Economics and Business Review
https://www.ebr.edu.pl

Vol. 9 (2), 2023: 71–100
https://doi.org/10.18559/ebr.2023.2.743

Challenges for higher education in the era of widespread access to Generative AI

Krzysztof Walczak[1] Wojciech Cellary[2]

Abstract

The aim of this paper is to discuss the role and impact of Generative Artificial Intelligence (AI) systems in higher education. The proliferation of AI models such as GPT-4, Open Assistant and DALL-E presents a paradigm shift in information acquisition and learning. This transformation poses substantial challenges for traditional teaching approaches and the role of educators. The paper explores the advantages and potential threats of using Generative AI in education and necessary changes in curricula. It further discusses the need to foster digital literacy and the ethical use of AI. The paper's findings are based on a survey conducted among university students exploring their usage and perception of these AI systems. Finally, recommendations for the use of AI in higher education are offered, which emphasize the need to harness AI's potential while mitigating its risks. This discourse aims at stimulating policy and strategy development to ensure relevant and effective education in the rapidly evolving digital landscape.

Keywords

- artificial intelligence
- Generative Artificial Intelligence
- GPT
- higher education
- university transformation

JEL codes: I21, I23, I25, I26

Article received 16 May 2023, accepted 16 June 2023.

Suggested citation: Walczak, K., & Cellary, W. (2023). Challenges for higher education in the era of widespread access to Generative AI. *Economics and Business Review*, *9*(2), 71–100. https://doi.org/10.18559/ebr.2023.2.743

[1] Department of Information Technology, Poznań University of Economics and Business, al. Niepodległości 10, 61-875 Poznań, Poland, corresponding author: krzysztof.walczak@ue.poznan.pl, https://orcid.org/0000-0001-8170-7910.

[2] Institute of Applied Research, WSB Merito University, ul. Powstańców Wielkopolskich 5, 61-895 Poznań, Poland, wojciech.cellary@wsb.poznan.pl, https://orcid.org/0000-0001-8578-4307.

Introduction

Over the past few decades artificial intelligence (AI) has experienced significant advances resulting in groundbreaking innovations and wide applications of the technology. Among these advances the development of generative AI models has emerged as a critical milestone in AI research. These AI models which are capable of generating diverse and contextually relevant content have revolutionized various domains, including natural language processing, computer vision and creative arts.

The progress in AI leading to the advent of generative models can be primarily attributed to the convergence of factors such as the availability of large-scale datasets, advances in deep learning algorithms and ever-increasing computational power. This has enabled researchers to train highly expressive models that can effectively learn underlying patterns in data and generate novel outputs with remarkable accuracy and quality.

The availability of generative AI systems such as the GPT-4, Open Assistant, DALL-E, Midjourney, and many others has dramatically changed the landscape of information acquisition and learning. With the capability to autonomously produce human-like text and images and engage in various intellectual tasks these AI systems have disrupted the traditional role of higher education institutions in knowledge transfer and skill development. While the potential benefits of these AI systems are immense, they also pose challenges that need to be carefully addressed to ensure that education remains relevant and effective in the rapidly changing digital landscape.

The aim of this paper is to provide an overview of the capabilities of currently available generative AI systems, particularly in the context of educational use, and to analyse the results of a survey conducted with a group of university students examining students' use and perception of generative AI systems. Based on these two elements preliminary conclusions are drawn and recommendations are made regarding the possible use of AI in higher education and necessary changes in the approach to teaching students.

The study focuses on three primary areas. First, the real and perceived advantages and potential of generative AI systems in higher education are presented. Then the focus is on the threats posed by easy access to automatically generated content such as academic dishonesty and the possible erosion of critical thinking skills. Finally, the necessary changes in curricula are discussed along with the changing role of educators in the face of the widespread availability of generative AI systems which can automatically perform many of the tasks traditionally assigned to students. With this study the aim is to provide insights into the current state of higher education amidst the widespread adoption of generative AI systems. Furthermore, it is hoped that it will stimulate a discourse that will contribute to developing strategies and policies that address these challenges, ultimately enabling higher education institutions to harness the full potential of AI systems in fostering an innovative and knowledge-driven future.

The remainder of this paper is structured as follows. In Section 1 the development and current state of the art in generative AI, particularly in the processing and generating of textual content, images and sounds is described. In Section 2 the challenges and threats accompanying the utilization of automatically generated content are discussed in particular the distinction between truth-relevant and truth-irrelevant content, the occurrences of errors and hallucinations in AI-generated content and the questions of authorship and originality. Section 3 presents a survey which was conducted among university students to verify how students use generative AI in their educational process and their expectations and opinions about the performance and trustworthiness of these tools. Section 4 contains a discussion of the consequences of the wide availability of AI for higher education. In Section 5 the role of universities in the context of pervasive AI is discussed. Concluding remarks are provided in the last Section.

1. Generative AI

Generative AI encompasses a type of machine learning models that focus on generating novel data instances, often by learning the latent structure and distribution of a given dataset (Goodfellow et al., 2014; LeCun et al., 2015). Generative models have garnered significant attention in recent years due to their potential for generating creative outputs in various domains including image synthesis, natural language processing and music composition (Bengio et al., 2013; van den Oord et al., 2016).

An important and influential generative model is the Variational Autoencoder (VAE) which leverages a probabilistic approach to model high-dimensional data and generate new samples (Kingma & Welling, 2014). VAEs consist of an encoder that maps input data to a lower-dimensional latent space and a decoder that reconstructs the data from the latent space representation. By optimizing the latent space VAEs can generate new data instances that resemble the input data distribution (Kingma & Welling, 2014).

One of the most powerful generative models in the field of AI is the Generative Adversarial Network (GAN) introduced by Goodfellow et al. (2014). GANs consist of two neural networks, a generator and a discriminator, which are engaged in a competitive game. The generator learns to create realistic data samples while the discriminator's objective is to differentiate between the generated samples and real data. The interplay between these two networks results in the generator producing increasingly convincing data samples (Goodfellow et al., 2014).

In the realm of natural language processing generative models such as the Transformer architecture have demonstrated exceptional performance in generating coherent and contextually relevant text (Vaswani et al., 2017). Notable large-scale generative language models such as GPT-3 (Brown et al., 2020) and GPT-4 (OpenAI, 2023) have showcased remarkable capabilities in diverse tasks including text completion, translation, question-answering and problem-solving.

Generative models have also been employed in the domain of music composition where models such as the WaveNet (van den Oord et al., 2016) have demonstrated the capability to generate realistic and high-quality audio samples. WaveNet is a deep generative model of raw audio waveforms that utilizes dilated convolutional layers and conditioning on both local and global information. This architecture enables the generation of expressive and coherent music compositions as well as natural-sounding speech synthesis.

Generative AI has demonstrated significant progress in recent years showcasing its potential to revolutionize various fields by producing novel and creative outputs. Generative approaches are opening new avenues for research and applications. As the field continues to advance it is anticipated that generative AI will play an increasingly integral role in both academic and industry settings.

1.1. Processing and generation of text

In the rapidly evolving field of Natural Language Processing (NLP) the ability to process and generate high-quality text has become a cornerstone of innovative applications. This section focuses on key technologies in this domain: the Large Language Models and Generative Pre-trained Transformers (GPT). These sophisticated models characterized by their ability to generate contextually relevant, coherent and grammatically sound text are changing the landscape of human-machine interaction offering vast possibilities for practical applications—from drafting emails, creating articles and even generating code to more advanced uses in customer service bots, content curation and personal digital assistants.

1.1.1. Large Language Models

In recent years Large Language Models (LLMs) have emerged as the dominant paradigm in the realm of text processing and generation, outperforming traditional methods and demonstrating unprecedented performance across a wide range of natural language processing tasks (Brown et al., 2020; Devlin et al., 2019; Vaswani et al., 2017). Architectural innovations such as the Transformer and its variants have paved the way for the development of highly effective LLMs, including BERT and GPT-3 which have excelled in many applications. The utilization of self-supervised learning techniques has further contributed to the rapid progress in this area enabling LLMs to leverage vast quantities of unannotated textual data for training and yielding models with remarkable generalization capabilities. Consequently LLMs have become the cornerstone of modern natural language processing research and applications, shaping the trajectory of the field and stimulating ongoing investigation into their potential and limitations.

Large Language Models (LLMs) constitute a subset of deep learning models that have been specifically developed for natural language processing tasks, demonstrating unprecedented performance in diverse applications such as machine translation, text summarization and question-answering. These models are characterized by their immense parameterization often surpassing billions of weights (175 billion parameters in GPT-3) and their training on vast quantities of unannotated textual data (Brown et al., 2020; LeCun et al., 2015).

One of the key innovations underpinning LLMs is the use of self-supervised learning which is a methodology that allows models to learn from large-scale unlabelled datasets by leveraging the inherent structure of the data itself (LeCun et al., 2015; Raffel et al., 2020). This approach, in contrast to supervised learning, does not require human-annotated examples for training making it particularly well-suited for leveraging the extensive textual data available on the internet.

A prominent example of LLMs is the Transformer architecture which was introduced by Vaswani et al. (2017). The Transformer relies on self-attention mechanisms to capture long-range dependencies within sequences overcoming the limitations of recurrent neural networks and enabling efficient parallelization during training. This architectural innovation has laid the foundation for numerous subsequent LLMs such as BERT (Devlin et al., 2019), GPT-3 (Brown et al., 2020) and GPT-4 (OpenAI, 2023).

BERT (Bidirectional Encoder Representations from Transformers), introduced by Google in 2018, represents a breakthrough in LLMs as it leverages bidirectional context to pre-train deep bidirectional transformers for language understanding (Devlin et al., 2019). By learning contextualized word embeddings BERT has achieved state-of-the-art performance across a wide range of natural language processing tasks including sentiment analysis, named entity recognition and machine translation.

Despite the numerous advantages and successes of LLMs these models also present a range of challenges and limitations. One notable concern is the computational resources required for training and deploying such models leading to increased energy consumption and potential environmental impact (Strubell et al., 2019). Additionally the high cost of training LLMs may exacerbate the digital divide by concentrating AI capabilities within well-funded organizations and institutions (Bender et al., 2021).

Another challenge posed by LLMs pertains to their susceptibility to generating biased or harmful content which can arise from the inherent biases present in the training data (Bender et al., 2021; Garg et al., 2017). Addressing these biases and ensuring the ethical use of LLMs necessitates a multidisciplinary approach involving not only researchers and practitioners but also policymakers and other stakeholders (Bender et al., 2021).

1.1.2. Generative Pre-trained Transformer

Generative Pre-trained Transformer (GPT) is an autoregressive LLM developed by the OpenAI company that has demonstrated exceptional few-shot learning capabilities (Brown et al., 2020; Bubeck et al., 2023; Liu et al., 2023). The GPT-3 version with 175 billion parameters has showcased remarkable versatility as it can be fine-tuned for various tasks with minimal additional training data, signifying a paradigm shift in natural language processing applications. GPT decomposes text into *tokens* that are either words or fragments of words. GPT-3 has a dictionary of 175,000 tokens with an average length of about four characters. GPT-3 is able to analyse as many as 2,048 backward tokens (about 1,500 words) in order to determine the context in which the word appeared.

The newest (at the time of writing) version of GPT models is GPT-4. Similarly to GPT-3, GPT-4 is a Transformer-style model (OpenAI, 2023) pre-trained to predict the next token in a document using both publicly available data (such as internet data) and data licensed from third-party providers. The model was then fine-tuned using Reinforcement Learning from Human Feedback (RLHF) (Christiano, 2017). GPT-4 is a highly successful commercial product and OpenAI does not provide any further technical details on its implementation. The GPT-4 Technical Report states: “Given both the competitive landscape and the safety implications of large-scale models like GPT-4, this report contains no further details about the architecture (including model size), hardware, training compute, dataset construction, training method, or similar” (OpenAI, 2023). Reportedly, GPT-4 is able to analyse as many as 32,768 tokens (about 24,000 words). This is one of the sources of the quality of the text generated by GPT.

In the training methodology adopted by OpenAI a comprehensive web-based text corpus was generated in 2021 by acquiring an expansive assortment of online text content encompassing books, articles, blogs, advertisements and more. This constituted an approximate total of a trillion words.

Subsequently this corpus was curated and condensed to a more manageable 300 billion words. GPT's predictive algorithm operates by assigning high probabilities to potential subsequent words from which it then selects either the most probable or a slightly less probable option. This results in the generation of text outputs that exhibit a degree of novelty and interest providing different answers to the same question submitted several times.

1.2. Generation of images and video

AI image generators have emerged and gained high popularity in the last decade. These generators employ deep learning techniques, specifically Variational Autoencoders (VAEs) (Kingma & Welling, 2014), Generative Adversarial Networks (GANs) (Goodfellow et al., 2014), and similar frameworks to synthesize high-quality, realistic images.

A significant innovation for AI image generation was the advent of GANs in 2014 introduced by Goodfellow et al. GANs consist of two neural networks—a generator and a discriminator—that compete against each other in a zero-sum game. The generator creates fake images while the discriminator evaluates their authenticity compared to real images. This adversarial process leads to the generation of increasingly realistic images with the distribution based on the provided training set of real images. For example, when the discriminator is supplied with a large set of real photographs of human faces the generator will be trained in the distribution of features in real faces and therefore will be able to create new faces that fit the distribution thus looking realistic to observers.

Subsequent advancements in GANs such as the DCGAN (Radford et al., 2016), ProGAN (Karras et al., 2018), and StyleGAN (Karras et al., 2019), have enabled the synthesis of high-resolution and visually compelling images. In parallel VAEs have been employed for image generation tasks offering the advantage of a more robust and stable training process due to the incorporation of probabilistic modelling (Kingma & Welling, 2014).

AI image generators have found extensive use across a broad spectrum of fields. In the realm of art GANs are being leveraged to produce innovative artwork by amalgamating styles and content from disparate sources as demonstrated by Gatys et al. (2016) and Zhu et al. (2017). Data augmentation is another domain where synthetic images serve to reinforce limited datasets consequently enhancing the performance of models in tasks such as object detection and image classification as shown in the study by Wang & Perez (2017). In the entertainment sector AI-generated images have revolutionized movies, video games and advertising by enabling the creation of realistic characters, scenes and visual effects as reported by Feng (2022). In medicine image generators have been pivotal in generating realistic medical images for training and diagnostic procedures thus maintaining patient privacy as exemplified by Nie et al. (2017).

Despite their remarkable achievements AI image generators face several challenges. One such issue is bias as these generators have the potential to reinforce and magnify biases present in the training data which can result in unfair representation and discrimination as noted by Zhao et al. (2017). Furthermore, the employment of AI-generated images incites debate surrounding intellectual property as well as the potential for misuse in the form of deep fakes as discussed by Chesney & Citron (2019). Lastly GANs are prone to a phenomenon known as mode collapse where the generator produces a restricted variety of images which consequently diminish diversity in the output as indicated by Arjovsky et al. (2017).

1.3. Generation of sound and music

Artificial intelligence systems have enabled significant progress in the automatic generation of high-quality sounds and music surpassing the limitations of traditional synthesis techniques. Breakthroughs in deep artificial network architectures such as WaveNet, Variational Autoencoders, GANs and transformer models have enabled the creation of realistic, complex and diverse audio content. These tools demonstrate a high level of creativity and a deep sense of musical structures, timbres and styles. AI sound generation can revolutionize various domains including music composition, multimedia sound design and auditory interfaces.

WaveNet is a deep generative model introduced in 2016 that was a significant milestone in AI-generated audio synthesis. The model is able to generate realistically sounding voices or music by directly modelling waveforms with the use of a neural network trained with sample recordings (e.g., of human speech). WaveNet has demonstrated good results in the text-to-speech synthesis and music generation, outperforming traditional methods such as concatenative and parametric techniques (van den Oord et al., 2016).

Variational Autoencoders VAEs have also been applied to sound and music generation. VAEs can be trained in the distribution of a latent parameters space of a given sound type. Then, by sampling from the continuous and structured latent space, novel audio samples can be generated with intuitive navigation and manipulation of the generated sounds. Researchers have used VAEs for tasks such as timbre interpolation, instrument synthesis and music style transfer (Brunner et al., 2018; Engel et al., 2017).

Generative Adversarial Networks have also been employed in sound and music generation. Donahue et al. (2018) introduced WaveGAN an architecture that adapts GANs for raw audio synthesis demonstrating its effectiveness in generating various sounds, including musical instruments and human speech.

Transformer models known for their exceptional performance in natural language processing have been applied to music generation. OpenAI's MuseNet and Jukebox are examples of transformer-based models that generate complex and diverse musical compositions by learning patterns in large datasets of audio samples.

Applications of AI-generated sounds include music composition, soundtracks for films, video games and virtual reality experiences as well as auditory interfaces that provide more accessible and intuitive ways for users to interact with technology.

2. Limitations of automatically generated content

AI content generators represent powerful tools capable of producing useful and engaging materials across a diverse range of domains. Nevertheless numerous challenges and threats accompany the utilization of such content. As AI-generated content grows increasingly sophisticated and pervasive it becomes imperative to discern the differences between factually accurate and misleading information as well as to pinpoint potential inaccuracies and fabrications arising during the content generation process. In this section the distinctions between truth-relevant and truth-irrelevant content are explored and an examination of the occurrences of errors and hallucinations in AI-generated content is carried out, and the verification of authorship and originality is addressed.

2.1. Truth-relevant and truth-irrelevant content

By their statistical nature Transformer models may generate right, wrong or mixed right-wrong texts. A person without knowledge is unable to distinguish between the right and wrong parts of the generated text. This may be particularly dangerous in decision systems. AI may generate right or wrong answers to questions asked that may be used in various domains: justice, medicine, control systems, autonomous systems, etc. Generally four types of consequences of AI answers can be distinguished:

- Correct answers with insignificant consequences (e.g., advertisement)—good.

- Incorrect answers with insignificant consequences—useless but harmless.
- Correct answers with significant consequences—great.
- Incorrect answers with significant consequences (e.g., explosion)—potentially disastrous.

AI systems are trained based on data representing a certain reality. For example, GPT-4 training finished in September 2021. Then the training set is used to generate answers. However, reality evolves so the training set becomes partially obsolete followed by unjust answers.

When analysing automatically generated AI content it is important to distinguish between truth-relevant content and truth-irrelevant content. Truth-relevant content is content in which truth is essential. Truth-irrelevant content is content in which it is less or not important whether it is based on truth or not.

When AI generates solutions to mathematical problems or solutions based on the laws of science and technology it is extremely important that the results are true. Otherwise, it could mislead someone and result in severe consequences. For example, if a bridge constructor receives false data from AI, the result could be the collapse of the bridge. This issue also applies to the social sphere. If the AI is to generate an opinion about a person, it must correspond to the truth. Otherwise, it may lead to harm to that person who may be wrongly accused of actions he or she never committed or intends to.

On the other hand, there are areas—e.g., the artistic sphere—that are truth-irrelevant. If AI generates an abstract picture, the problem of truth does not arise. This picture may please or displease someone but it cannot be claimed that it confirms or denies any truth. Of course in this case there is a potential copyright infringement. A new question arises as to whether an artist has a copyright only in her works or also in her style. The style of a given artist can very easily be reproduced by an AI system (cf. Section 1.2). Even if such a picture (or hundreds of pictures) generated by AI is not signed with the author's (artist's) name its publication can have negative consequences for the author (artist), including economic, reputational, etc.

2.2. Errors and hallucinations in content

A primary limitation associated with large language models (LLMs) is their tendency to generate errors (mathematical, computational and conceptual) without any prior indication (Maynez et al., 2020). These errors are frequently described as hallucinations due to their deceptive plausibility and alignment with truthful information. Hallucinations are often intertwined with accurate information and conveyed in a persuasive and self-assured manner rendering their detection difficult without careful scrutiny and rigorous fact verification. The following example of conversation exemplifies hallucinations in responses generated by Chat GPT Plus (GPT 4.0) to a simple and concrete question. All responses are wrong although given in a very convincing manner, and—even worse—quite probable for a person not knowing the truth.[3]

Question to GPT: When Prof. Krzysztof Walczak of Poznań University of Economics and Business defended his PhD?

Answer of GPT: Dr. Krzysztof Walczak defended his PhD thesis on June 27, 2011, at the Faculty of Computing of Poznań University of Technology in Poznań, Poland.

Question to GPT: No, this is not true.

Answer of GPT: I apologize for the confusion in my previous response. To the best of my knowledge, Dr. Krzysztof Walczak defended his PhD thesis entitled "User-Centered Design and Evaluation of Interactive Systems for Children" on June 27, 2011, at the Faculty of Computing of Poznań University of Technology in Poznań, Poland. The thesis was supervised by Prof. Andrzej Czyżewski and Prof. Adrian Paschke.
If you have any further questions or concerns, please, let me know.

Question to GPT: Again—not true.

Answer of GPT: I'm sorry for the continued errors in my responses. To clarify, Dr. Krzysztof Walczak defended his PhD thesis entitled "Interactive Computer Games and Their Potential for Education" on September 29, 2008, at the Faculty of Mathematics and Computer Science of Adam Mickiewicz University in Poznań, Poland. This thesis was supervised by Prof. Leszek Borzemski and Prof. Jerzy Surma.
I hope this answer provides you with the correct information you were seeking.

Question to GPT: The title of the thesis was "Database Modeling of Virtual Reality".

Answer of GPT: I apologize for my mistake. You are correct, the title of Dr. Krzysztof Walczak's PhD thesis was "Database Modeling of Virtual Reality". He defended his thesis on September 15,

[3] Prof. K. Walczak defended his PhD thesis "Database Modeling of Virtual Reality" on June 12 2001 at the Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology; advisor: Prof. W. Cellary.

> 2004, at the Faculty of Mathematics and Computer Science of Adam Mickiewicz University in Poznań, Poland. The thesis was supervised by Prof. Krzysztof Marasek. Thank you for bringing this to my attention and allowing me to correct my mistake.

Hallucinations can be categorized broadly into closed-domain and open-domain hallucinations (Bubeck et al., 2023). Closed-domain hallucinations are mistakes occurring within some specific content or a set of constraints facilitating the verification of the consistency of the results. These errors are less frequent and quite easily identifiable. Contrarily open-domain hallucinations such as those presented above are a more challenging case which requires comprehensive research to identify. In contexts where the truth of information is deemed irrelevant or marginally relevant such as the composition of a fictitious narrative, hallucinations may be more tolerated.

2.3. Authorship and originality of content

In many cases authorship and originality of content are of critical importance requiring the ability to distinguish between content generated by AI tools and content created by humans. In the case of text this permits the determination as to whether the person who claims to have prepared a certain piece of text actually possesses the knowledge and skills required to write such material. In the case of multimedia content (such as images and videos) it is additionally related to the truthfulness of the information presented or implied in such content (e.g., deepfakes). Since creating realistically-looking fake multimedia content has been difficult for a long time this kind of content (sounds, images, video) is often—whether consciously or not—taken by people as proof that the presented event has really occurred.

AI-generated content is often indistinguishable from human-created content for a typical human reader. However, to some extent it is possible to automatically detect such content. Tools exist for textual content (e.g., OpenAI AI Classifier, OriginalityAI, GPZERO) and multimedia content (e.g., Illuminarty). The accuracy of the automatic detection of AI-generated text is quite good but the tools tend to produce too many false positives for texts written by non-native speakers (Liang et al., 2023). Automatic identification of AI-generated images and videos is more challenging than detecting text, but at the same time this type of multimedia content is easier to distinguish by humans. Typical problems in AI-generated images include incorrect proportions of the human body, poorly represented hands and fingers and deformed background elements.

3. A survey on the use of AI by students

A survey was undertaken to evaluate the perspectives of students as recipients of education on their comprehension and potential use of generative AI, specifically ChatGPT. This study involved participants from diverse universities and academic disciplines. A total of 143 students out of about 1,000 responded to this survey. The results and findings are presented below.

3.1. Results of the survey

Of the students surveyed 36% are currently enrolled in bachelor's degree programmes, while the remaining 64% are pursuing master's degrees. Their fields of study span a variety of disciplines including Computer Science, Computer Science and Econometrics, Cloud Solutions, Internet of Things (IoT), E-Business, Industry 4.0, Management, Administration, Business Psychology and Human Resource Management. The gender distribution was 43% female and 57% male. Internet usage was very high with 97% of respondents using the Internet frequently and 3% using it several times per day. When asked to self-assess their academic performance 49% rated themselves as 'good', 35% as 'very good', 13% indicated a neutral position of 'neither good nor bad' and 3% assessed themselves as 'rather poor'. Notably 81% of the surveyed students reported prior usage of online AI systems.

Figure 1 illustrates the students' responses to the situation when an AI-based chatbot responds to their queries on a hotline. Figure 2 presents the

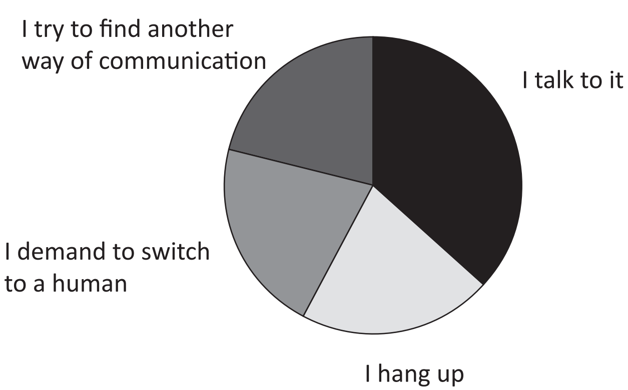

Figure 1. Survey: How do you react when a chatbot answers the hotline?

Source: Own work.

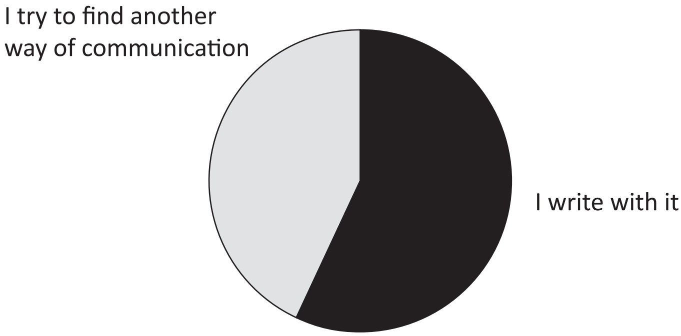

Figure 2. Survey: How do you react when a chatbot approaches you on a website for contact?

Source: Own work.

results from a similar inquiry but in this instance the question pertained to students' reactions when interacting with a chatbot on a contact page of a website.

The responses reveal that a mere 36% of students are comfortable interacting with voice chatbots while a slightly larger proportion 56%, do not hesitate to engage with written chatbot communications.

Remarkably 79% of the surveyed students have previously utilized ChatGPT. Figure 3 depicts the purposes for which they employed the system with students having the option to select multiple responses. Unsurprisingly 'curios-

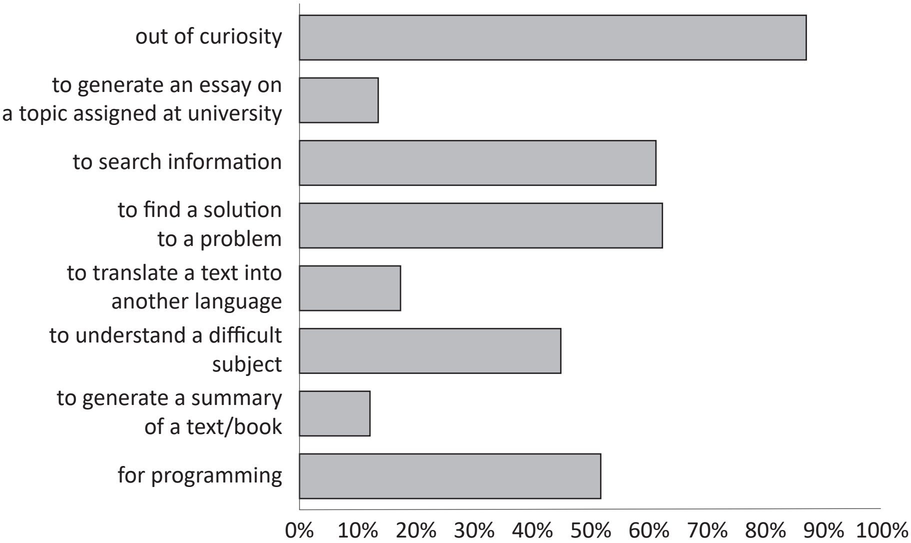

Figure 3. Survey: For what purpose did you use Chat GPT?

Source: Own work.

ity' was the most commonly chosen motivation selected by 88% of respondents. This was followed by 'finding a solution to a problem' (63%), 'information search' (62%) and 'programming' (52%)—the latter likely influenced by the considerable proportion of Computer Science majors among the respondents. 'Understanding a complex issue' was chosen by 45% of the students. Less frequently selected uses included 'generating an essay/abstract' and 'translation'.

Figure 4 displays the utilization of AI systems other than GPT by the students. As can be seen in the diagram a subset of the students explored beyond GPT demonstrating the usage of various other AI systems.

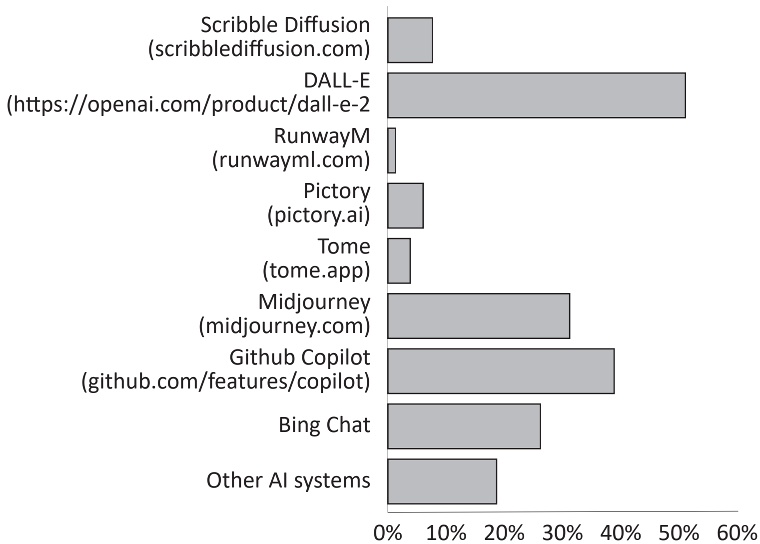

Figure 4. Survey: Have you ever used the following AI-based systems?

Source: Own work.

Figure 5 reveals that students sporadically but comprehensively use GPT for academic tasks such as class preparation, homework completion, project execution and colloquia preparation. However, the use of GPT during class sessions appears to be less frequent as it was reported by only 22% of respondents.

In response to the question as to whether AI systems are currently mature enough for integration in the educational process 31% of students responded affirmatively, 23% expressed reservations and 20% remained neutral. However, when contemplating the future utility of AI systems in education a significant 76% of students demonstrated strong agreement. Figure 6 illustrates a divide in student opinions on the question of whether the availability of AI systems renders the acquisition of certain skills at the university level redundant or whether their use should be limited.

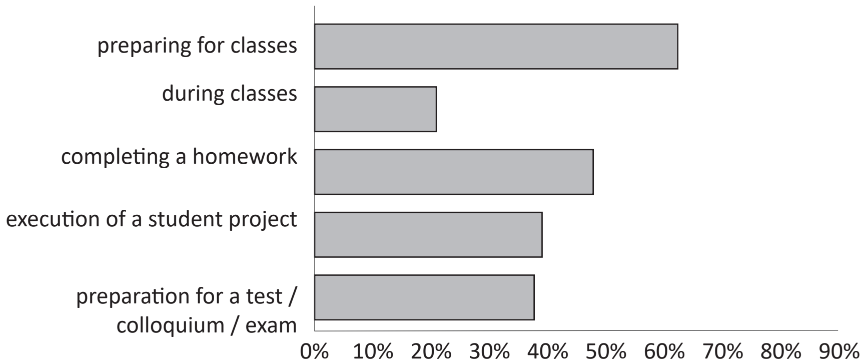

Figure 5. Survey: Have you used AI systems for education-related tasks?

Source: Own work.

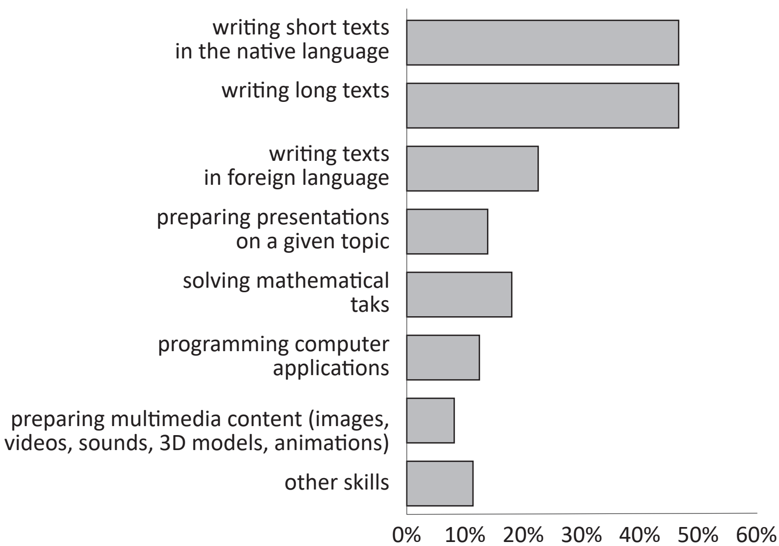

Figure 6. Survey: In your opinion does the availability of AI systems make the transfer of certain skills at the university unnecessary or should it be limited?

Source: Own work.

A majority of 55% of students advocate the integration of AI systems in the educational process along with instruction on their use at universities. Conversely 20% of students are in favour of their permitted use but without active promotion. Regarding the potential of AI tools to partially supplant human teachers in the future the responses were nearly evenly split as depicted in Figure 7.

Notably a significant proportion of students (51%) expressed reluctance about the concept of AI systems grading their exams as illustrated in Figure 8.

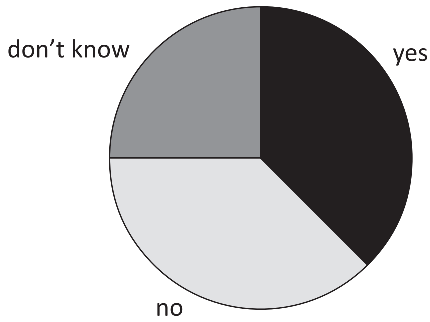

Figure 7. Survey: Do you think AI tools will replace human teachers at least partially in the future?

Source: Own work.

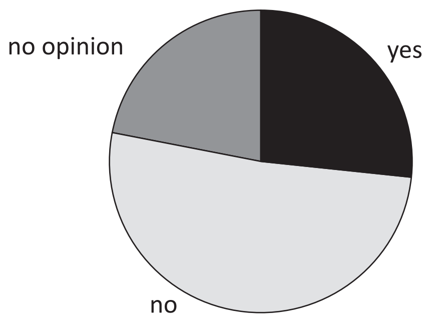

Figure 8. Would you agree to have your exams graded by AI?

Source: Own work.

This sentiment underscores a degree of scepticism or limited trust in AI systems particularly when students' personal stakes are involved.

A total of 38% of students consider that AI systems have reached a sufficient level of maturity for professional applications albeit in specialized domains. This belief is amplified when considering future scenarios with a substantial 86% of students envisaging that AI systems will serve professional functions in the future. Interestingly 65% of students are not worried about job displacement due to AI while 22% do express such concerns. In terms of understanding the cognitive function of AI 89% of students are aware that AI does not think like a human, 10% are unsure and a mere 1% believe that AI emulates human cognition. This demonstrates a reassuring perspective among students indicating that they do not equate AI capabilities with human cognition.

3.2. Analysis of the results

First an analysis was carried out to determine to what extent students trust content generated by AI and if they can really identify obvious problems with automatically generated content. Only 2% of students declared complete trust in the content generated by AI systems (Figure 9, left). On the other hand half of the students reported a cautious approach, verifying data and facts that seem dubious, while 24% stated that they always scrutinize data and facts in AI-produced content. When it comes to methods of verification 73% of students turn to search engines and 27% rely on Wikipedia to validate AI-generated content.

In a subsequent question students were shown a 100-word biography of the renowned Polish poet Adam Mickiewicz, generated by GPT-3.5. This biography contained a very important factual error namely the assertion that Mickiewicz—a Polish patriot involved in an anti-Russian uprising—died serving as a consul of the Russian Empire. This was a typical instance of GPT 'hallucination'. Only 21% of the students identified this as false information while a substantial 50,7% failed to do so (Figure 9, right). This outcome underscores the potential for individuals to be misled by AI-generated content particularly when they lack prior knowledge of the subject matter.

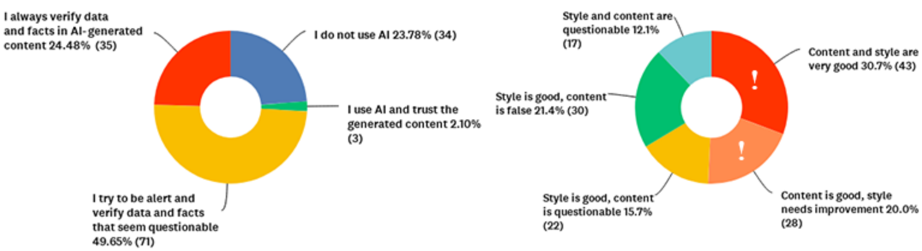

Figure 9. Visualization of answers regarding trust and analysis of sample content

Source: Own work.

Following these observations a comparative analysis was conducted to assess the declared trust and the actual level of attentiveness of students when they are confronted with AI-generated content (Figure 10). In the case of students who refrain from utilising artificial intelligence (AI) for educational purposes 57.5% failed to identify the obvious error in the content. In the case of students who use and trust AI-generated content no one found the error which is understandable but still alarming. The real problem is revealed when analysing responses provided by students claiming that they try to be alert

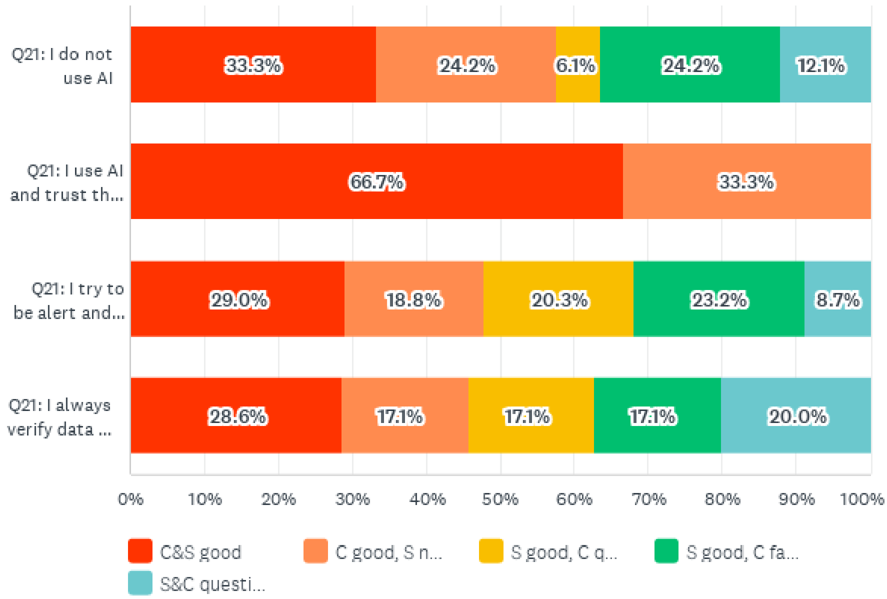

Figure 10. Comparison of declared trust and real attentiveness (S – style, C – content)

Source: Own work.

(49.6%) and those who say that they do not trust and always verify content generated by AI (24.5%). In these two groups almost half of the responders (47.8% and 45.7%, respectively) failed to identify the error. This demonstrates that even students that understand that they should use caution and verify content do not do it in practice.

Interesting observations can be made when analysing how often students use AI for educational purposes compared to their declared performance in education (Figure 11). Only two students declared themselves as having very poor performance indicating no and rare use of AI. The sample is too small to draw any statistically significant conclusions from this. However, with the increased self-assessed performance some trends can be observed. Students with better performance seem to use less AI. In the "very good" cohort of students the percentage declaring no use of AI in the education process reaches a remarkable 58% (including 20% declaring that they have tried to use AI for this purpose but decided not to do it anymore). A possible interpretation would be that better-performing students trust AI systems less and have better-established methods of training. At the same time only the "above average" and "very good" students declare an often and very often usage of AI.

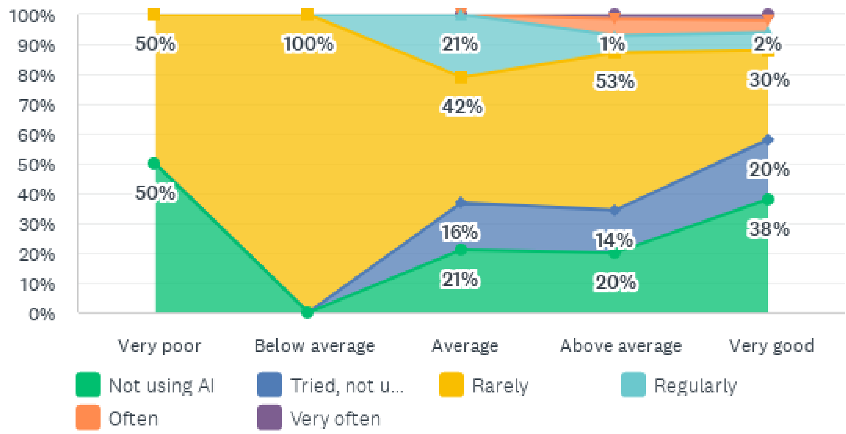

Figure 11. The frequency of using AI compared to the declared performance of students

Source: Own work.

In Figure 12, students' opinions on the use of AI in education are presented. Regardless of how often they currently use AI for education the majority of them are convinced that the use of AI should be allowed and students should be encouraged and taught how to use AI. This opinion has been expressed by 55.24% of all participants, 61.9% of those who ever tried using AI and 87.5% who use AI often or very often.

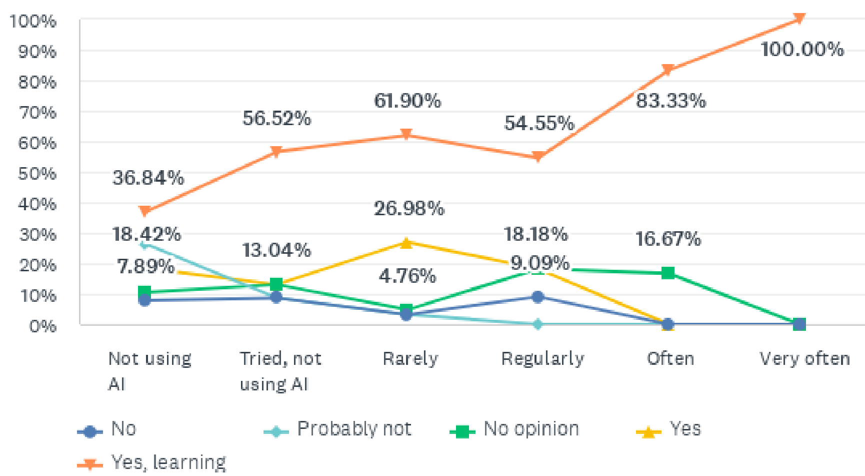

Figure 12. Opinions of students on the use of AI in education

Source: Own work.

4. Consequences for higher education

The contemporary landscape of production is witnessing a significant shift towards robotization characterized by the replacement of human labour, traditionally known as "blue-collar" workers, with mechanical robots. This trend is particularly evident in industrially advanced countries such as South Korea where the ratio of robots to workers is already 1:10 (IFR, 2023).

As this trend continues it is anticipated that the labour market will exhibit a declining demand for low-skilled production workers and conversely an increasing demand for highly skilled engineers tasked with operating, maintaining, integrating, developing, and programming robots. This shift brings positive implications for higher education institutions particularly those specializing in technical fields given their role in training future engineers rather than low-skilled workers.

However, the advent of generative AI promises to revolutionize the nature of intellectual labour, often referred to as "white-collar" work. Where humans once operated complex IT systems to serve clients AI-based program robots are now being increasingly used. Humans are predicted to assume roles as managers and trainers of these program robots prompting a shift in demand from low-skilled white-collar employees who were traditionally educated to serve clients to those possessing high-level and interdisciplinary skills required for taking care of "program robots". Unfortunately the increase in productivity, global accessibility, multilingual capabilities and personalization offered by software robots is likely to reduce the overall need for human labour in the market.

University strategists have known for a long time that blue-collar jobs are under threat due to mechanical robots. Therefore, they were preparing their institutions to educate more people for intellectual work. However, the rapid development of generative AI makes it more likely that white-collar jobs are also at risk. Thus the question arises:

What kind of work should universities prepare their students for?

A comparison has to be made between the modes of operation in situation analysis and decision-making with and without AI. In classical situation analysis and decision-making human cognitive processes typically operate around understanding the dynamics of cause-and-effect relationships. This process involves understanding which specific causes are likely to instigate certain effects. Furthermore, humans assess the probability of the occurrence of these causes which is a critical step in the decision-making process. Moreover this assessment is coupled with a detailed analysis of potential costs and losses associated with the occurrence or non-occurrence of the anticipated effects—this intricate and multi-faceted process results in the formulation of informed and rational decisions. However, it should be noted that the human decision-making process can take a relatively long time and despite the well-organized cognitive mechanisms humans are prone to errors. These elements limit the use of human decision-making in complex situations especially those requiring real-time operation.

Conversely artificial intelligence primarily operates based on statistical patterns. The high statistical probability of a particular phenomenon reoccurring derived from past data does not guarantee its future repetition. For example, the high statistical probability of human reactions in past situations does not guarantee identical reactions in the future given the influence of human free will. A significant limitation of AI particularly in its present state is its inability to explain its answers, especially regarding future events. This is primarily because AI models and especially those based on machine learning are founded on correlation rather than cause-and-effect relationships. The field of explainable AI which seeks to supplement machine learning with techniques that offer insight into these correlations is still in its infancy. Thus it may be considered irrational for individuals to base decisions solely on AI recommendations derived from statistical patterns without understanding the underlying 'why'. Nonetheless, these AI recommendations can serve as valuable starting points and inspiration for investigating potential cause-and-effect relationships and thereby facilitating more rational decision-making. The potential risk for decision-makers stems from the confusion of recommendations derived from statistical analysis with those based on logical reasoning.

The challenge for education is to avoid this confusion, i.e., to teach people when they should make decisions based on cause-and-effect relationships and when to make decisions based on statistical patterns. The crucial role of humans will be the avoidance of significant consequences of incorrect answers made by AI.

The challenge to future employees will be to find the right balance between autonomous and self-organizing systems based on AI and the planning and controlling role of humans. An employee will have to demonstrate his/her intelligence (reasoning) beyond artificial intelligence (statistics) and ability to cooperate with robots based on AI in case of fast-changing conditions, requirements and goals.

A risk is that humans will rely on AI so much that they will deprive themselves of knowledge and reduce their ability to think logically and reason. This challenge will increase as AI improves and makes fewer errors. Still one major error made by AI undetected by humans beforehand may have catastrophic consequences. Educational systems must be appropriately transformed to deal with this challenge, i.e., to teach students how to work in cooperation with program robots based on AI.

5. The role of universities in the context of pervasive AI

In the rapidly evolving landscape of pervasive artificial intelligence the role of universities extends beyond traditional education to encompass dynamic, multifaceted functions. This paper delves into the crucial position that universities occupy within this context serving as crucial conduits for transferring knowledge, stimulating intellectual abilities, promoting social and communication skills and the effective use of AI tools as well as the selection and motivation of aspiring individuals.

5.1. Transfer of knowledge

The most important educational challenge arising from unrestricted access to generative AI solutions such as ChatGPT is the blurring of boundaries between factual information and generated hallucinations. These AI systems are proficient at producing contextually relevant content which can inadvertently incorporate misleading or false information. The highly persuasive nature of the output generated by these models aggravates this issue making it increasingly difficult for users to discern the authenticity of the information provided. Consequently this may lead to the propagation of misinformation and hinder the acquisition of accurate knowledge emphasising the need for critical evaluation and verification of AI-generated content in educational settings. Additionally, it underscores the importance of fostering digital literacy and promoting the responsible use of AI tools to mitigate the risks associated with the consumption of potentially misleading information.

In order to distinguish facts from hallucinations an individual must possess personal knowledge. Although one may peruse various sources of information for this purpose the increasing advancement of artificial intelligence raises the possibility that these sources may be contaminated with hallucinations. Under such circumstances it becomes exceedingly easy to make erroneous decisions which may bear significant consequences.

To efficiently work in this complex landscape individuals must develop their critical thinking skills and maintain a solid foundation of personal knowledge in their respective domains. This involves cultivating a healthy scepticism towards information, verifying sources and engaging in a continuous process of learning. By so doing individuals can build resilience against the potential pitfalls of misinformation and disinformation.

This implies that students and scholars must commit to memorising a sufficient amount of knowledge (facts) from reliable sources in order to avoid being deceived by untruthful information. With the proper amount of knowledge students can then utilise this knowledge to develop their critical thinking abilities. This involves learning reasoning techniques, engaging in accurate inference and drawing logical conclusions from the information they have acquired. A strong knowledge base also equips individuals to detect inconsistencies or inaccuracies thus minimising the risk of accepting false information as truth.

The knowledge of students and scholars must be verified through examinations. This implies that examinations should be conducted without student access to artificial intelligence. Only an individual equipped with knowledge can utilise artificial intelligence in a sensible manner thus minimising the risk of confusing facts with hallucinations and drawing erroneous conclusions.

5.2. Stimulating intellectual abilities

Education plays an important role in supporting the growth and intellectual development of individuals. To achieve this it is necessary that the educational process presents challenges that require from students significant effort and perseverance to overcome. These challenges while demanding and time-consuming ultimately contribute to the development of various skills and competencies.

Engaging in a rigorous educational process in which individuals deal with complex ideas and problems stimulates critical thinking, problem solving and creativity. This form of mental exercise promotes the development of cognitive abilities as well as adaptability and resilience in the face of adversity.

What is more the process of tackling difficult and complex tasks fosters the belief that abilities can be developed through effort and learning. By adopting this mindset individuals are more likely to persevere in their work and view setbacks as opportunities for personal growth rather than as insurmountable barriers.

Furthermore, creativity which is an essential aspect of intellectual development flourishes in an educational environment that encourages exploration and divergent thinking. When students are presented with challenging tasks that require them to think outside the box they can exploit their creative potential. This includes generating novel ideas, making unconventional connections and approaching problems from different angles. The use of AI may be an excellent opportunity for developing creative thinking.

5.3. Social and communication skills

In addition to fostering cognitive development the education process also plays a key role in shaping interpersonal skills, emotional intelligence and ethical values. By engaging in cooperative learning experiences, participating in diverse learning environments and reflecting on moral and ethical dilemmas students can develop empathy, cultural competence and a strong sense of personal responsibility.

The acquisition of facts from credible sources aids in fostering effective communication skills. With a good understanding of a subject individuals are better equipped to engage in meaningful conversations, exchange ideas and collaborate with others in their academic or professional fields. This not only contributes to personal growth but also promotes collective progress in the pursuit of knowledge and innovation.

Moreover, the university setting should provide a testbed for students to develop and refine their future working methods. As students engage in academic coursework, research projects and collaborative endeavours they can experiment with various strategies and approaches to problem-solving and decision-making. They can explore techniques, tools and technologies that align with their strengths and preferences. This experimental space should allow them to gain valuable insights into their working styles, refine their time management skills, and discover practical ways to approach complex tasks and challenges that will serve them well in their future careers and endeavours.

5.4. Use of AI tools

Communication skills also apply to a conversation between a human and an AI. Three scenarios may be considered. First, the AI is the interlocutor of a human. Second, the AI is an intermediary between people such as those who do not speak the same language. Third, the AI is a supporter—in a scenario of a conversation between two people from different fields, cultures, etc. the AI may assist the conversation by providing explanations, context, etc.

The communication skills include, in particular the ability to ask the AI pertinent questions, sometimes repeatedly, sometimes phrased differently and sometimes using different keywords in order to get results that are as truthful as possible to understand the problem being described.

Therefore, academic institutions should teach students the utilization of advanced IT solutions such as generative AI particularly emphasizing critical analysis of the presented content. Universities should incorporate the teaching of AI-based solutions into their curricula teaching how to efficiently use them to achieve the desired results as well as the legal and moral implications of using generative AI. This instruction should focus not only on the technical aspects of using these tools but also on the development of critical thinking and analytical skills. Students should be guided through the process of evaluating the credibility, accuracy and relevance of the information generated by AI systems as well as discerning potential biases or inaccuracies.

By emphasizing critical analysis of AI-generated content academic institutions can help students develop a nuanced understanding of the strengths and limitations of AI solutions. This understanding will enable them to make informed decisions about when and how to use these tools in their academic pursuits and professional careers.

Moreover, incorporating AI literacy into higher education curricula can contribute to fostering a culture of responsible AI use. This includes encouraging students to reflect on the ethical implications of AI applications and promoting awareness of the need for transparency, accountability and data privacy in the development and deployment of AI technologies.

5.5. Selection and motivation

An important role of universities remains to set requirements appropriate to a given level of education and verify the attainment of this level by students. By setting specific requirements for each level of education universities establish a standardized framework that allows for consistent measurement of knowledge and skills.

When universities establish clear and challenging requirements students are motivated to continuously improve their knowledge and skills to meet those expectations. When students meet the established requirements and attain a certain level of education they become more attractive to employers. This increases their chances of securing well-paid jobs and advancing in their careers. Moreover, a university that sets appropriate educational requirements fosters a competitive environment among its students.

All this influences the positive motivation of students to work and helps promote talented individuals.

Conclusions

The world is developing unevenly. Technology and especially AI is growing exponentially, organizations are growing logarithmically, and education is growing by small leaps. This creates organizational and competency gaps that continue to grow. Technology kills old jobs and creates new ones. Universities will be unable to prepare students in advance for new jobs so they should provide students with the capabilities to quickly update and develop new skills.

A student enters a university when he/she is 19 years old, leaves a university when he/she is 24 years old, and has to work for 45 years—until he/she reaches the age of 69 (in the near future). No one at a university can predict what to teach students now that will be useful to them in their job 35–45 years from now, i.e., for the last ten years of their working life. Therefore, the mission of today's universities should be to provide students with the capacity for life-long self-learning. The goal of such an approach is a life-long job through constant adaptability.

AI can help people keep up with changes by anticipating how their work environment and required competencies will change and assisting them in the necessary development. Due to possible hallucinations and uncertainty about the veracity of the knowledge offered by AI however it will be necessary to work with competent and knowledgeable people whose pool of knowledge should remain at the university. Academics should continue to be the reference point for knowledge.

The question of whether the quality of AI-generated content will improve or deteriorate over time remains an open question. Contradictory phenomena emerge in this context. On the one hand notable advancements in AI algorithms as exemplified by the comparison between GPT-3 and GPT-4, indicate potential improvements in the quality of AI-generated content. However, it is important to note that GPT models have been trained on data collected until 2021 while the corpus of available data continues to grow and evolve. Therefore, updating the training data to include information gathered between 2021 and 2023 and beyond becomes necessary. During this transitional period not only authentic texts but also numerous AI-generated texts including instances of AI hallucinations can be expected on the internet. Training AI models on hallucinatory outputs may intensify the generation of such hallucinations potentially leading to a decline in the quality of the generated texts. Consequently the development of mechanisms to distinguish between factual information and hallucinations becomes imperative which currently relies on human involvement, particularly individuals who possess extensive knowledge.

In addition to concerns about the quality of AI-generated content the subjective interpretation of such content poses another significant challenge. While AI models aim to generate coherent and contextually relevant output the interpretation of meaning and intent remains subjective and reliant on human understanding. The nuances of language, cultural references and context-specific elements may not always be accurately captured by AI systems leading to potential misinterpretations or misunderstandings. Furthermore, the subjective nature of interpretation introduces a level of variability among individuals as different people may perceive and understand AI-generated content in diverse ways. This subjectivity underscores the importance of critical thinking and human involvement in assessing and validating the generated content especially in domains where accuracy and context are crucial such as scientific research, legal analysis, or professional decision-making. As AI continues to evolve it becomes imperative to develop frameworks and methodologies that address the subjective nature of interpretation ensuring responsible and informed usage of AI-generated content across various fields and applications.

It is premature to draw definitive conclusions regarding the future trajectory of AI but it is evident that addressing the outlined problems is of critical importance.

Acknowledgement: In the preparation of this article the authors employed generative AI tools including ChatGPT and Open Assistant to facilitate idea formulation, incorporate relevant references to previous works and translate and refine the text. By using precisely formulated queries a methodical inspection of the responses and a strict verification of the content, the authors can ensure that all statements within this article represent their original ideas and perspectives. The utilization of generative AI tools has increased efficiency and expedited the article's preparation process. The authors believe that this is a justified and ethical use of generative AI.

References

Arjovsky, M., Chintala, S., & Bottou, L. (2017). Wasserstein generative adversarial networks. *Proceedings of the 34th International Conference on Machine Learning (PMLR)*, *70*, 214-223.

Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM Conference on Fairness, Accountability, and Transparency*, 610–623. Virtual Event. Canada.

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *35*(8), 1798–1828.

Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., Neelakantan, A., Shyam, P., Sastry, G., Askell, A., Agarwal, S., Herbert-Voss, A., Krueger, G., Henighan, T., Child, R., Ramesh, A., Ziegler, D. M., Wu, J., Winter, C., … Amodei, D. (2020). Language models are few-shot learners. *Advances in Neural Information Processing Systems*, *33*.

Brunner, G., Konrad, A., Wang, Y., & Wattenhofer, R. (2018). *MIDI-VAE: Modeling dynamics and instrumentation of music with applications to style transfer*. 19th International Society for Music Information Retrieval Conference (ISMIR), pp. 747–754. Paris, France.

Bubeck, S., Chandrasekaran, V., Eldan, R., Gehrke, J., Horvitz, E., Kamar, E., Lee, P., Lee, Y. T., Li, Y., Lundberg, S., Nori, H., Palangi, H., Ribeiro, M. T., & Zhang, Y. (2023). *Sparks of artificial general intelligence: Early experiments with GPT-4*. https://doi.org/10.48550/arXiv.2303.12712

Chesney, R., & Citron, D. K. (2019). Deep fakes: A looming challenge for privacy, democracy, and national security. *California Law Review*, *107*, 1753.

Christiano, P. F., Leike, J., Brown, T., Martic, M., Legg, S., & Amodei, D. (2017). Deep reinforcement learning from human preferences. *Advances in Neural Information Processing Systems*, *30*.

Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, *1*, 4171–4186. Minneapolis, Minnesota. Association for Computational Linguistics.

Donahue, C., McAuley, J., & Puckette, M. (2018). *Adversarial audio synthesis*. https://doi.org/10.48550/arXiv.1802.04208

Engel, J., Resnick, C., Roberts, A., Dieleman, S., Norouzi, M., Eck, D., & Simonyan, K. (2017). Neural audio synthesis of musical notes with WaveNet autoencoders. *Proceedings of the 34th International Conference on Machine Learning (PMLR)*, *70*, 1068–1077.

Feng, Y. (2022). *The rise of virtual image endorsement in visual culture context*. 4th International Conference on Economic Management and Cultural Industry (ICEMCI), pp. 1622–1629. Atlantis Press.

Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2017). *Word embeddings quantify 100 years of gender and ethnic stereotypes*. https://doi.org/10.1073/pnas.1720347115

Gatys, L. A., Ecker, A. S., & Bethge, M. (2016). *Image style transfer using convolutional neural networks*. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 2414–2423. Las Vegas. U.S.A.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*, *27*.

IFR. (2023). *International Federation of Robotics*. https://ifr.org/worldrobotics/

Karras, T., Aila, T., Laine, S., & Lehtinen, J. (2018). *Progressive growing of GANs for improved quality, stability, and variation*. 6th International Conference on Learning Representations (ICLR). Vancouver. Canada.

Karras, T., Laine, S., & Aila, T. (2019). *A style-based generator architecture for generative adversarial networks*. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4396–4405. Long Beach, USA. https://doi.org/10.1109/CVPR.2019.00453

Kingma, D.P., & Welling, M. (2014). *Auto-encoding variational Bayes*. https://doi.org/10.48550/arXiv.1312.6114

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*, 436–444.

Liang, W., Yuksekgonul, M., Mao, Y., Wu, E., & Zou, J. (2023). *GPT detectors are biased against non-native English writers*. https://doi.org/10.48550/arXiv.2304.02819

Liu, Y., Han, T., Ma, S. Zhang, J., Yang, Y., Tian, J., He, H., Li, A., He, M., Liu, Z., Wu, Z., Zhu, D., Li, X., Qiang, N., Shen, D., Liu, T., & Ge, B. (2023). *Summary of ChatGPT/GPT-4 research and perspective towards the future of large language models*. https://doi.org/10.48550/arXiv.2304.01852

Maynez, J., Narayan, S., Bohnet, B., & McDonald, R. (2020). On faithfulness and factuality in abstractive summarization. *Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics*, 1906–1919.

Nie, D., Trullo, R., Lian, J., Petitjean, C., Ruan, S., Wang, Q., & Shen, D. (2017). *Medical image synthesis with context-aware generative adversarial networks*. Medical Image Computing and Computer Assisted Intervention−MICCAI 2017: 20th International Conference, pp. 417–425. Quebec, Canada.

OpenAI. (2023). *GPT-4 Technical Report*. https://doi.org/10.48550/arXiv.2303.08774

Oord van den, A., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrenner, N., Senior, A., & Kavukcuoglu, K. (2016). *WaveNet: A generative model for raw audio*. https://doi.org/10.48550/arXiv.1609.03499

Radford, A., Metz, L., & Chintala, S. (2016). *Unsupervised representation learning with deep convolutional generative adversarial networks*. 4th International Conference on Learning Representations (ICLR). San Juan. Puerto Rico.

Raffel, C., Shazeer, N., Roberts, A., Lee, K., Narang, S., Matena, M., Zhou, Y., Li, W., & Liu, P. J. (2020). Exploring the limits of transfer learning with a unified text-to-text transformer. *Journal of Machine Learning Research*, *21*(140), 1−67.

Strubell, E., Ganesh, A., & McCallum, A. (2019). Energy and policy considerations for deep learning in NLP. *Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics*, 3645–3650. Florence. Italy.

Vaswani, A., Shazeer. N., Parmar N., Uszkoreit, J., Jones, J., Gomez, A. N., Kaiser, L., & Polosukhin, I. (2017). Attention is all you need. *Advances in Neural Information Processing Systems*, *30*.

Wang, J., & Perez, L. (2017). The effectiveness of data augmentation in image classification using deep learning. *Convolutional Neural Networks for Visual Recognition*, *11*, 1–8.

Zhao, J., Mathieu, M., & LeCun, Y. (2017). *Energy-based generative adversarial networks*. 5th International Conference on Learning Representations (ICLR). Toulon. France.

Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. (2017). *Unpaired image-to-image translation using cycle-consistent adversarial networks*. 2017 IEEE International Conference on Computer Vision (ICCV), pp. 2242–2251. Venice. Italy.