sciendo

Volker Brühl

Generative Artificial Intelligence – Foundations, Use Cases and Economic Potential

A key technology driving digital transformation is artificial intelligence (AI), which has gained a sudden momentum and a high degree of public attention with the initial release of the chatbot ChatGPT (where GPT stands for generative pretrained transformer) on 30 November 2022. ChatGPT is an AI-based application developed by Open AI that anyone can experiment with and use. Despite lots of deficiencies, ChatGPT demonstrates the capabilities of an AI-based application for everybody. Within a very short period, ChatGPT has triggered discussions about the economic and social impact of AI and whether the rapid diffusion of AI technologies is desirable. Nevertheless, it is quite apparent that AI is and will continue to evolve into one of the key meta technologies that will shape the future of industry and society in the coming years, if not decades.

This article provides a brief overview of the foundations of generative AI (GAI), including machine learning, important players in this emerging market, possible use cases and the economic potential as of today. Finally, it looks at the current status of the forthcoming European Artificial Intelligence Act (AI Act), which will be an important milestone in developing a regulatory framework for trustworthy AI in Europe and beyond.

Machine learning – A brief overview

There are a number of quite diverse definitions for the term "artificial intelligence". For instance, it may be understood as a generic term for technologies and systems with the ability to perform tasks that would otherwise require human intelligence. This presupposes certain skills that can be roughly broken down into "perceiving", "reasoning/decision-making", "acting" and "learning" (e.g. Russell and Norvig, 2022). According to these basic generic elements of an AI system, AI is often divided into different fields of technologies such as "speech recognition", "image/video recognition", "natural language processing", "computer vision" and "robotics". The perceived structured or unstructured data has to be transformed into knowledge ("knowledge representation"), forming the basis to find the optimal solution in a given environment ("knowledge reasoning and optimisation"). Performing a physical act (e.g. through actuators of robots) or providing digital information (e.g. text) usually leads to a change of the environment, which may feed into a learning process of the respective AI system. Hence, a key element of AI systems is usually a type of machine learning. Machine learning (ML) is a subset of AI that seeks to enable machines to perform tasks in an optimised way through experience (Mitchell, 1997).

For instance, ML is supposed to improve decision-making, forecasting or classification problems (see e.g. Murphy, 2022; Russell and Norvig, 2021). In data science, the concept of machine learning involves using statistical learning and optimisation methods that let computers analyse datasets and identify patterns (UC Berkeley, 2022).

It should be noted that ML systems do not operate on programmed solution algorithms but build a model themselves by learning from the available data. Machine learning algorithms ensure that the results of actions or changes in environmental conditions are used to optimise the system's performance in an iterative process. Hence, models based on machine learning must be trained and tested with given data sets before they can reasonably be used to draw conclusions from new data. ML is applied in many fields such as speech, image or natural language recognition. It may also be integrated into robotics, smart factories or smart home applications. Depending on the data availability and method of training, two basic learning categories – "supervised learning" and "unsupervised learning" – can be distinguished (see e.g. Goodfellow et al., 2016; Mitchell, 1997; Murphy, 2022).

Supervised learning refers to the development of a prediction model that is trained with the help of given input and known output data. By comparing the model-based predictions with the correct outputs, prediction errors can be identified and the model can gradually be optimised. As the volume of training data increases, prediction errors are reduced. Learning algorithms used in the context of supervised learning include decision tree techniques, regression analysis, support vector machines, discriminant analysis and k-nearest neighbours algorithms. These

Volker Brühl, Center for Financial Studies, Frankfurt am Main, Germany.

algorithms can be used for tasks such as classification or forecasting. Common use cases include spam filters, fraud detection systems, recommendation engines and speech or image recognition systems. Many virtual assistants, such as Apple's Siri or Amazon's Alexa, are trained with supervised learning algorithms to communicate with users through a natural language interface.

In unsupervised learning settings, algorithms analyse unlabelled data. The objective is to discover hidden patterns, relationships, or structures within the data. Common unsupervised learning techniques include Bayesian networks, hidden Markov models and clustering algorithms such as K-means or hierarchical clustering. Typical applications are clustering processes, e.g. for customer segmentation or in healthcare to gain a better understanding in diagnosis, prevention and treatment of diseases.

"Reinforcement learning" is a further category of machine learning in which intelligent agents attempt to optimise an output through a given incentive system. A particular form of reinforcement learning has been used in GAI by integrating human feedback during the training phase (reinforcement learning from human feedback).

Self-supervised learning refers to a category of machine learning in which the system trains itself by first structuring the given unlabelled data, then using this structure to further optimise the output of another task. As such, this process involves transforming the unsupervised problem into a supervised problem by auto-generating the labels.

Artificial neural networks (ANN) are used as a form of machine learning, both in the field of supervised and unsupervised learning. The way that the human nervous system transmits information serves as the conceptual model for ANN. Like biological neural networks, ANNs consist of a large number of artificial neurons (units). ANNs can be classified according to various criteria, such as the number of hidden layers. Conventional ANNs contain only a few hidden layers, while so-called deep ANNs contain numerous hidden layers. These ANNs, also known as deep learning models, require a high volume of training data in order to achieve a satisfactory prediction quality due to the large number of connections between the neurons. Different learning algorithms are used in deep learning, such as convolutional neural networks, long short-term memory networks, recurrent neural networks, generative adversarial networks, multilayer perceptrons, deep belief networks or restricted Boltzmann machines. Artificial neural networks are now being used or tested in many different fields. These include, for example, natural language processing and speech or image recognition. Common business applications include quality management, pro-

duction or sales planning, maintenance processes and credit rating systems. ANNs also play a role in research and development, e.g. in autonomous driving or biotechnology.

Generative AI – What sets it apart?

Generative AI is a relatively new field of AI that has gained considerable public attention, especially since the release of ChatGPT in November 2022, which is easy to use and demonstrates the power of generative pretrained transformer systems (Cao et al., 2023). GAI is a category of AI systems that is capable of generating new text, images, videos or programming code in response to instructions (prompts) entered by the respective user. GAI applications such as the chatbot ChatGPT (OpenAI/Microsoft) are based on the large language model (LLM) GPT-3/GPT-4. This is similar to alternative solutions including Bard (Google) that use their foundation model LaMDA (Language Model for Dialogue Applications). LLMs are deep learning networks trained on huge amounts of text data to understand the syntax and semantics of human languages in different contexts. GAI systems apply "transformer" models, which are advanced LLMs using the "attention mechanism" to help identify the most informative parts in a text by perceiving associations and meanings of words and sequences of words.

Compared to traditional LLMs, transformers get a better and faster understanding of a task by analysing all words in a text simultaneously (rather than sequentially) while dynamically adjusting the "attention" (i.e. the perceived relevance of the word) during the solution process (Vaswani et al., 2017). The amount of training data and the number of parameters the neural network attempts to optimise are important factors for the performance of the system, for example in terms of comprehending text, generating answers to given questions, recognising images or generating code. For instance, GPT-3 is trained on 570 GB of text and can optimise up to 175 billion parameters to solve a specific task (Brown et al., 2020). Recent breakthroughs in GAI have been made possible by advances in computer hardware, especially in graphics processing units, enabling massive parallel processing of data in machine learning models.

Competitive landscape and expected market potential of generative AI

Table 1 provides a brief overview of the current important players in the field. Although this is only a snapshot that will certainly change in the coming years, it is striking that a substantial portion of the players belong to Big Tech companies. Most of these have their own GAI ac-

Table 1
Competitive landscape in generative AI (selection)

Category	Tool	Company	Country
Chatbots	ChatGPT	OpenAI (Microsoft)	USA
	Bard	Google	USA
	Bing AI	Microsoft	USA
	HuggingChat	Hugging Face	USA
	Jasper AI	Jasper	USA
	ChatFlash	neuroflash	Germany
Large language models	GPT-4, GPT-3.5	OpenAI (Microsoft)	USA
	Claude 2	Anthropic	USA
	Luminous	Aleph Alpha	Germany
	cohere command	Cohere	Canada
	PalM2	Google	USA
	LLaMA	Meta	USA
Code generation	AlphaCode	DeepMind (Google)	USA
	GitHub Copilot	GitHub(Microsoft)	USA
	Tabnine	Tabnine	Israel
	Open AI Codex	Open AI	USA
	Codebase	MutableAI	USA
	Replit AI	Replit	USA
	Codacy	Codacy	Portugal
Image/Videos	DALL-E2	OpenAI (Microsoft)	USA
	Imagen	Google	USA
	Stable Diffusion	Stability AI	UK
	Synthesia	Synthesia	UK
	Midjourney	Midjourney	USA
	Openart	OpenArt	USA

Source: Author's own analysis.

Table 2
Use cases of generative AI

Modality	Fields of application	Use cases (examples)
Text/Audio	Customer communication	Content writing, personalised customer care, sales, marketing
	Data Analytics	Customer behaviour, segmentation, profiling
	Virtual Assistants	Sales support, technical support, process documentation
	Publishing	Editorial services, translations, creative writing
Images/Videos	Image recognition	Face recognition, cybersecurity services
	Image generation	Marketing, PR, Non-Fungible-Tokens (NFT)
	Video generation	Media production, PR
	Cross Media	E-commerce, E-learning
Code	Code generation	Agile software development/ engineering
	Code optimisation	Rapid prototyping, accelerated system integration
	Code testing	Quality assurance
3D	Virtual Reality	Product development with digital twins
	Augmented Reality	Industrial maintenance
	Metaverse	Gaming, new work, social metaverse

Source: Author's own analysis.

tivities, but they have also added to their portfolio through acquisitions of upcoming start-ups in the field. Examples include the investment of Microsoft in OpenAI and the takeover of DeepMind by Google.

It can be expected that further takeovers of promising start-ups on the list will follow. This is another warning signal underlying the severe antitrust issue related to the dominant market position of Big Tech companies and their deep pockets (see e.g. Brühl, 2023). European companies are, with a few exceptions, not among the leading players in a field that will most likely be a major area of growth in the next decade. This becomes especially clear when we take a look at a list of important use cases that can already be identified and which covers basically all industries and many components of their value chain (Table 2).

Potential fields of application cover not only the automation and acceleration of homogeneous workflows (e.g. text generation and processing in administrative functions). GAI can also support creative and technically complex activities, e.g. in IT, marketing or product development. Hence, the diffusion of GAI may lead to redundancies of low-skilled employees while at the same time the need for highly skilled personnel will further increase. Research on potential productivity enhancements through the use of GAI tools is still at an early stage. However, a recent study by MIT researchers found that the productivity of skilled employees could be improved by up to 37% when using chatbots such as ChatGPT in their daily writing routines (Noy and Zhang, 2023). Therefore, the smart adoption of AI tools could help to mitigate the shortage of skilled workers, especially in ageing societies. At the same time, human-machine interaction will become a key part of many people's work, including in corporate functions such as accounting, controlling or human resources, which were considered to be less affected than, e.g. product development, industrial manufacturing, logistics or maintenance.

At this early stage of development, it is very hard to seriously estimate the future market potential of GAI as it is

Figure 1
Market potential generative AI until 2030
in bn USD
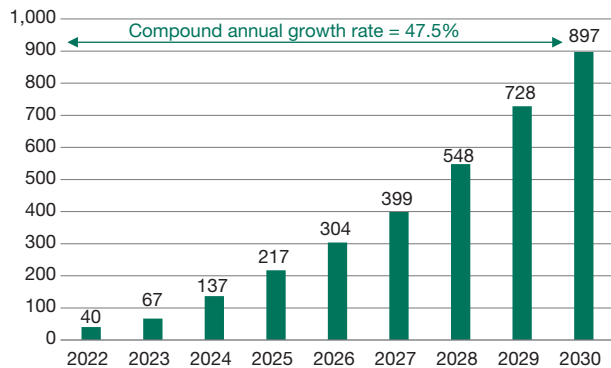
Source: Bloomberg Intelligence (2023), own calculations.

currently unclear how quickly new use cases will be created and how soon users will be ready to adopt them in their private or professional environment. Nevertheless, if we look at the recent publication of Bloomberg Intelligence (2023), we see that the market may experience enormous growth (Figure 1).

The expected annual growth rate of approximately 47.5% until 2030 covers both incremental revenues from GAI infrastructure (e.g. AI servers, AI storage solutions, computer vision and conversational AI devices) as well as specialised GAI assistant software.

Regulatory framework

The European AI Act (European Commission, 2021) was published in its final draft by the European Commission as early as 2021. After long and intensive discussions, a political agreement was reached on 8 December 2023. Therefore, it can be expected that the EU AI Act will soon be officially adopted by the EU Parliament and the EU Council followed by the publication in the EU's Official Journal to enter into force. It should be noted that the majority of the Act's provisions will apply after a two-year grace period with some exceptions particularly with regard to prohibitions and foundation models.

The forthcoming AI Act will establish a European regulatory framework for providers and users of AI systems. The obligations will depend on the level of risk associated with the respective systems. The regulation will differentiate between uses of AI that create an unacceptable risk, a high risk, a limited risk and a low or minimal risk. Some AI systems may fall into the prohibited "unacceptable risk" category as they could potentially harm or disadvantage people, e.g. through cognitive behavioural manipulation, social

scoring or (with very few exceptions) real-time biometric identification systems in publicly accessible spaces.

Another category of AI systems shall cover "high risk" applications that could negatively affect safety or fundamental human rights. Such AI systems include AI technologies used in critical infrastructures, safety components of products, educational or vocational training, certain aspects of employment, essential private and public services (e.g. credit scoring) or border control management. The approval of such high-risk AI systems will be subject to strict obligations, including

- adequate risk assessment and mitigation systems

- high quality of the datasets feeding the system to minimise risks and discriminatory outcomes

- logging of activity to ensure traceability of results

- detailed documentation providing all information necessary on the system and its purpose for authorities to assess its compliance

- clear and adequate information to the user

- appropriate human oversight measures to minimise risk

- high level of robustness, security and accuracy.

Operators of high-risk AI systems need to undergo a conformity assessment procedure conducted by competent institutions and, if needed, implement corrective actions before putting the system into service. Each member state has to designate or establish at least one national competent authority to serve as the national supervisory authority, the notifying authority and the market surveillance authority.

However, most GAI applications, such as chatbots, will probably be categorised as "limited risk" AI systems, which are primarily subject to transparency obligations. Hence, users have to be made aware when they are interacting with a machine so they can decide whether or not to continue using the AI system. In addition, the AI system must be designed in such a way that it prevents the generation and diffusion of illegal content. Finally, AI systems considered to be of minimal or no risk, such as many gaming apps or spam filters, will largely remain outside the scope of the AI Act. Many AI systems currently used in the EU will fall into this category.

It should be noted that, in particular, the regulatory treatment of foundation models (e.g. large language models)

underlying many downstream AI applications has been the subject of much debate. Leading AI scientists have demanded in an open letter of concern (KIRA, 2023) to the German government that foundation models need to be included in the AI Act due to their potentially high safety risks. Alternative proposals have advocated for leaving foundation models outside the scope of the AI Act and addressing their risks with a system of self-regulation rules.

The political agreement achieved stipulates a differentiated approach to the regulatory requirements of foundation models that belong to the category of general-purpose systems or models (GPAI). All providers of GPAI have to adhere to specific transparency requirements on training data and technical documentations. If GPAI are designated as encompassing systemic risk, additional obligations like e.g. model evaluations, risk assessments, testing and reporting have to be fulfilled.

Conclusions

Given that Europe lags behind in many areas of digital technology (e.g. AI in general, digital platforms, blockchain technology or cloud computing), it is worrying that once again the European corporate sector is about to miss out on an important digital technology. The potential risks include becoming even more dependent on the well-known Big Tech companies, as they will certainly combine GAI solutions with their existing technology franchise, e.g. cloud providers, search engines or social media platforms. Furthermore, we have to be aware that AI in general may dramatically change our way of creating, processing and consuming content of any kind. The already significant potential to distribute "fake news" as well as discriminatory, racist or manipulative content will be multiplied by these new technologies. Besides, the chances of prosecuting over illegal content on social media, for example, could become slimmer when the producer or owner of such harmful content is a machine. The enhanced risks of hidden plagiarism and infringements of intellectual property rights are obvious.

This is not to argue against investing heavily in these new applications of GAI, but we must not neglect the technological, social and political risks associated with them while harnessing the economic benefits. It is crucial that the regulatory framework keeps up with the accelerated development of digital technologies. Therefore, the upcoming AI Act, together with the recently adopted Digital Markets Act and the Digital Service Act, will be an important milestone in developing a regulatory framework for trustworthy AI in Europe and beyond. As AI becomes more advanced, humans face the challenge of comprehending how the algorithm came to a certain result. Many AI applications lack transparency in the internal calculation processes, making the generated results more or less a "black box" for the users. However, if even the data scientists who create the algorithms are not able to precisely understand and explain the AI processes applied, trust in the output and the controllability of the system is diminished. Trustworthiness and explainability of AI systems are essential features in order to establish, monitor and enforce a regulatory framework for AI. Without these characteristics, organisations cannot adopt a responsible approach to AI development that requires the system to be compliant with the regulatory standards and ensures that those affected by a decision are able to challenge or change that outcome. Hence, the AI Act shall support European start-ups and established technology firms in participating in these fast-growing markets, enabling society to benefit from these new technologies, while simultaneously avoiding irresponsible risks, ensuring public safety and preserving human rights. The implementation of transparent and clear guardrails for AI systems that balance opportunities with the elimination or mitigation of risks could be a chance for the European technology sector to catch up with its global competitors.

References

Bloomberg Intelligence (2023, 1 June), Generative AI to Become a $1.3 Trillion Market by 2032, Bloomberg announcement.

Brown, T., B. Mann, N. Ryder, M. Subbiah, J. D. Kaplan, P. Dhariwal, A. Neelakantan, P. Shyam, G. Sastry, A. Askell, S. Agarwal, A. Herbert-Voss, G. Krueger, T. Henighan, R. Child, A. Ramesh, D. Ziegler, J. Wu, C. Winter, C. Hesse, M. Chen, E. Sigler, M. Litwin, S. Gray, B. Chess, J. Clark, C. Berner, S. McCandlish, A. Radford, I. Sutskever and D. Amodei (2020), Language models are few-shot learners, *Advances in neural information processing systems*, 33, 1877-1901.

Brühl, V. (2023), Big Tech, the Platform Economy and the European Digital Markets, *Intereconomics*, 58(5), 274-282, https://www.intereconomics.eu/contents/year/2023/number/5/article/big-tech-the-platform-economy-and-the-european-digital-markets.html (2 January 2024).

Cao, Y., S. Li, Y. Liu, Z. Yan, Y. Dai, P. S. Yu and L. Sun (2023), A Comprehensive Survey of AI-Generated Content (AIGC), A History of Generative AI from Gan to ChatGPT, *arXiv*, Cornell University.

European Commission (2021), Regulation of the European Parliament and of the council laying down harmonised rules on Artificial Intelligence and amending certain union legislative acts, COM(2021) 206 final.

Goodfellow, I., Y. Bengio and A. Courville (2016), *Deep Learning*, MIT Press.

KIRA Center for AI Risks & Impacts (2023, 28 November), The EU AI Act needs Foundation Model Regulation, https://www.foundation-models.eu/ (18 January 2024).

Mitchell, T. M. (1997), *Machine Learning*, McGraw-Hill.

Murphy, K. P. (2022), *Probabilistic Machine Learning: An Introduction*, MIT Press.

Noy, S. and W. Zhang (2023), Experimental Evidence on the Productivity Effects of Generative Artificial Intelligence, Working paper.

Russell, S. J. and P. Norvig (2022), *Artificial Intelligence – A Modern Approach*, 4th edition, Pearson.

UC Berkeley (2022), What is Machine Learning?, https://ischoolonline.berkeley.edu/blog/what-is-machine-learning/ (14 July 2023).

Vaswani, A., N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser and I. Polosukhin (2017), Attention Is All You Need, *Advances in Neural Information Processing Systems*, 30.