The role of Artificial Intelligence in future technology

Frederik Schittny

I. INTRODUCTION

Since Artificial Intelligence (AI) is most broadly defined as intelligence exerted by computers or other artificial machines [1, p. 29], the classification of what is an AI system and what is not heavily depends on the underlying definition of intelligence [1, p. 31]. Because AI is a field of research in computer science rather than a specific technology [2, p. 1], it includes a variety of subfields like learning, reasoning, knowledge representation, and data analysis [2, pp. 20-30].

The introduction of AI as a computer science research field took place in 1956 during the Dartmouth Summer Research Project on Artificial Intelligence [3]. However, theoretical considerations of artificial and machine-based intelligence date back earlier, for example, in philosophy [4, pp. 5-6] or science fiction [5, pp. 29-40]. Notable early forms of AI research predating the Dartmouth Research Project include Turing's considerations on machine-based intelligence [6], as well as McCulloch and Pitts introduction of artificial neurons as mathematical structures [7].

The development of AI over the following decades up to today was not characterized by continuous progress, but rather by alternating phases of so-called AI booms and AI winters [8, 9]. The AI booms were phases of consecutive research breakthroughs like the 1960s, in which research focused on the first uses of neural networks and language processing [1, pp. 65-69] or the AI boom from 1980 to 1987 with the widespread use of expert systems [1, pp. 71-74]. More modern phases of AI boom include the years between 2000 and 2011 in which AI was widely incorporated into consumer applications by big technology companies [10], as well as the 2010s in which the main focus of AI development was on deep learning and commercial big data analysis [1, 9, pp. 77-79, p.5]. The phases in between, often referred to as AI winters, were periods of reduced academic progress, frequently accompanied or initiated by a lack of funding, insufficient computational power, and a lack of public interest [8, 1]. The exact classification varies, but the phases from 1974 to 1980 and 1987 to 2000 are often considered the two major AI winters [11, 1].

Today, AI systems are used in various fields including finance [1, p. 1845], medical applications [1, p. 84], manufacturing [2, p. 293], gaming [2, pp. 45-46], logistics [1, pp. 82-83] and military applications [1, pp. 86, 1031]. Despite the widespread use of AI in consumer applications like text autocompletion and correction, voice assistants, or search and recommendation algorithms, and their importance in the everyday life of many people, these systems are often not recognized or perceived as AI applications [12].

II. RECENT ADVANCEMENTS

The last boost of AI development was mainly driven by AI models based on the transformer architecture introduced in 2017 [13]. Especially the introduction of GPT-3 by OpenAI in 2020 [14] marked the beginning of a variety of different tools like chatbots [15, 16], AI-powered search engines [17] as well as tools generating audiovisual materials based on text prompts [18, 19]. Transformer-based AI models were quickly incorporated into numerous consumer software products and saw rapid adoption by end users [20, pp. 437-451]. This advancement marks the milestone of AI models being able to mass generate texts and audiovisual content, which is difficult to recognize as non-human generated by most people [21]. The extreme popularity of these tools makes this milestone comparable to other AI milestones with high media attention, like the chess AI Deep Blue beating world chess champion Garry Kasparov in 1997. In contrast to other consumer software products mentioned above, the new transformer-based models are highly recognized as AI technology by most consumers [22, pp. 33-37].

III. CURRENT PROBLEMS AND NEAR-FUTURE DEVELOPMENT

The quickly advancing development of AI applications has highlighted several technical and societal problems. Solving these problems is the current effort of a variety of companies and research institutions. Since partial solutions and solution approaches are foreseeable or already under discussion, a look at current problems allows for a prediction of the near-future development of AI.

Technical problems include the reliability and explainability of AI models' outputs, as current models can experience so-called hallucinations [23] and struggle with logic and temporal coherence of information [24]. Solving this problem, ideally by making the models' processes comprehensible [1, pp. 1312-1315], would allow for AI models to be used

in more critical and high-risk scenarios. Another technical problem is the high energy demand of training processes, which poses a potential risk for efforts toward climate protection [25]. Smaller, more efficient models could not only alleviate the concerns about the climate impact of AI but would also allow on-device usage with less powerful hardware, often referred to as edge computing [26], making AI applications more accessible and easier to adapt in many situations.

Besides technical problems, there are also highly discussed societal concerns with current AI developments. This includes algorithmic biases discriminating against certain groups, mostly caused by biased training data [1, pp. 1812-1818]. Other concerns are copyright [27] and privacy issues [1, pp. 1807-1812] caused by the mass data collection necessary for the training of large AI models, as well as misinformation issues [22, p. 55] caused by the new ability of AI models to produce realistic media content. While these problems do not directly impede the further adoption and development of current AI technology, they undermine the trust and reputation of AI as a technology in the view of many users. Therefore, specialized regulation and law enforcement, as well as industry standards, are necessary to prevent a potential future backlash against AI.

IV. FROM MULTIMODAL AI TO AGI

The long-term development goal stated by many companies and researchers is Artificial General Intelligence (AGI) [28, 29]. The definition is once again dependent on the conception of the term intelligence, but is generally understood to describe a system possessing human-like intelligence rather than just imitating it [1, pp. 88-89].

One approach to AGI is training increasingly large and capable multimodal models [30]. These can not only process and output one form of data like text but several different ones, including audiovisual information and the ability to perform actions in digital or even real-world environments [31]. While multimodality is a recent concept and can be designated a near-future development, there are not only discussions about whether current models show first features of AGI [32, pp. 92-95] but also doubt from experts whether these models understand their outputs in a human-like way [33].

If AGI can match or surpass human intelligence, this new technology will present both new benefits and challenges. AGI could prove to be more reliable, versatile and capable than current AI models and could potentially take over many tasks performed by humans today, yielding improved productivity and efficiency in many areas. On the other hand, it would drastically amplify existing concerns like job displacement [1, p. 1821-1826] and the potential for weaponization [1, pp. 1803–1807] and could pose an existential threat to humanity [34]. These issues become especially concerning in the case of self-improving systems [1, pp. 1832-1833].

Estimations of when AGI will be achieved vary widely [35, 36, 37]. This is partly due to the vague definition of AGI and the lack of expressive tests. Early tests for identifying intelligent machines, like the Turing test, were arguably passed by programs not falling under the AGI definition [1, p. 1796]. The original approach of modelling human neurons has not proven effective in producing intelligence yet, but is further pursued through research on whole-brain simulation [38]. However, this and other approaches are still constrained by technical limits like computational power and availability of training data [2, pp. 681-703]. Besides this, the inadequate understanding of intelligence itself makes developing approaches to AGI difficult [22, pp. 24-29].

V. CONCLUSION

The near-future development of AI can be predicted with relative certainty based on research and first approaches to current problems. These point towards AI being further adopted into everyday life and taking over more tasks as availability and reliability improve. This progress could be slowed down either by missing regulations and unsolved societal problems impacting the reputation and acceptance of AI or by consequent overregulation.

The long-term goal of AGI is already specified; however, a clear definition, working approaches, and standardized testing methods do not yet exist. Up until now, the development of AI has been characterized by alternating phases of AI booms and AI winters, and it is unlikely the current boom phase will continue indefinitely.

REFERENCES

[1] Stuart Russel and Peter Norvig. *Artificial Intelligence: A Modern Approach*. fourth edition. Pearson, 2021.

[2] George F Luger. *Artificial Intelligence: Structures and Strategies for Complex Problem Solving*. sixth edition. Pearson, 2009.

[3] James Moor. "The Dartmouth College Artificial Intelligence Conference: The Next Fifty Years." In: *AI Magazine* 27 (Jan. 2006), pp. 87–91.

[4] Mariusz Flasiński. "History of Artificial Intelligence". In: *Introduction to Artificial Intelligence*. Springer International Publishing, 2016, pp. 3–13.

[5] Tim Taylor and Alan Dorin. *Rise of the Self-Replicators*. sixth edition. Springer Cham, 2020.

[6] Alan M. Turing. "Computing Machinery and Intelligence". In: *Mind* LIX (1950), pp. 433–460.

[7] Warren Mcculloch and Walter Pitts. "A Logical Calculus of Ideas Immanent in Nervous Activity". In: *Bulletin of Mathematical Biophysics* 5 (1943), pp. 127–147.

[8] Gary Yang. "AI Winter and its lessons". In: *The History of Artificial Intelligence*. University of Washington, 2006, pp. 17–21.

[9] Michael Haenlein and Andreas Kaplan. "A Brief History of Artificial Intelligence: On the Past, Present, and Future of Artificial Intelligence". In: *California Management Review* 61 (July 2019), p. 000812561986492.

[10] Matthew Hoy. "Alexa, Siri, Cortana, and More: An Introduction to Voice Assistants". In: *Medical Reference Services Quarterly* 37 (Jan. 2018), pp. 81–88.

[11] Amirhosein Toosi et al. "A Brief History of AI: How to Prevent Another Winter (A Critical Review)". In: *PET Clinics* 16.4 (Oct. 2021), pp. 449–469. ISSN: 1556-8598.

[12] Pegasystems. *What Consumers Really Think About AI: A Global Study*. 2017.

[13] Ashish Vaswani et al. "Attention is All you Need". In: *Advances in Neural Information Processing Systems*. Ed. by I. Guyon et al. Vol. 30. Curran Associates, Inc., 2017.

[14] Tom B. Brown et al. *Language Models are Few-Shot Learners*. 2020.

[15] Google LLC. *Gemini Models*. URL: https://deepmind.google/technologies/gemini/ (visited on 05/12/2024).

[16] OpenAI Inc. *Introducing ChatGPT*. URL: https://openai.com/index/chatgpt/ (visited on 05/12/2024).

[17] Perplexity. *Getting Started with Perplexity*. URL: https://www.perplexity.ai/hub/getting-started (visited on 05/14/2024).

[18] Midjourney Inc. *Midjourney*. URL: https://www.midjourney.com/home (visited on 05/14/2024).

[19] OpenAI Inc. *DALL·E 3*. URL: https://openai.com/index/dall-e-3/ (visited on 05/14/2024).

[20] Nestor Maslej et al. *Artificial Intelligence Index Report 2024*. 2024.

[21] Joel Frank et al. *A Representative Study on Human Detection of Artificially Generated Media Across Countries*. 2023.

[22] Michael L. Littman et al. *Gathering Strength, Gathering Storms: The One Hundred Year Study on Artificial Intelligence (AI100) 2021 Study Panel Report*. 2022.

[23] Lei Huang et al. *A Survey on Hallucination in Large Language Models: Principles, Taxonomy, Challenges, and Open Questions*. 2023.

[24] Meiqi Chen et al. *Learning To Teach Large Language Models Logical Reasoning*. 2023.

[25] Payal Dhar. "The carbon impact of artificial intelligence". In: *Nature Machine Intelligence* 2.8 (Aug. 2020), pp. 423–425. ISSN: 2522-5839.

[26] Shuiguang Deng et al. "Edge Intelligence: The Confluence of Edge Computing and Artificial Intelligence". In: *IEEE Internet of Things Journal* 7.8 (2020), pp. 7457–7469.

[27] Christopher T. Zirpoli. *Generative Artificial Intelligence and Copyright Law*. 2023.

[28] OpenAI Inc. *Planning for AGI and beyond*. URL: https://openai.com/index/planning-for-agi-and-beyond/ (visited on 05/20/2024).

[29] Google LLC. *About Google DeepMind*. URL: https://deepmind.google/about/ (visited on 05/20/2024).

[30] Nanyi Fei et al. "Towards artificial general intelligence via a multimodal foundation model". In: *Nature Communications* 13.1 (June 2022), p. 3094. ISSN: 2041-1723.

[31] Tadas Baltrušaitis, Chaitanya Ahuja, and Louis-Philippe Morency. *Multimodal Machine Learning: A Survey and Taxonomy*. 2017.

[32] Sébastien Bubeck et al. *Sparks of Artificial General Intelligence: Early experiments with GPT-4*. 2023.

[33] Kyle Mahowald et al. *Dissociating language and thought in large language models*. 2024.

[34] Alexey Turchin and David Denkenberger. "Classification of Global Catastrophic Risks Connected with Artificial Intelligence". In: *AI and Society* 35.1 (2020), pp. 147–163.

[35] Melanie Mitchell. "Debates on the nature of artificial general intelligence". In: *Science* 383.6689 (2024), eado7069.

[36] Baobao Zhang et al. *Forecasting AI Progress: Evidence from a Survey of Machine Learning Researchers*. 2022.

[37] Ragnar Fjelland. "Why general artificial intelligence will not be realized". In: *Humanities and Social Sciences Communications* 7.1 (June 2020), p. 10. ISSN: 2662-9992.

[38] Klaus M. Stiefel and Daniel S. Brooks. "Why is There No Successful Whole Brain Simulation (Yet)?" In: *Biological Theory* 14.2 (June 2019), pp. 122–130. ISSN: 1555-5550.