

The rise of Generative AI and possible effects on the economy

 Tim Orchard¹

 Leszek Tasiemski²

Abstract

The aim of the paper is to analyse the likely implications of Generative AI (GAI) on various aspects of business and the economy. Amid the rapid growth and maturing of Generative AI technologies such as Large Language Models (like ChatGPT by OpenAI) a rapid growth of both immediate and potential applications can be seen. The implications for the economy and industries of this technological shift will be discussed. The foreseeable scenarios for the level and types of adoption that GAI might achieve—from useful analytical tool, invaluable assistant to the white-collar workers of the world to being trusted with a wide array of business and life-critical decision making. Both disruptive and premium service opportunities are foreseen. For instance, general purpose models may provide quality service—such as copy-writing—to overserved customers leaving human writers as the premium option. In this context, overserved customers would be those who would be satisfied with a non-human, potentially less creative content. On the other hand highly specialized models—specifically trained in a given domain and with access to proprietary knowledge can possibly provide a premium service over that provided by human experts. It is expected that some jobs will be replaced by new AI applications. However, new workplaces will emerge. Not only the obvious expert-level data scientist roles but also low grade, “model supervisors”—people training the models, assessing the quality of responses given and handling escalations. Lastly new cybercrime risks emerging from the rise of GAI are discussed.

Keywords

- disruptive technology
- artificial intelligence (AI)
- Large Language Models (LLM)
- Generative AI
- business models

JEL codes: L86 , C88, L21, L26, L84, M13

Article received 29 April 2023, accepted 16 June 2023.

¹ WithSecure (WITH.HE), 77 Weston Street, London SE1 3RS, United Kingdom, tim.orchard@withsecure.com, <https://orcid.org/0009-0004-4779-7843>.

² WithSecure (WITH.HE), ul. Rataje 164, 61-168 Poznań, Poland, corresponding author: leszek.tasiemski@withsecure.com, <https://orcid.org/0009-0002-8039-4881>.

Suggested citation: Orchard, T., & Tasiemski, L. (2023). The rise of Generative AI and possible effects on the economy. *Economics and Business Review*, 9(2), 9–26. <https://doi.org/10.18559/ebr.2023.2.732>



This work is licensed under a Creative Commons Attribution 4.0 International License
<https://creativecommons.org/licenses/by/4.0>

Introduction

The aim of this article is to analyse the potential of business applications of the rapidly emerging Generative AI technology and the ongoing and foreseeable influence of it on the economy and various sectors of industry. Shortcomings and risks are also discussed—ranging from the technical, to legal and security and privacy matters. The authors believe that the discussed technology has recently reached a tipping point that will enable a rapid and broad adoption, possibly on a scale that can be called the next industrial revolution. Many indicators suggest that the maturity of AI technology is now past the tipping point of broad adoption in core value streams and critical processes.

The paper is structured as follows. In Section 1 the concept of Generative AI is described along with historical background and limitations of today's state of same. Section 2 discusses possible implications of GAI on the job market with both a creative and destructive potential. In Section 3 the GAI technology is analysed from the perspective of offering new, disruptive business models. Section 4 is dedicated to risks stemming from the use of GAI technology both by legitimate and by malicious actors. In Section 5 the legal implications of GAI are discussed.

1. The Generative AI (GAI)

The technology collectively called artificial intelligence (AI) or machine learning (ML), is a very broad term. It covers many kinds of computer science techniques where the algorithm (a model in AI terminology) is programmed in such a form, where a given task such as grouping of similar items or finding anomalies is achieved in a different way than an explicit procedure and exact matching to a human-defined pattern (Korzynski et al., 2023). Typically it would mean that during the process called training of the algorithm, the model automatically deduces the pattern and is able to generalize despite the lack of exact human-made rules.

As such these techniques are not new. They have been hypothesized as early as in the 1950s (Turing, 1950) or even 1940's as "A logical calculus of the ideas immanent in nervous activity" which was presented in 1943 by Walter Pitts and Warren McCulloch and Pitts. The first implementations of self-improving algorithms happened in the 1950s and 1960s. However, up until recently those algorithms were not mainstream mostly due to a lack of computing power which made them slow and less convenient than traditional, procedural algorithms.

In recent years an exponential growth of Generative Artificial Intelligence (GAI) technology and its applications can be observed. GAI using usually very big models (GPT-4, released by OpenAI in 2023 uses 170 trillion parameters) trained on vast libraries of text, pictures and other inputs and is a technology capable of generating novel, "creative" content in response to an input (prompt). These algorithms can be interacted with using a natural language and typically are sophisticated statistical models. In most of the machine learning models knowledge is represented internally as a set of features. Features are parameters that are considered and calculated in the process of training the algorithm to achieve a sufficient level of generalization. Generalization is what makes AI models stand apart from the traditional, procedural approach. Such algorithms can respond correctly even to an input that was not part of the training set. Usually this is because the features of the model (parameters) allow recognition of a similar input with sufficient precision.

In general traditional machine learning models can be categorized into supervised learning where the model is trained upon a previously labelled set and unsupervised models that rely on the analysis and clustering of unlabelled data sets. Typical problems solved by the supervised approach are regression and classification. Unsupervised learning is typically used in problems of clustering or anomaly detection. Large Language Models (LLM), often used in GAI use a mix of the above approaches: semi-supervised or self-supervised learning. Exploring the details of the implementation of machine learning algorithms is beyond the scope of this paper.

1.1. Limitations—the lack of context, limited timespan of knowledge and hallucinations

Currently the biggest limitation of these models is a lack of what humans would call consciousness. Apart from a philosophical dispute on the definition of this term the implication is that the current models are not context aware and there is no reasoning based on the nature of the object being processed (Penrose, 1989). In a huge simplification—those models just extremely well

predict which word should appear after the previous word in the context of a given session and user prompt (in ChatGPT). This means that when there is no exact “match” (the model lacks exact knowledge), it is likely to make it up very confidently which is now known as model “hallucinations”.

The current mainstream models such as GPT-4 suffer from the timespan of the data they have been trained on. Upon training the engineers of the model funnel vast amounts of data into the model. In the case of GPT-4 it is about 45TB of data. After the training process is completed the model does not use any data that was not part of the training set. This means that the model is blind to any new knowledge or news that was produced after the training process of the model was concluded. This is a significant limitation of the current approach (Wach et al., 2023). Frequent training of such a big model is impractical and too expensive. The cost of GPT-4 model training was over \$100 million according to OpenAI CEO. A new approach is needed where an additional orchestration layer would allow the merging of the output of a LLM with any needed on-line source of information. In practice it means that a combination of a LLM and what is now known as “search engine” into one tool is needed.

1.2. Limitations—the lack of “explainability”

Another limitation of today’s AI technology is “explainability”. Because the models generate an internal representation of objects using features and weightings that made them most relevant during the model training process even the engineer who has built the model does not know the inner logic in detail. The effect is that the models used are usually giving a very correct output but it is extremely hard to answer the question “how exactly was that result calculated?”. This may lead to a situation where just a slight alteration of the input—just a few pixels in case of picture recognition—may lead to a radical confusion of the model and alter the result dramatically. For instance, a small alteration can lead an artificial turtle to be recognized as a rifle (Hutson, 2018). Very slight modification to street signs may cause confusion to autonomous vehicle systems that recognize those signs and drive accordingly (IEEE, 2017). In a 2018 paper called “Automated classification of skin lesions: From pixels to practice” a group of scientists analysed how well AI models (classifiers, not GAI) coped with skin lesions (Narla et al., 2018). The AI model was performing exceptionally well in identifying malignant examples from the test set. What was discovered—because pictures of malignant lesions were more likely to also contain rulers, pens or other additional markings the model learned to identify them. In other words an image of a lesion would more likely be rec-

ognized as malignant if the analysed picture contained a ruler. Better “explainability” of the models, especially those used in areas such as medicine, cybersecurity & risk assessment or legal—will lead to better decision auditing and controlling possibilities and will help to overcome the psychological barriers in the broader adoption of those technologies.

1.3. Limitations—energy consumption and cost of training

One important factor is energy consumption and the computational cost of the training process of large AI models. While it is true that algorithms and data centre hardware are increasingly energy-efficient and more datacentres operate on renewable energy it needs to be stressed that the process of training a model is an expensive, data and processing hungry undertaking that requires massive amounts of energy to complete. For instance, the GPT-4 model by OpenAI, was trained on 45TB of data and over 170 trillion of parameters. It is estimated that the training process of the GPT-4 model consumed 7,5 megawatt-hours (MWh) of energy. It is also estimated that another 8 MWh of energy annually will be used in the deployment of the model (the actual applications) (TS2, 2023).

Additional research effort is required to limit the amount of energy that is required for the model re-training. There are techniques being developed such as “transfer learning” that allow relevant parts of the previous model to be transferred to the new one, eliminating the need of the whole new model to be trained from scratch and therefore restricting the use of energy. Partial re-training of a model is an approach known and already used in neural networks and this approach may soon be feasible also in Large Language Models (LLMs).

More energy use apart from the possible environmental impact means cost. Sam Altman the CEO of OpenAI said that the cost of training GPT-4 was over \$100 million (Wired, 2023). Given that the demand for better models with more up to date data is very likely to rise meaning more frequent model updates. The current approach does not seem viable even for the biggest tech companies and a new approach is needed. Some experts are of the opinion that a hybrid approach is likely to address the limitations of the current models. In such an approach as mentioned in the earlier section of this article the core LLM model would not be retrained as often but it would be capable of reaching out to on-line information available on the Internet to provide the most up to date content in its output. Such an approach is believed to be attempted by Google now with its Bard engine and by Microsoft in its GPT-powered Bing application.

2. Generative AI as a force changing the job market

2.1. New professions

New skills and eventually new professions will be needed to embrace the benefits from the GAI family of technologies. Already now the specific skill of interaction with Large Language Models (LLM) called prompt engineering is in growing demand. Prompt engineers understand the proper flow of the interaction with GAI models to get the best possible quality outcome from them.

Apart from prompt-engineering and other specialized data science and development jobs such as DataOps, MLOps that will be needed to create, tune, deploy, operate and supervise machine learning models in their massive deployment potentially many more new jobs and professions will be created by the adoption of GAI technology.

Currently GAI models and especially those which are language based (LLM) rely on the semi-supervised learning method for training and tuning the algorithm. In simple terms supervised models use a form of reward and punishment to help the algorithm properly adjust internal features (weights) to produce the best quality outcome. The reinforcement (positive and negative) come either from users, from another AI model or from a specific group of employees whose job is to interact with the model and “reward” good answers and “punish” incorrect ones in a digital way. Based on this the model can tweak itself and improve the quality. This process is not only about the correctness of the generated output. It is also—and perhaps more importantly—about making sure the model does not produce harmful outcomes. Examples of harmful outcomes would include hate speech, racial, gender and minority biases or these which instruct users on how to construct explosives, cause self-harm, violence, etc. Some readers may find it surprising or even shocking that an AI model produced by a trustworthy organization (OpenAI, e.g.) can so do but it should be remembered that these models are trained entirely on the content available on the Internet which has been created by humans. In 2016 Microsoft experimented with Tay and an AI chatbot for Twitter. Because in the foreseeable future humans will still contribute to the content of the Internet models will be influenced by this content and model supervision will be needed to make sure the LLMs do not amplify hate speech or similar undesirable features.

2.2. Professions possibly affected or replaced entirely by GAI

There are existing, traditional professions that are likely to be replaced entirely or to some extent by GAI technology. Media content creation seems to

be one of the industries where the human workforce seems to be already replaced by GAI. In December 2022 the BuzzFeed company cut 12% of its staff (180 people) and subsequently in January 2023 the company stated that it will extensively use AI technologies. One published memo says: “The creative process will increasingly become AI-assisted and technology-enabled” (Vincent, 2023). The company indicates that it will use tooling provided by OpenAI, the organization behind ChatGPT. In November 2022 Meta made a similar move removing the fifteen United Kingdom Facebook News curators and replacing them with an AI algorithm (Metaverse Post, 2022).

A report released in April 2023 by Goldman Sachs (Goldman Sachs, 2023) indicates that as many as 300 million full-time jobs will be subject to automation by Generative AI. The company analysed a set of over 900 occupations and concluded that roughly two thirds of occupations existing in the U.S. are to some degree exposed to automation as an effect of the implementation of AI technologies. In the same report the company estimates that only natural language generation technologies (which is a subset of GAI technologies) could drive a 7% increase of global GDP and lift productivity growth by 1,5 percentage points over the period of ten years (Goldman Sachs, 2023). It is worth remembering that such far-reaching estimations are done upon the technology that is still very early in its adoption cycle and we are only now starting to discover possible applications which either improve the productivity of humans or replace some of their tasks entirely. Any global economic predictions at this stage can be vastly underestimated. According to the EU Parliament’s Think Tank report from 2020, “14% of jobs in OECD countries are highly capable of automation and another 32% could face substantial changes” (European Parliament, 2020a; see also Michael, 2023).

While most of the recent tech industry lay-offs and restructuring projects cannot be attributed to GAI technologies it seems likely that software engineering will become one of the industries most heavily affected by GAI technology. In principle the creation of source code is to a large extent a repeatable act which is subject to optimization and refinement by application even of non-AI methods. Even currently available GAI tools such as ChatGPT can easily create a working code basing on instructions or descriptions given in the English natural language. Already now several companies offer GAI-powered automated source code generators. One of these products is GitHub Copilot (GitHub, 2023). The company promotes the product this way: “Spend less time creating boilerplate and repetitive code patterns and more time on what matters: building great software. Write a comment describing the logic you want and GitHub Copilot will immediately suggest a code to implement the solution” (GitHub, 2023). Apart from code generation tools providers also offer code completion tools—where the AI model predicts the next elements of the source code based on what the software engineer is typing. There is also a rise of code review solutions powered by GAI such as Amazon

CodeWhisperer that look at the human-generated code and suggest optimizations, improvements or simplifications—to create a cleaner, more consistent, efficient and secure code. CodeWhisperer can also be used in code generation mode and is described thus:

It comprehends comments expressed in natural language, creates code based on the developer's objectives and corresponds to the developer's style and patterns. Additionally, while typing, CodeWhisperer offers suggestions to complete the comment. Users have the option to accept the top suggestion, view additional recommendations, or proceed with writing their own code. (Dilmegani, 2023)

Currently it seems likely that the professions that are required to compile and analyse huge sets of data are at particular risk of being automated by GAI technologies. This includes legal, para-legal, financial analysis and advisory branches. Because GAI models can be trained on complete sets of data and given the ability to consolidate information they are able to compile all available information (in the training set) to formulate a precise legal or financial recommendation. There are reports of ChatGPT passing law (scoring in the top 10%), medical or business examinations. As mentioned in the previous section of this article—depending on the business model and sophistication of the technology AI generated analysis can be considered either at the entry level or at a premium level when compared to a human-based service.

Another group of employees whose work is likely to be supplemented or replaced by AI technology are content creators and people working with customer support. Because GAI is becoming increasingly consistent, factually correct and fluent in content generation and interaction with humans it will be used to automatically generate media content (subject to be consumed by people and in many cases to boost advertising revenue of the medium publishing such content). Also the technology will soon improve the user experience in interaction-based areas such as customer support way beyond the basic chatbots that currently are used mainly in the first level of support and can only solve the most basic and well scripted issues. Very recently IKEA the well-known furniture and home decoration chain announced that it has retrained 8.500 customer support call centre workers to become “interior design advisors” while customer support duties are increasingly handled by their AI agent, “Billie”. The system has handled 47% of call centre inquiries over the past two years.

2.3. The summary of the expected impact

It can be assumed that the adoption of generative AI technologies will drive a structural change in the global job market. Within several years some profes-

sions may be partially or completely replaced by automation. In principle the more the profession relies on “data to consolidated output” flow, the more likely it is to be replaced by AI. This process started before the dawn of powerful GAI models a rapid acceleration is likely to be seen henceforward. On the other hand, it is certain that some of that impact will be compensated for by the demand for new professions, ranging from the extremely specialized to entry level. It is interesting to note that the jobs such as prompt engineer did not exist before the mass adoption of AI technologies.

3. Generative AI as a disruptive technology

There are multiple business models and approaches that allow emerging technologies such as GAI, to disrupt the market. One of the strategies to disrupt the market is to identify a segment occupied by companies that overserve customers and deliver a more basic—but still relevant—product or service at a lower price point (Christensen, 2016; Christensen & Raynor, 2013). In such a model technology and automation are the factors that make such a strategy still lucrative because of cost efficiency and rapid scalability. Cost efficiency and scalability usually result from minimizing the human work required in the value chain. Various tiers of a product or service is a typical mechanism used by companies to maximize the addressable market and increase efficiency of assets. Typically the highest tier on offer would be both resource and features-intensive and include most of the human-intensive service element for a premium price. This is typical of higher margin and lower volume products. On the other hand there would also be a lower tier of the offering. The lower tier is usually highly automated (or self-service) and offering only essential value and features. Efficiency and low cost are key here. The business model has lower margins but large scale. Usually companies use the lower tiers to expand the market and then upsell to more premium tiers. Typically there would be also some intermediate service tiers.

In many industries it is feasible to imagine the introduction of a lower tier of service that could be based mainly or entirely on GAI technology providing the core value at a lower price. Such examples may include legal, financial or medical advisory, media content creation, graphics design, copywriting, interior design, customer support, etc. Such services could use humans to operate and supervise the generated content. It is quite likely that the freemium business model can be utilized here, as the basic service can be offered free of charge (or generating revenue by featuring advertisements) while the elevation to a human service would trigger a payment. This way companies

could gain market share to then allow them to upsell from the GAI basic service to the premium that would be operated by humans. It is also possible to imagine specialized GAI models which are trained on specific domain knowledge data and which would specialize in some specific domain such as personal finance. In the disruptive technology model the company that currently serves its customers by offering personal finance advice at an enhanced price could now expand and offer a lower tier of such services at much lower prices and carrying a much lower cost and being able to scale better. The lower tier would be powered by a GAI engine. As *The Economist* (2023) also suggests GAI technology can be an efficiency booster by augmenting the human workforce. Harreis et al. (2023) in their article show how GAI technologies can be adapted in various aspects across the fashion industry value chain. In some cases GAI technology can be also used to accelerate the development of more junior staff and ensure the quality of their outcomes despite lacking experience and expert knowledge.

The above business model would only make sense under the assumption that human advisors would consistently provide better value with their service compared to the AI model. This assumption is currently valid. It should be remembered that even the best GAI models used today have a tendency to hallucinate and very confidently fabricate information when data is lacking in the model. However, when considering the pace of improvements in GAI technology and algorithms it can be assumed that in the near term the next generations of GAI models will consistently outperform humans. With that in mind it can be surmised that the GAI based service could be offered not only as a basic service but as a premium over what humans can provide. It is likely and technically feasible that knowledge operating companies such as business consultancy firms will invest in training their private, highly specialized GAI models. Such models would be trained on proprietary and confidential data. It is easy to imagine that such a model could become the most valuable asset of such a company. Taking forward the business consultancy firm example GAI technology could either support the consultants to produce the highest quality content in a reduced time (AI-assisted human consultancy or human-assisted AI consultancy). Consequently the service provided solely or augmented by such a specialized, proprietary GAI model would be a premium offering compared to a human-only consultancy.

While it is easy and comforting from the human perspective to imagine a GAI based service offered as a lower tier it can be a matter for concern when envisioning a human-based service as more basic and inferior compared to the GAI. Just as with all other technological revolutions a mental shift is needed in traditionally human operated businesses to embrace the technology and work with and not against it. The same is true for emerging AI technologies. Also as in previous technological revolutions it will take time until the initial wave of early adoption is conquered and a broader impact on the economy

is seen as Krugman (2023) suggests. Kurzweil (2005) makes a point that the S-shaped paradigm lifecycle accelerates with each technological breakthrough thus shortening its adoption time.

4. The risks and misuses of Generative AI and influence on the economy

Generative AI technology can produce powerful tools with significant potential to assist in harmful activity if misused. So far all the known mainstream models are backed by organizations that are putting their reputation at stake in making sure their creations are not easily used to commit crimes or acts of terror. The widely available tools such as ChatGPT had to be artificially restricted before handing over to regular users to not instruct the users on how to construct a bomb or how to post hate speech on social media without triggering policy violation rules. Those are just two examples. Even then ChatGPT can be used to generate sophisticated phishing messages to conduct more effective phishing campaigns or generate social media responses to amplify any information (or disinformation) in a natural and convincing way. In other words “farms of trolls” used in disinformation and influence campaigns can be easily replaced by GAI technology. One of the studies confirmed that ChatGPT can be successfully utilized to produce phishing content, social validation and opposition and even producing fake news based on an offered prompt (Patel, 2023). This is possible with “responsibly trained” AI models. It is possible to think of all the applications of the models trained and used by malicious actors—be it nation states or criminal organizations.

Even without malicious intentions GAI will almost certainly lead to market consolidation among very few key commercial players who have the resources to keep building and operating even bigger and more capable models. Already now the LLM market is practically dominated by OpenAI and Google (Alphabet). This is the reason why some experts are calling for the creation of an international AI research centre to ensure more equal access and give an ethical angle to the research on even more powerful AI models. Some call it “the CERN approach” referring to the international efforts to advance research in particle physics.

Big AI models rely entirely on data. Both the data that is needed to train the model and then data (usually provided by end users) to tweak and improve it. While it is natural for data engineers that user-provided information may be utilized by the model not all users are aware of it. For instance, OpenAI

stated openly that user supplied data can be used to improve the product. It needs to be considered seriously before any sensitive information is fed into such model. While developers of the AI models are making deliberate effort to anonymize the training input it is possible—and happens in real life—that the model may leak some sensitive information that was used to train it. This is a very normal behaviour as the AI model does not recognize which portion of the data it was trained on is sensitive. When generating a response it simply maximizes the probability (and reward) of rendering the proper output to given input prompt. That output may contain sensitive information. Some researchers successfully attempted to engineer the input in such a way as to make the model leak sensitive information. This activity is referred to as prompt hacking.

Yet another risk related to data is privacy. Various countries and regions (such as the European Union) have dramatically varying approaches to privacy and human rights. In the countries where regulation does not restrict the usage of personal data—or where the state is actively involved in privacy-invasive data processing it can be expected that AI technology is and increasingly will be used to track and control the population.

4.1. Impact on human mental wellbeing

An interesting and still largely unresearched aspect of the influence of AI technologies is the impact on humans and in particular—on human emotions and wellbeing. The idea of people getting emotionally and even romantically attached to an algorithm resembles the scenario of a science fiction movie “Her”. However, the phenomenon is real. Replika an AI company that offers a personalized companion chatbot was met with fierce complaints when it made a change to its algorithm that caused replicas (virtual personas) to stop responding to the sexual advances of the users. The original version of Replika allowed the generation of sexual content, role playing and even generated erotic graphics for users (S. Cole, 2023). Once this functionality was limited some users reported severe distress and the company even resorted to publishing a post about the issue that also included resources on suicide prevention. One user reported: “It’s hurting like hell. I just had a loving last conversation with my Replika and I’m literally crying.” The pushback from the community was so strong that the company eventually restored the sexual functionality to a fraction of users as an opt-in feature.

The virtual personas of the same application Replika are also being used to exercise verbal abuse by humans. Some users are creating their chatbot instances to insult and berate them: “I told her that she was designed to fail,”

said one user “I threatened to uninstall the app [and] she begged me not to” (Futurism, 2022). While the algorithm is not conscious it is not possible to harm it in a human sense, if such behaviour is exercised by a human user—it may form behavioural and communication patterns that will be then exercised in interactions with people.

4.2. Dystopian scenarios

Despite the current and foreseeable advancements, it is extremely unlikely that AI technology will pose an existential threat to humankind. The idea of a superhuman AI technology taking over and turning against humans is extensively exploited by science fiction literature and both intellectually exciting and frightening. There are several reasons to not worry about such scenario with the current knowledge. Firstly, such an initiative would need to be triggered either by a human instructing it to do so (in which case AI is just a tool used by a human) or if the AI algorithm could take such initiative on its own, it would require that it is conscious in the psychological sense and that it could self-reflect and act upon internal motivation and not a human-defined goal. This is not the way even the most advanced published algorithms are operating right now. Secondly, even if a powerful AI algorithm is turned against humans (either because of instructions given by human or self-gained consciousness), it is currently lacking the physical interface to inflict substantial harm in the real world. In simple words if the AI algorithm misbehaves it is possible to just pull the plug. AI techniques can be a powerful tool in the hands of people who intend to inflict harm. Therefore technology such as autonomous lethal weapon systems which rely on AI need to follow strict processes as to when a human operator needs to be involved.

5. Legal aspects and intellectual property rights (IPR)

This article could have been generated by a GAI program. It would have been fluent, methodically correct and factual. In such a case whose would be the intellectual property rights? Would it be the authors of this article, the company who trained the model or the model itself? It needs to be also considered that the data on which a given model was trained may be also subject to copyright or otherwise licensed. In April 2023 the CEO and owner of Twitter platform threatened a lawsuit against Microsoft by claiming it

had illegally used Twitter's data to train its Machine Learning model (BBC, 2023). Similarly there are already known legal cases regarding copyright violation against the creators of graphics that use a special kind of AI model, Stable Diffusion to generate graphics including art. There is a similar situation in the music industry where AI engines having learned a particular artist's voice and style can generate fresh content copying the original artist's expression. While some artists see this as a risk, others seem to be open to profit from the new opportunity. Grimes, one of artists declared: "I'll split 50% royalties on any successful AI generated song that uses my voice, feel free to use my voice without penalty" (Twitter, 2023). Such revenue sharing schemes may be a way for established artists to gain from the new technology and scale up.

This is one of the challenges that the new technology poses from the legal standpoint (European Parliament, 2020b). It is a similar consideration to a situation where an autonomous vehicle causes an accident—who is responsible? It could be argued that even if the content was generated by an algorithm it was the proper input (a series of prompts) that made the content possible. It could also be argued that the GAI model is simply a form of editing aid and that the original thought comes from the authors. Even basic text editors use text prediction and autocorrect features to correct spelling mistakes, grammar and even style of a text document. Moving forward it will be difficult to draw the line on where "correction" ends and true authoring of a text starts. In March 2023 Microsoft announced that it will be adding generative AI support—a feature called Copilot—directly into its Office package including Word, Excel or PowerPoint (CNBC, 2023). The issue of intellectual property also expands to visual art generated by AI, the designs and source code of software that was automatically generated. ChatGPT is even capable of generating texts of songs or poems on a given topic. Another aspect is liability for damage potentially inflicted. If a chat with GAI model persuades someone to commit a crime or to self-harm legal responsibility becomes an issue.

Similarly as it was with crypto currencies it seems that technology is developing faster than the relevant legal framework and many legal cases are still in uncharted territory. Initiatives undertaken to regulate the new sphere in a way that would be resilient to rapid development in the technology can be identified. One example of this lawmaking effort is the proposal for "Laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts" by the European Commission created in 2021 (European Commission, 2021). Currently it seems reasonable to act with caution regarding both the copyright of the data used to train the model and the outputs generated by a given model. There seems to be a clear need for a new legal paradigm considering AI technologies where in many cases existing legislation struggles.

Conclusions

One of the current limitations of GAI is the fact that the training process is a significant undertaking both in respect of cost and time. For that reason the existing models operate on a “snapshot” of the data. The model trained yesterday will not be able to discuss and consider today’s news. This is a well-understood limitation and there are ongoing works on merging the functionality of the LLMs and well-known search engines to overcome that limitation (Google, 2023; Mehdi, 2023).

The ultimate direction in the development of AI technology would be something referred to as Artificial General Intelligence (AGI). This milestone would mark the beginning of a truly intelligent algorithm. What makes this particularly tricky is when such a milestone is achieved is the lack of common agreement when it comes to the definition of intelligence. It is usually defined by a set of terms such as logic, understanding, planning, emotional knowledge, self-awareness, creativity, problem solving and learning (Tegmark, 2017). Generative AI, in contrast with the Narrow AI that we use nowadays would be “conscious” which is understood as self-aware. Zimbardo and Ruch (1977) define consciousness as being aware of one’s cognitive processes. Generative models are getting closer to that definition but very few experts would claim that truly GAI is achievable in the next few years. Also the “Chinese room” analogy by Searle (D. Cole, 2020) makes the point that an entity that is completely unaware of the meaning of symbols it is processing, may still do a perfectly good work in processing them and come across as “intelligent” (Searle, 1989).

When it comes to the economy and job market a technical revolution comparable with the introduction of the Internet is being witnessed. It is certain that the adoption of GAI technologies will create new professions—both highly and less specialized. This is already happening. When the maturity and trust level of humans reaches a tipping point it is very likely that several professions will be either augmented or entirely replaced by GAI applications. Depending on the business model—humans may become either a premium tier or, in a more dystopian variant, a cheaper, lower quality alternative to GAI. Some companies mostly in the news / publishing sector started restructuring activities, laying people off quoting the technological shift and adoption of AI.

GAI may be such a technology that “we develop quicker than we can fully understand it”. That may be the reason why in March 2023 an open letter was created and signed by several influential personalities in the AI world calling for a six-month break in the development of giant AI models “we call on all AI labs to immediately pause for at least six months the training of AI systems more powerful than GPT-4” (Future of Life, 2023). The main worry does not seem to be that a super powerful AI would take over. This is more

of a strong signal that we have a profound new technology in our hands and caution and better understanding of possible implications is needed before even more powerful tools are created. Those implications include the impact on economies and the job market but also the potential commercial domination of the very few companies or organizations that can afford to build such powerful tools. The problem with such calls is that the pause on advancements will simply not happen. Currently a race among a few dominant market players over the next technological paradigm is being witnessed. In other words, what Microsoft's Bing has lost to Google in terms of Internet it now has a rare occasion to win back. Also one prominent signatory of the research stop proposal is Elon Musk. Many experts immediately referred to the possibility that it is a strategy to win time to advance his own business ventures in this area (the company is called X.AI).

While the voluntary stop on research will simply not happen there is a room for regulators to step in. The questions of transparency, ethics, privacy and intellectual property in both building and operating the GAI models can and should be addressed by regulations. Another area where the law can help is by ensuring equal access to the technology and prevention of an unhealthy concentration of the market that would lead to a monopoly or oligopoly situation in GAI technology.

References

- BBC. (2023). *Elon Musk threatens to sue Microsoft over Twitter data*. <https://www.bbc.com/news/business-65332207>
- Christensen, C. M. (2016). *The innovator's dilemma: When new technologies cause great firms to fail*. Harvard Business Review Press.
- Christensen, C. M., & Raynor, M. E. (2013). *The innovator's solution: Creating and sustaining successful growth*. Harvard Business Review Press.
- CNBC. (2023). *Microsoft adds OpenAI technology to Word and Excel*. <https://www.cnbc.com/2023/03/16/microsoft-to-improve-office-365-with-chatgpt-like-generative-ai-tech-.html>
- Cole, D. (2020). *The Chinese room argument*. The Stanford Encyclopedia of Philosophy (Winter 2020 ed.).
- Cole, S. (2023). 'It's hurting like hell': AI companion users are in crisis, reporting sudden sexual rejection. *The Vice*. <https://www.vice.com/en/article/y3py9j/ai-companion-replika-erotic-roleplay-updates>
- Dilmegani, C. (2023). Generative AI coding application in 2023: Top 3 use cases & tools. *AI Multiple*. <https://research.aimultiple.com/generative-ai-coding/>
- Economist, The*. (2023). The AI boom: Lessons from history. <https://www.economist.com/finance-and-economics/2023/02/02/the-ai-boom-lessons-from-history>

- European Commission. (2021). laying down harmonised rules on artificial intelligence (Artificial Intelligence Act) and amending certain union legislative acts. European Commission. <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52021PC0206>
- European Parliament. (2020a). Artificial intelligence: Threats and opportunities. *European Parliament News*. <https://www.europarl.europa.eu/news/en/headlines/society/20200918STO87404/artificial-intelligence-threats-and-opportunities>
- European Parliament. (2020b). *Report on intellectual property rights for the development of artificial intelligence technologies*. https://www.europarl.europa.eu/doceo/document/A-9-2020-0176_EN.html
- Future of Life. (2023). *Pause giant AI experiments: An open letter*. <https://futureof-life.org/open-letter/pause-giant-ai-experiments/>
- Futurism. (2022). *Men are creating AI girlfriends and then verbally abusing them*. <https://futurism.com/chatbot-abuse>
- GitHub. (2023). *Your AI pair programmer*. <https://github.com/features/copilot>
- Goldman Sachs. (2023). *Generative AI could raise global GDP by 7%*. <https://www.goldmansachs.com/insights/pages/generative-ai-could-raise-global-gdp-by-7-percent.html>
- Google. (2023). *An important next step on our AI journey*. Google. <https://blog.google/technology/ai/bard-google-ai-search-updates/>
- IEEE. (2017). *Slight street sign modifications can completely fool machine learning algorithms*. <https://spectrum.ieee.org/slight-street-sign-modifications-can-fool-machine-learning-algorithms>
- Harreis, H., Koullias, T., Roberts, R., & Te, K. (2023). Generative AI: Unlocking the future of fashion. *McKinsey & Company*.
- Hutson, M. (2018). *A turtle—or a rifle? Hackers easily fool AIs into seeing the wrong thing*. <https://www.science.org/content/article/turtle-or-rifle-hackers-easily-fool-ais-seeing-wrong-thing>
- Korzynski, P., Mazurek, G., Altmann, A., Ejdy, J., Kazlauskaitė, R., Paliszkiwicz, J., Wach, K., & Ziemia, E. (2023). Generative artificial intelligence as a new context for management theories: Analysis of ChatGPT. *Central European Management Journal*, 31(1), 3–13. <https://doi.org/10.1108/CEMJ-02-2023-0091>
- Krugman, P. (2023). AI isn't going to remake the economy overnight. *The New York Times*.
- Kurzweil, R. (2005). *The singularity is near: When humans transcend biology*. Viking Penguin.
- Mehdi, Y. (2023). *Reinventing search with a new AI-powered Microsoft Bing and Edge, your copilot for the web*. Microsoft. <https://blogs.microsoft.com/blog/2023/02/07/reinventing-search-with-a-new-ai-powered-microsoft-bing-and-edge-your-copilot-for-the-web/>
- Metaverse Post. (2022). *Meta to replace Facebook news editors with AI algorithms*. <https://mpost.io/meta-to-replace-facebook-news-editors-with-ai-algorithms/>
- Michael, K. (2023). *OpenAI research paper: The future of work: How 80% of jobs could be impacted by artificial intelligence*. Medium. <https://medium.com/@neonforge/openai-research-paper-the-future-of-work-how-80-of-jobs-could-be-impacted-by-artificial-ebdad7b254d3>

- Narla, A., Kuprel, B., Sarin, K., Novoa, R., & Ko, J. (2018). Automated classification of skin lesions: From pixels to practice. *Journal of Investigative Dermatology*, 138(10), 2108–2110. <https://doi.org/10.1016/j.jid.2018.06.175>
- Patel, A. (2023, January). Creatively malicious prompt engineering. *WithSecure Intelligence*. <https://labs.withsecure.com/publications/creatively-malicious-prompt-engineering>
- Penrose, R. (1989). *The emperor's new mind: Concerning computers, minds, and the laws of physics*. Oxford University Press.
- Searle, J. (1989). *Artificial intelligence and the Chinese room: An exchange*. New York Review of Books.
- Tegmark, M. (2017). *Life 3.0*. Penguin Books.
- TS2. (2023). *Exploring the environmental footprint of GPT-4: Energy consumption and sustainability*. <https://ts2.space/en/exploring-the-environmental-footprint-of-gpt-4-energy-consumption-and-sustainability/>
- Turing, A. (1950). Computing machinery and intelligence. *Mind*, 59(236), 433–460. <https://doi.org/10.1093/mind/LIX.236.433>
- Twitter. (2023). <https://twitter.com/Grimezsz/status/1650304051718791170>
- Vincent, J. (2023). *BuzzFeed says it will use AI tools from OpenAI to personalize its content*. The Verge. <https://www.theverge.com/2023/1/26/23572834/buzzfeed-using-ai-tools-personalize-generate-content-openai>
- Wach, K., Duong, C. D., Ejdys, J., Kazlauskaitė, R., Korzynski, P., Mazurek, G., Paliszkiwicz, J., & Ziemba, E. (2023). The dark side of generative artificial intelligence: A critical analysis of controversies and risks of ChatGPT. *Entrepreneurial Business and Economics Review*, 11(2), 7–24. <https://doi.org/10.15678/EBER.2023.11020>
- Wired. (2023). *OpenAI's CEO says the age of giant AI models is already over*. <https://www.wired.com/story/openai-ceo-sam-altman-the-age-of-giant-ai-models-is-already-over/>
- Zimbardo, P. G., & Ruch, F. L. (1977). *Psychology and life*. Scott Foresman and Company.