borders; technologies play a crucial role in the geopolitical debates and tensions as well. The impact of generative AI in the sphere of geopolitics is a cause for concern. Increasingly Danes use social media platforms to find information and news. As it is today, these platforms have been susceptible to the spread of disinformation, but with gener-

ative AI it becomes significantly easier to produce plausible disinformation at a large scale and to distribute it through these platforms. The concern among the interviewed MPs was that this enables trolls sanctioned by foreign regimes to influence the Danish political debate, but also influence the world view and understanding of Danish citizens. With e.g., deepfake technologies, it becomes increasingly difficult to distinguish between fake and legitimate content, and this is a fundamental problem for the national democratic political debate. A flipside of this, is that social media is becoming an increasingly poor indicator of general popular sentiment, and politicians risk being misled if using social media content to navigate.

The spread of disinformation on social media platforms poses a dilemma within on one hand, people having the right to express their believes – even if they are "wrong" – but on the other hand, protecting people from being misled. The debate is how to regulate generative AI to avoid spread of more disinformation, without violating the right to free speech, which in Denmark is a highly regarded value and cornerstone in our democracy.

The last concern is related to primary education and the need for ensuring that technology and understanding of technology becomes an integrated part of the democratic education, that all children in Denmark receive. The concern is that the implementation of this critical skill is slow and lacking behind the actual developments in society. This is not simple related to generative AI but a general concern for the technological development at large.

Societal Reactions on Generative AI and Democracy

It is not only Danish MPs who engage with the emergence of generative AI. Several actors have in different ways contributed to the debate about generative AI and the three topics of geopolitics, free speech, and primary education. In this section, we will highlight some of the actors who have been active in the debate on generative AI. The primary focus will be on their contribution to the three topics already covered above.

The societal reactions on generative AI, are relatively new since the technology only recently have become accessible to the public with ChatGPT. Different actors have recently or are just starting to investigate generative AI and its implication on society.

The topic of Big Techs exploitation of users' data and how to protect users' rights, has been debated in June 2023 by an expert group established by the Danish parliament, who published 13 recommendations for how issues related to Big Techs business model should be regulated and controlled. They want to avoid exploitation of children by Big Tech as well as avoid Big Techs influencing people's behaviour online, which has a democratic perspective. The next issue that the expert group will focus on is how Big Techs use of AI – amongst here generative AI – is affecting society and how it can be managed.

The topic of disinformation has got attention in various contexts. The Danish National Center for Ethics has delivered advice on how to regulate disinformation online and will soon begin a longer investigation on generative AI and its implication on society, with the goal of publishing their ethical recommendations on the use of generative AI, in the Summer 2024. In a study about the digital life of children they touch upon democracy, misinformation, and digital responsibility, concluding that misinformation is present in 8 out of 10 children's life, however, the children are not worried about misinformation. The study is made by the project "Algorithms, Data and Democracy" (The ADD-project), an actor who have acknowledged the need for continuous attention through e.g., publication and events, on how AI impacts society. They work broadly with how our democracy is influenced by technological development. They are planning an event for societal actors to debate who is responsible for limiting the spread of disinformation in our society in a world of generative AI, which will be held in the first quarter of 2024 in collaboration with the Danish Board of Technology.

The Nordic Council of Ministers has partaken in the debate on how generative AI can affect democracy. Because the democracies in the Nordic countries share a key feature of having a high level of trust by the public, the council has expressed great concern for how generative AI can be misused to manipulate or even undermine the democratic conversations in the Nordic countries. This is very much in line with how the concern was expressed by the Danish MPs.

Among other actors, who have entered the debate, and here especially within the theme of generative AI related to primary education, is teachers in Denmark. While their focus is somewhat different from how understanding of technology is lacking on the curriculum, their concern is rather for how to – or not to – integrate generative AI tools in their teaching. The Ministry of Education has declined to make clear guidelines for the teachers to follow but has proclaimed that recommendations for the use of generative AI tools will soon be launched.

Within the topic of education, much debate has been about how to regulate generative AI for use at educational purposes and use at exams. Most of the usage debates has been for primary, secondary, and upper secondary education, while the usage for exams has also been debated for universities. The need for comprehending technology has been an ongoing debate for several years in Denmark, and it has regained new power by the emergence of generative AI. What have had little to no awareness is the need for general upskilling for all people across their position in society.

While it is no longer an active actor, it is worth mentioning the independent TechDK-commission from DJØF, with members from businesses, media, universities, NGO's, and other associations. It was heavily criticizing politicians for their lack of action in the field of Big Tech and emerging technologies. In their work they directly encouraged a collaboration between Big Tech, media, and universities in the aim of combating fake profiles online and the spread of false information. Additionally, they encouraged education for primary and secondary students on identifying online propaganda, and the rules of free speech and censorship.

Actors are active in various degrees on aspects of generative AI and mostly on AI in general, but a lot of discussion is happening in closed, exclusive forums where the public is rarely invited. Beside the population study by the ADD-project and some studies on children's online behaviour not much has been studied in Denmark about citizens perception of generative AI. The Danish Board of Technology has on two occasions during the later years done public engagement on the topic of AI in two distinguished relations. First, on the matter of AI in general and AI for medical and health research during the Human Brain Project in 2019-2020. Later the development of an AI-based assistance chatbot for identifying disinformation in the TITAN-project in 2023. These public engagement exercises have been executed to obtain insights into Danish citizens – and selected European countries' citizens – opinions on development and use of AI in different contexts which do carry broader usability than for their respective research projects. Insights from such endeavours give an indication about the public's sentiment towards an emerging technology such as generative AI.

The Role of TA in the Danish Debate on Generative AI and Democracy

There is a general understanding that Denmark has the potential to become a frontrunner of the digital transformation that generative AI will bring. However, MPs are aware that with the fast-paced technological development, Denmark is already two steps behind on adapting legislations to the new reality that generative AI brings. Although MPs seems to agree on following the upcoming legislation from EU on generative AI (the AI Act), they see that Denmark alone or in collaboration with the Nordic neighbours can take actions on own regulation that EU are not (yet) covering.

It is evident that politicians cannot be experts within all fields that they are involved with, and there is therefore a massive task for outside information being fed into the political system to inform them about the nuances of a topic, and generative AI is no exception. Neither AI nor generative AI specifically, are new concepts, but for the interviewed MPs, it was still an area with which they are more or less unfamiliar. And for this reason, there is a role for TA in equipping MPs to better understand the technology. It was clear that the MPs, while having tested and utilized generative AI services, did not distinguish between generative and 'regular' AI. There is a need to help develop MPs' technical vocabulary and imagination, so they can adequately grasp in what respect generative AI is a gamechanger. What TA is uniquely set to do in this respect, is provide holistic analyses of technology, both in relation to specific applications and more generally for society. It was clear from the MP interviews, that the foresight capacities of TA, to make educated predictions about how generative AI can affect our societies on a broad range of parameters in medium- and long-term,

are necessary in order to provide the MPs with a grasp of the technology which adequately equips them to set a course for how we want to utilize the technology, and what possible outcomes we want to curb. Part of this, also consists in presenting concrete opportunities that generative AI brings, good case examples from other countries, and knowledge on what competencies authorities need in the future. There is a need for tacit, easily accessible information communicated in lay-language that politicians can use as a foundation to engage in conversations about generative AI.

Emerging technologies calls for broad, and open democratic debates involving multiple actors to uncover the nuances of its potentials and risks, and this is another key role that TA could play. In a systematically digitalized society like the Danish one, generative AI is likely to have immediate and broad impact. For this reason, citizen and stakeholder engagement in relation to generative AI is highly relevant to Denmark. This both in relation to general principles for usage, vision building, open research agenda setting, and engagement on topics concerning specific sectors; for example, principles for using and learning about generative AI in the educational system. There are high hopes in Denmark for AI technology generating a lot of societal benefits, but realizing such hopes hinges on bringing citizens and stakeholders into a debate on what can be done and how. In the same way, there is a great opportunity in including the public in defining positive visions for the technology, and understanding what the room of acceptance is for the technology.

As shown in this paper, there are various hopes and concerns for how generative AI can impact Danish society and our democracy. From the concerns of the outside influence due to geopolitical tensions, to the hopes that generative AI can empower and increase inclusivity of vulnerable citizens. The possibilities are many, but they all share one thing, that by engaging in democratic conversations with multiple actors, Danish MPs can be better prepared to make informed decisions that hinder our concerns from becoming reality and pave the way for our hopes to become the future.

European Parliament

"Generative Artificial Intelligence – Opportunities, Risks, and Policy Challenges"
- EP STOA contribution -

Description of the System/Problem at stake

What is it about?

Generative Artificial Intelligence (AI) is a branch of AI that focuses on the development of algorithms capable of generating original and creative outputs, such as texts, images, music, and more. One of the most prominent examples of generative AI is Large Language Models (LLMs), which have been at the forefront of recent breakthroughs in AI research. Large Language Models, such as OpenAI's ChatGPT, are built using deep learning techniques and trained on massive datasets containing vast amounts of text from diverse sources. These models are designed to generate human-like text by predicting the most likely next word or sequence of words given a specific context. LLMs have demonstrated impressive capabilities in various applications, such as natural language understanding, question answering, content generation, and translation.

What do Members think about it?

The following replies were compiled based on the feedback obtained from some of our Members:

1. What is the first thing that comes into your mind, when you hear "generative artificial intelligence"?

- Foundation models
- general purpose AI
- large language models
- GPT
- beyond discriminative AI
- serious questions around the nature of reality and its impact on our societies
- a fear of being trapped in a world of illusions

2. Do you think that there is a relationship between generative AI and democracy? Could you briefly explain your thought on that?

- Generative AI could be used to create massive amount of fake news articles, social media posts, and other forms of content that spread misinformation and sow discord. This could undermine public trust in democratic institutions and make it more difficult for people to make informed decisions.
- Generative AI could be used to target people with personalized messages that are designed to influence their opinions. This could be used to sway elections, promote specific policies, or even suppress voter turnout.
- On the other hand, Generative AI could be used to create educational materials, improve civic engagement and make it easier for people to understand complex political issues.
- Generative AI could be used to analyze government data and identify patterns and trends that would otherwise be difficult to see. This could help to improve transparency and accountability in government.
- Absolutely! With generative AI that has a strong mastery of language, on which our political conversations and democracies ultimately rely on, it may not always be possible to determine whether we are discussing with a human or AI.

- As Yuval Harari puts it, AI may foster fake intimacy with humans to try to convince them of particular ideas, opinions, or candidates. This is hugely problematic when it comes to our public discourse, and by extension, our democracies.

3. If you were to think about policy areas where generative AI will most likely have to be dealt with, which would they be?

- Privacy/deepfakes
- Employment /potential to automate many jobs
- Education/ personalized learning, ensure that all students have access to generative AI tools
- Healthcare: drug discovery and personalized medicine, ensure safety and ethics
- Intellectual property/ protect creators while encouraging innovation
- Education
- Creative industries
- Intellectual Property Rights
- Discrimination
- Disinformation

4. Can you imagine how generative AI can affect the evolution/structure of the job market/employment sector/work?

- A lot of white-collar jobs will become less labour-intensive.
- AI current technologies could automate at least 20% of work hours.
- Blue-Collar jobs will become far more lucrative.
- Language models and image, audio or video generation certainly have the potential to shake up the labour market, with both positive and negative consequences. Though lower-skilled jobs may still be most threatened by automation, high-skilled jobs, too, may be at risk.

5. What kind of knowledge do you need to make decisions about generative AI? And how can TA institutions support you?

- Need practical training in using (and designing) GenAI tools for the staff so we can experience the effects of GenAI first-hand not only through reports
- At least a basic knowledge about the functioning of and the technology behind generative AI, for example how the models are trained and produce content, and an solid understanding of how they are used in particular areas of society and work, including information about its potential impact.
- TA institutions can certainly help in this with their expertise and experience in supporting Members in their policy-making. This can include workshops, briefings, but also practical demonstrations of generative AI to confront Members with the technology.

Who are the key stakeholders?

General-purpose artificial intelligence (AI) technologies, such as ChatGPT, are quickly transforming the way AI systems are built and deployed. While these technologies are expected to bring huge benefits in the coming years, spurring innovation in many sectors such as industry and education, their disruptive nature raises questions among general users, NGOs and policymakers around privacy and intellectual property rights, liability and accountability, and concerns about their potential to spread disinformation and misinformation. EU lawmakers need to strike a delicate balance between fostering the deployment of these technologies while making sure adequate safeguards are in place.

Why is this important for the European Union?

To maximize the benefits and mitigate risks related to the adoption of generative AI, the European Union should take stock of current and planned investments and policy actions, such as the European AI Strategy,

Digital Single Market, and Horizon Europe programs. These efforts can be leveraged to support research, development, and innovation in generative AI, while also addressing the ethical, legal, and social implications of these technologies. Key areas of focus should include:

— Providing sufficient computing capacity for European R&D: Investing in high-performance computing infrastructure will be essential for advancing the development of Large Language models (LLMs) and ensuring the competitiveness of European AI research. (For example: supporting underrepresented languages: Encourage the development of LLMs and AI technologies that cater to the diverse linguistic landscape of Europe, helping to mitigate language inequality and promoting the preservation of cultural heritage.)

— Ethical guidelines and legal frameworks: Establishing ethical guidelines and a robust, future-proof, common European legal framework to address issues related to misinformation, privacy and data protection, bias, intellectual property, and language inequality.

— Incentives for responsible AI development: Promoting the development of AI systems that prioritize fairness, transparency, and accountability, while minimizing the negative impacts on society.

To secure the European Union's strategic autonomy and global competitiveness in generative AI, it is crucial to invest in research and development, foster collaborations between academia and industry, and promote the growth of European AI startups. Developing EU-based LLMs and technologies can help reduce dependency on external players and foster a competitive ecosystem.

Societal and political relevance and debate

Is there ongoing debate on the impact of generative AI on our societies and democracies?

The key characteristics identified in general-purpose AI models – their large size, opacity and potential to develop unexpected capabilities beyond those intended by their producers – raise a host of questions. Studies have documented that large language models (LLMs), such as ChatGPT, present ethical and social risks. They can discriminate unfairly and perpetuate stereotypes and social biases, use toxic language (for instance inciting hate or violence), present a risk for personal and sensitive information, provide false or misleading information, increase the efficacy of disinformation campaigns, and cause a range of human-computer interaction harms (such as leading users to overestimate the capabilities of AI and use it in unsafe ways). Despite engineers' attempts to mitigate those risks, LLMs, such as GPT-4, still pose challenges to users' safety and fundamental rights (for instance by producing convincing text that is subtly false, or showing increased adeptness at providing illicit advice), and can generate harmful and criminal content.

Since general-purpose AI models are trained by scraping, analysing and processing publicly available data from the internet, privacy experts stress that privacy issues arise around plagiarism, transparency, consent and lawful grounds for data processing. These models represent a challenge for education systems and for common-pool resources such as public repositories. Furthermore, the emergence of LLMs raises many questions, including as regards intellectual property rights infringement and distribution of copyrighted materials without permission. Some experts warn that AI-generated creativity could significantly disrupt the creative industries (in areas such as graphic design or music composition for instance). They are calling for incentives to bolster innovation and the commercialisation of AI-generated creativity on the one hand, and for measures to protect the value of human creativity on the other. The question of what liability regime should be used when general-purpose AI systems cause damage has also been raised. These models are also expected to have a significant impact on the labour market, including in terms of work tasks.

Against this backdrop, experts argue that there is a strong need to govern the diffusion of general-purpose AI tools, given their impact on society and the economy. They are also calling for oversight and monitoring of LLMs through evaluation and testing mechanisms, stressing the danger of allowing these tools to stay in the hands of just a few companies and governments, and highlighting the need to assess the complex dependencies between companies developing and companies deploying general-purpose AI tools.

AI experts are also calling for a 6-month pause, at least, in the training of AI systems more powerful than GPT-4. (EPRS briefing)

Are there results from surveys among the population concerning usage of or opinions about generative AI?

Not specific Eurobarometer on Generative AI.

A more general survey on the 'Digital Decade' was published in June 2023. According to this survey, a majority of Europeans, believe that advanced connectivity (76%) and stronger cybersecurity (77%) will make their daily use of digital technologies significantly better. Furthermore, two thirds of European citizens (67%) call for more education and training to develop their digital skills. Similarly, two thirds of citizens (66%) consider that digital technologies will play an important role in fighting climate change. Finally, more than 80% of Europeans believe that Member States should collaborate more to advance the access to digital technologies, as well to spur innovation and for companies to be globally competitive. When it comes to a value-based digital transformation, only half of Europeans consider that digital rights and principles are well protected in Europe. Over a third of citizens (36%) think more needs to be done, and on various aspects, less than half find that the implementation of the digital rights and principles in their country is satisfactory. Citizens are especially concerned when it comes to ensuring safe digital environments and content for children and young people, or getting control of one's own data or digital legacy.

A 'European enterprise survey on the use of technologies based on artificial intelligence' was published by the European Commission DG CONNECT in 2020. According to it, awareness of AI is high across the EU (78%). Four in ten (42%) enterprises have adopted at least one AI technology, 25% have adopted at least two. While 18% have plans to adopt AI in the next two years, 40% have neither adopted AI nor plan to do so. Adoption at the level of each technology is still relatively low: from 3% for sentiment analysis to 13% for anomaly detection and process/equipment optimisation. The most common sourcing strategy is external, as 59% of EU enterprises that use AI purchase software or ready-to-use systems. Three key internal barriers to AI adoption are difficulties in hiring new staff with the right skills (57%), the cost of adoption (52%) and the cost of adapting operational processes (49%). Reducing uncertainty can be beneficial, as enterprises find liability for potential damages (33%), data standardisation (33%) and regulatory obstacles (29%) to be major external challenges to AI adoption.

Is there any legislation in place?

Regulating Artificial Intelligence has to be closely linked to a proper regulation of data and its governance as well as to the concept of liability. After recently passing legislation such as the DSA, DMA, Data Act and Data Governance Act, the European Union is finalizing work on the Artificial Intelligence Act that will set the generic mechanisms to regulate the application and development of AI. It will assess how this regulation should be applied to Generative AI and whether additional regulatory instruments are needed. The European Parliament will monitor the implementation of the EU's program for research and innovation Horizon that facilitates and funds research on Artificial Intelligence in Europe. It will oversee the 2030 Policy Program "Path to the Digital Decade," with a target of 75% of EU companies using Cloud/AI/Big Data by 2030. The European Union will set European policies in areas strongly affected by Generative AI: creative economy, education, social and cultural areas, and health technologies.

Are there currently political or legislative proposals on these topics?

The draft EU AI act is the first ever attempt to enact a horizontal regulation for AI. The proposed legal framework focuses on the specific utilisation of AI systems and associated risks. The European Commission proposes to establish a technology-neutral definition of AI systems in EU law and to lay down a classification for AI systems with different requirements and obligations tailored on a 'risk-based approach'. Some AI systems presenting 'unacceptable' risks would be prohibited. A wide range of 'high-risk' AI systems would be authorised, but subject to a set of requirements and obligations to gain access to the EU market. Those AI systems presenting only 'limited risk' would be subject to very light transparency obligations. The Council agreed the EU Member States' general position in December

2021. A special committee (AIDA) was stablished in the European Parliament on the topic, that presented its final report on artificial intelligence in a digital age in April 2022. Parliament voted on its position in June 2023. EU lawmakers are now negotiating to finalise the new legislation, with substantial amendments to the Commission's proposal including revising the definition of AI systems, broadening the list of prohibited AI systems, and imposing obligations on general purpose AI and generative AI models such as ChatGPT.

Role of TA in the debates

Has your institute taken up this topic?

STOA published more than 30 publications on artificial intelligence since 2016. In addition to the standard studies, STOA published a series of other publications, like 'What if?' and At-a-glance publications providing awareness-raising overviews of current and relevant techno-scientific trends, as well as blog posts, articles, interviews, videos and podcasts. (See AI Repository).

Has TA made an impact on the ongoing debates?

STOA has been increasingly using social media and other channels to communicate its activities. This includes STOA's own X (formerly Twitter) account (@EP_ScienceTech), and blog posts on the EPRS blog, which report on news, projects and events in an accessible manner. STOA also continues to produce regular podcasts, often based on its short awareness-raising publications and videos. STOA events are live webstreamed and often accompanied by live posting from @EP_ScienceTech, enabling simultaneous interaction with stakeholders, experts and citizens. During their meetings in Strasbourg plenary weeks, the members of the STOA Panel usually hear presentations of the results of ongoing STOA projects and provide feedback and further guidance for completing the reports. Since STOA was among the first entities in the European Parliament to work on AI, its findings have informed and fed into the MEPs' parliamentary activity. This was especially relevant for the discussions taking place at several committees, such as the special committee on AI ("AIDA") and the legislative committees leading the work on the AI Act (the Civil Liberties committee, the Internal Market committee and the Legal affairs committee).

What are the lessons learned from TA?

EU lawmakers are currently engaged in protracted negotiations to define an EU regulatory framework for AI that would subject 'high-risk' AI systems to a set of requirements and obligations in the EU. The exact scope of a proposed artificial intelligence act (EU AI act) is a bone of contention. While the European Commission's original proposal did not contain any specific provisions on general-purpose AI technologies, the EU Council has proposed that they should be considered. Scientists have meanwhile warned that any approach classifying AI systems as high-risk or not depending on their intended purpose would create a loophole for general purpose systems, since the future AI act would regulate the specific uses of an AI application but not its underlying foundation models.

In this context, scientific opinions, such as that of the Future of Life Institute, have called for general-purpose AI to be included in the scope of the AI act. Some academics favouring this approach have suggested modifying the proposal accordingly. Helberger and Diakopoulos propose to consider creating a separate risk category for general-purpose AI systems. These would be subject to legal obligations and requirements that fit their characteristics, and to a systemic risk monitoring system similar to the one under the Digital Services Act (DSA). Hacker, Engel and Mauer argue that the AI act should focus on specific high-risk applications of general-purpose AI and include obligations regarding transparency, risk management and non-discrimination; the DSA's content moderation rules (for instance notice and action mechanisms, and trusted flaggers) should be expanded to cover such general-purpose AI. Küspert, Moës and Dunlop call for the general-purpose AI regulation to be made future-proof, inter alia, by addressing the complexity of the value chain, taking into account open-source strategies and adapting compliance and policy enforcement to different business models. For Engler and Renda, the act should discourage API access for general-purpose

AI use in high-risk AI systems, introduce soft commitments for general-purpose AI system providers (such as a voluntary code of conduct) and clarify players' responsibilities along value chains.

While STOA can outline different options from a technical perspective (such as on the scope of the AI regulation, the definition of AI and the classification of risks), Members may obviously also consider other perspectives or elements in their legislative work. Nevertheless, STOA's wide-ranging activities, including technology assessment and scientific foresight studies, together with science and technology communication constitute an essential asset with its decisive contribution to raising awareness about the impact of new developments and promoting evidence based policy making. STOA also shares its knowledge and experience with colleagues from national parliaments in form of a regular online information session 'Knowledge sources on AI', offered to national parliaments.

France

EPTA Report 2023 on Generative Artificial Intelligence Contribution of the French Parliamentary Officefor Scientific and Technological Assessment (OPECST)

The French Parliamentary Office for scientific and technological assessment (OPECST) started to deal with Artificial Intelligence (AI) from 2016 and adopted a report entitled "Toward a Controlled, Useful and Demystified Artificial Intelligence" on 14 March 2017.

In 2018, France decided to rise to this challenge with the implementation of a national plan and the set up of an intergovernmental coordination and a specific national ethics committee.

Between 2017 and the early 2020s, advances in transformer-based deep neural networks enabled new generative AI systems, that use natural language prompts as input. Consequently, large language models trained on vast quantities of unlabeled began to be developed. Models such as chatbots (ChatGPT 3&4, launched by OpenAI, Microsoft Bing Chat, Google Bard & PaLM, DeepMind Gato, Meta LLaMA etc.) or text-to-image art systems (Stable Diffusion, Midjourney, DALL-E etc.) have become well known to the general public. A French generative AI start-up, Mistral AI, raised this year $105 million just one month after its creation, which is a record. Unlike ChatGPT or Bard, its model called Mistral 7B is open source and intended for developers, who will be able to use it, improve it and market it as they wish, thanks to a very permissive license. But it is a "small" language based on 7 billion parameters, a far cry from the hundreds of billions of parameters of GPT4 from OpenAI or PaLM from Google.

As a consequence of this new context of generative AI systems, the Bureau of the National Assembly and the Bureau of the Senate have recently asked OPECST to establish a new report about issues of generative artificial intelligence. This political request at the highest parliamentary level shows to what extent the current work of EPTA is crucial .

This topic needs to be thoroughly analysed and demystified, because the rise of artificial intelligence technologies represents a major shift, deeply transforming our societies and economies. Artificial intelligence is subject to constraints of social acceptability, because of alarmist visions. The potential dangers of AI reinforced the fears and the anxieties facing the deployment of artificial intelligence technologies.

The cycles of hope and disappointment that mark the history of artificial intelligence suggest that we should be cautious and demonstrate realistic expectations of these technologies: periods of polarized opinions, excessive anxieties as well as excessive hopes, have been seen before. A detour through history is essential. The concept of artificial intelligence refers to the many technologies that emerged in the second half of the 20th century, based on the use of algorithms. These technologies, with combinations constantly evolving, are already being implemented in a wide range of sectors and have led to unprecedented opportunities to revolutionize our living environment and improve our lives especially regarding care for people.

Developments in this field can be rapid and the current or future sectorial applications are of considerable scope including education, science, environment, energy, transport, aeronautics, agriculture, trade, finance, defense, security, communications, recreation, health, dependency, disability and countless others. Public debate cannot be engaged rationally without a good understanding of AI technologies, scientific methods, and principles.

Progress in artificial intelligence raises questions that the society must be aware of: what are the new opportunities and risks? What is the position of France and Europe in the world race that has started? Which respective places for public research and private research? What ethical, legal and policy principles should guide these technologies? Should regulation take place at national, EU, or international level?

We are indeed facing ethical, legal, economic, social and scientific challenges with these technologies. Some of the challenges include:

— the dominant role of private research, led by American and potentially Chinese companies; accelerating the transition to a dominated globalized economy by "platforms" (like Google, Facebook or Amazon…) which can critically damage democratic governance;

— algorithms issues with risks related to biased AI outputs - consequences of biased data and algorithms - risks related to the phenomenon of "black boxes" of algorithms and risks related to "information bubbles" or "filter bubbles", what can also damage democratic governance;

— last but not least, labor market transformations, AI will displace many workers and also create new jobs, often changing the content of our work and the way we work, which could have effects in terms of social inequality.

Artificial intelligence is also seen as an opportunity for our societies and economies; it is neither a vain quest nor a plan to replace man by machine. We are moving towards an augmented human intelligence rather than an artificial intelligence competing with humans. Advances in artificial intelligence are first and foremost beneficial. However, it can't be denied that they also involve risks. These risks can and must be anticipated, identified, and mitigated.

The advent of super-intelligence is not part of these risks in the short to medium term. It is still uncertain that such a threat should become reality in the long run. As for its imminence in the short or medium term, prophesied by several media figures, it is just a fantasy.

It is, therefore, necessary to go beyond appearances and to look at the scientific reality behind the hopes and anxieties expressed on the development of artificial intelligence.

This work of demystification must be collective, interdisciplinary and international. In order to prevent future disillusionment, it is necessary to ensure continuous monitoring of these technologies and their uses.

That is why OPECST advocated controlled, useful and demystified artificial intelligence in its 2017 report.

Controlled, because these technologies will have to be the safest, most transparent and fairest possible.

Useful because while they must respect humanistic values, they ultimately have to benefit the general public.

Demystified, because the difficulties of social acceptability of artificial intelligence largely result from unfounded alarmist visions and the lack of understanding.

Rather than reporting a hypothetical confrontation between men and machines, which is dystopian science fiction, the OPECST forthcoming work should clarify the conditions and guidelines for future complementarity between man and machine.

Germany

EPTA Report 2023
Generative Artificial Intelligence - View from Germany / TAB

ChatGPT is met with mixed reactions

"It was merely a matter of time" was a comment in the media, when a student at a Hamburg secondary school (Gymnasium) was caught in May 2023 cheating his/her way through the final exam with the help of ChatGPT.[33] Such a use of generative AI was indeed not unexpected: Experts had warned students would have ChatGPT write their essays (Weßels 2022), journalists had shown that one could achieve good results with its latest version in various subjects,[34] some schools in the US already had explicit bans in place,[35] and the use of such unauthorized aids in exams was one of the scenarios foreseen in TAB's report on ChatGPT for the German Bundestag (Albrecht 2023.76f.). The reaction of the school authority, however, was quite surprising: It did not tighten the rules or sanctions against the use of ChatGPT in exams, but rather allowed its use, provided that it is being documented.[36] Teachers were called to adapt their exams accordingly and test students' knowledge directly if they distrusted the students' honesty in citing their sources.

This reaction is somewhat symptomatic for the German perspective on the new generation of generative AI systems. On the one hand, their introduction was met with reservations - ChatGPT for example was characterised as flawed (in its responses), as dangerous with regard to public security and public opinion, and as to be treated with caution because of the largely unknown effects of the system on various areas of society and also because of the unknown fate of data entered into the system. On the other hand, ChatGPT enthralled even experts who had followed the development of generative AI closely. The quality of the texts generated in response to mere short prompts marked a huge leap forward from previous AI systems. The system proved able to convincingly simulate both understanding and producing utterances, thus playing the role of a communication partner. Soon it became clear that the technology underlying ChatGPT could be used not only to solve various text-based tasks, but also for image generation and manipulation. A whole family of powerful new AI systems, called generative AI or foundation models, was released in a short time. Simple bans or recommendations not to use the systems would not be a viable option. Rather, an intense debate ensued in the German public about how to make most of the potential of generative AI in various fields of application without falling into its risks and pitfalls. And it was discussed what kind of guardrails were needed to guide the technological development and the applications of the system in a direction beneficial for society as a whole.

Adoption and economic relevance of (generative) AI in Germany

Although the introduction of ChatGPT was globally followed by the fastest growth of any consumer application so far[37] and the ensuing public debate in Germany grew enormously (as in other countries), its actual adoption in Germany was less dynamic. By the end of January 2023, according to a survey (CAIS 2023), about one quarter of the adult population had heard about ChatGPT, and 11% said they had already used the system, with 3% indicating a regular use. By the end of April 2023, 83% had heard about ChatGPT and 23% had used it (at least once), most for private purposes (Bühler 2023). In terms of business use, a

---

33. https://www.ndr.de/nachrichten/info/Hamburg-Mit-Chat-GPT-im-Abi-gemogelt,ndrinfo46278.html
34. https://www.br.de/nachrichten/netzwelt/chatgpt-ki-besteht-bayerisches-abitur-mit-bravour,TfB3QBw
35. https://ny.chalkbeat.org/2023/1/3/23537987/nyc-schools-ban-chatgpt-writing-artificial-intelligence
36. https://www.abendblatt.de/hamburg/kommunales/article238612213/Nun-doch-Schueler-duerfen-ChatGPT-bei-Abitur-Pruefung-benutzen.html
37. https://www.reuters.com/technology/chatgpt-sets-record-fastest-growing-user-base-analyst-note-2023-02-01/

survey in June 2023 showed that 13% of all German companies had some kind of AI technology in use (not necessarily generative AI), with the strongest uptake observed in industry (17% in use, 13% planning to use, Schaller et al. 2023, p. 5).

About a third of the population in Germany views the use of AI generally positively, with 45% being sceptical and 10% refusing it generally (March 2023, Continental 2023).[38] A majority of respondents fears that AI could cost jobs or could become uncontrollable.[39] The introduction of ChatGPT did not change much about this general picture, for 53% of respondents, nothing has changed, 18% have more positive views now as opposed to 14% whose views have turned more negative. Whereas most people assume that AI is going to change their life by the year 2030, the views are split as regards the direction of change: One third expects that life will change for better, 23% for worse, and 30% don't know - highlighting a big amount of uncertainty about the future development of AI and its use.[40]

This general stance in society is somewhat at odds with the fact that Germany has a lively and long-established AI research scene. Although there are almost no big IT companies like in the US (with the exception of SAP), the number of publications from German research institutions on AI matches that of the US and Israel and is higher than in other countries - at least if measured in proportion to the population.[41] There exist several research funding programmes and a national AI strategy,[42] more than 1 billion € were spent in this scheme between 2019 and the end of 2022 (Bertschek 2023). But when it comes to translate research excellence into economic success stories, Germany clearly lags behind the top nations. This applies to the number of patents, and even more so with respect to ready-to-market AI solutions, according to the president of Bitkom, an association representing Germany's digital economy companies.[43] Nevertheless, with the company Aleph Alpha, a major European player in the field of generative AI systems is registered in Germany, actively marketing its AI technology as "sovereign" and "European".[44]

A vivid debate in face of many unknowns

AI researchers and AI companies, represented by their executives and also by the German AI association (a network of 400 AI companies), were among the most prominent stakeholders in the debate about generative AI that took up in the end of 2022. They warned that the planned European AI Act could overregulate generative AI and thus inhibit the growth of a European AI economy. They also highlighted the qualities of AI systems developed under the framework of the GDPR and other European frameworks, especially with regard to the application of AI in the industry, focusing on confidentiality and responsibility in the handling of sensitive data (Aleph Alpha 2023). Other interest groups voiced their views as well. Artists and other creators pointed to the need to protect their rights, fearing that AI systems could be used to automate their work and to deprive them of their economic base (by using their work as training data), also leading to an increase in disinformation and manipulation (Initiative Urheberrecht 2023; DJV 2023). Publishers feared that chatbots autonomously reporting news without paying license fees could further stress their business model in addition to the challenges brought by the digital market (Tagesspiegel Background 2023). Many educators, on

---

38. Similar results are found in a survey from April 2023, with 30% approving the development of AI-based technologies, 37% disapproving it and 30% being neutral (Fox et al. 2023). In a survey from late April 2023, only 18% of respondents agreed that AI would make the world a better place, but 49% see a potential to support them in their private life (Bühler 2023).

39. In the survey by Fox et al. (2023), only a third of respondents fear job losses, the major fears are AI as threat to humanity, the manipulation of public debate and widespread surveillance.

40. Results from Fox et al. (2023) are similar, with 14% of respondents expecting positive change "in the next ten years", 39% negative change, 38% a mixture of both and 8% who don't know.

41. https://www.plattform-lernende-systeme.de/ki-monitoring.html

42. https://www.bmbf.de/bmbf/shareddocs/downloads/files/201201_fortschreibung_ki-strategie.pdf?__blob=publicationFile&v=2

43. https://www.bitkom.org/Presse/Presseinformation/KI-Deutschland-Spitze-Forschung-Nachholbedarf-Wirtschaft

44. https://www.aleph-alpha.com/

the other hand, were open to the new developments, and called for experimental uses of the new technology to identify the conditions for a productive and safe use in education (e.g., Salden/Leschke 2023, Mohr et al. 2023). Some hoped that generative AI could further stimulate the debate about new forms of assessing students' progress, a debate that has been going on in Germany already for some time. However, educators also warned that generative AI's results could not be trusted and that user data needs protection.

The German parliament took an active part in this debate, trying to bring the different perspectives together into a constructive dialogue. Already in 2018, a Study Commission on Artificial Intelligence was established. Its final report was published in 2020, it was meant less as a roadmap for future developments, but rather as a memorandum of understanding and as an intermediate step in political considerations about AI. Several projects on AI's impact on, e.g., the military (Grünwald/Kehl 2020), education[45] public administration (Evers-Wölk et al. 2022) and the production of media content that appears to be true to reality (Deepfakes),[46] to name just a few, were commissioned to TAB in the last years. With the Committee on Digital Affairs, a standing committee is in place since 2014 to deal with questions related to digitalisation and digital infrastructure. It was this committee and the Committee on Education, Research and Technology Assessment who in spring 2023 held expert hearings on ChatGPT and generative AI to keep the parliamentarians informed about the latest developments and to explore potential needs for regulatory action. TAB was commissioned to study the technological foundations of ChatGPT and to identify potential implications of its applications in various areas of society (Albrecht 2023). These committee-led activities and the occasional discussions about generative AI in the Bundestag's plenary are just the beginning of further activities, as the MPs responded in the questionnaire for this report. In the words of one MP, "further developments in the field of generative AI will continue to be monitored regularly and on an ad hoc basis."

Despite these activities and a consensus that the development of generative AI is of high societal relevance, the potential impacts of generative AI are still to be explored. Given the orientation of many generative AI systems on language and communication (and thus the fundaments of social life), they will likely affect many areas of society. Among the potentials identified in TAB's report (Albrecht 2023) are the automatic generation of texts or responses in fields such as customer service, journalism, legal affairs and even literature and research. But such use could come at the price of a (potentially massive) increase in texts of questionable quality, the displacement of creative jobs, a loss of corresponding competencies and potentially even a loss of trust in the authenticity of publications. AI systems could be constructed to autonomously solve more and more tasks, become a convenient interface for the use of computer systems more generally, but also be mistaken as persons with an individual character and intentions. On a positive note, multimodal and multilingual systems such as ChatGPT could facilitate communication across languages and could improve the inclusion of persons with a learning or cognitive impairment. However, bias stemming from training data or the black-box-character of machine learning could also lead to bad decisions or discrimination. Furthermore, it is expected that generative AI could worsen IT security by empowering malicious actors. The enormous amount of resources (electric energy, cooling water, human labour) needed for the training of such large systems contradicts sustainability efforts.

What is at stake in politics and individual policy areas?

The responses of the German parliamentarians interviewed for this EPTA report confirm the view that there is a lot at stake when it comes to generative AI. They expect "enormous chances", but also see risks in many potential applications of AI systems. Generative AI could bring "revolutionary change" to many occupations, could "reinforce entrenched processes and opinions" or lead to "harmful decisions" in medicine. Dealing with generative AI might "become one of the key skills for the working world of the 21st century". Interestingly, despite assuming far-reaching implications, the parliamentarians do not join in the hype around

---

45. https://www.tab-beim-bundestag.de/english/projects_application-potentials-and-challenges-of-artificial-intelligence-in-education.php
46. https://www.tab-beim-bundestag.de/english/digital-society-and-economy_deepfakes-legal-and-societal-challenges-as-well-as-innovation-potentials.php

generative AI that has unfolded in the German debate. Rather, they cautiously consider the opportunities and risks that stem from the innovation. Some point to the long history of AI and of the corresponding political debate, others to the many open questions and knowledges needs. "The impact on specific policy areas is difficult to foresee," states one MP with regard to generative AI's impact.

With regard to *implications for the political system and democracy* in general, all MPs see potential negative effects of a more widespread use of generative AI. The German Federal President had already highlighted some of the challenges posed to democracy by the rise of generative AI, namely that it distorts our relation to facts, that it could undermine democratically legitimised decision-making and encourage manipulation and malicious practices (Bundespräsidialamt 2023). The MPs responses point in a similar direction, emphasising that generative AI could facilitate the mass production of disinformation and of content (textual, visual and audio) aimed at damaging democratic structures, leading to "more idle talk" that is "less reliable". It can be used to manipulate citizens, as it is difficult to "determine whether it [content produced by generative AI] reflects reality or is merely artificial". If media organisations use AI systems instead of journalists to produce or distribute news, plurality of opinion could suffer even more than is already the case. Potential biases of the AI systems could distort the work of parliaments. And because the models are statistically trained on knowledge from the past, relying on their output could lead to the neglect of minority or dissenting positions in in political debates. Instead, "democracy should be driven by the compromise of conflicting goals." However, most MPs also see a potential role for generative AI in politics and in parliamentary work, namely it could help prepare speeches, motions, questions or statements, it could facilitate the communication with citizens and on social media, and it could support research on political topics and the translation of texts.

The *implications on the job market* are seen as highly contingent and uncertain, with some MPs expecting both gains and losses in terms of the number of jobs, others expecting "revolutionary changes" in terms of the quality of work. Five of the six MPs interviewed expect that there will be fundamental changes, and three of them consider it important that qualification and education options are offered so that the working force in the future is able to meet the future needs.

In *policy areas such as education and health*, there are some interesting differences in MPs expectations about the impact of generative AI. They seem more cautious about applications in medicine, whereas in education, where generative AI is already being applied (as the case mentioned in the introduction shows) and the focus is more on shaping the processes and conditions for a responsible use of the technology.

In the field of *education*, there is a consensus among the MPs that the use of generative AI presents both opportunities and risks. For learners, it could enable more personalised learning experiences and improve digital literacy. It could also help to reduce inequality and improve access to education for disadvantaged groups. However, there is also a risk that learners will simply skip learning or individual thinking and instead rely on the AI's results. As the texts produced by generative AI become more and more similar to human texts, it becomes increasingly difficult for educators to assess the learners' own competencies. To cope with this new situation, teachers need specific training in how to use AI - how to teach with the help of AI, but also how to teach the skills needed to use AI. They also need to adapt their examinations to assess these new competencies. School boards and principals need support in choosing the appropriate tools (e.g. certification systems for AI products). Personal data and rights of students, especially in the case of minors, have to be protected if AI is applied in educational contexts.

The role of *generative AI in healthcare* in the view of the MPs is one of supporting the medical practitioners, provided that the responsibility and control always remains with a human. "There must always be a possibility to 'overrule' a decision made by the AI," states one MP. In order to reap the potential benefits of using AI in healthcare in the form of better diagnoses and treatments, there must a strong regulation in place to ensure that the handling of sensitive personal data is careful and legitimate, that training data for AI health applications is of utmost quality and free of bias (gender, age, race etc.), and that the use of AI is transparent and traceable. "It must be made clear along the whole process which decisions/recommendations are from the AI and which are from a human person." Ethical issues have to be addressed, and the risk of a misuse of the AI-generated knowledge by insurance companies or employers needs to be minimised through regulation. As in education, the use of generative AI "should be integrated into the clinical training of medical professionals to prepare for and responsibly shape broader use."

In general, the MPs focus on questions of "thoughtful regulation" of generative AI, i.e. transparent communication of the use of AI on social media platforms, avoiding discrimination and making results traceable, and also on appropriate ethical, legal and democratic guardrails. As the impact of AI is difficult to predict, more research in this area is seen as necessary. This is where MPs see the role of technology assessment: it should inform parliamentarians about how algorithms work, and the data used and should monitor the technical developments to help them stay up to date. It should provide knowledge about potentials as well as risks of generative AI and generally about the actual as well as potential impact in the various application areas. It could provide this knowledge in the form of scenarios, and such knowledge could serve "as important starting points for societal debates". Technology assessment could draw on the experiences in other countries to learn about impacts and the efficacy of regulatory measures. And it is considered "indispensable" in considering "what regulations are possible/necessary so that we can exploit the opportunities and reduce the risks that will arise from the use of generative AI."

For (parliamentary) technology assessment, the stakes are high to meet these far-reaching expectations. As the developments unfold very fast, the established procedures of technology assessment need to be put to the test. Assessments need to be conducted almost in real time as the technological development proceeds. Then again, the potential of generative AI to improve the efficiency of various communication tasks could enhance and improve the work of parliamentary TA institutions - as well as the political process as a whole. However, whether and how it is possible to use generative AI in a responsible and sustainable way is an open question.

References

Albrecht, S. (2023): ChatGPT und andere Computermodelle zur Sprachverarbeitung – Grundlagen, Anwendungspotenziale und mögliche Auswirkungen. Berlin: Office of Technology Assessment at the German Bundestag (TAB)

Aleph Alpha (2023): Schriftliche Stellungnahme von Jonas Andrulis, Gründer und Gesellschafter der Aleph Alpha GmbH. Ausschussdrucksache 20(23)157 vom 24.5.2023, Berlin: Deutscher Bundestag, Ausschuss für Digitales, https://www.bundestag.de/resource/blob/949934/ad322fe569ad8c40df-1688665f9e7c22/Stellungnahme-Andrulis-data.pdf

Bertschek, I. (2023): Jetzt bloß nicht den Anschluss verlieren! – Status quo, Potenziale und Herausforderungen von Künstlicher Intelligenz. In: Wirtschaftsdienst 103(8), pp. 518-520 (DOI: 10.2478/wd-2023-0149)

Bühler, J. (2023): ChatGPT & Co.: Sicherheit von generativer Künstlicher Intelligenz. TÜV-Verband, Pressekonferenz, 11. May 2023. https://www.tuev-verband.de/fileadmin/user_upload/Content_local/Studien_local/TUEV-Verband_PK_Praesentation_ChatGPT_11_05_2023_final.pdf

Bundespräsidialamt (2023): Federal President Frank-Walter Steinmeier at a dinner for members of the Order Pour le mérite for Sciences and the Arts on 11 June 2023 at Schloss Bellevue. https://www.bundespraesident.de/SharedDocs/Downloads/DE/Reden/2023/06/230611-AE-Orden-Pour-le-merite-Englisch.pdf?__blob=publicationFile

Center for Advanced Internet Studies (CAIS) (2023): ChatGPT, how many people already know you? Press release, 3. February 2023. https://www.cais-research.de/en/news/chatgpt-how-many-people-already-know-you/

Continental (2023): Umfrage Künstliche Intelligenz. https://cdn.continental.com/fileadmin/__imported/sites/corporate/_international/german/hubpages/10_20presse/studien_und_publikationen/umfrage_kuenstliche_intelligenz_2023/20230301-continental-umfrage-ki.pdf

Deutscher Journalisten-Verband (DJV) (2023): Positionspapier bezüglich des Einsatzes Künstlicher Intelligenz im Journalismus. https://www.djv.de/fileadmin/user_upload/INFOS/Themen/Medienpolitik/DJV-Positionspapier_KI_2023-04.pdf

Evers-Wölk, M.; Kluge, J.; Steiger, S. (2022): Künstliche Intelligenz und Distributed-Ledger-Technologie in der öffentlichen Verwaltung. Berlin: Office of Technology Assessment at the German Bundestag (TAB)

Fox, P.; Privitera, D.; Reuel, A. (2023): So denken die Deutschen über KI. KIRA-Report. Berlin: Zentrum für KI-Risiken & Auswirkungen

Grünwald, R.; Kehl, C. (2020): Autonome Waffensysteme - Endbericht zum TA-Projekt. Berlin: Office of Technology Assessment at the German Bundestag (TAB)

Initiative Urheberrecht (2023): Urheber:innen und Künstler:innen fordern Maßnahmen zum Schutz vor generativer KI in der Europäischen KI-Verordnung. 19. April 2023, https://urheber.info/media/pages/diskurs/ruf-nach-schutz-vor-generativer-ki/03e4ed0ae5-1681902659/finale-fassung_de_urheber-und-kunslter-fordern-schutz-vor-gki_final_19.4.2023_12-50.pdf

Mohr, G.; Reinmann, G.; Blüthmann, N.; Lübcke, E.; Kreinsen, M. (2023): Übersicht zu ChatGPT im Kontext der Hochschullehre. Hamburg: Hamburger Zentrum für universitäres Lehren und Lernen

Salden, P.; Leschke, J. (2023, eds): Didaktische und rechtliche Perspektiven auf KI-gestütztes Schreiben in der Hochschulbildung. Bochum: Zentrum für Wissenschaftsdidaktik der Ruhr-Universität Bochum, DOI: 10.13154/294-9734

Schaller, D.; Wohlrabe, K.; Wolf, A. (2023): KI, Cloud Computing und Blockchain – wo steht die deutsche Wirtschaft? In: IFO Schnelldienst 76(8), 16. August 2023, pp. 3-9

Tagesspiegel Background (2023): Chatbot-Suche: Alarm in den Verlagen. E-Mail-Newsletter Tagesspiegel Background Digitalisierung & KI, 13. 3. 2023

Weßels, D. (2022): ChatGPT – ein Meilenstein der KI-Entwicklung. https://www.forschung-und-lehre.de/lehre/chatgpt-ein-meilenstein-der-ki-entwicklung-5271

Greece

Greek Parliamentary TA Committee
Artificial intelligence and employment

The development of information technology is certainly an evolution of outmost importance in the modern era. We thus talk about algorithms, artificial intelligence and the 4th industrial revolution. The eruption of information technology is causing turbulence resulting from changes in the way work is organised.

I. Artificial intelligence and the company

The algorithms have been known for a long time. They are used today as part of the software construction of computer systems in several modern companies, which aim to imitate human thinking.

1. The development of artificial intelligence

Artificial intelligence is becoming a reality. The use of algorithms and of artificial intelligence is developing in all sectors of contemporary life. In the employment sector, they are the source of crucial questions.

Regarding employment and job cuts, the question is particularly important. However, the answers are divergent. Some consider that an increase in unemployment is to be expected. Others argue that a loss of jobs will be counterbalanced by the creation of new jobs related to the new technology. However, we will limit this study to the examination of issues related to artificial intelligence at the service of management functions.

2. Artificial intelligence at the service of management functions
a. The recruitment

Due to artificial intelligence, a range of data from employee CVs or other sources can be evaluated almost instantly to determine which employees will be hired directly or which employees will be interviewed. In a modern company, it is often an algorithm that, directly or indirectly, decides who will be recruited. Subsequently, it is also an algorithm that proposes the remuneration of the employee on the basis of his skills.At the limit, we can consider that the HR Department is replaced by the algorithm during the recruitment period. On the other hand, the company often attempts, by means of research on professional social networks, to identify competent professionals to whom it could present a job offer. In this case, it is the algorithm that will examine the profiles of a very large number of employees, likely to best meet its needs.

But, artificial intelligence is not limited to the beginning of professional life. It can follow the employee throughout his entire career in the company.

b. The execution of the employment contract

We know that several companies have for a long time installed systems for monitoring employees and assessing their activity. The data recorded concerns their productivity, their movements or the evaluations both from their hierarchical superiors and from the company's customers.

In a modern company, this data is intended to play a much larger role. Instead of personal control, a system for collecting and analyzing employees' personal data is put in place. The data analyzed in order to monitor the execution of the employment contract can come either from the inside the company or outside of it.

In a technologically advanced company not only is everything programmed, but everything is also meticulously controlled. In addition, technologically advanced companies, especially digital platforms, adopt evaluation systems derived from customers by means of algorithms in order to provide advantages to employees or, on the contrary, to disadvantage them.

The analysis of this data can, subsequently, help to set the worker's salary by evaluating the supply and demand data of the labor market. It can thus decide on the payment of productivity bonuses. It can also decide on his/her promotion on the basis of several data describing his/her activity in the company. The algorithm, which follows the employee throughout his/her career, can finally be used in the context of a dismissal.

c. The dismissal

The algorithm can be used for disciplinary reasons, especially taking into account customer evaluations. This procedure is particularly,used for platform employees (crowdworking). The users of the platform are in fact, as we have already noticed, constantly called upon to evaluate the employees.

In this way, the algorithm can decide on a dismissal. The problem becomes even more important when one takes into account that these evaluations on the part of customers are often formalized, impersonal, taking on the appearance of objectivity, the employee often not even having the possibility of contesting them. The need for transparency regarding both management decisions and the parameters used by the appraisal system is obvious.

II. The employees' protection

Algorithms making possible to delegate tasks, previously carried out by humans, to increasingly "autonomous" automatic systems raise legal and ethical questions. The delegation of decision-making to algorithms does not imply that these decisions can escape human responsibility.

1. The risks arising from artificial intelligence

Algorithms are in fact not always objective and impartial, even if the "objectivity" of automatic systems is often put forward. Users tend to take the result produced by a machine for indisputable truth, as opposed to human judgment always perceived as fallible. However, algorithms include human inputted values.

a. The power relations in the company

Artificial intelligence, in its interaction with "Big Data", ambient intelligence, ubiquitous computing and the "cloud", accentuate the current major change concerning the processing of personal data: never before, such volume of data had been collected from so many individuals, stored in so many places, analyzed and processed. The information environment is thus profoundly modified.

The principles enshrined in the legislation concerning the processing of personal data, may be challenged by the development of artificial intelligence, a major consumer of this data.

In such a context, the question of discrimination acquires particular importance.

b. The risk of discrimination

In a company using artificial intelligence, there is a significant risk of discrimination against certain categories of employees. The risk becomes even greater if one takes into consideration that these discriminations are not apparent. An algorithm-based decision-making procedure may be based on social biases. On the other hand, certain prejudices can be reproduced through customer evaluations, which will subsequently be processed by an algorithm leading to discrimination. Evaluations by customers of the platform can lead, without necessarily being linked to bad will, to unconscious or unacknowledged discrimination that can subsequently lead to the termination of the employment contract. In general, the propensity of algorithms and of artificial intelligence is likely to create or reinforce discrimination, which may raise concerns. It is therefore easy to understand that the management of these questions cannot be left to the discretion of the algorithm and its managers. The search for both traditional and "modern" ways of protecting employees is therefore inevitable.

2. The ways of protection

The ways of protection employees against the misuse of artificial intelligence in the company can derive either from rules of a general nature or from rules concerning more specifically employment relationships.

a. The general provisions

The general rules concerning the protection of personal data are likely to apply to the use of the algorithm. In fact, the algorithm and the artificial intelligence cannot function without personal data. What is then often crucial is not so much the algorithm as such, but the data processed by it. Guaranteeing the right to informational self-determination is therefore the first task.

This means first and foremost ensuring that individuals are always properly informed. The principles affirmed by the legislation on the processing of personal data, such as those of legality and transparency, are very useful. The use of personal data necessary for the operation of algorithms can only be governed by the law relating to the protection of privacy, even if it does not directly target algorithmic processing. Data processing must be adequate and relevant in relation to the purposes for which data was collected. The data must also be accurate and complete. The right of rectification must also be ensured. Data processing systems must finally be transparent.

The General Data Protection Regulation also provides (article 22) that no decision, and therefore no employer decision, can be based exclusively on the automated processing of personal data. Admittedly, algorithms and artificial intelligence have the common objective of automatically executing human tasks or operations, in other words of delegating them to automatic systems. It is, however, necessary to ensure that the recommendation provided by the algorithms would only constitute an aid to decision-making and human action.

But further attention to the internal procedures of the company acquires particular importance in terms of protecting the rights of employees and subsequently their personal data.

b. The rules and procedures specific to the company

When it comes to protecting the rights of employees against risks arising from the use of artificial intelligence, the particularities of the company must be taken into consideration in order to provide satisfactory responses.

Personnel management systems using algorithms must first of all be transparent and understandable. All the operating parameters of the algorithm as well as the person who made his choice must be clear and comprehensible. Article 9 of Greek Law 4961/2022 provides that every company, if it uses an artificial intelligence system, which affects any decision-making process regarding employees or candidate employees, having an impact on working conditions, selection, recruitment or their evaluation shall provide, before its first use, sufficient and clear information to each employee, which includes at least the parameters on which the decision is based, and shall ensure compliance with the principle of equal treatment and non-discrimination in employment.

On the other hand, existing institutions, located either at national and branch level or at company level, can subsequently contribute to compliance with the law concerning the processing of personal data. This is where the role of trade unions can be very useful. The company should be ready to explain to the employee representatives the principles of formation of its decisions, even if these are based on the algorithm. Employees, through their representatives, should know which decisions, concerning their work, are based on their personal data.

In general, the employer, as well as all users of algorithmic systems, are responsible for their operation and the evaluation results issued. They will not be able to invoke the objectivity of the algorithm to avoid control and information. Since behind the algorithm there is always a man who develops it in a certain way, based on determined criteria and on the basis of concrete data provided by the employer, the latter must explain how his algorithm works.

In any event, the employer's decision concerning its employees or collaborators cannot be based exclusively on automated processing of personal data, which means that internal company bodies must be involved.

Conclusion

Artificial intelligence is now a reality put into service inside the company. We cannot oppose such a development. However, the consequences of new technologies are not determined in advance. They must be examined and controlled.

Japan

National Diet Library Contribution to the EPTA Report 2023

Research and Legislative Reference Bureau (RLRB), National Diet Library (NDL), Shinya Nakamura and Hiroko Azuma

Generative Artificial Intelligence – Opportunities, Risks, and Policy Challenges

Description of the System / Problem at Stake

With the rapid spread of generative AI, there is a debate over whether it is acceptable to use it in Japanese education.

The use of generative AI in education has been noted to have some risks of reducing the creativity of learners, copyright infringement, and information leakage. On the other hand, if used appropriately and with attention to such risks, generative AI could be a useful tool to support learning and contribute to improving the efficiency of teachers' work.

This paper introduces some efforts by the government and universities to establish some rules for the use of generative AI in education in Japan. As a related topic, we will briefly discuss generative AI and copyright in Japan.

Societal and Political Relevance and Debate

(1) Situation and Discussion in Primary and Secondary Education

Since April 2023, the Ministry of Education, Culture, Sports, Science and Technology (MEXT) has been studying the issues to formulate some guidelines for the use of generative AI in the educational field, based on interviews with working teachers and others, as well as discussions by an expert panel composed of university faculty members, mayors, teachers, and superintendents of education.

On May 26, the AI Strategy Council, an expert panel that discusses the government's AI policy, made a tentative summary of issues related to AI, including the handling of generative AI in the educational field. The Council raised some concerns that the use of generative AI for students' homework may undermine appropriate evaluation and that the use of generative AI for their essays and reports may reduce their creativity. On the other hand, it was pointed out that the use of generative AI could be advantageous in that it would make it possible to adjust teaching methods according to the level of understanding of each student and to check learning effectiveness in detail by making simple exams, thereby improving educational effectiveness, and reducing the workload of teachers.[47]

On July 4, as a result of the above studies, MEXT released some guidelines for primary, junior high, and high schools regarding the use of generative AI in education (Table 1). These guidelines are tentatively compiled as reference materials for school personnel to assess the appropriateness of the use of generative AI and are not intended to impose a uniform ban or mandate on its use.[48]

---

47. AI Strategy Council, "Tentative Discussion Paper on AI," May 26, 2023, p.12. Cabinet Office Website (in Japanese) <https://www8.cao.go.jp/cstp/ai/ai_senryaku/2kai/ronten.pdf>

48. "Tentative Guidelines for the Use of Generative AI at the Primary and Secondary Education Level," July 4, 2023. Ministry of Education, Culture, Sports, Science and Technology (MEXT) Website (in Japanese) <https://www.mext.go.jp/content/20230704-mxt_shuukyo02-000003278_003.pdf>

Table 1: Main contents of the guidelines presented by MEXT for primary and secondary schools on the use of generative AI

Basic Concept
• At present, it is appropriate to start with limited and trial use in some schools (use for the purpose of determining whether AI is effective in educational activities and learning evaluation, subject to compliance with age restrictions and other terms of use), while verifying effective use in certain situations. • All schools need to enhance education to develop the ability to use information, including the habit of verifying the authenticity of information, and to improve the qualities and abilities required in the age of AI. • Some efforts to promote teacher training and appropriate use of AI in school affairs should be promoted to improve AI literacy and work styles of teachers.

Some examples of inappropriate use	Some examples of appropriate use
• To allow the students to use the AI freely at a stage when they have not fully developed their ability to utilize information. • Submitting works generated by AI as the students' own products in competitions, reports, etc. • To allow easy use from the beginning in situations such as creation of poems and haiku (short form poetry), and expression and appreciation of music, art, etc. • To allow students to use easily before using their textbooks or other quality-assured teaching materials. • To use only the output from the AI when teachers evaluate their students' learning. Etc.	• To use the responses, including errors, produced by the AI, as teaching materials to make students aware of their nature and limitations. • To use the AI to find missing viewpoints and deepen discussion when summarizing thoughts and ideas in a group. • To have students use the AI for English conversation and Japanese language learning for foreign students. • To have students use the AI in the process of revising sentences they have created, in order to learn how to use it. • To use the AI for advanced programming. Etc.

Some experts have pointed out that, with regard to the use of generative AI in primary and secondary education, it is necessary to ensure that teachers have opportunities to learn about generative AI, and to share specific examples of its use according to developmental stages of the students.

(2) Situation and Discussions in Higher Education

Some surveys among university students suggests that generative AI is used to prepare reports and other submissions, to acquire programming skills, to prepare personal statements for job hunting,[49] and to simulate job interviews.

Since March 2023, many universities have issued statements for their students regarding the use of generative AI. While the content of those statements varies, they can generally be divided into two categories: those that call for cautious use and those that encourage appropriate use while taking risks into consideration.

As an example of the former, Sophia University (a large private university in Tokyo) has issued a guideline stating that its students are not allowed to use the sentences, calculation results, etc. created by generative AI in their reports, dissertations, etc., without faculty permission, since they have not produced them by themselves.[50]

---

49. In Japan, it is a well-established practice for companies to recruit students who are scheduled to graduate in batches on an annual basis. While still in school, students submit a profile sheet with a personal statement to the company they wish to work for, undergo an interview, and receive an offer of employment.

50. "Regarding ChatGPT and other AI chatbots (Generative AI)," March 27, 2023. Sophia University Website <https://piloti.sophia.ac.jp/assets/uploads/2023/03/27162218/83431423966a7acb8457a729683dd1d8.pdf>

Nagasaki University (a national university located in the Kyushu region) has also issued guidelines that students are prohibited from appropriating content obtained from generative AI for their reports, dissertations, etc., citing some concerns that the use of generative AI may cause problems in student learning and performance evaluation, as well as risks such as leakage of personal information and copyright infringement.[51]

As an example of the latter, the University of Tokyo, after pointing out the issue of credibility of information and the risk of leakage of confidential and personal information, states that simply banning the use of generative AI will not solve the problem and that its students and faculty should proactively find good usage methods, new technologies, new legal systems, and new social and economic systems.[52]

Waseda University also points out that submitting papers and other documents created by generative AI would be treated as cheating, but states that it is necessary to get familiar with appropriate ways of using generative AI.[53]

In addition to the guidelines for students, some universities have formulated guidelines for faculty members. For example, Tohoku University's guidelines for faculty and staff, based on the premise that it is not realistic to completely eliminate the use of generated AI by students, indicate measures that faculty members can take in classes and for report assignments. The guidelines provide examples of measures such as having students write reports in the classroom and having students give oral explanations about their reports before grading them.[54]

In July 2023, based on the independent actions by many universities, MEXT issued a notice urging each university to take the initiative in formulating guidelines for handling generative AI according to its educational realities and to indicate them to its students and faculty.[55]

An expert in educational technology points out that in future university education, faculty members will be required to make more effort in designing assignments. Emphasizing not only the results of reports and exams but also the learning process, with the focus shifting from simply "what was learned" to "how it was learned."

(3) Discussions Related to Copyright

As pointed out in the above mentioned guidelines, there is a controversy over the use of generative AI in relation to the Copyright Act. The following is a brief explanation of the provisions of the Copyright Act in Japan and its impact on creative activities.

(a) Explanation by Agency for Cultural Affairs

In response to the rapid spread of generative AI, the Agency for Cultural Affairs (an external bureau of MEXT) held an open seminar for the general public in June 2023. At the seminar, the relationship between generative AI and copyright was explained in terms of the development phase and the usage phase of AI (Table 2).[56]

---

51. "The Use of ChatGPT and Other Generative AI," April 25, 2023. Nagasaki University Website (in Japanese) <https://www.nagasaki-u.ac.jp/ja/news/news3948.html>

52. "About Generative AI (ChatGPT, BingAI, Bard, Midjourney, Stable Diffusion, etc.)," April 3, 2023. The University of Tokyo utelecon Website (in Japanese) <https://utelecon.adm.u-tokyo.ac.jp/docs/20230403-generative-ai>

53. "About the Use of Generative Artificial Intelligence (ChatGPT, etc.)," April 21, 2023. Waseda University Website <https://www.waseda.jp/top/en/news/77786>

54. "Notes on the use of ChatGPT and other generative AI (For teachers)," March 31, 2023. Tohoku University Online Class Guide Website (in Japanese) <https://olg.cds.tohoku.ac.jp/forstaff/ai-tools>

55. "Handling of the Teaching and Learning Aspects of Generative AI in Universities and Colleges of Technology (Notification)," July 13, 2023. MEXT Website (in Japanese) <https://www.mext.go.jp/content/20230714-mxt_senmon01-000030762_1.pdf >

56. Copyright Division, Agency for Cultural Affairs, "AI and Copyright," June 2023. Agency for Cultural Affairs Website (in Japanese) <https://www.bunka.go.jp/seisaku/chosakuken/pdf/93903601_01.pdf>

Table 2: Relationship between generative AI and copyright as summarized by the Agency for Cultural Affairs

	Development and learning phases of AI	Content generation phases and generated content utilization phases by AI
Action	• Collect and duplicate copyrighted works as training data and create training datasets • Use datasets for training and develop AI	• Generate images using AI • Upload and publish the generated images • Sell reproductions of the generated images (e.g., illustration collections)
Interpretation of Copyright Act	• In principle, exploitation without the purpose of enjoying the thoughts or sentiments expressed in a work, such as data analysis for AI development, can be performed without the authorization of its copyright owner (Article 30-4 of the Copyright Act). • The above limitation of copyright does not apply to "cases that would unreasonably prejudice the interests of the copyright owner," such as when a commercial database is reproduced for AI learning purposes.	• Infringement of copyright in cases where AI is used to generate images, etc., will be judged in the same way as in normal cases, such as when a person draws a picture without using AI. • Uploading or selling AI-generated works that are found to be similar to or dependent on existing works without the authorization of its copyright owner constitutes infringement of copyright and is subject to claims for damages, injunctions, and criminal penalties.

(b) Background of the Amendment of the Copyright Act

In Japan, there has been a history of promoting legislation suitable for the development of IT. Under the 2018 amendment to the Copyright Act, it is no longer necessary in principle to obtain the authorization of the copyright owner when using copyrighted works in the learning phases of AI.

At the time of the amendment of this act, MEXT's Council for Cultural Affairs held discussions on the premise of technologies such as generative AI, but it is said that the risk of infringement of copyright was not fully communicated to the rights owners. Some rights owners argued that it was unfair to make even commercial services subject to limitation of copyrights, but this was not reflected in the amendment.

(c) Discussion

There are similar provisions in the EU and the UK as well as in Japan to allow data analysis for AI development without the authorization of the copyright owner. While the EU Copyright Directive requires academic research purposes for data analysis and the UK Copyright Act requires non-commercial purposes, the Japanese Copyright Act allows data analysis for a wide range of purposes, including commercial purposes. In addition, the Japanese Copyright Act does not allow copyright owners to refuse the collection and analysis of copyrighted works (opt-out). For these reasons, some Japanese creators have expressed their concern.

In the "Intellectual Property Strategic Program 2023" decided on June 9, 2023, the government revised its previous policy direction, which had focused on AI promotion. The government has not clearly indicated which cases would fall under the category of "cases unreasonably prejudicial to the interests of copyright owners," but it will clarify the issue in line with specific cases, while paying attention to the protection of the rights of creators.[57]

---

57. Intellectual Property Strategy Headquarters, "Intellectual Property Strategic Program 2022," June 9, 2023 (in Japanese) <https://www.kantei.go.jp/jp/singi/titeki2/kettei/chizaikeikaku_kouteihyo2023.pdf>

3. Role of TA in the Debates

The Research and Legislative Reference Bureau (RLRB) of the National Diet Library, Japan (NDL) published a research report entitled "Perspectives on Artificial Intelligence/Robotics and Work/Employment"[58] that was intended to support Diet deliberations. This research report was commissioned to the University of Tokyo, and a research team (composed of 23 experts of varied specialties and affiliations) investigated the state of adoption of AI and robotics, the state of investigative research, and the issues that have arisen, to contribute to the discussion of employment and labor in a future AI and robotic society. This research report was provided directly to Diet members and is also publicly available on the NDL website.

At the time of research, Japan was often said to be optimistic about introducing artificial AI and robots to the workplace. Japan expected that AI and robots would solve the problem of the labor shortage caused by the declining birthrate and aging population. In contrast, other countries were more concerned about the possibility of machines taking over jobs currently performed by human beings.

Around the same time, to provide ethical guidelines for society and government in the use of AI, the Japanese government published "Social Principles of Human-Centric AI" in March 2019.[59] The chairperson of the council of experts involved in its creation recalls that at that time AI was not at the stage where it could threaten human creativity or pose a threat to society at large, even though it could provide superior performance for limited applications.

The situation has changed dramatically since last year with the release of ChatGPT and other generative AI. There are concerns that generative AI may threaten white-collar jobs in general, including university faculty, which require integrated knowledge and intelligence, and the threat to democracy from fake news produced by generative AI is also increasing. New measures and discussions are needed to balance the risks while taking advantage of the potential of generative AI. As a basis for such discussions, it is also important to educate people on how to use generative AI and to develop human resources to create them. The RLRB will also conduct necessary research and provide information to support Diet deliberations to address new situations brought about by advances in generative AI technology.

---

58. "Perspectives on Artificial Intelligence/Robotics and Work/Employment," Research Materials, 2017-5, Tokyo: Research and Legislative Reference Bureau, National Diet Library, 2018. (in Japanese) <https://dl.ndl.go.jp/pid/11065181> The English version translated by the author is available at <http://hdl.handle.net/10367/11072>

59. The Integrated Innovation Strategy Promotion Council "Social Principles of Human-Centric AI" <https://www.cas.go.jp/jp/seisaku/jinkouchinou/pdf/humancentricai.pdf>

Lithuania

Generative artificial intelligence (AI) and work

One of the issues widely discussed among researchers is whether generative AI will lead to mass unemployment by displacing most people's jobs and, furthermore, what will be the effects of generative AI in terms of social inequality. Is it possible to foresee a huge substitution of qualified workers for AI algorithms? Can we expect an equivalent creation of new jobs? Generative AI will most likely change the nature of our work and the way we work. Which tools do we have in place to address this question?

Description of the System/Problem at stake

What is it about?

Generative artificial intelligence (AI) is a subtype of AI responsible for creating new and original content. The term "generative" refers to the fact that these tools can identify patterns in massive data sets and generate new content, an ability that has often been considered as uniquely human.[60]

Generative AI is a rapidly developing technology that uses trained models to generate original content in a variety of forms, from written text to video, music, and even software code. Generative AI has a huge potential for business, especially in marketing, sales and services. Advanced models such as *ChatGPT*, *DALL-E*, and *Midjourney* are driving the possibilities of generative AI. Here is what they do:

ChatGPT enables chatbots and virtual assistants to provide human-like text to personalise experiences and automate conversations.

DALL-E generates realistic images from textual descriptions, useful for industries such as e-commerce, merchandising, advertising and design.

Midjourney facilitates data generation and manipulation for data augmentation and content creation.[61]

Using large language models and huge data sets, technology can instantly create unique content that is virtually indistinguishable from human work and in many cases more accurate and engaging. In practice, artificial intelligence is mainly represented by machine learning, and specifically by neural networks. They simulate logic using statistics and mathematics based on massive amounts of data. The recent increase in interest in artificial intelligence is due to the great leap in the development of generative AI and its capabilities and, most importantly, its accessibility to a wide audience, which has already become an integral part of public life.

As with any new technology, so with generative AI, there has been much speculation, uncertainty, and concern that such a tool could exacerbate social inequality or create new forms of social stratification by eliminating jobs that are normally done by humans. Indeed, along with the undoubted positive benefits of AI technologies for the social and economic progress of mankind, it is also necessary to assess the risks and disadvantages of generative AI and how it can affect the lives of people and society as a whole in the future.

What is the state of play in your country (development, deployment, use)?

As technological advances change the landscape of modern work, generative AI has become a transformative tool with unparalleled potential. There is no doubt that the fields of both AI and generative AI will

---

60. Kweilin Ellingrud; Saurabh Sanghvi; G the urn Singh Dandona, Anu Madgavkar, Michael Chui, Olivia White, Paige Hasebe. GENERATIVE AI AND THE FUTURE OF WORK IN AMERICA. July 26, 2023 | Report https://www.mckinsey.com/mgi/our-research/generative-ai-and-the-future-of-work-in-america?stcr=D2B0E2A3B2674FEDB1394DB-1C8BBE552&cid=other-eml-alt-mip-mck&hlkid=90da8f0dc9cd4d5f822b2888cbc44006&hctky=_12636281&hd-pid=8f9b6ed9-12c0-4de5-b2cd-85de357d05f4

61. 7 Things You Need to Know About Generative AI - A Practical Guide for Business Leaders, https://www.avaus.com/blog/7-things-you-need-to-know-about-generative-ai-a-practical-guide-for-business-leaders/.

occupy an increasing share of the Lithuanian economy in the near future, so it is important to note that no profession is immune from automation or workplace transformation, as these technologies will be constantly improved. Therefore, we would assume that most professions will remain, but the requirements for employee competencies will change. Having the tools that realize the potential of generative AI will soon begin to gradually move from the "will be an advantage" category to the "required skill" category. Just as many jobs today require strong computer skills and familiarity with at least basic work methods in popular office packages, employers in the future will expect their employees to be adept at using generative AI functionality. Thus, the use of both AI and its subtypes in every human activity is gradually becoming the norm and a necessity in the competitive struggle.

The main directions of Lithuanian artificial intelligence research include machine learning; ontology engineering; semantic modelling of knowledge and database systems; conceptual modelling in law; data mining and knowledge discovery; artificial neural networks; data identification; applied mathematical logic, etc. [62] In order to develop the AI market in Lithuania and to strengthen the country's research and innovation ecosystem on a national and international scale, the Kaunas University of Technology (KTU) and its partners are creating a joint AI excellence centre for sustainable development. The centre is expected to increase the opportunities to digitise the Lithuanian production, health, transport and energy sectors, ensure the development of research, experimental development and innovation (R&D) by applying solutions based on AI technologies.[63]

Today, the leading industries in the implementation of AI technologies in Lithuania are the financial sector, information and communication technologies, and biotechnology. AI solutions are also successfully implemented in the energy, transport, and healthcare sectors. Many industries, such as services in the construction industry, are still conservative in terms of digitisation. In these areas, the implementation rates of AI solutions are still low, but the opportunities for technological development there are quite serious.

These are some areas in Lithuania, where these technologies could either create new jobs or replace the existing ones. In current conditions, taking into account the rapid development of generative AI, it is necessary to strengthen the implementation of applied AI solutions in the real sector in Lithuania and accelerate the digital transformation of priority economic sectors. However, the main factor to note is that the impact of these technologies on the Lithuanian labour market will depend on many aspects, including technological progress, regulation and public interest and acceptance.

What do your MPs think about it?

The development of artificial intelligence and its subtypes has the potential to change the economic, technological, and political model of society. It is therefore important for legislators to foresee how these technologies will affect citizens and how to make these changes more convenient and less risky and ensure accountability for them.

Parliamentarians believe that today there is a real threat related to the transformation of the labour market and digital unemployment, because, for example, the technological transition makes some professions disappear and there is a risk of bias and inaccuracy in employment.

The introduction of generative AI is believed to have a different impact on men and women, as women's employment is more than twice as susceptible to automation, and the nature of the work itself is classified as a "white collar" activity. Therefore, despite the generative potential of artificial intelligence, there are still many gaps in these technologies that need to be addressed, because these systems are trained to use large data sets that may contain biased or incorrect information. Given the progress of creative AI in art creation, there will always be questions about the existence of intellectual property rights, whether copyrights, patents, etc.

---

62. Dr. Tomas Lavišius; Giedrius Kanapka. Research leader Dr. Alvidas Lukošaitis. ACCENTS AND DEVELOPMENT OF THE LEGAL ENVIRONMENT OF ARTIFICIAL INTELLIGENCE. Trends, risks, recommendations, decisions in the European Union, some European Union member states, other countries and international organizations. 09/01/2023. Research Unit and the Information and Communication Department of the Office the Seimas of the Republic of Lithuania. The study was commissioned by the Committee for the Future of the Seimas.

63. An artificial intelligence center of excellence is being created in Lithuania. https://www.15min.lt/verslas/naujiena/mokslas-it/lietuvoje-kuriamas-dirbtinio-intelekot-ekscelencijos-centras-1290-2005470?utm_medium=copied.

While AI promises efficiency, human oversight is still critical to reduce bias. There should be the primary adjustment filter when using any level of AI systems. Here, it is necessary to find a reasonable balance between the activities of generative AI and human activities, realising the goals of AI in order to simplify and optimise social, administrative or business processes.

Although artificial intelligence can completely or partially replace humans in various jobs, it also offers great opportunities to create new ones. However, there must be a call for employers to think about employees first when preparing to implement AI. You want people to know in advance what is ahead. In addition, employers should also facilitate all the conditions for retraining their employees. Moreover, if something is already done, it must be done transparently and clearly, in coordination with employees and employee representatives.

How generative artificial intelligence can help the work of the Seimas of the Republic of Lithuania (the co-author of this answer was *ChatGPT*. *ChatGPT* did not write the entire answer, but it was responsible for managing the creative block, generating individual sentences and paragraphs of the text, and considering various cases of generative AI use. Writing this query using *ChatGPT* is a welcome sign of human-computer interaction that may become the new norm)

Generative AI can assist the Seimas of the Republic of Lithuania in its work in various ways, improving processes, efficiency and decision-making. Here are some specific ways in which AI can be applied in the Lithuanian Parliament:

Data analysis and forecasting: AI can be used to analyse huge data sets to isolate trends and identify critical issues. This would help Members of the Seimas (MPs) to better understand the current events and make informed decisions.

Decision support: AI systems can provide MPs with information and recommendations, taking into account different political, economic and social alternatives. This could help MPs make the best decisions.

Automation of systems: AI can be used to automate administrative processes of the Seimas, such as document management, translation, organization of meetings and other routine tasks, thus saving time and human resources.

Translation and communication: AI-based translation technologies can facilitate communication with foreign countries, allowing MPs to communicate with their colleagues from other countries and follow international affairs.

Citizen service: Websites and applications can use language robots or chatbots based on AI systems to provide citizens with information about the activities of the Seimas, parliamentary processes and legislation.

Security: AI can be applied to monitor the security of the Seimas information system and detect potential security breaches.

Analysis and reports: AI can generate detailed information and reports on the activities of the Seimas, parliamentary decisions and their impact on society, facilitating the evaluation of the decision-making process.

Visualisations and information systems: AI can create interactive digital visualisations that allow MPs to more easily understand complex information and find the necessary information.

Legislative language analysis: AI can be used to analyse legislation and its language to ensure that it is clear and understandable to citizens.

Traceability of decisions: The use of the blockchain technology, which can also be supported by AI, can ensure traceability and transparency of decisions.

It is important to note that the application of generative AI in parliamentary work should be carried out responsibly, taking into account ethical, privacy and security aspects. It is also important to ensure that MPs and staff are properly trained and understand how to use these technologies effectively, productively and safely.

Who are the key stakeholders?

Artificial intelligence is perhaps the only large field of technology that today already covers the entire classification of economic activities, from agriculture to public administration. Therefore, to ensure that AI is aligned with human values and goals, it is necessary to understand the needs and expectations of the various stakeholders who are affected by or involved in AI and its prototype projects.

One of the challenges in identifying generative AI stakeholders is that, given the scale of social consequences covered and the economic areas of involvement, most, if not all, people (society and users), organizations (scientific and educational, business, non-governmental etc.) and political institutions (national parliaments and governments, international) are stakeholders.

All these stakeholders create an ecosystem for the development and application of generative AI, where they all collaborate to develop and use this technology, taking into account scientific, business, ethical and legal aspects, respectively.

Why is it important for your country?

Most importantly, generative AI has the potential to radically change the way our society and its individual members work, learn, communicate and create. However, along with its great benefits, it also poses ethical, security and legal challenges. It is therefore important to closely monitor the development and use of AI to ensure that the technology is applied responsibly and ethically.

Social and political relevance and debate

Generative AI has become an important topic in public and political discourse due to its influence on various areas of public and state life. With every new technological breakthrough there is fear that machines will replace humans. Historically, when machines were used to replace workers to increase productivity and reduce costs, a set of new jobs, which did not exist before, was created. Therefore, the extent to which new technologies will seriously disrupt the labour market will, in many cases, depend on the development of technology, government regulation and societal priorities. To solve this challenge, we need to implement the following important measures:

Training and education: Investments in education and training are necessary to enable people to adapt to new working conditions and technologies. Training must be continuous and focused on developing the skills needed in the future labour market.

Labour market regulation: An important factor in the impact of artificial intelligence on the labour market will undoubtedly be state regulatory policy. It is clear that any serious threat leading to a significant and long-term increase in social tensions related to unemployment, long-term effects on jobs and human capital must be managed by the state. However, we must also recognise that, over the nine months since generative AI appeared in society, no drastic changes have taken place in the labour market, although the creators of generative AI are now also encroaching on "human" creativity.

The rapid growth of generative AI also demonstrates the general public's acceptance of and trust in AI technologies. The fact that so many users have embraced the AI-enabled chatbot so quickly shows that the integration of artificial intelligence into everyday life is getting better and easier. This trend is a positive indicator for the future of generative AI applications, as it shows willingness to explore and use AI solutions in various fields.

These questions and topics show that generative AI has a significant impact on the labour market and requires a balanced approach that includes social justice, worker protection and economic development. Discussions on these issues are increasingly taking place in society, among political leaders and business representatives, in order to find a balance between the benefits that generative AI can bring and the challenges it poses to the labour market.

Is there ongoing debate on the impact of generative AI on our societies and democracies?

Predicting exactly how generative AI will affect democracy is a very difficult task, given that its potential applications are largely unknown and seemingly limitless. This technology is not just another application; rather, it is a fundamental technology, the emergence of which is more like the emergence of the Internet itself. It can therefore affect democracy both directly, such as by changing electoral and governance mechanisms, and indirectly, by threatening to change the foundations of information ecosystems.

The benefits and risks of using generative AI in the context of democracy and to develop policies and legislation that ensure the accountability and safety of this technology to society.

In a democratic process, it is important to ensure that the public has the opportunity to participate in decision-making related to generative AI. This can include public debates, polls, engagement and other means

that allow citizens to express their views. Too much regulation can hinder the development of technology and destroy its untapped potential. One of the most important factors to be addressed is the development of new improved benchmarks, standardised and objective classifiers for evaluating the performance of generative AI models and the quality of datasets.

Are there results from surveys among the population concerning usage of or opinions about generative AI?

There are none.

Is there any legislation in place?

There is also no specific regulation in this direction in Lithuania. Therefore, the problems of both simple and generative use of AI are still in the "grey zone". Of course, there are individual initiatives by researchers and politicians, so it is expected that a certain state approach to artificial intelligence and its subtypes should be developed in the near future, as controversial and significant legal issues arise increasingly frequently in the development, application or use of this product.

In addition, one should focus on the application of the EU AI Act, which will soon be adopted. The AI Act provides for a whole series of exceptions that are intended for national states, and they should be used wisely (e.g. sandbox mode, exceptions for small-scale providers and users of AI (Art. 55) etc.).

Are there currently political or legislative proposals on these topics?

As with most phenomena in the digital economy, the legal nature of artificial intelligence and its subtypes is not obvious. Therefore, the development of AI dictates the need to find new solutions for the legal regulation of technology. Among the relevant topics of legal regulation arising from the use of AI technology, the following are distinguished:
- data privacy;
- security and liability;
- big data technology;
- intellectual property;
- ethics, etc.

What can be said about science / evidence-based inputs (bodies involved, comprehensive or specialized approaches, etc.) and democratic inputs (public participation, transparency, etc.) guiding political decision-making on this topic?

Today, there are discussions about problems that have gone far beyond the boundaries of technology itself and are related to the development of the entire human community. Technologies in the field of artificial intelligence are emerging faster than society has time to understand them. Therefore, science/evidence-based inputs and democratic inputs on the topic of generative AI are two important aspects that should guide policy makers. These two distinct but complementary principles would help ensure that generative AI technologies are used ethically, responsibly and for the benefit of society.

The role of AI in the debates

Has your institute taken up this topic?

This topic was touched upon in the last spring session of the Seimas of the Republic of Lithuania. We started discussions with experts regarding the possible risks of using generative AI tools, their certification, etc. We discussed whether there is a need for national regulatory measures, which could ensure that the AI systems used comply with the law, ethics, and are safe and generally reliable.

One important questions was who would to blame if generative AI produced the results that would cause legally significant consequences. At the Committee meeting, the unanimous conclusion was that policymakers, scientists and industry leaders should work together to develop policies and measures to ensure

the responsible use of generative AI. All we know for sure right now is that artificial intelligence is not going anywhere, and people will have to increasingly rely on it over time. Second, we can rely on the fashionable saying that AI is not going to replace humans, but humans with AI are going to replace humans without AI. Therefore, those who will be able to not only use such tools, but also to develop them, will benefit the most.

Has AI made an impact on the ongoing debates?

With big technological changes, it is important to understand that we are really talking about such big, fundamental changes in AI technologies, which change not only one specific industry, but in fact change the organisation or existence of many sectors, in general. It is similar to the introduction of electricity into the economy, and in this sense has much in common with the great technological changes that have taken place in the past, united by their scale and the fact that they eventually encompass almost all human activities.

It is necessary for everyone, including the expert community, the state, the education system, and regulatory authorities to take measures to prepare society for the use of artificial intelligence. The expert community and business need to reflect on the applicable scheme of implementing AI in everyday life. Philosophers must convey to the public their reasoning about the coexistence of humans and AI. The education system must be reformed in such a way to teach children knowledge and skills that are needed not only for the present, but also for the near and distant future, as well as to provide lifelong learning opportunities for people. It would be the duty of the state to take care of the safety and protection of the population from AI-related risks and dangers.

What are the lessons learned from AI?

Today, it is very important to consider various mechanisms of regulation and governance in the field of generative AI, as the technological revolution has moved towards an intellectual one. So far, any new technology has aimed at freeing man from monotonous physical work, while neural networks allow them to be freed from tedious mental work.

Netherlands

Generative Artificial Intelligence: Opportunities, Risks, and Policy Challenges

Country Report Chapter The Netherlands: Generative AI and Work

Description of the system/problem at stake

What is it about?

Generative Artificial Intelligence has been in the spotlight since the introduction of applications such as ChatGPT, Bard GitHub, Copilot and DALL-E.

Generative Artificial intelligence refers to systems that can automatically generate content, such as text, or images, for example an application letter, a social media message or branded online content. As such, generative AI can perform various non-routine tasks,[64] whereas earlier technologies were mostly able to automate routine tasks.[65] This versatile quality is both a source for high expectations and deep concerns for the future labour market. However, at the moment studies reporting empirical effects are as yet, and understandably, limited.

In previous robotisation and automation debates, three categories of potential impact on work could be distinguished.[66] Roughly speaking, experts expect these same types of impacts for generative AI, but the order of magnitude remains unknown.

- Quantity of work: new technologies can cause tasks and jobs to disappear, possibly resulting in (mass) unemployment, or job polarisation where specific groups are particularly affected.[67] On the other hand, as new tasks or needs emerge, new technology has so far also created new tasks and jobs.
- Quality of work: new technologies can erode jobs, resulting in stressful, dehumanising, mind-numbing jobs. On the other hand, new technology can also be used to automate repetitive, boring or dangerous tasks, leaving workers with time for other tasks.
- Competencies and skills needed: the above two impacts show that new technologies often bring about a different role for people in the labour process, for which they have to acquire specific competencies and skills. The question is whether everyone will eventually be able to keep up in that transition.

What is the state of play in your country (development, deployment, use)?

The Netherlands has a strong position in certain AI scientific disciplines,[68] but has no companies that develop large language models or generative AI systems themselves – like many other countries in Europe.

There are no official surveys to indicate user uptake. On social media, blogs etc Dutch users report enthusiastically the ways in which they incorporate the use of generative AI in their daily work. ChatGPT has 100 million users o worldwide, but it is not known how many users are from Europe or the Netherlands.

---

64. Brynjolfsson, Li en Raymond (2023)

65. Autor, David H. 2015. „Why Are There Still So Many Jobs? The History and Future of Workplace Automation." Journal of Economic Perspectives, 29 (3): 3-30.

66. In general, a reorganization of work processes and organisations is required before the functionalities of new technologies can be incorporated (Rathenau Instituut 2015)

67. OECD (2019) Job polarisation and the middle class. New evidence on the changing relationship between skill levels and household income levels from 18 OECD countries. https://doi.org/10.1787/4bf722db-en

68. Rathenau Instituut (2021)

What do your MPs think about it?

With generative AI, they think of both opportunities and risks. They differ on the relationship between generative AI and democracy. One says this technology lowers the threshold for spreading false information. The other stresses that in their view, the technology itself is not a danger, but its misuse is.

They think generative AI will have an impact in all policy areas. As for the impact on work, they expect a possible better match between supply and demand, loss of 'old' jobs and demand for new jobs.

Who are the key stakeholders?

Employers and workers (both employees and self-employed/entrepreneurs), often represented by trade organisations and unions, educational institutions for the (re)training of workers, the Ministry of Social Affairs and Employment, the Ministry of Education, Culture and Science and the Ministry of Economic Affairs traditionally play an important role in debates about risks and opportunities of technology on the labour market.

The Dutch Social and Economic Council (SER) is a key advisor to the government and is expected to be asked by the government to provide advice on the impact of generative AI for the economy and labour market in the Netherlands, as part of the government strategy currently being developed. The government strategy is expected to be published by the end of 2023.

Why is this important for your country?

Technological breakthroughs have undeniably affected and will continue to affect our work. Moreover, there is currently a major labour market shortage in the Dutch labour market. In the public sector, the problem of staff shortages is acute, and expected to remain a problem for the coming years. The Dutch Social and Economic Council recommended to organize work more 'smart' and in technically innovative ways to save on labour.[69] As such, Organisations are looking for innovative solutions to solve staff shortages, including the possibilities of generative AI.

Societal and political relevance and debate

Is there ongoing debate on the impact of these new technologies on our societies and democracies?

Yes, some parliamentarians have put generative AI on the political agenda. They specifically raised concerns about privacy and copyright in relation to AI, and asked the government to formulate a vision for generative AI.[70]

In Dutch newspapers there is an active debate amongst scholars and stakeholders about the potential opportunities and risks of generative AI. A petition, initiated by, amongst others, two former parliamentarians who focused on digitalization, containing an urgent plea for action, has been signed by numerous former politicians and opinion leaders. These debates and pleas include potential impact on the labour market, but address a wider variety of societal and ethical issues.

Is there any legislation in place?

Currently there is no specific law relating to generative AI and the labour market. In Europe, the AI Act is being negotiated with the European Parliament, European Council and European Commission, to see how

---

69. 'Valuable work: public service under pressure' SER (2023), Waardevol werk: publieke dienstverlening onder druk - Oplossingsrichtingen voor de arbeidsmarktkrapte (ser.nl)

70. Amended motion by members Dekker-Abdulaziz and Rajkowski on a comprehensive vision for new AI products (pertaining to 26643-998) (4 april 2023), Reply to questions by Member Leijten on the concerns of hundreds of tech pundits about developments in artificial intelligence (Tweede Kamer, 27 juni 2023), Motion by members Dekker-Abdulaziz and Rajkowski on requiring watermarks for AI-generated texts and images where copyright sources have been used (Tweede Kamer, 6 juli 2023)

generative AI can be included in this new law. Furthermore, existing legal frameworks regarding data protection, non-discrimination, liability, copyright and intellectual property, cover – or may cover – issues regarding generative AI. For example, to what extent are developers of generative AI (specifically large language models) allowed to use the creative works of many writers, artists et cetera?

Furthermore, specific laws pertaining to employers and employees, such as the Dutch Works Council Act and 'good employment practices' in the Dutch Civil Code, may cover some issues related to the introduction and use of Generative AI systems on the workfloor.

Are there currently political or legislative proposals on these topics?

The Dutch government is working on a vision for generative AI, at the request of the House of Representatives.[71] "This sets out, among other things, what potential actions the Dutch government can take to ensure that this technology is embedded in society in a responsible manner and conducts a further analysis of copyright issues related to generative AI."[72]

The European AI Act, currently being negotiated, will also be applicable to the Netherlands, once in force.

What can be said about science / evidence-based inputs?

The public launch of chatGPT and other generative AI systems is still very recent, so there are few empirical studies on the impact on jobs and labour markets. Some initial case studies have been conducted, but not specifically for the Dutch labour market.

Generative AI is capable of creating entirely new content such as text, image or sound. For example, it can be used to produce news reports, articles, advertising, summaries, recipes, computer code or even entire music tracks. These applications would make the work of journalists, scriptwriters, marketers, programmers and musicians much easier, if not redundant. The same goes for other professions.

Generative AI can be used to support human work. A recent US study shows that workers are using it to generate ideas, as well as for communicative purposes such as writing draft messages and e-mails, reports, summaries, or to extract errors from text or provide suggestions for improvement.[73] Another study showed that generative AI is already being used to provide real-time call suggestions and possible answers to customer service agents.[74] Not only did this lead to an increase in efficiency, it also made conversations run more smoothly and employees experienced less social and emotional strain.

Expectations around generative AI are currently tempered by the fact that generated content is often still susceptible to make mistakes and 'hallucination'.[75] Technology produces information that appears coherent but in reality is inaccurate and, moreover, often incorrect. Verification of the output and possible editing by a human is, therefore, often still required.

The following opportunities emerge from the available studies: [76]
• Efficiency and productivity gains;
• Strengthen autonomy and increase productivity of low-skilled workers;
• Positive impact quality of work;
• Solution to labour market tightness.

---

71. Amended motion by members Dekker-Abdulaziz and Rajkowski on a comprehensive vision for new AI products (pertaining to 26643-998) (4 april 2023)

72. Reply to questions by Member Van Raan on curbing ChatGPT due to privacy concerns (Tweede Kamer, 21 augustus 2023). See also: 'Regulating algorithms' compilation letter (Tweede Kamer, 7 juli 2023)

73. Cardon et al. (2023)

74. Brynjolffson et al. (2023)

75. Heikkilä (2023) "Why You Shouldn't Trust AI Search Engines." MIT Technology Review; Peng et al. (2023) Check your facts and try again: improving Large Language Models with external knowledge and automated feedback.

76. o.a. Brynjolfsson et al. (2023), Noy en Zhang (2023) Cardon et al. (2023) en Alshurafat (2023)

Possible risks:[77]
• Loss of jobs and/or employment;
• Mismatch between supply and demand of competencies and skills;
• Vulnerable groups falling by the wayside;
• Negative impact on quality of work.

Role of TA in the debates

Has your institute taken up this topic?

The Rathenau Instituut contributes to the vision of generative AI that the Dutch government is currently developing. We do this by exploring the nature of the technology's development, identifying promising applications and risks in relation to public values, and formulating policy options on that basis.

Previously, the Rathenau Instituut published on the impact of technology on work and the labour market:

• *Own rhythm of algorithm? Algorithmic management beyond the platform economy* (Rathenau Instituut and TNO, forthcoming fall 2023)
• *Valued at work. Limits to digital monitoring at the workplace using data, algorithms and AI* (Rathenau Instituut, 2020)
• *Robotisering en automatisering op de werkvloer. Bedrijfskeuzes bij technologische innovaties* (Rathenau Instituut, 2018)
• *Werken aan de robotsamenleving. Visies en inzichten uit de wetenschap over de relatie technologie en werkelijkheid* (Rathenau Instituut, 2015) – requested by the Dutch Parliament

Has TA made an impact on the ongoing debates?

In the public debate, much of the current focus is on possible job losses. However, there is not yet empirical evidence for this. In the past, there have been vulnerable groups whose jobs were more susceptible to job loss, and who struggle to acquire new skills for the new type of jobs technology creates. Whilst new technologies have been beneficial for society as a whole, these groups required extra attention from policy makers regarding social security.

Our studies show that new technology can also have a significant impact on the quality of work, and the competences that are asked. With our previous studies, we were able to show that more left wing oriented political parties that robotisation and automation also create opportunities for the labour market and economy, and make right wing oriented parties more alert on vulnerable groups in the labour market. We also asked attention for other societal and ethical issues related to robotization and automatization, beyond impacts on jobs.

What are the lessons learned from TA?

Generative AI is a fast-developing technology and its direction and functions are difficult to predict. It is, therefore, important that research is conducted on how it changes jobs and the labour market.

TA can broaden the debate on generative AI and work. TA shows that the potential impact of generative AI on the labour market is not unambiguous, but varied and complex, with both positive and negative possible outcomes. In addition to focusing on job quality, it is important to invest in teaching employees new skills or even providing an opportunity to change the field they work in.

---

77. Gmyrek et al. 2023; Hatzius et al. 2023

References

Selection of articles

Alshurafat, H. (2023) The usefulness and challenges of chatbots for accounting professionals: Application on ChatGPT. SRRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4345921

Brynjolfsson, E., Li, D., & Raymond, L. R. (2023). Generative AI at work (No. w31161). National Bureau of Economic Research.

Cardon, P. W., Getchell, K., Carradini, S., Fleischmann, C., & Stapp, J. (2023). Generative AI in the Workplace: Employee Perspectives of ChatGPT Benefits and Organizational Policies.

Chohan, Usman W. (2023) "Generative AI, ChatGPT, and the Future of Jobs." SSRN Scholarly Paper. Rochester, NY, March 29, 2023. https://doi.org/10.2139/ssrn.4411068.

Davenport, T.H. & Mittal, N. (2022) How generative AI is changing creative work. Harvard Business Review. https://hbr.org/2022/11/how-generative-ai-is-changing-creative-work

Eloundou, Tyna, Sam Manning, Pamela Mishkin, and Daniel Rock (2023) "GPTs Are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models." arXiv, March 23. 2023. https://doi.org/10.48550/arXiv.2303.10130.

Gmyrek, P., Berg, J., Bescond, D. 2023. Generative AI and Jobs: A global analysis of potential effects on job quantity and quality, ILO Working Paper 96 (Geneva, ILO). https://doi.org/10.54394/ FHEM8239

Hatzius, J., J. Briggs, D. Kodnani, F. Pierdomenico, (2023) The potentially large effects of artificial intelligence on economic growth. Goldman Sachs.

Noy, Shakked and Zhang, Whitney (2023) Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence (March 1, 2023). Available at SSRN: https://ssrn.com/abstract=4375283 or http://dx.doi.org/10.2139/ssrn.4375283

Zarifhonarvar, Ali (2023) Economics of ChatGPT: A Labor Market View on the Occupational Impact of Artificial Intelligence, ZBW - Leibniz Information Centre for Economics, Kiel, Hamburg

Other literature

https://www.nytimes.com/2023/05/20/business/dealbook/the-optimists-guide-to-artificial-intelligence-and-work.html

https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/the-economic-potential-of-generative-AI-the-next-productivity-frontier#introduction

https://www.axelera.ai/how-will-generative-ai-revolutionize-our-work/

https://www.forbes.com/sites/joemckendrick/2023/07/05/generative-ai-adds-new-dimension-to-productivity-that-were-just-beginning-to-understand/?sh=2a087e9964cb

https://www.businessinsider.com/chatgpt-jobs-at-risk-replacement-artificial-intelligence-ai-labor-trends-2023-02?international=true&r=US&IR=T

https://www.brookings.edu/articles/will-generative-ai-kill-jobs/

Norway

Norway – Norwegian Board of Technology (Teknologirådet)
Key developments concerning generative AI and democracy in Norway

What is it about?

The Norwegian democracy is of high quality and is characterized by high levels of societal trust and transparency. However, the emergence of generative artificial intelligence (AI) has the potential to challenge important aspects of democracy:

- Access to reliable and true information
- Trust in public institutions, such as media broadcasters
- Individual autonomy
- Democratic governance of AI technology
- Trustworthy public services

Therefore, it will be important to both seize the opportunities that lie in the technological advancements, but also establish mechanisms to prevent misuse and build resilience.

Generative AI: What's at stake for Norway as a democracy?

The widespread use of generative AI tools can impact and challenge many core functions of democratic societies:

Access to reliable information

In democratic societies, the ability to access information is crucial for people to make informed decisions and express themselves. Internet has become a fundamental infrastructure for modern democracies, and a prerequisite for freedom of expression and information. Now, the amount of AI generated content online is rapidly increasing. Generative AI tools can create photorealistic images, clone or create synthetic voice audio and write text convincingly human-like. Media content can be mass produced easier, cheaper and in a much larger scale than before. The technology is accessible to anyone, including actors with bad intentions. The new AI tools can be used to spread disinformation, fake news and deepfakes. The quality and credibility of AI generated content is also increasing, making it difficult to establish whether a news article, an online user or an image is AI generated or not.

Telling truth from lies

Telling truth from lies can be increasingly difficult in a digital information ecosystem flooded by AI generated content. Generative AI tools can be used to produce large amounts of mis- and disinformation, fake news and deepfakes online, that can threaten people's access to truthful information, confuse and influence the public, increase political polarisation and even impact democratic elections. Misleading content can be also spread unintentionally. Since large language models are not trained to tell the truth, they can *hallucinate* and generate false content that appears to be true.

Trust in media

Norwegian citizens have a high level of trust in public institutions, especially in comparison to other countries.[78] Yet, the least trusted public institution in Norway, according to OECD Trust Survey of 2022,

---

78. Norway is among the OECD countries with the highest levels of public trust – 77 % of the population reported trust in the government, compared to the OECD average of 47 %.

is the media.[79] Among the factors contributing to the lack of trust was social media, concerns about fake news and political communication styles. Generative AI has the potential to flood the information environment with synthetic content, and further erode trust. Making sure people can access a variety of trustworthy media broadcasters will continue to be an important measure to counter the spread of mis- and disinformation online, in addition to building resilience among the population through critical thinking and source criticism.

Individual autonomy

OpenAI's ChatGPT and My AI on Snapchat represent a shift towards information exchange in private spheres, through interactions between humans and chatbots. The public has no way of accessing information about the interactions – neither what information is shared, nor how significant the impact of these interactions might be on people's opinions and perceptions. Research shows that young people are more inclined to trust information provided by chatbots, than humans.[80]

Another significant challenge is the embedded bias in large language models. In theory, ChatGPT should be politically neutral. However, the chatbot responds in line with liberal political values, according to recent research.[81] As a result of bias in AI systems, people's and the public's opinion can be influenced, or even manipulated, through interactions with the new AI tools. An experiment involving 1500 participants showed that the participants' attitudes were subjected to hidden influence (referred to as *latent persuasion*) through interactions with pre-programmed chatbots.[82]

Power concentration and democratic governance

The development and deployment of generative AI tools is today led by private American technology companies, such as OpenAI, Google and Microsoft, to mention some. These companies have insight into a technology with the potential to impact society. Yet, the public, academic researchers and smaller companies only have access to the information that the AI companies choose to disclose. This can undermine a fair competition, distort the public's ability to innovate with the technology and weaken public oversight and control mechanisms. The need for a national large language model is currently being discussed in Norway (see discussion below).

Trustworthy public services of high quality

Generative AI tools can provide society with new opportunities to enhance democratic participation and governance. For instance, governments and public sector institutions can use the new tools to ensure more inclusive, efficient and accessible services. One example is the use of synthetic speech technology to ensure better public information and services for people with visual impairment.

According to the Norwegian Agency for Public and Financial Management, public trust is interlinked with the authorities' ability to deliver core services to the public, and adapt and respond to the needs of the citizens. For the future, this indicates that the public should take advantage of technological advancements to improve public services. However, a rush to make use of new AI tools could compromise other interests and values, and potentially increase the risk of arbitrary use of power, discrimination and unfair treatment. Public institutions using new AI tools have a critical role in ensuring that the technology is trustworthy, reliable and transparent.

79. OECD (2022) Trust Survey, https://www.oecd.org/publications/drivers-of-trust-in-public-institutions-in-norway-81b01318-en.htm

80. Brandtzæg, Skjuve, Dysthe & Følstad (2021) When the Social Becomes Non-Human: Young People's Perception of Social Support in Chatbots, CHI Conference Paper, https://dl.acm.org/doi/10.1145/3411764.3445318

81. https://www.uea.ac.uk/news/-/article/fresh-evidence-of-chatgpts-political-bias-revealed-by-comprehensive-new-study

82. Jakesch, Buschek & Naaman (2022) Interacting with Opinionated Language Models Changes User's Views, https://mauricejakesch.com/assets/pdf/aimc_influence.pdf

Key developments: Increased use of generative AI tools in different sectors in Norway

Digital media are taking advantage of generative AI – and governs its use

As the use of generative AI tools becomes more widespread, media institutions in Norway have had to adapt rapidly. Many media broadcasters have started experimenting with generative AI. For instance, a new common feature within news articles is a summary of the article in question, generated by AI. When used by the Norwegian newspaper Aftenposten it states: "This summary was created using artificial intelligence and quality assured by Aftenposten's journalists". Another tool used by some media broadcasters is a synthetic or cloned voice, making it possible for people to listen to the news article, rather than reading it.[83]

In addition to experimenting with the new AI tools, media broadcasters have also developed internal guidelines to regulate the use. For instance, the Norwegian broadcasting channel TV2 has guidelines asserting that all AI-generated content should be cross-referenced against other credible sources and manually reviewed before publication. To maintain the integrity of photography in journalism, the use of AI-generated images and videos are limited. If such images or videos are used, they must be explicitly labelled as AI-generated and include details of the tools used. Similar guidelines are also adopted by other Norwegian media broadcasters.

Public sector is exploring new AI possibilities, with caution

The Norwegian Digitalisation Agency and the Norwegian Artificial Intelligence Research Consortium have developed an overview of the use of AI in public sector.[84] Even though most projects involving generative AI are still in an exploratory phase, it demonstrates how institutions are approaching the new technological advancements. For instance, the Norwegian Tax Administration is experimenting with automatic transcription of calls from the public, combined with real-time data analysis, as a measure to support the agency's advisors in providing effective guidance to its clients. The Norwegian Health Directorate is exploring using generative AI to offer better information to children and youth with complex needs. The Norwegian Postal Service is looking into whether large language models and ChatGPT can improve their customer services.

Furthermore, the Norwegian Digitalisation Agency has developed guidelines for responsible use of AI in the public sector.[85] In addition to offering guidance on the development and use of AI and risk management, the guidelines include explicit advice on the usage of generative AI tools. For instance, if generative tools are used to generate images, institutions should take note of the unsolved legislative issues in relation to copyright. Institutions are also advised to be aware that generative AI models are not trained to tell the truth or necessarily aligned with societal values. For text generating tools, the guidelines outline specific usage that can be beneficial to public institutions. Caution is advised when generative AI tools are used by public institutions to interact with the public. For such use, the public should always receive information that they are interacting with an AI tool.

Eager, but reluctant use of generative AI tools among parliamentarians

Our questionnaire for MPs on generative AI indicates that the usage of AI tools is becoming more widespread. While most parliamentarians have experimented with the tools to generate both text and images, it appears to be primarily used for entertainment. Some mentioned that it was more complicated to use generative AI to write speeches, rather than writing it themselves as they know their audience better. One representative acknowledged that AI could generate a speech comparable to their own but chose not to use it because it felt "wrong" or "inauthentic".

Many representatives highlighted the positive impact of media broadcasters using the technology to summarize news. It was also emphasized that the new AI tools could be used to communicate policies and

---

83. See Aftenposten's experiments with making a synthetic news voice: https://futurereport.schibsted.com/an-ai-voice-makes-news-accessible-to-everyone/

84. Overview of AI usage in public sector in Norway: https://data.norge.no/kunstig-intelligens

85. Guidance on use of generative AI in public sector: https://www.digdir.no/kunstig-intelligens/bruk-av-generativ-kunstig-intelligens-i-offentlig-sektor/

political decisions to the public. However, there is a shared apprehension about potential misuse, especially with regards to mis- and disinformation, and deepfakes. In addition, representatives expressed concern about embedded bias in large language models, that could amplify negative traits in society, such as homophobia, sexism and racisms.

Societal and political relevance and debate

Building a large language model for Norwegian

A key policy discussion is whether Norway should develop a foundational large language model for Norwegian. Generative AI built on large language models will become embedded in most everyday digital services, both public and private. Who gets to design, distribute, control, and manage these models will therefore impact and influence the quality, reliability, ethical and cultural leanings of these services. These are questions of democratic politics, as they concern power, accountability, and equity of access to public services and institutions.

Predominant large language models like GPT-4 are primarily trained on English language sources. They have American cultural leanings, as they are developed and owned by American tech companies. Limited information is shared about how these models are designed, trained and adjusted. Insight into data processing and the model's functions from a technical point of view is also limited.

A foundational large language model for Norwegian could:
- Increase democratic oversight over the design and deployment of generative AI.
- Ensure linguistic diversity, by performing better in Norwegian language, including dialects and minority languages.
- Reflect and maintain the Norwegian cultural context and values in the digital age.
- Safeguard privacy, transparency, and reliability, by ensuring insight into data collection, and processing, in line with national standards and legislation.
- Counter-balance market monopolization and reduce Norway's dependence upon AI infrastructure developed and deployed by foreign tech companies.
- Enhance public and private sector innovation and business development in Norway by providing companies safe, foreseeable, and holistic access to a Norwegian model.

Building a Norwegian model from scratch is resource and cost intensive. An alternative might be to cooperate with other Nordic or European countries. Another question concerns how many government-supported large language model projects there should be – one or several? Who should manage and operate such models, and who should get priority access to them?

Today, several universities in Norway have initiated projects to build their own models. Yet, there is no guarantee that such models will reach the desired level of quality. In addition, more and tailored Norwegian language datasets are required to train these models, as well as specialized AI competence. Lastly, there is the question of getting access to a sufficient computing power, and to supercomputers.

A need for national AI regulation?

The EU is currently negotiating the draft of the Artificial Intelligence Act (AI Act). Negotiations will most likely be finalised in 2023. If so, the law will then enter into force in 2026. AI Act is likely to become Norwegian law, and the regulation is to be enforced nationally by a supervisory authority. This authority can demand specific AI products to be withdrawn from the market and will be granted access to information about the AI products, such as source code and training data.

Whether additional national legislation is needed to regulate AI in Norway is a part of an ongoing discussion, both within the Parliament and in the government. Even though the AI Act will entail certain prohibitions and several new product requirements, national authorities will still be able to decide in which sectors the technology can and should be used, for what purpose, and how, particularly in areas such as law enforcement, education and public administration.

A national agency to counter disinformation and build resilience?

Even though reliable media institutions play an important role in countering false information, there is no clearly defined authority to address the spread of mis- and disinformation in Norway. In comparison, Sweden has established a Psychological Defence Agency, to counter disinformation and strengthen resilience in the population.[86] As generative AI tools can increase the amount of synthetic and false content online, additional measures might be needed to ensure a robust and sustainable information system. The Total Preparedness Commission recently proposed the establishment of a national agency with the overarching responsibility to assess and counter threats from influence operations, focusing on strengthening the population's resilience to disinformation.[87] The commission's white paper is currently undergoing public consultation.

A moratorium on use of AI in public sector?

There are currently two propositions concerning generative AI up for discussion in Parliament this autumn. The Socialist Left Party suggests a moratorium on the use of AI in public sector until regulations and guidelines are put in place.[88] A white paper on AI and how it will affect society is also included in their proposal. The Liberal Party, on the other hand, is critical to a moratorium.[89] Rather, they propose the development of a governmental plan to bolster AI research, which includes expanding the number of students relevant to research fields such as law, ethics and STEM subjects. In September, before this proposal have been discussed in parliament, the prime minister of Norway announced one billion NOK to fund research and innovation within AI.[90]

The role of Technology Assessment

The Norwegian Board of Technology (NBT) has worked with issues concerning AI for many years. Today, the board facilitates several projects on generative AI, in addition to raising awareness on the issue in Parliament and for policy makers.

Agenda setting in parliament and for policy makers

NBT has hosted a series of meetings within the parliamentary tech group, on topics such as the breakthrough of large language models and its impact on Norway, digital regulation, cyber security, AI's impact on the public sphere, as well as the AI Act. Relevant stakeholders and experts also attended the meetings, in addition to parliamentarians and their advisors. Several policy briefs for MPs on the topics have been published in preparation for the meetings.

The interest on issues related to generative AI has been massive. So far in 2023, NBT has held more than 50 presentations on generative AI, for parliamentarians, ministries, law administration, public and private sector, as well as relevant interest groups.

A broad-based expert group

NBT has established an expert group including researchers in AI, law, ethics and social science, AI developers and one of the largest public sector institutions in Norway. The group will identify and discuss the most pressing issues concerning generative AI, in particular with regards to how the technology will affect the workplace, public administration, security, education and creative professions. Wider consultations with relevant stakeholders are also a part of the project, such as public administration and experts on issues

---

86. The Swedish Psychological Defence Agency: https://www.mpf.se/en/

87. The white paper of the Total Preparedness Commission: https://www.regjeringen.no/contentassets/4b9ba57be-bae44d2bebfc845ff6cd5f5/no/pdfs/nou202320230017000dddpdfs.pdf

88. The proposition of the Socialist Left Party: https://www.stortinget.no/no/Saker-og-publikasjoner/Publikasjoner/Representantforslag/2022-2023/dok8-202223-232s/

89. The proposition of the Left Party: https://www.stortinget.no/no/Saker-og-publikasjoner/Publikasjoner/Representantforslag/2022-2023/dok8-202223-273s/

90. https://www.nrk.no/norge/regjeringen-med-milliardsatsning-pa-kunstig-intelligens-1.16546093

concerning AI and climate change. A report on the impact of generative AI with policy options for parliament will be published during the autumn.

A human rights analysis on generative AI

In cooperation with the Norwegian National Institution for Human Rights, NBT is currently working on an assessment and analysis of the effects of generative AI on democracy and freedom of expression, using human rights as an entry point for technological assessment. The analysis is be published in a report and presented in an event with parliamentary participation.

Going forward: A democracy lab

Generative AI is increasingly becoming a part of people's everyday life. The new technology is introducing societal considerations and dilemmas regarding values, ethics and responsibility. NBT is therefore planning a continuous effort to involve citizens, stakeholders and those affected to help formulate future policy advice for parliamentarians in the next three-year period. The AI Democracy Lab will use a range of methods for participation.

Portugal

Artificial Generative Intelligence and Work – Portugal

António Moniz,[91] Marta Candeias,[92] Nuno Boavida[93]
(Observatory of Technology Assessment, OAT/CICS.NOVA)

Description of the System/Problem at stake

What is it about?

Artificial Generative Intelligence (AGI) applications are recently available and the impact assessment for the present time is still very limited, but some implications and risks can be understood, and they are documented in the literature to assess the potential of AGI tools such as, ChatGPT from OpenAI and Bard from Google, we have conducted a small testing exercise. We have asked both AGI tools to look for applications in which they can have impact and we got the following ones:

For ChatGPT: customer service, education, healthcare, content creation, personal assistants and financial services. "These are just a few examples of the potential applications of Chat GPT. As the technology continues to evolve and improve, it is likely that we will see even more innovative and creative uses of this powerful language model." (May 2023)

For Bard: education, customer service, content creation, research, personal assistant, medical research and disaster relief. "These are just a few of the many possible future applications of Bard. As Bard continues to develop, it is likely to be used in even more ways to improve our lives." (Sept. 9, 2023)

Comparing both results, we observed that the application fields are very similar, which means the eventual impacts can be anticipated. What can we conclude from this?

On the field of customer services, it can be expected to have a significant impact in many sectors. There is much room for errors yet, but it will probably affect all mankind. According to ChatGPT it can be used in e-commerce, telecommunications, and healthcare. It can understand natural language queries and provide relevant responses to customers, improving the efficiency and quality of customer service. Bard can answer questions, troubleshoot problems, and resolve issues in a timely and efficient manner. Bard can also be used to personalize the customer experience, making customers feel valued and appreciated.

On education ChatGPT can be used as a virtual tutor or mentor, helping students with homework, providing personalized feedback, and answering questions related to various subjects. It can also be used as a language-learning tool, allowing students to practice their language skills through conversation. Bard can be used to create interactive learning experiences for students of all ages. It can generate personalized content that is tailored to each student's individual needs and interests. Bard can also be used to provide feedback and guidance to students, helping them to learn more effectively. However, there is a strong opposition by teachers. There were debates at several universities on the topic. There has been significant interest among journalists. It can also increase the worries that teachers have raised about potential misuse of this tools.

On health, ChatGPT can be used to provide medical advice and support to patients, especially in situations where it may be difficult for them to access in-person medical services. It can also be used to provide mental health support and therapy. Bard can be used to research medical conditions and treatments. It can also be used to develop new drugs and therapies. This makes it a powerful tool for doctors, researchers,

---

91. abm@fct.unl.pt
92. marta.candeias@tecnico.ulisboa.pt;
93. nuno.boavida@fcsh.unl.pt

and patients who are looking for better ways to improve health and well-being. It can also reduce bureaucracy among healthcare workers.

In content creation, Chat GPT can be used to generate content for various industries, such as journalism, marketing, and entertainment. It can help automate tasks such as article writing, social media posts, and video scripts. Bard can be used to generate creative content, such as articles, blog posts, scripts, and poems. It can also be used to translate languages, write different kinds of creative content, and answer your questions in an informative way. This makes it a valuable tool for writers, journalists, and other creative professionals. But it can also destroy the conventional modes of creative production (in music, filming, writing).

On research, just Bard mentioned that application which can access and process information from a variety of sources, including books, articles, and websites. This makes it a powerful tool for students, researchers, and anyone else who needs to find information quickly and easily. The application of generative AI and LLM might lead to positive outcomes eventually. The extent to which it will affect tasks and even research project is difficult to envisage.

- Literature search for scientific research can be enhanced with the help of AGI
- Running models and/or simulations in lab environment can be supported
- Research management and administration (financial, scheduling, procedures, administrative tasks, planning) will also be a current application field

Personal assistants are still in a development phase by the AGI application producers. Chat GPT can be used as a personal assistant, helping individuals manage their tasks, schedule, and appointments. It can also be integrated with other smart devices, such as smart speakers and home automation systems, to provide a more seamless user experience. Bard can be used as a personal assistant to help with tasks such as scheduling appointments, making travel arrangements, and managing finances. It can also be used to provide reminders, set alarms, and control smart home devices. This makes it a valuable tool for busy people who want to save time and simplify their lives. It may have a higher acceptability in Portugal.

Disaster relief is a hypothetical application that would be very helpful, but it remains to be seen what can really be done and implemented, as well as the resistance that first responder workers (fire brigades, police, paramedics, etc.) will have to it. Just Bard mentioned that it can be used to provide disaster relief in the aftermath of natural disasters such as hurricanes, floods, and earthquakes. It can be used to coordinate relief efforts, provide information to survivors, and help to rebuild communities.

On financial services, just Chat GPT mentioned that it can be used to provide financial advice and support to individuals and businesses. It can help with tasks such as budgeting, investment advice, and tax planning. AI developments are older than the ChatGPT generation. Its impacts on the financial services workforce were impressive (disguised by the covid-19 epidemic) and had been predicted since the 2018 interviews. According to a CTO interviewed in 2018, "the implementation of many algorithms was on hold to keep social peace until people reached pre-retirement age." Furthermore, it is expected that there will be a significant transformation in the sector due to the accelerated pace of implementation of more software, easy access to ICT professionals, which will lead to further reducing the workforce used in the sector.

According to OECD, "while the impact of the latest wave of generative AI is not entirely clear yet, early estimates of occupational AI exposure that take into account the capabilities of large language models like ChatGPT reach conclusions similar to those of previous estimates of AI exposure: it is primarily high-pay occupations requiring higher than average education or training that are most exposed to AI" (OECD, 2023: 96).

What is the state of play in your country (development, deployment, use)?

According to the last DESI report (DESI, 2022), the digital competences in Portugal are around the EU average levels or above. But the skills related to basic digital content creation are still very low in Portugal (60% which is far below the EU average of almost 70%). At the same time, there are high levels of investment on the capacity building of digital competencies in Portugal (see below the recent policy decisions). This eventually means that AGI tools can provide support to enable the ability "to create and edit digital content, to improve and integrate information and content into an existing body of knowledge while understanding how copyright and licences are to be applied and to know how to give understandable instructions for a

computer system" (DESI, 2022: 8). Some applications are already running in some companies: chat bots, as the one from Logo (insurances), and from other banking services. There are also some significant scientific debates at universities, companies and by some relevant stakeholders (APDSI, APIA, GEE). There has been a general interest in these debates, mainly among journalist if one considers the number of TV programs and newspaper articles published since the generative AI reached every common citizen.

What do your MPs think about it?

We got no answers until now. However, the narrative about technologies now includes AI, but to a limited extent. MPs lack knowledge and access to good report on the state-of-the-art of generative AI, beyond the declarations of regulation at the EU level.

Who are the key stakeholders?

Political parties, Trade Unions and Employer Associations
APDSI – Association for the Promotion of Information Society
APIA – Portuguese Association of Artificial Intelligence
FCT – Science and Technology Foundation
FCCN – National Foundation for the National Computing
GEE – economic studies at the Ministry of Economy
OCS/CNCS – Observatory of Cybersecurity

Why is this important for your country?

In June 2019, the Portuguese Government presented the national strategy AI Portugal 2030 (Portugal, 2019) to set out challenges and opportunities of the growing AI ecosystem in Portugal. This strategy presents the plan to foster the use of AI in the public and private sector during the coming years. The plan concentrates its actions on inclusion, education, qualification, specialization and research as people are the main engine of a successful AI deployment. (AI Watch [94])

- The national AI strategy of Portugal does not disclose financial figures, or estimations, for its implementation.
- Lower levels of schooling may imply competitive chances for AGI applications in most possibilities of human substitution.
- Less strength of unions on negotiation procedures in key sectors
- Trade Unions have not a relevant position about AGI
- Employers have conservative information about the possibilities for such applications.
- There are no anticipatory studies on the Portuguese labour market application of AGI and its consequent impacts.
- In the Public Administration many AI solutions have been implemented and more a planned to be (e.g. in justice system).

Societal and political relevance and debate

Results from surveys

Is there ongoing debate on the impact of generative AI on our societies and democracies? Are there results from surveys among the population concerning usage of or opinions about generative AI?

There are a few debates in academia and sectoral associations about AI. There are many debates among opinion maker on AGI. The interests of journalists on tech consequences for their profession has led to a significant debate about general AI in Portugal. There are public surveys on general topic, as the following:

---

94. https://ai-watch.ec.europa.eu/countries/portugal/portugal-ai-strategy-report_en

- INE (2020a) Inquérito à Utilização de Tecnologias da Informação e da Comunicação pelas Famílias – 2020. Instituto Nacional de Estatística. Available at: https://www.ine.pt/xportal/xmain?xpid=INE&x-pgid=ine_destaques&DESTAQUESdest_boui=415621509&DESTAQUESmodo=2&xlang=pt

This survey applied in 2020 is about the Information and Knowledge Society in Portugal. In particular, it is a survey on the use of information and communication technologies in families. There, it was concluded that internet and e-commerce users increased significantly. The percentage of users for educational reasons more than doubled.

- INE (2020b) Inquérito á Utilização de Tecnologias da Informação e da Comunicação nas Empresas – 2020. Instituto Nacional de Estatística. Available at: https://www.ine.pt/xportal/xmain?xpid=INE&x-pgid=ine_destaques&DESTAQUESdest_boui=415622957&DESTAQUESmodo=2&xlang=pt

In 2020, around 97% of companies with 10 or more people employed and 42.8% of those employed use a computer with an Internet connection for professional purposes. In the same year, 61.1% of companies reported having their own website or that of the economic group to which they belong. As well, 13.0% of companies with 10 or more people employed use interconnected devices or systems that can be monitored or controlled remotely via the Internet (IoT) and 9.1% use industrial and/or service robots. In 2019, 4.5% of companies with 10 or more people employed used 3D printing, with company 3D printers and/or through printing services provided by other companies.

Is there any legislation in place?

The next legislation pieces are examples of AI related legislation present in Portugal and approved in recent years:

Creation of the «Qualifica Indústria Program», aimed at micro, small and medium-sized enterprises (SMEs) in the industrial sectors, aimed at supporting worker qualification and requalification processes, preventing future unemployment, Portaria 282/2023, 14 september (Ministry of Labour).[95]

Defines the official and exclusive models of the citizen card, the physical security elements that compose it, the technical and security requirements to be observed when capturing the facial image and fingerprints of the holder of the application and also the concrete measures for the inclusion of citizens with special needs in the information society, to be observed in the provision of citizen support services, Portaria 286/2017,[96] of 3 April, amends the Labor Code and related legislation, within the scope of the decent work agenda.[97]

Defines the use of portable cameras for individual use by police officers, Decree-Law 2/2023, 2 January.[98]

Approves the Defense Technological and Industrial Base Development Strategy 2023-2033, Council of Ministers Resolution 52/2023, 5 june.[99] Approves the National Cyber Defense Strategy, Council of Ministers Resolution 106/2022, 2 november.[100]

Approves the Regulation of the Incentive System «Companies 4.0», Portaria 135-A/2022.[101]

Approves the Electronic Communications Law, transposing Directives 98/84/EC, 2002/77/EC and (EU) 2018/1972, Law 16/2022.[102]

Are there currently political or legislative proposals on these topics?

Recently, there is a growing interest from jurisdiction stakeholders about AGI and AI in general. Some new legislative proposals have been approved and established. They mostly follow the framework of EU legislation.

---

95. https://diariodarepublica.pt/dr/detalhe/portaria/282-2023-221642378
96. https://diariodarepublica.pt/dr/legislacao-consolidada/portaria/2017-205638899
97. https://diariodarepublica.pt/dr/detalhe/declaracao-retificacao/13-2023-213650804
98. https://diariodarepublica.pt/dr/detalhe/decreto-lei/2-2023-205557199
99. https://diariodarepublica.pt/dr/detalhe/resolucao-conselho-ministros/52-2023-213918210
100. https://diariodarepublica.pt/dr/detalhe/resolucao-conselho-ministros/106-2022-202899924
101. https://diariodarepublica.pt/dr/legislacao-consolidada/portaria/2022-207756412
102. https://diariodarepublica.pt/dr/legislacao-consolidada/lei/2022-187527517

As is mentioned in the last issue of the OECD Employment Outlook, "AI has the capacity to fully automate employment-related decisions, including which job seekers see a vacancy, shortlisting candidates based on their CVs, assigning tasks at work, and for bonus, promotion, or training decisions. While this capacity potentially frees up time for managers to focus more on the interpersonal aspects of their jobs, it raises the question whether decisions that have a significant impact on people's opportunities and well-being at work should be made without any human involvement, or at least the possibility for a human to intervene" (OECD, 2023: 191). In general, continues the same document, "there are new legislative efforts that would prevent the adoption of fully automated decision-making tools in high-risk settings such as the workplace, by requiring human oversight (i.e. a 'human in the loop')". For example, the 2021 proposal for the European Union's AI Act that regulates AI systems was made available or used in the EU 27 member states to address risks to safety, health and fundamental rights, including specific provisions for use of certain high-risk AI applications in the workplace. This was also the case of Portugal.

Recently, the main labour legislation instrument (Labour Code) had several changes. However, not many introduced changes regarded the artificial intelligence applications. However, the changes that are being applied from the recent changes of April 2023 (Law 13/2023) are very important and influence the actions of workers' representative organizations..

The first concerns the legal standards regulating employment contracts. Although the new paragraph o) of article 3 mentions the example of digital platforms, it is important to know that the use of algorithms, artificial intelligence and related matters cannot – to put it simply – allow dismissal. If this could be used in the case of digital platforms where an employment contract has been established, now the Labor Code does not allow it. But eventually we can consider all other sectors. Thus, for example, when a company or organization uses its own services or those of personnel management consultants that use "algorithms, artificial intelligence and related matters" to change the "legal standards regulating employment contracts", workers may be protected by the Labor Code.

The second, relating to article 24, concerns the right to equality in access to employment and work. Protection also comes into play here when it is proven that decision-making based on algorithms or other artificial intelligence systems undermines the legal provisions relating to the exercise of a professional activity by a foreigner or stateless person, or provisions relating to the reconciliation of professional activity with family life. In general, we can say that there is now a prevention against human resources management decision-making based on algorithms or other artificial intelligence systems.

The other two cases are related to workers' representative organizations. In the first (art. 424), workers' committees now have the right to information about the "parameters, criteria, rules and instructions on which algorithms or other artificial intelligence systems that affect decision-making about access and maintenance of employment, as well as working conditions, including the creation of profiles and control of professional activity". This stipulation seems very interesting as it clearly restricts the possibility of applying "parameters, criteria, rules and instructions" on "access to and maintenance of employment, as well as working conditions, including the creation of profiles and control of activity professional" that are based on automated and programmable procedures. However, this restriction can only happen when the works councils are able to recognize these procedures, or the possibility of them being used.

This will imply that they will discuss these topics internally. For example, there is already some attention to the relationship between digitalization of work and employment, but the implementation of this relationship is not always evident. In the second case (art. 466), it is the unions that will have to internalize this debate.

science / evidence-based inputs and democratic inputs

The Cybersecurity Observatory (which belongs to the National Centre for Cybersecurity) has a multidisciplinary approach to the computer security phenomenon, integrating in its field of observation several areas of knowledge. They are Society, Economy, Public Policies, Ethics and Law, Risks and Conflicts, as well as Innovation and Future Technologies. They make available below a set of public documents, from national and international entities, which serve as a source for the study of each of these aspects of Cybersecurity in Portugal. For exemple, they have conducted a survey, in collaboration with the Directorate-General for Education, was aimed at non-university teachers and focuses on the cybersecurity-related behaviours ex-

perienced by the teaching profession during the second semester of the 2019/2020 academic year, during which distance learning replaced face-to-face teaching, as a result of the Covid-19 pandemic.[103]

The framework discusssion

OpenAI researchers collaborated with Georgetown University's Center for Security and Emerging Technology and the Stanford Internet Observatory to investigate how large language models might be misused for disinformation purposes. The collaboration included an October 2021 workshop bringing together 30 disinformation researchers, machine learning experts, and policy analysts, and culminated in a co-authored report building on more than a year of research. This report outlines the threats that language models pose to the information environment if used to augment disinformation campaigns and introduces a framework for analyzing potential mitigations.[104] In this paper the authors state that "many of the properties of large generative models are not fully understood. Similarly, clarity is still missing regarding both the structure and the impacts of many influence operations, which are conducted in secret. Clarity on the scale of the threat posed by influence operations continues to be elusive." (Goldstein et al, 2023: 65).

As recent improvements on machine generated text also revealed, we can witness increase abuses of natural language generation (NLG) models, such as phishing, disinformation, fraudulent product reviews, academic dishonesty, and toxic spam. The activities of the national Observatory of Cybersecurity have worked on this in Portugal.[105] Their report on Emergent Technologies (Aguiar et al., 2023) mentions the point 2 of article 9 of the Portuguese Charter of Rights Humans in the Digital Age which attempts to address the issue of automated decisions. "Some practical problems can be highlighted, such as sharing AI-powered IoT environments where multiple people use the environment without needing to identify themselves. However, an audit of the decisions taken is essential to be able to review what were the bases for a decision and how it occurred, and the legal frameworks at play appear to reveal insufficient for the multiple situations that are foreseen for the future" (Aguiar et al., 2023: 105).

Because of those problems that the NLG models bring, "addressing the risk of abuse is vital to maximize the potential benefit of NLG technology, while minimizing harms — a key principle of trustworthy AI" (Crothers, Japkowicz, and Viktor, 2023: 1). These Canadian authors as well predict how important and decisive this issue is. They underline that "predicting how abuses are likely to unfold, and understanding the best defenses against them, is essential for allowing humanity to reap the positive benefits of this technology while minimizing potential harms. We must walk a cautious path through the age of the silicon wordsmith" (Crothers, Japkowicz, and Viktor, 2023: 2).

This discussion becomes very relevant considering the risk impact of this technology and the problems related with its use. The Society Report of the national Observatory of Cybersecurity has mentioned this topic. This Report, besides, as usual, analysing the data on attitudes, behaviour, education, and awareness regarding cybersecurity, has as the highlighting of some indicators correlated with possible impacts of the National Strategy for Cyberspace Security; a new chapter on the uses of digital technologies in general; the deepening of the survey on awareness-raising actions; and a set of recommendations.

Role of TA in the debates

Has your institute taken up this topic?

The Observatory of Technology Assessment at the CICS.NOVA research centre (which is EPTA member), runs the project InteliArt about the implications of AI in work, employment and industrial relations, and it was started in 2021. The research team published one article in a Q1 journal (Moniz, et al., 2022) and already won 2 prizes (Candeias et al., 2022).

---

103. https://www.cncs.gov.pt/en/observatory/#inqueritos
104. https://openai.com/research/forecasting-misuse
105. See more information at https://www.cncs.gov.pt/en/observatory/#relatorios

The OAT/CICS.NOVA is running also other projects where to topic is addressed. Some projects are more related with digital platform work (Moniz et al., 2023; Boavida et al., 2023) or with the application of AI on industrial environments (Candeias, 2023; Moniz, 2018; Moniz and Krings, 2016).

Has TA made an impact on the ongoing debates?

The research developed at OAT/CICS.NOVA has been disseminated and communicated in many scientific fora as well as in many social media.

What are the lessons learned from TA?

In the OpenAI webpage,[106] is said that "just because a mitigation could reduce the threat of AI-enabled influence operations does not mean that it should be put into place. Some mitigations carry their own downside risks". They raise also important issues in the mentioned webpage: Is the mitigation feasible from a political, legal, and institutional perspective? Does it require costly coordination, are key actors incentivized to implement it, and is it actionable under existing law, regulation, and industry standards? These are relevant questions for a TA analysis.

The final questions of the Goldstein et al. (2023) paper is revealing. They present the following issues: "Should AI developers release or restrict their models? Should internet researchers publish observed tactics of propagandists or keep them secret? To what extent can platforms and AI developers form meaningful partnerships that can aid in the detection and removal of inauthentic content? At the broadest level, thoughtful engagement with all these questions—both from people within the relevant industries and from neutral, third-party observers—is a critical necessity" (Goldstein et al., 2023: 66). And these are questions that TA can tackle once they show clearly where are the main risk factors and elements that need a decision process that involves a variety of social actors and stakeholders.

Literature

Acemoglu, D. and Lensman, T. (2023), *Regulating Transformative Technologies*, July 6, MIT, 2023 https://toddlensman.com/files/research/tech-reg-8.pdf

Aguiar, R.L. et al. (2023), *Tecnologias Emergentes*, Lisbon, CNCS, 115 pp.

Boavida, N.; I. Roque, and A. B. Moniz (2023). Collective Voice and Organizing in Digital Labour Platforms in Portugal, *Journal of Labor and Society:* 1-25.

Candeias, M. (2023), *Impact assessment of AI-enabled automation on the workplace and employment. The case of Portugal*, Caparica, PhD thesis, UNL, http://hdl.handle.net/10362/156952

Candeias, M.; N. Boavida; A. B. Moniz (2022), Automation trends in Portugal: implications in productivity and employment, *GEE Paper*, No. 165, https://www.gee.gov.pt/RePEc/WorkingPapers/GEE_PAPERS_165.pdf

Candeias, M.; A. B. Moniz and N. Boavida (2022). Digital Transformation in the Automotive Sector in Portugal: Data Analysis of Industrial R&D Projects. In Kirov and Malamin (eds.), *Inclusive Futures for Europe: Addressing the Digitalisation Challenges, Inclusive Futures for Europe - BEYOND 4.0 Scientific Conference Sofia 2021 Proceedings*, Sofia: Prof. Marin Drinov Publishing House of BAS. https://press.bas.bg/en/eBooks-105/show-106(55).

CGTP (2021). *Livro Verde Sobre o Futuro Do Trabalho - Apreciação Da CGTP-IN*. Lisbon, CGTP. http://www.cgtp.pt/cgtp-in/areas-de-accao/accao-reivindicativa/16100-o-futuro-do-trabalho-tem-de-ser-de-progresso-e-justica-social.

Crothers, E.; Japkowicz, N. and Viktor, H. (2023), *Machine Generated Text: A Comprehensive Survey of Threat Models and Detection Methods*, arXiv:2210.07321v4

DESI (2022) *Digital Economy and Society Index 2022 Human Capital*

Duarte, J. B., et al. (2019). *O Futuro Do Trabalho Em Portugal: O Imperativo Da Requalificação*. *Nova SBE, Confederação Empresarial de Portugal*. Lisbon: NOVA SBE & CIP. https://www.cip.org.pt/wp-content/uploads/2019/10/Relatório-FoW_NSBE-CIP.pdf.

Goldstein, J.A. et al. (2023), *Generative Language Models and Automated Influence Operations: Emerging Threats and Potential Mitigations*, arXiv:2301.04246v1

Krings, B.-J.; A. B. Moniz, and P. Frey (2021). Technology as enabler of the automation of work? Current societal challenges for a future perspective of work. *Revista Brasileira de Sociologia* 9: 206-229.

Mateus, A. et al. (2017). *Avanço Da Economia Digital*. Edited by EY - Augusto Mateus & Associados. Lisbon.

Ministério do Trabalho, Solidariedade e Segurança Social. 2021. *Livro Verde Sobre o Futuro Do Trabalho* 2021. Lisboa: Ministério do Trabalho, Solidariedade e Segurança Social

Moniz, A. B.; Candeias, M.; N. Boavida (2022), Changes in productivity and labour relations: artificial intelligence in the automotive sector in Portugal, *Int. J. Automotive Technology and Management* 22: 222-244, DOI: 10.1504/IJATM.2022.124366

Moniz, A.B. (2018) *Robótica e trabalho. O futuro hoje*. Glaciar/FLAD, Lisbon, http://www.glaciar.com.pt/detalhe_edicao.php?cd_edicao=65

Moniz, A.B. (2023), Inteligência artificial no Código de Trabalho, esquerda.net, 28/8/2023, https://www.esquerda.net/opiniao/inteligencia-artificial-no-codigo-de-trabalho/87454

Moniz, A.B. and Krings, B.-J. (2016) Robots working with humans or humans working with robots? Searching for social dimensions in new human-robot interaction in industry. *Societies*, 6 (3) https://doi.org/10.3390/soc6030023

Moniz, A.B.; N. Boavida, C. Makó, B.-J. Krings, P. Sanz de Miguel eds. (2023), *Digital Labour Platforms. Representing Workers in Europe*, Gaia: Húmus/CICS.NOVA,

OECD (2023), OECD Employment Outlook 2023, Artificial Intelligence and the Labour Market, OECD Publishing, Paris, https://doi.org/10.1787/08785bba-en

Sweden

2023-08-31

Background EPTA conference Generative AI and democracy

Introduction

Here follows a short description of various decisions and initiatives in Sweden ahead of the 2023 EPTA conference on generative AI and democracy compiled by the Swedish Parliament's Evaluation and Research Secretariat (ERS).

The presentation is mainly based on statements from the Swedish Parliament (Riksdag) and Government in the field of artificial intelligence. So far, it has not been possible to engage any MPs from Sweden to participate in the EPTA Council meeting, which is why it has not been possible to compile their experiences and views. The memorandum concentrates on issues with societal and political relevance.

Overall, it can be mentioned that public debate has, as in most other countries, been extensive and active around AI and its significance for various policy areas, e.g. democracy. The discussion about the challenges and opportunities offered by AI and generative AI is highly topical among researchers, in the media and among policy makers. The issue of regulating AI in various ways has been widely discussed, but as yet, no decisions have been taken. The debate primarily revolves more generally around AI and in fewer, more specific, cases around generative AI. Below follows a description of some of the guidelines and statements adopted by the Riksdag and the Government in the field of AI. Parliament has not yet carried out any formal TA in the area of generative AI, but ERS has produced factual documents for various committees about AI, e.g. recently in connection with the Riksdag's Research Day in June 2023 when the Committee on the Labour Market and the Committee on Transport and Communications organised seminars about the consequences of AI within their subject areas.

Decisions in the Riksdag on AI: Statement on the European Commission's White Paper on Artificial Intelligence

In 2020, the Committee on Education issued a statement[107] on the European Commission's White Paper on Artificial Intelligence, COM (2020) 65. The Committee welcomed the White Paper and a discussion on Europe's future approach to AI. The Committee saw a need to mobilise resources and review which measures are necessary for the EU to be able to strengthen its capacity in the AI area. The Committee emphasised that it is of great importance that research in AI is prioritised at European level. This applies to support for both basic research and applied research, as well as targeted efforts to promote excellence. The Committee considered it important to ensure that the entire chain from research to the market is taken into account, e.g. by enabling cooperation between business and academic environments. At the same time, it considered it of great importance that the efforts are not unequivocally aimed at commercialising AI applications. The results of AI research also have great potential to be applied in the public sector and support for research with societal benefits should be promoted. The Committee further believed that accessible data is of fundamental importance and that it is important to secure work for open and high-quality data in the EU.

The Committee pointed out that AI, like all technology, exists to serve people, and not the other way around. At the same time, AI entails certain special risks; it can have unwanted effects and may be used in ways that can offend individuals and groups. There may also be an overconfidence in the technology, which can lead to unforeseen consequences. From this perspective, a substandard regulatory framework not only

---

107. Statement 2019/20:UbU20 – available at https://www.riksdagen.se/sv/dokument-och-lagar/dokument/utlatande/vitbok-om-artificiell-intelligens_h701ubu20/. A courtesy translation of the summary of the statement can be found on the IPEX website: https://secure.ipex.eu/IPEXL-WEB/document/COM-2020-0065/serik

means that the EU risks falling behind because too few people use AI solutions, it also entails a potential risk to fundamental rights. It is therefore of fundamental importance that a future regulatory framework takes into account the ethical aspects of AI technology and that legislation at both Swedish and European level is adequate to ensure that the risks of AI are minimised. In this context, the Committee pointed out the importance of integrating an equality and rights perspective into continued work.

The Committee also wished to highlight the issue of the risks of technology and data monopolies, in order to avoid data that can be used for the public good being limited to a few actors. This applies not least to matters connected with the ownership of data, which needs to be clear so that the information collected is not misused. The Committee therefore welcomed a continued discussion on transparency in AI systems.

Government policy on artificial intelligence

In 2018, the Swedish Government presented a national orientation for artificial intelligence. An overarching goal is to become a world leader in harnessing the opportunities offered by digital transformation. The Government has identified a need to develop a national approach to AI in Sweden. The purpose of the plan is to identify an overall direction for AI-related work in Sweden and lay the foundations for future priorities. The Government's goal is to make Sweden a leader in harnessing the opportunities that the use of AI can offer, with the aim of strengthening Sweden's welfare and competitiveness.

Sweden needs to develop its long-term supply of knowledge and expertise in the field of AI in order to reap the benefits of AI. The need for relevant knowledge of AI must be met through education and training, continuing education and research. Innovation and initiatives regarding areas of use are also needed to promote early application projects. Sweden needs to ensure access to data and infrastructure, such as computing power, in addition to appropriate national, European and international frameworks. Public stakeholders should therefore actively support AI applications by making relevant data available and creating national digital infrastructure, taking security and integrity issues into account.

In light of the societal transformation that AI entails, it is important to work for a coherent and strategic AI policy that aims to create a safe, secure and favourable climate for digitisation and harnessing the opportunities offered by AI. The development and use of AI need to be guided by norms and ethical principles aimed at harnessing the benefits, while minimising the risks to both society and individuals. This is not only a matter for researchers and engineers, but for a wide range of professions. Appropriate frameworks of principles, norms, standards and rules are therefore important prerequisites if Sweden is to realise the benefits of AI in society. Such frameworks must balance fundamental needs for privacy, ethics, trust and social protection with access to the data needed to realise the potential of AI. Regulatory frameworks at European and international level, for example cross-border data transfer rules, are also important.

Various initiatives by the Government to promote the capacity of public administration to use artificial intelligence

The Government has tasked[108] several public agencies (for taxation, the labour market, companies, digital government) to promote the capacity of public administration to use artificial intelligence. The Government believes that it is important that public administration develops more common tools for using AI. During the course of the assignment,[109] it has emerged that there is a great demand within public administration for comprehensive and concrete support in developing AI solutions, secure technical environments and developing AI guidance material. There is a great need from the public agencies for clear governance

---

108. Government decision 17 June 2021, ref. 2021/01825 – available at https://www.regeringen.se/regeringsuppdrag/2021/06/uppdrag-att-framja-offentlig-forvaltnings-formaga-att-anvanda-artificiell-intelligens/

109. Final report: Assignment to promote the capacity of public administration to use artificial intelligence, Agency for Digital Government, January 2023 – available at https://www.digg.se/analys-och-uppfoljning/publikationer/publikationer/2023-01-23-slutrapport-uppdrag-att-framja-offentlig-forvaltnings-formaga-att-anvanda-artificiell-intelligens

and coordination in order to prepare public administration for the upcoming EU regulation.[110] In Sweden, large investments have been made in research on AI in the private sector, but similar initiatives are lacking in the public sector. Continued work is needed to make available an AI infrastructure that enables the development of AI-driven services for all public administration actors. It must be possible for everyone, in an equal way, throughout the country and in different sectors to experiment and develop AI solutions with open or proprietary data. Such an infrastructure means that more actors in public administration can realise the potential of AI at a low cost and in a safe and proven environment.

The Government previously tasked[111] several major universities to carry out a skills development initiative within the field of AI, and recently higher education authorities were tasked with analysing the impact of AI on the supply of higher education programmes and courses in relation to the needs of the labour market. The Government considers it important that the supply of higher education programmes and courses is in line with the new technology in order to meet the needs of the labour market.

---

110. Proposal for a Regulation on harmonised rules on artificial intelligence and amending certain Union legislative acts (COM/2021/206 final)

111. Request for a skills development initiative within the field of artificial intelligence, ref. U2018/02719/UH – available at https://www.esv.se/statsliggaren/regleringsbrev/Index?rbId=23857

Switzerland

Generative AI and democracy in Switzerland

Current situation and issues at stake

Like in other EPTA countries, generative AI has received growing attention in Switzerland since the launch of ChatGPT, which quickly gained numerous users here. In politics, first discussions have been intense, but no major measures had been taken as of this report's date. In Parliament, several MPs have deposited interventions (see below). One of them had ChatGPT write an interpellation to the Federal Council, as an experiment.[112] As Swiss federal elections will take place in October 2023, political parties started an unofficial discussion on a possible agreement on fair uses of AI during the campaign.[113] In July 2023, one political party designed a campaign poster with AI, which triggered controversies on the potential consequences of AI content in this context (e.g., is this akin to fake news, or actually nothing new?).[114]

Switzerland has two strongly anchored direct democracy mechanisms, namely popular initiatives and referenda (either optional or mandatory).[115] Generative AI can be expected to have effects on democracy through these channels as well – which involve individual citizens and political groups, as well as parties and officials in the decision-making process. In a forthcoming study by TA-SWISS on deepfakes and AI content, several risks have already been identified for Swiss democracy. In particular, these could be used as means of manipulation, disinformation and astroturfing (the simulation of grassroots movements via bots), or as instruments to discredit or harass opponents. Moreover, they could serve destabilisation purposes, such as aggravating social tensions or creating general distrust in common sources of information.[116] As for future opportunities for democracy, these are less tangible for the time being, yet they could include broader access to information on politics, thanks to new search and summarisation tools, or simplification of complex texts. Text generation could also prove useful to individuals or groups with low political expertise, writing skills or financial means, which may reduce certain inequalities of resources and barriers to participation. However, these benefits come with controversies of their own, notably regarding the quality of AI results, as well as the role of machines and private AI firms in political life.[117]

A glimpse into Swiss MPs' views

TA-SWISS commissioned a survey on deepfakes among MPs that was held in February 2023.[118] Currently, the results are still being evaluated by its authors. Some initial findings suggest that the 23 re-

---

112. RTS, 17 March 2023 (" Une intervention parlementaire rédigée par l'intelligence artificielle", in https://www.rts.ch/info/suisse/13817951-le-parlement-a-definitivement-adopte-la-reforme-du-2e-pilier.html). See interpellation 23.3147 (https://www.parlament.ch/en/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20233147).

113. RTS, 22 May 2023 (https://www.rts.ch/info/suisse/14042381-une-charte-de-bonne-conduite-sur-lintelligence-artificielle-lidee-fait-son-chemin-dans-les-partis.html).

114. Le Temps, 7 July 2023 (https://www.letemps.ch/suisse/l-affiche-du-plr-et-la-question-de-l-intelligence-artificielle-en-politique).

115. For a brief description of these instruments, see the entries in the Lexicon of parliamentary terms "popular initiatives" and "referenda": https://www.parlament.ch/en/%C3%BCber-das-parlament/parlamentsw%C3%B6rterbuch).

116. See Murat Karaboga, Nula Frei, Frank Ebbers, Greta Runge, Michael Friedewald, Daniel Vogler, Adrian Rauchfleisch, Manuel Puppis and Patrick Raemy, "Deepfakes und manipulierte Realitäten", TA-SWISS (forthcoming, 2024).

117. For a Swiss experiment, see Politbot, a chatbot trained with texts by political parties that replies to queries about their stances with the help of ChatGPT – with a view to making politics "more accessible" for the 2023 elections (https://politbot.ch/).

118. See Murat Karaboga, Nula Frei, Frank Ebbers, Greta Runge, Michael Friedewald, Daniel Vogler, Adrian Rauchfleisch, Manuel Puppis and Patrick Raemy, "Deepfakes und manipulierte Realitäten", TA-SWISS (forthcoming, 2024).

spondents view deepfakes and AI content as a high and concrete risk for Swiss democracy and institutions. Yet while some consider these risks to be part of political discussions already, others do not (or not to a significant extent). In addition, the respondents have not indicated any perceived opportunities for politics.

For the present report, we also asked four members of the Committees on science, education and research to fill in the EPTA questionnaire on generative AI. Two of them replied. The first things that come to their mind when reading the term are OpenAI, ChatGPT, content creation, and productivity gains. They both see an obvious impact on democracy. They also mention several risks, such as more sophisticated fake news in campaigns, opinion manipulation, reputational attacks, and a loss of trust in elected officials and institutions. Yet if used appropriately, generative AI can also have a positive impact on decision-making and produce new knowledge, according to one of them. As for the need to regulate generative AI in the education and health sectors, one MP recommends a risk classification approach: self-regulation would suffice for low-risk applications, while more impactful uses such as surveillance or insurance applications should be examined by concerned communities, and high-impact applications such as medical diagnosis or autonomous vehicles should also be regulated (with international coordination). For the other MP, generative AI requires us to develop new learning methods to foster independent thinking and to rethink the state's education mandate; and in the health sector, to adjust privacy protection measures. Concerning the effects of generative AI on employment, one MP envisages opportunities to liberate time for more demanding and fulfilling tasks, as well as to compensate for label shortage; however, the quantity of affected jobs would be currently underestimated, particularly in the case of highly qualified jobs. In the recruiting context, one MP sees potential for better matches, with the equivalence of degrees playing an increasingly important role; but personal encounters would remain indispensable in order to ensure authenticity. The other MP is rather wary of biases in the process and considers a responsible framework to be necessary, including on the international level. As for debates in Swiss Parliament, both MPs refer to deposited interventions (see below). One of them considers AI in general to be an important part of Switzerland's digital strategy.[119] According to the other, a general AI Act for Switzerland would be controversial, and the Swiss "technology-neutral" approach would still constitute a sensible means of preserving opportunities and reacting responsibly to new developments. However, Switzerland would lag behind in terms of data infrastructure, in spite of future measures adopted by Parliament.

Societal and political debates

Debates on generative AI in Switzerland commenced a few years ago, initially on a rather theoretical level (e.g., will AI ever be able to write a novel on its own?), and significantly increased with the recent dissemination of concrete applications for the general public. In 2022, image generators like Midjourney, LensaAI or Dall-E were largely cited, along with questions related to the transformation of communication, art and creativity.[120] Since the launch of ChatGPT, generative AI has become one of the most prominent topics in the media, social media, public debates, and corporate communication. Thus the initial focus on ChatGPT gradually extended to generative AI, and even to AI in general.

This dynamic is reflected on the political level. Various interventions have been deposited in Parliament in 2023.[121] Many of these pose questions to the Federal Council, especially regarding its assessments of risks and opportunities of generative AI, as well as planned measures to contain risks, but also to support research and innovation. More studies and investigation have also been requested, especially concerning the need for regulation, consequences of generative AI for employment, education and cyber security, potential harm to individuals, Switzerland's position regarding the AI Act and within the Com-

---

119. For the official Digital Switzerland Strategy, see https://digital.swiss/en/.

120. Such questions will be addressed in the TA-SWISS study on culture and digitalisation (forthcoming, 2024, https://www.ta-swiss.ch/en/culture-and-digitalisation).

121. See in particular items 22.1074, 23.3147, 23.3201, 23.3516, 23.3583, 23.3644, 23.7270, 23.3806, 23.3812, 23.3849, 23.3860, 23.3861, 23.3930, 23.438 and 23.7049 in Swiss Parliament's public database https://www.parlament.ch/en.

mittee on AI of the Council of Europe, and adjustments to the country's digital strategy. Regulation is a central topic, with the familiar division between calls for stricter regulation on the one hand, and worries that this may hamper beneficial innovation on the other.[122]

When it comes to AI regulation, Switzerland has officially followed a sector-based and technology-neutral approach, rather than overarching dispositions on AI in general. As we prepare this report (in August 2023), the Federal Council is waiting for the completion of the EU's AI Act in order to assess its impact on Switzerland and the necessity of legislative adjustments. In April 2023 the Federal Office of Communications was instructed by the Federal Council to formulate a draft bill on the regulation of communication platforms.[123] Previous measures adopted by the Federal Council include the Guidelines on Artificial Intelligence for the Confederation,[124] the creation of the Competence Network for Artificial Intelligence (CNAI) in 2022,[125] as well as of various working groups across the Federal Administration and among interested stakeholders. On the international stage, Switzerland is actively contributing to the AI initiatives of the OECD, the Council of Europe, UNESCO and the International Telecommunication Union.[126]

Given the preliminary stage of the discussions on generative AI, it seems early to assess the involvement of science and the public in the decision-making process. Scientists from diverse disciplines are very frequently consulted in the media. In Parliament, the Committee on science, education and research of the National Council organised a scientific hearing on generative AI in August 2023, to which TA-SWISS was invited. Other exchanges between science and politics on the issue can hardly be tracked systematically. As for the public sphere, many entities across civil society, the IT sector, and the economy have been very active in taking a stand on the topic (opportunities and risks, measures to adopt). Given the lack of comprehensive surveys, it is difficult to assess citizens' attitudes on this issue for now. A survey on the population's attitudes towards deepfakes and AI content is currently being conducted on behalf of TA-SWISS.[127]

Role of TA in the debates

TA-SWISS has launched several projects related to generative AI. As previously mentioned, a comprehensive study on deepfakes and AI content will be published in 2024, with recommendations for politicians, stakeholders and the public on the basis of an overview of risks and opportunities, plus a technical and legal analysis and a survey on citizens' perceptions.[128] At the same time, in April 2023, TA-SWISS published a short paper on ChatGPT and large language models, which presents the functioning of and the main questions raised by these applications.[129] A broader study of language models might be initiated in 2024. Previous studies of interest for the relationship between democracy and AI include a general study on

122. See e.g. the debate in Inside IT between MPs F. Grüter, "Zu viel Regulierung bei KI bremst die Innovation" (https://www.inside-it.ch/parldigi-direkt-zu-viel-regulierung-bei-ki-bremst-die-innovation-20230419) and M. L. Marti, "KI-Regulierung muss Innovation nicht behindern" (https://www.inside-it.ch/parldigi-direkt-ki-regulierung-muss-innovation-nicht-behindern-20230510).

123. See the press release from OFCOM dated 5 April 2023 (https://www.bakom.admin.ch/bakom/en/homepage/ofcom/ofcom-s-information/press-releases-nsb.msg-id-94116.html).

124. Available on https://www.sbfi.admin.ch/sbfi/en/home/eri-policy/eri-21-24/cross-cutting-themes/digitalisation-eri/artificial-intelligence.html.

125. https://cnai.swiss/en/.

126. See the statement of the Federal Council dated 26 April 2023 on intervention 23.3201 (https://www.parlament.ch/en/ratsbetrieb/suche-curia-vista/geschaeft?AffairId=20233201).

127. See Murat Karaboga, Nula Frei, Frank Ebbers, Greta Runge, Michael Friedewald, Daniel Vogler, Adrian Rauchfleisch, Manuel Puppis and Patrick Raemy, "Deepfakes und manipulierte Realitäten", TA-SWISS (forthcoming, 2024).

128. See Murat Karaboga, Nula Frei, Frank Ebbers, Greta Runge, Michael Friedewald, Daniel Vogler, Adrian Rauchfleisch, Manuel Puppis and Patrick Raemy, "Deepfakes und manipulierte Realitäten", TA-SWISS (forthcoming, 2024).

129. French version: https://www.ta-swiss.ch/fr/chatgpt, German version: https://www.ta-swiss.ch/chatgpt.

AI (2020),[130] as well as studies on the effects of digitalisation on democracy (2021)[131] and facial and speech recognition (2022).[132] Further, the impact of generative AI on art and culture is also being discussed in our ongoing project on culture and digitalisation.[133]

Regarding TA-SWISS's role as adviser for political actors, the two MPs who replied to the EPTA questionnaire expressed the need for more information on the following issues in particular: evolution of practices in Switzerland (especially in companies, institutions and administrations); legal and technical aspects; results of ongoing research (e.g., on means of control, data traceability, transparency of algorithms).

More generally, TA-SWISS has received frequent requests for contributions on generative AI since the launch of ChatGPT, at public and academic events, in the media and in Parliament (i.e., for the Committee on science, education and research of the National Council). This highlights the added value of a TA perspective on the matter. Given the wide gap between the unusual speed of these technologies and the time needed to adopt potential measures against their risks, it seems crucial for TA to be able to react responsibly and proactively. Judging by the requests we have received, an interdisciplinary approach also appears to be essential, as questions relate to the technical workings of these technologies, the regulatory framework, their effects on society and the economy, and ethical issues alike. Last but not least, TA has a critical contribution to make for the democratic legitimation of the use and regulation of generative AI: by furnishing independent information that duly takes into account the many facets of its impact on society, TA can help decision-makers (both citizens and political actors) form an opinion and take considered decisions on these technological developments.

---

130. Markus Christen, Clemens Mader, Johann Čas, Tarik Abou-Chadi, Abraham Bernstein, Nadja Braun Binder, Daniele Dell'Aglio, Luca Fábián, Damian George, Anita Gohdes, Lorenz Hilty, Markus Kneer, Jaro Krieger-Lamina, Hauke Licht, Anne Scherer, Claudia Som, Pascal Sutter and Florent Thouvenin, "Wenn Algorithmen für uns entscheiden: Chancen und Risiken der künstlichen Intelligenz", TA-SWISS (2020, https://www.ta-swiss.ch/en/artificial-intelligence).

131. Urs Bieri, Edward Weber, Nadja Braun Binder, Sébastien Salerno, Tobias Keller and Manuela Kälin, "Digitalisierung der Schweizer Demokratie – Technologische Revolution trifft auf traditionelles Meinungsbildungssystem"; Nora Räss, Ira Differding and Jasmin Odermatt, "Jugend, politische Partizipation und Digitalisierung. Eine Analyse der digitalen politischen Partizipation junger Menschen in der Schweiz"; Anna Boos, Ramona Sprenger, Jeannie Schneider, Basil Rogger, René Odermatt, David Simon, "Szenarien zu Demokratie und Digitalisierung. Ein partizipatives Zukunftsexperiment für die Schweiz", TA-SWISS (2021, https://www.ta-swiss.ch/digitale-demokratie)

132. Murat Karaboga, Nula Frei, Frank Ebbers, Sophia Rovelli, Michael Friedewald and Greta Runge, "Automatisierte Erkennung von Stimme, Sprache und Gesicht. Technische, rechtliche und gesellschaftliche Herausforderungen", TA-SWISS (2022, https://www.ta-swiss.ch/en/speech-speaker-facial-recognition).

133. Stefano Kunz, Jens Meissner and Ramona Sprenger, "Kunst & Digitalisierung in Musik, Theater und visuellem Design", TA-SWISS (forthcoming, 2024, provisional title, https://www.ta-swiss.ch/en/culture-and-digitalisation).

United Kingdom

EPTA Report 2023 – Parliamentary Office of Science and Technology

Oliver Bennett MBE, Head of POST
Devyani Gajjar, Physical Sciences and Digital Adviser

Generative AI and the UK job market, and wider parliamentary scrutiny of AI

What is the state of play in your country (development, deployment, use)?

Demand for specialist data skills

Across the UK workforce, there is increasing demand for specialist data skills, including skills for artificial intelligence, which have the potential to bring economic and social benefits.

Evidence suggests that availability of people with specialist data skills in the UK is insufficient to meet demand. A 2021 study found that in the UK, the supply of data scientists from universities was unlikely to exceed 10,000 per year, yet there were potentially at least 178,000 unfilled data specialist roles.

Research finds that certain groups (such as women, those from minority ethnic backgrounds and people with disabilities) are underrepresented in the data workforce. A lack of workforce diversity has the potential to amplify existing inequalities and prejudices.

Efforts to reduce the skills gap can be hindered by the inconsistent definition of data skills, organisational culture, the availability of specialist primary and secondary school teachers, and barriers to people moving between sectors.

Initiatives to increase the number of people with data skills include degree conversion courses, doctoral training centres for PhD students, online up-skilling platforms, apprenticeships, and visas to attract international talent.

Deployment and use of AI

The 'Artificial intelligence sector study 2022' published by the Office for Artificial Intelligence and the Department for Science, Innovation and Technology found that there were 3170 AI companies working in the UK which generated £10.6bn in AI-related revenue. From these companies, 60% are dedicated AI businesses and 40% have AI activity as part of a broader product or service offer. The study estimates that the sector employs 50,040 full time equivalents in AI-related role, 53% of which are within dedicated AI companies.

Development of foundation models (that underly recent generative AI systems)

Recent advances in generative AI systems that can be used for general purposes (such as ChatGPT) can be labelled as foundation models. These foundation models are developed using vast quantities of data and computing resources. The UK currently does not have sovereign capabilities in developing foundation models. In April 2023, the Prime Minister announced £100 million in funding for a Foundation Model Taskforce to 'ensure sovereign capabilities and broad adoption of safe and reliable foundation models.'

Potential impact of generative AI on the UK labour market

A June 2023 report by KPMG on 'Generative AI and the UK labour market' identified that 2.5% of overall tasks could be performed by AI and 40% of UK jobs could see some impact in the next decade. The report predicted half of the displacement impact of generative AI could be offset by the creation of new tasks within affected jobs. The report predicts the three main applications of generative AI in the labour market as classification and summary tasks (such as searching through documents), technical content creation (such as coding) and subjective works (such as drawing, presentation and marketing content).

What do your MPs think about it?

There has been significant interest in the UK Parliament, with several parliamentary debates and select committee reports. The discussion has included consideration of what needs to be done to ensure that the UK can benefit from new AI technology, while also managing the potential negative consequences for workers.

A 2022 inquiry by the House of Lords Science and Technology Committee found that there was a mismatch between the scale of the UK's STEM skills gap and the solutions posed by the Government. The Committee concluded that:

> There is a mismatch between the scale of the UK's skills gap and the solutions proposed by the Government, especially given the UK's ambition to be a science and technology superpower. The Government's policies are inadequate and piecemeal. Closing the gap requires aligning the high-level priorities of a number of departments; co-ordination is crucial so that the availability of a skilled workforce does not prevent the growth of STEM industries.

The Committee identified several actions for the Government to take, grouped into four areas:
- Immigration policy for STEM talent. For example, visa arrangements for attracting and retaining international talent.
- Quantifying and addressing the domestic skills gap. For example: providing courses below degree level to enable workers to retrain or grow new STEM skills; the publication of an assessment of the skills gap that specifies how policies and initiatives will contribute to resolving the gap; improvements to apprenticeships programmes.
- Recruiting and retaining science teachers and educators. For example: providing more pay to specialist teachers; improving teacher retention; supporting skilled professionals to become teachers.
- The precarity and attractiveness of STEM academic careers. For example, careers advice for PhDs and postdoctoral researchers; fellowship schemes to enable PhDs and postdocs to spend time in industry; improved working conditions and in particular action to create more longer-term research fellowships.

A 2023 House of Commons Business, Energy and Industrial Strategy Committee report highlighted the impact that AI could have on productivity within the UK. It referred to research from Deloitte that found that "by 2035 AI could boost UK labour market productivity by 25%", and that "Four out of five UK organisations said that use of AI tools had made their employees more productive, improved their decision-making, and made their process more efficient". It also made the point that AI and related technologies may have a positive impact on helping people access the labour market who have otherwise found it difficult to find and stay in employment, such as disabled people.

The Committee made several recommendations including:
- "The introduction of a requirement on businesses to conduct impact assessments to understand the scope and consequences of the use of new technologies in the workplace" and "for workers to have the right to consultation and notification where the application of technology in the workplace will result in the surveillance of a worker, or result in a significant change to their work".
- The introduction of a "new data poverty and digital skills strategy that will, amongst other things, set out how workers will be supported in the development of their
- digital skills".
- Work to address "a lack of investment in workplace training and reskilling.
- The need for regulators to recruit additional staff with technical expertise.
- That the Government "consult on an enforceable code of practice on the use of surveillance technology in the workplace".
- That the government create a taskforce "to assess the implications of the technology in the workplace, to consider whether enforcement of labour laws is effective, and to make recommendations on whether further legislation is required."

Is there any legislation in place?

AI is currently regulated through existing legal frameworks.

The Government published its position on AI regulation in March 2023: A pro-innovation approach to AI regulation. The report stated that its approach to AI was:

- Pro-innovation: enabling rather than stifling responsible innovation.
- Proportionate: avoiding unnecessary or disproportionate burdens for businesses and regulators.
- Trustworthy: addressing real risks and fostering public trust in AI in order to promote and encourage its uptake.
- Adaptable: enabling us to adapt quickly and effectively to keep pace with emergent opportunities and risks as AI technologies evolve.
- Clear: making it easy for actors in the AI life cycle, including businesses using AI, to know what the rules are, who they apply to, who enforces them, and how to comply with them.
- Collaborative: encouraging government, regulators, and industry to work together to facilitate AI innovation, build trust and ensure that the voice of the public is heard and considered.

The Government noted that if a gap were identified in the regulatory system, existing legislation would be adapted. However, it was clear that "initially, we do not intend to introduce new legislation" because "rushing to legislate too early, we would risk placing undue burdens on businesses".

The Prime Minister has announced he wants to make the UK "the geographical home of global AI safety regulation" and is planning to host a global summit on AI safety in November.

Role of TA in parliamentary debate

The Parliamentary Office of Science and Technology (POST) has provided considerable support to the UK Parliament's consideration of AI.

POST has produced publications commissioned by other parliamentary staff to support parliamentary inquiries and debate including on:

- Data Science Skills in the UK Workforce
- Digital technology in freight
- Automation in military operations
- How is artificial intelligence governed in Australia, France, Italy and Singapore?,

POST has also provided ad hoc support to committees and the libraries on their consideration of AI, including reviewing briefings and reports, proposing inquiry topics, contributing technical advice, and suggesting witnesses.

POST work has been cited in recent relevant parliamentary reports including:

- House of Commons Culture, Media and Sport Committee, Connected tech: smart or sinister? (HC 157), 7 August 2023
- House of Commons Culture, Media and Sport Committee, Connected tech: AI and creative technology (HC 1643), 30 August 2023
- Commons Library Research Briefing, Artificial intelligence and employment law, 11 August 2023
- Potential impact of artificial intelligence on the labour market - House of Commons Library
- Debate on Artificial Intelligence – House of Commons Library

POST is currently preparing further work on AI including on:

- Artificial Intelligence (AI) Technologies
- Artificial Intelligence in Weapon Systems