ARTIFICIAL INTELLIGENCE AND THE NEW CHALLENGES

Loredana MOCEAN*
Miranda-Petronella VLAD**

Abstract

This paper discusses the impact and legal framework of Artificial Intelligence (AI), with a focus on the use of AI-based generative technology and virtual assistants such as ChatGPT.

In today's technological context, generative AI is both an opportunity and a challenge in areas such as artistic creation, machine learning, and human-machine interaction. AI legislation varies by country and region, addressing issues such as accountability, transparency, safety and human rights.

The European Union (EU) and the United States have proposed or adopted legislative initiatives on AI, including the EU on Artificial Intelligence (AI Act). It regulates the use of AI in the EU, the stability of high requirements for systems with risk imposing obligations on developers and operators.

The benefits of using generative technologies include 24/7 availability, scalability, efficiency and customization. However, there are also disadvantages, such as a lack of empathy and understanding of the human context, the risk of spreading incorrect information, and data security and privacy concerns.

Despite these advantages, ChatGPT technology and virtual assistants based on generative language models are useful and beneficial in various contexts, and users should be aware of the limitations and challenges associated with their use.

Key words: *Artificial Intelligence, Generative, Risks, Responsible development*
JEL Classification: *[L63, L86]*

1. Introduction and legislative background

In the current age of technology, generative Artificial Intelligence (AI) has been a turning point, bringing with it both promise and great challenges in areas such as artistic creation, machine learning, and human-machine interaction. AI penetrates more and more into what is traditionally (Buning, 2018). AI legislation varies by country and region. In general, governments and international organizations are trying to develop regulations to address issues of accountability, transparency, safety and human rights in the context of the use of AI technologies. The law is not immune to the future of society (Pica, 2023).

* Associate professor, "Babeș-Bolyai" University, Cluj-Napoca, Romania, loredana.mocean@ubbcluj.ro

** Associate professor, "Dimitrie Cantemir" Christian University, Bucharest, Romania, miranda.vlad@cantemircluj.ro

Here are some examples of issues that AI legislation can regulate:

i. **Legal liability:** Laws can determine who is responsible for damages or negative consequences caused by AI systems and can regulate how they can be held liable.
ii. **Data protection and privacy:** Legislation may impose strict standards for the collection, storage and use of data in the context of artificial intelligence to protect individual privacy and rights.
iii. **Human rights and ethics:** Regulations may exist to ensure that AI systems respect human rights and ethical norms, such as avoiding discrimination and protecting inclusion.
iv. **Cyber safety and security:** Legislation may impose requirements to ensure the safety and security of AI systems to prevent unauthorized access or use for harmful purposes.
v. **Algorithmic transparency and accountability**: Laws can require transparency in AI decision-making processes and require explanations for the decisions made by these systems.
vi. **Oversight and regulation of military and law enforcement use:** There are concerns about the use of AI in military and law enforcement applications, and some countries may have specific regulations to regulate these areas

In 2021 and 2022, the European Union (EU) and other major jurisdictions, such as the United States, proposed or adopted legislative initiatives on artificial intelligence. It is important to monitor legislative and regulatory developments in this area, as they may influence how organizations and technology developers approach and implement AI solutions.

The Regulation on Artificial Intelligence (AI Act) proposed by the European Commission in April 2021 is a broad legislative framework that regulates the use of AI in EU. Here are some important aspects of this regulation:

i. **Category of high-risk AI systems:** The Regulation sets specific requirements for AI systems considered high-risk, such as monitoring and control systems, real-time facial recognition technologies and AI systems used in healthcare.
ii. **Obligations for developers and operators:** The AI Act requires developers and operators of high-risk AI systems to comply with strict requirements regarding transparency, accountability, as well as the quality and reliability of AI systems.
iii. **Compliance and assessment:** The Regulation establishes a framework for the assessment and certification of compliance of high-risk AI systems, as well as for the continuous surveillance of compliance after commissioning.
iv. **Penalties:** The AI Act provides for penalties for non-compliance, which may include substantial fines and other corrective measures.

v. **Supervisory authority:** It is proposed to create an AI supervisory authority at the EU level, which would ensure the implementation and enforcement of the regulation in the member states.

These are just a few key aspects of the AI Act Regulation. The proposed regulation is designed to address concerns about the safety, transparency and accountability of the use of AI technologies in the EU and to promote ethical and responsible development of artificial intelligence.

On 12 July 2024, the EU updated the AI Act with a reference regulation aimed at standardizing AI rules across the EU. The law prohibits certain AI practices and imposes regulations on high-risk AI systems, transparency-related AI, and general purpose AI (GPAI) models.

The regulations will be phased in: prohibited practices from 2 February 2025, GPAI obligations from 2 August 2025 and high-risk AI obligations and transparency obligations from 2 August 2026. However, there are exceptions for high-risk AI systems and GPAI models already on the market before these dates, postponing the application of the Law under certain conditions[1].

Here are also some examples of US state-level laws and initiatives related to the use of AI technologies in various fields:

1. **California**:
   - **Algorithm Transparency Act (AB-13):** Passed in 2019, this law requires government agencies to provide transparency about the use of algorithms in decision-making, particularly as it relates to the criminal justice system.
   - **AI-Assisted Healthcare Act (AB-1546):** Passed in 2018, the law regulates the use of artificial intelligence in healthcare systems, including transparency, certification and accountability requirements.
2. **Washington**:
   - **Justice Algorithm Analysis Act (HB-1655):** Passed in 2021, this law requires state law enforcement agencies to conduct analysis and provide transparency about the algorithms used in their decision-making process.
   - **Digital Rights Protection and Government Ethics Act (SB-5527):** Enacted in 2020, this law establishes requirements for transparency and accountability in the use of AI technologies by government agencies.
3. **Illinois**:
   - **Algorithm Transparency in Lending Act (HB-1815):** Enacted in 2021, this law regulates the use of algorithms in credit

[1]https://www.insideprivacy.com/artificial-intelligence/eu-artificial-intelligence-act-published/

decision-making and calls for transparency and accountability in the process.

There are many other state-level legislative initiatives in the US aimed at regulating the use of AI technologies in various fields. It is important to monitor legislative developments in each state to better understand the impact and implications of these regulations.

Regarding AI and rules for different levels of risk, there are various approaches and practices depending on the specific context and scope. Here are some important things to consider.

2. AI Act: different rules for different levels of risk

The new rules set obligations for providers and users depending on the level of risk from artificial intelligence. Although many AI systems present minimal risk, they must be evaluated.
The rapid development of AI indicates the absence of uniform approaches to assessing the risks and prospects associated with the use of AI (Matyuk, 2020).

Unacceptable risk

AI systems with unacceptable risk are systems considered a threat to humans and will be banned. They include:

- Cognitive-behavioral manipulation of individuals or vulnerable groups: voice-activated toys that encourage dangerous behavior in children, movies with a high degree of aggression
- Social score: classification of people according to behavior, socio-economic status or personal characteristics
- Biometric identification and classification of natural persons
- Real-time and remote biometric identification systems such as facial recognition

Legislatively speaking, the existence of several exceptions is desired, but these are also for the purpose of applying the law: remote biometric identification "in real time", for serious cases, and "subsequent", after the passage of a significant period of time, in the case of serious crimes.

High risk

A violation of fundamental rights is also identified in the following situations:

i. Violation of EU product safety legislation, here are included toys, aviation, cars, medical devices and elevators
ii. AI systems in eight specific areas that will have to be registered in an EU database:
    - Management and operation of critical infrastructure
    - Education and vocational training
    - Employment, human resources management
    - Access to independent activities

- Access and use of essential private services, as well as public services and benefits
- Law enforcement
- Management of migration, asylum and border control
- Assistance in the interpretation and application of the law.

High-risk components must be evaluated before they are placed on the market and the evaluation continues throughout their life cycle.

Limited risk

This type of risk requires certain minimum standards of transparency so that users can make informed decisions. Once the information is accessed, users may have the option to decide whether they wish to continue using it or not. The most conclusive examples are AI systems that produce or manipulate images, audio or video content (deepfakes).

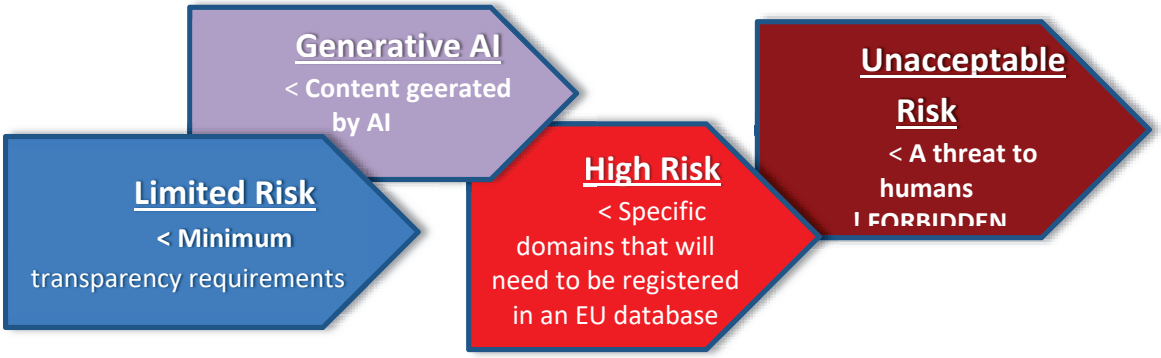

Figure 1. Risk categories

3. Use of AI-based generative technologies

More and more systems that we interact with without knowing, use AI techniques since 2015 and earlier (Garrido, 2015). The simplest applications in the generative domain are those of the chatbot type. They are software applications that aim to mimic human conversation through text or voice interactions. Chatbots are online, and they use artificial intelligence systems that are able to maintain a conversation with a user in natural language, simulating how a human would behave as a conversation partner.

Such technologies use aspects of natural language processing and deep learning processes. The most well-known processing models are Chat Generative Pre-trained Transformer (GPT-3, GPT-3.5 and GPT-4), the first being launched in 2022, by OpenAI, which is an artificial intelligence research laboratory headquartered in San Francisco, USA. They are capable of generating text like human interlocutors, and have numerous applications, including translations from different languages, language modeling and text generation for chatbots.

Immediately after its launch, ChatGPT was a huge success, and by March 2024 it had become the fastest growing consumer software application in history, reaching over 100 million users. At the same time, the OpenAI company ended up being valued at 29 billion dollars.

The free version of ChatGPT was called GPT-3.5, and could be accessed by anyone on the company's website by logging in.

Afterwards the company also launched a new version, ChatGPT-4 available only for a fee.

Chat GPT 4o, the newest version of the AI platform, has the ability to perceive emotions and even give math meditations.

Chat GPT is trained using a transformation algorithm and large amounts of text data and is able to generate responses to prompts or questions based on what it has learned.

ChatGPT uses a variant of the Transformer architecture, which is a type of neural network architecture specifically designed for processing sequential data such as text. The model is trained using a variant of Transformer called "Reinforcement Learning from Human Feedback" (RLHF).

Initially, it goes through a supervised fine-tuning process where human AI trainers provide conversations while playing both the role of the user and that of an AI assistant.

These conversations are used to train the model to predict the next response. The model is then fine-tuned using a technique called Proximal Policy Optimization (PPO) to optimize the generation of better answers based on comparison data from human AI trainers who rank different answers generated by the model.

4. Advantages of generative systems

There are several advantages associated with using ChatGPT, or virtual assistance technology based on generative language models such as this.

Here are some of them:

i. **24/7 Availability:** ChatGPT can be accessed anytime, anywhere and is available 24/7. It does not require breaks or free time and can be used to receive assistance or information at any time of the day.

ii. **Scalability:** The ability to handle a large number of simultaneous interactions is a major advantage of ChatGPT. It can serve a large number of users simultaneously without affecting the quality of responses or response time.

iii. **Efficiency:** Can quickly respond to user inquiries and provide information or support in an efficient manner. This can reduce the time it takes to find answers or solutions and improve the user experience.

iv. **Personalization:** Through repeated interaction with users, ChatGPT can learn their individual preferences and needs, providing more personalized responses and recommendations over time.

v. **Flexibility:** Can be configured and adapted to respond to a wide variety of questions and requests, from fields as varied as customer support, education, entertainment.
vi. **Low cost:** Deploying a ChatGPT-based virtual assistant can be more economical than hiring and training human staff to provide support or information to customers.
vii. **Accessibility:** Being available online, ChatGPT is accessible to anyone with an internet connection and a compatible device, without the need to be physically present at a specific place or time.
viii. **Improving customer experience:** By providing quick and accurate responses, ChatGPT can help improve customer experience and increase customer satisfaction with a product or service.
ix. **Accessibility and easy of use:** ChatGPT is accessible to anyone with an internet connection, being easy to use on different platforms.
x. **Prompt responses and 24/7 availability:** ChatGPT can provide prompt responses and is available 24/7 without the need for human intervention to obtain information or assistance.
xi. **Continuous learning capability:** Due to its ability to learn from a large volume of data and interactions, ChatGPT can improve the quality and accuracy of its responses.
xii. **Various applications:** ChatGPT can be used in a variety of fields and applications including customer support, education, entertainment and more.
xiii. **Saving time and resources:** Using ChatGPT can save time and resources for organizations and individual users by eliminating the need for human interaction for certain tasks or repetitive questions.
xiv. **Customization and adaptability:** ChatGPT can be tailored and customized to meet the specific needs and preferences of users or organizations.
xv. **Privacy and anonymity:** Users can interact with ChatGPT without revealing their identity or other personal information, providing an increased degree of privacy and anonymity.
xvi. **Multi-Language support:** ChatGPT can respond in a variety of languages, facilitating communication and interaction with users from different regions of the world.

These are just some of the advantages associated with using ChatGPT or other virtual assistants based on generative language models. However, it is also important to consider the limits and challenges associated with their use, such as the need to ensure the accuracy and relevance of responses, as well as the protection of user data and privacy.

5. Disadvantages of generative systems

Although ChatGPT technology and virtual assistants based on generative language models have many advantages, there are also some disadvantages and challenges associated with their use:

- **Lack of empathy and understanding of human context:** ChatGPT cannot feel emotions or fully understand the emotional or social context of its interlocutor, which can lead to responses that are perceived as lacking empathy or insensitive in certain situations.
- **Limitations in accuracy and consistency of responses:** Although ChatGPT is trained on large and diverse data sets, it may sometimes provide incorrect or inconsistent responses, especially for more specialized or contextual information.
- **Risk of propagation of incorrect information or bias:** If the model is trained on datasets containing incorrect information or biases, there is a risk that these will be propagated in the responses provided by ChatGPT.
- **Data security and privacy:** The use of virtual assistants may involve the transmission of sensitive or personal information, which raises concerns about the security and privacy of such data.
- **Dependence on technology and loss of human contact:** With prolonged use of virtual assistants, there may be a risk that people may lose touch with real human interactions and become too dependent on technology for problem solving or social interaction.
- **Need for supervision and management:** Virtual assistants like ChatGPT require constant supervision and management to ensure that the responses provided are appropriate and that interactions with users are positive and helpful.
- **Limitations in ability to learn and adapt:** Although ChatGPT can learn from repeated interactions with its users, it may have difficulty adapting to rapid changes or learning from less frequent or unusual interactions.

Despite these drawbacks, ChatGPT technology and generative language-based virtual assistants can still be extremely useful and beneficial in a variety of contexts, provided that users are aware of the limitations and challenges associated with their use and take appropriate measures to to manage them.

There are disadvantages associated with ChatGPT technology and virtual assistants:

- **Inability to provide non-verbal support:** ChatGPT relies solely on text to communicate and cannot provide non-verbal support or non-verbal signals such as facial expressions, tone of voice or body language, which may limit its ability to communicate effectively in certain situations.
- **Risk of abuse or misuse:** ChatGPT may be misused or abused to generate inappropriate content, spam, or manipulate or mislead users, which may have negative consequences for their experience.

- **Limitations in anticipating user needs:** may have difficulty anticipating user needs or questions in advance and providing useful information before it is explicitly requested.
- **Dependence on quality of training and input data:** The performance of ChatGPT can be affected by the quality of training data and data sets used to improve the algorithm. If the data is incomplete, inaccurate or biased, this may affect the quality and relevance of the answers provided.

These are some of the additional disadvantages associated with ChatGPT technology and virtual assistants based on generative language models. It is important to consider these aspects and take appropriate measures to manage and minimize their impact in practical implementations.

In terms of human development, excessive use or dependence on generative technologies can present some disadvantages:

- **Reduction of interpersonal communication skills:** The individual's interpersonal communication skills may be diminished. Face-to-face interaction is crucial for developing skills such as expressing emotions, interpreting non-verbal signals and developing empathy.
- **Decreased focus and concentration:** Digital technologies, including virtual assistants, can compete with other important activities that require focus and attention, such as studying, work or social interactions. This can lead to a decrease in the level of attention and concentration.
- **Increased social isolation:** Dependence on technology can lead to social isolation.
- **Reduced problem-solving skills:** Dependence on technology to find quick answers can undermine an individual's problem-solving and critical thinking skills. This can hinder the development of the ability to analyze, synthesize and evaluate information independently.
- **Increasing dependence on technology:** Excessive use of AI can lead to dependence on it, which can have a negative impact on an individual's mental and physical health. It can lead to anxiety, depression and other mental health problems.
- **Limiting creativity and imagination:** It is important to notice how these technologies can affect different aspects of life and take steps to use technology responsibly and in a balanced way.

Conclusions

AI tools such as ChatGPT and OpenAI among others, are likely to transform the way we live, work, and interact with each other (Ajami, 2023).

Our paper provides a broad look at the impact of legislation and regulation on AI, focusing mainly on the European Union and the USA. It highlights the rapid evolution of AI legislation in various countries and regions.

Legislative initiatives, such as the AI Act proposed by the European Commission, reflect growing concerns about the accountability, transparency and safety of the use of AI technologies.

AI legislation covers issues such as legal liability, data protection and privacy, human rights and ethics, cyber security and algorithmic transparency. These aspects are essential to promote an ethical and responsible development of artificial intelligence.

The importance of different regulation depending on the level of risk of AI systems is highlighted. High-risk systems are subject to more stringent requirements, while low-risk systems are subject to minimum transparency requirements.

Also, the paper describes the usefulness and positive impact of generative technologies, such as ChatGPT, in various fields. These technologies offer a number of advantages, including 24/7 availability, scalability, efficiency and personalization, but they also come with certain disadvantages, such as lack of empathy and data security.

Excessive use or dependence on technologies such as ChatGPT can have negative effects on human development, including reduced interpersonal communication skills, decreased attention span, and increased dependence on technology.

In conclusion, the paper highlights the importance of regulation and responsible use of AI, while also recognizing the benefits and challenges associated with AI technologies, including generative technologies such as ChatGPT. The Regulation (EU) 2024/168 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on Artificial Intelligence will apply from 2 August 2026 with some exceptions (cf. Art. 133 AI Act)[2].

Bibliography:

Articles

1. Buning, M., 2018, ”Artificial intelligence and the creative industry : new challenges for the EU paradigm for art and technology by autonomous creation”, *Research handbook on the law of artificial intelligence*, Northampton, Edward Elgar Publishing, pp. 511-535
2. Garrido, A., 2015, ”New Challenges of Artificial Intelligence”, *Conference: ICTAMI 2015*
3. Pica, L., 2023,”The New Challenges of Artificial Intelligence, Profiling and Bigdata Analysis by Tax Administrations: Will the Right to Meet these New Challenges Be Shown?” *JusGov Research Paper Series*, No. 2023-04
4. Matyuk, Y., 2020,” Artificial Intelligence: New Challenges and Prospects”, *Advances in Law Studies* 8(5), pp. 42-48

[2] https://ai-act-law.eu/

5. Riad A. Ajami, Homa A. Karimi, 2023, ”Artificial Intelligence: Opportunities and Challenges”, *Journal of Asia-Pacific Business*, Volume 24, 2023 - Issue 2, pp. 73-79

Web resources

https://ai-act-law.eu/

https://www.insideprivacy.com/artificial-intelligence/eu-artificial-intelligence-act-published/

https://www.europarl.europa.eu/topics/ro/article/20230601STO93804/legea-ue-privind-ia-prima-reglementare-a-inteligentei-artificiale

https://cursdeguvernare.ro/ce-poate-face-chat-gpt-12-moduri-in-care-poate-fi-utilizat.html

https://www.namepepper.com/chatgpt-users

https://playtech.ro/stiri/ce-este-chat-gpt-si-cum-functioneaza-de-fapt-690862