

ОБЩЕСТВЕНИ КОМУНИКАЦИИ И ИНФОРМАЦИОННИ НАУКИ **PUBLIC COMMUNICATIONS AND INFORMATION SCIENCES**

ARTIFICIAL INTELLIGENCE: HISTORY, TYPES AND BOUNDARIES

Christian Funk

University of Library Studies and Information Technologies

Abstract: *This article deals with the question of what is artificial intelligence. Inspired by the great success of Chat GPT by the organization open AI and the media discussions about the challenges and risks of artificial intelligence, this paper provides an overview of the emergence, types and limits of artificial intelligence. To develop the study, sources from the academic environment and sources from the IT industry were used based on literature research. After an overview of the chronology of the origins of artificial intelligence, where the reader will get an overview of the history of Artificial Intelligence and that science has its root decades back. The article explains the most important technological trends and their effects. The different development streams will be explained, as well as the different types, like natural language processing or neuronal networks. Finally, the article provides an assessment of the opportunities and risks of the new technology. The reader will get an overview of the history of Artificial Intelligence and that science has its root decades back.*

Keywords: *Gen AI, Natural Language Processing (NLP), Neural Networks – Deep Learning, strong AI hypothesis, weak AI hypothesis*

INTRODUCTION

Since November 30, 2022, (generative) Artificial Intelligence is the new hype topic in media, economy and politics. On that Wednesday in November 2022, the company Open AI released the ChatBot Chat GPT, which can communicate with users on a text basis based on artificial intelligence.

The adoption rate for this service on the Internet broke all known records up to that point. Five days after its release, one million users were already counted (Lindner 2023).

By comparison, the two most successful platforms on the Internet at the time, Instagram and Spotify, took two and a half and five months, respectively, to reach this number of users. The rapid dissemination has made the possibilities and potential of artificial intelligence accessible to a broad mass of people in a very simple way. At the same time, the first experts have started warning about the dangers of AI. “Reducing the risk of annihilation by AI should be a global priority alongside other risks of societal magnitude, such as pandemics or nuclear war” (Kelly 2023). The call was issued by the San Francisco-based U.S. nongovernmental organization Center of AI Safety. Co-signer Sam Altman founder of OPEN AI.

Why should AI pose a threat?

Artificial intelligence (AI) has become one of the most disruptive technologies of modern times, revolutionizing industries and changing our daily lives. At its core, AI is about developing intelligent systems that can replicate the cognitive processes of human learning, thinking and problem-solving.

In the following pages, I will explore these questions and illuminate how AI has evolved over time, what types exist, and what the basic mode of action is. And whether the “capabilities” of AI have such a great influence on the economy and society as is prophesied in the media.

As already mentioned above, the topic is currently on everyone’s lips, but the beginnings of this

discipline and the fundamental considerations that are still valid date back several decades.

In this paper the path of evolution of AI will be shown and a brief explanation what AI is will be given. Based on that we will see how risky the technology might be.

RESEARCH METHODOLOGY

For this paper, a comprehensive research methodology incorporating both qualitative and quantitative approaches was adopted. The methodology primarily revolves around a systematic literature review to provide a holistic understanding of AI's applications, challenges, and societal implications.

A systematic literature review was conducted to identify and analyze (ideally) peer-reviewed studies, scholarly articles, and commercial web sources related to AI. This involved defining specific research questions, establishing inclusion and exclusion criteria, and systematically searching electronic databases such as Google Scholar.

The selection criteria for studies and data sources encompassed several key considerations:

Relevance: Studies and sources were selected based on their relevance to the research objectives and questions.

Credibility: Emphasis was placed on peer-reviewed articles, reputable reports, and scholarly publications from recognized institutions and experts in the field of AI.

Currentness: When it came to Technology newer publications were preferred.

RESULTS

The history of AI is not a new phenomenon. It began decades ago. A brief overview of the most important milestones in AI development:

The year 1943 is cited as the “birth year” of artificial intelligence. In that year, McCulloch and Pitts presented the first mathematical model of an artificial neuron, which served as the basis for the development of neural networks (Kaplan 2017).

1950: Alan Turing, in his article “Computing Machinery and Intelligence”, proposes the idea that machines might be able to imitate human thought (Kaplan 2017).

1956: The Dartmouth Conference is held, at which the term “artificial intelligence” is used for the first time; the term was coined by Professor John McCarthy. Researchers present their ideas and developments in the field of AI at this conference (Kaplan 2017).

1960s: Expert systems based on knowledge representation and inference are developed and used in various application areas (Kaplan 2017).

1970s: Research in AI is slowed by a lack of progress in development and limited computing power (Kaplan 2017).

1972 French scientist Alain Colmerauer invents the logic programming language PROLOG (Ertel, 2009)

1980s: Advances in computer hardware and algorithm development lead to further development of expert systems and neural networks (Kaplan 2017).

1990s: Machine learning becomes an important area in AI research. Data-driven algorithms are developed to detect patterns in large amounts of data. Pearl, Cheeseman, Whittaker, and Spiegelhalter bring probability theory into AI with Bayesian networks (Kaplan 2017) .

2000s: Deep Learning, a special machine learning method, is developed. Deep Learning is used to solve complex problems, such as speech recognition (Kaplan 2017).

2010s: AI systems are used in various application areas, such as robotics, medicine, and autonomous driving. Discussions about the ethical and societal implications of AI are also increasing (Stecking 2000).

The development of AI is characterized by ups and downs and is driven by advances in computer hardware and algorithm development.

If we look at Ertel's graphical representation of the chronological development of artificial intelligence, we realize that there were already at the beginning different approaches under the common term artificial intelligence.

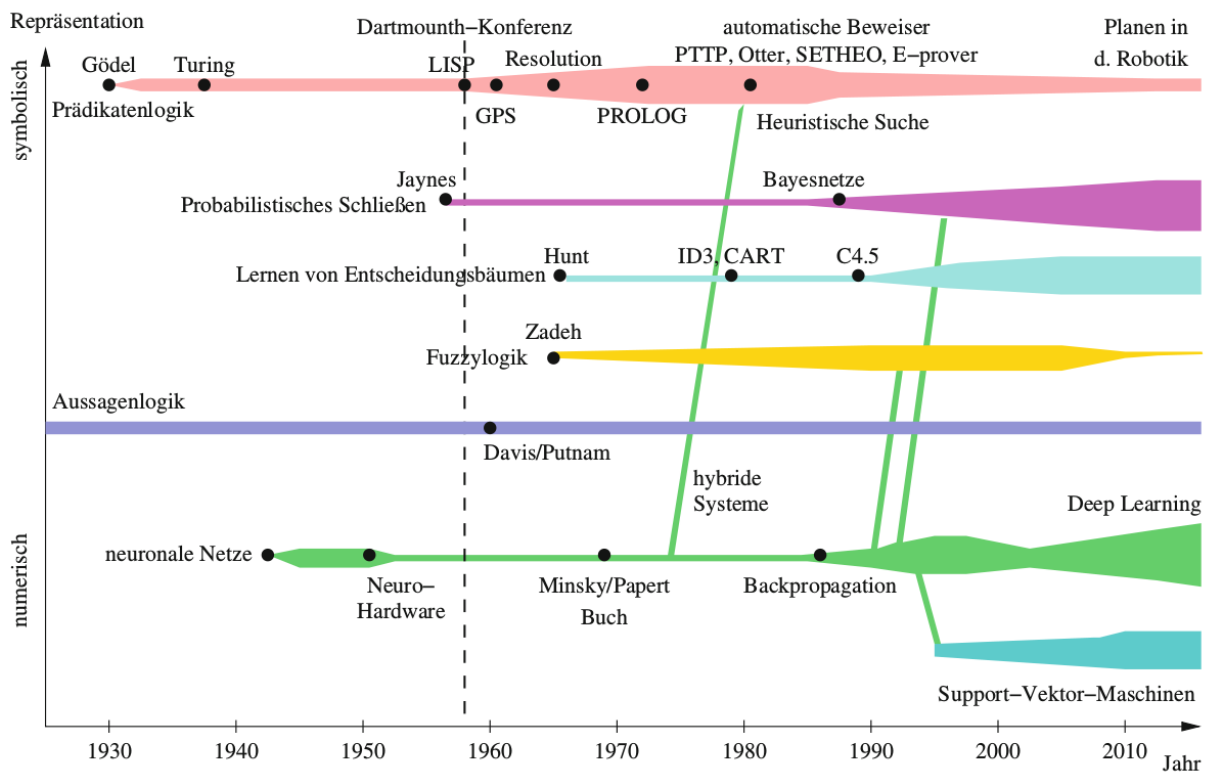


Fig. 1. Overview of development stream (Ertel, 2009)

Types of Artificial Intelligence

Figure 1 shows the different types of technologies and methods associated with Artificial Intelligence, as well as their evolutionary history. From top to bottom, the different AI approaches can be broadly categorized as follows:

- Natural Language Processing (NLP) – pink. cf (Bosse 2020)
- Dialogue robots & expert systems (bots, chatbots...) – purple and turquoise cf (Bosse 2020)
- Knowledge based systems (fuzzy logic) – yellow cf (Bosse 2020)
- Neural Networks – Deep Learning – green cf (Bosse 2020)
- Latest development Generative AI (Chat GPT) – not in the picture

What is Natural Language Processing NLP?

“Natural Language Processing is a theoretically motivated range of computational techniques for analyzing and representing naturally occurring texts at one or more levels of linguistic analysis for the purpose of achieving human-like language processing for a range of tasks or applications” (Liddy 2001).

IBM is giving a quite pragmatic explanation and makes the topic a bit more tangible: Natural Language Processing (NLP) refers to the branch of computer science – and artificial intelligence (AI) in particular – that is concerned with giving computers the ability to understand text and spoken language in the same way that humans do. NLP combines computational linguistics – rule-based modeling of natural language – with statistical machine learning and deep learning models.

These technologies enable computers to process human speech in the form of text or voice data and “understand” its full meaning, including interpreting the speaker’s or author’s intentions and feelings.

NLP drives computer programs that translate text from one language to another in real time, respond to spoken commands, and quickly summarize large amounts of text. Anyone who uses digital assistants such as Amazon’s Alexa or Apple’s Siri, voice-to-text dictation software, customer service chatbots, and other consumer features has very likely encountered Natural Language Processing. But NLP is also increasingly playing a role in enterprise solutions that help streamline business operations, increase employee productivity, and simplify business processes. cf from German (IBM 2024).

What are Dialog robots or expert systems (bots, chatbots)

An expert system is a computer system that uses artificial intelligence (AI) algorithms to simulate the decision-making capabilities of a human expert. The four main components of an expert system are the knowledge base, search or inference engine, knowledge acquisition system, and user interface or communication system. The basic goal of an expert is to duplicate human experts and replace them with problem-solving. There are five main types of expert systems: rule-based systems, frame-based systems, fuzzy systems, neural systems, and neuro-fuzzy systems (Lourdusamy & Gnanaprakasam 2023).

Nowadays, expert systems are no longer considered an active area of AI research... There are several reasons for this: The most important is the drastic increase in computing power, storage space, and networks, which means that vast amounts of data are available in easily accessible electronic formats (Kaplan 2017).

That's why from here on we continue with the system called ChatBot, which represents a further development of the expert systems. A chatbot is a computer program that can mimic human conversation by using voice commands, text dialogues, or both. Chatbots can be integrated into any messaging service and are useful for providing easy and quick communication for users. Good chatbot development is dependent on the chatbot algorithm and the implementation approach employed by the chatbot developer. Today's chatbot landscape is wide (Lourdusamy & Gnanaprakasam 2023).

Also for the ChatBots IBM is offering a hands-on explanation: A chatbot is an application that uses artificial intelligence to converse with humans in natural language. Users can ask questions, to which the system responds in natural language. It can support text input, audio input, or both (IBM 2024).

The terms chatbot, virtual assistant, and conversational agent are sometimes used as synonyms. Chatbots tend to support simpler conversations and more customized tasks. For example, it can tell you whether or not it will rain tomorrow. A conversational agent, on the other hand, might determine that what you really want to know is what to wear tomorrow.

Definition Fuzzy Logic

Is defined as vague logic, fuzzy logic; a domain of logic that allows semantic interpretation of statements that cannot be classified as unambiguously true or false (e.g. "Peter is tall."). Discrete truth values (true and false or 1 and 0) are replaced by a continuous range (usually interval from 0 to 1). For values from this range propositional logic operations are defined. Fuzzy Logic is used for example in the examination of address data (Gillenkirch 2018).

As an example, the correct name of a street can be "Bahnhof Straße". The use of Bahnhofstrasse, Bahnhofstr. or Bahnhof Str. semantically conveys approximately the same information and would be filtered out in an automated 0 or 1 decision – when using an appropriately "calibrated" fuzzy logic, this fuzziness would be accepted and a still sufficiently accurate result would be delivered.

Definition of Neural Networks – Deep Learning

Neural networks, also known as artificial neural networks (ANN) or simulated neural networks (SNN), are a subfield of the machine learning (ML) discipline and are at the heart of deep learning algorithms. Their name and structure are modeled after the human brain, and they mimic the way biological neurons send signals to each other. Artificial neural networks (ANNs) consist of a node layer that contains an input layer, one or more hidden layers, and an output layer. Each node or artificial neuron is connected to another and has an associated weight and threshold. If the output of any individual node is above the threshold set for it, that node is activated and sends data to the next layer in the network. Otherwise, no data is passed to the next layer in the network.

Neural networks use training data to learn and improve their accuracy over time. However, once these learning algorithms are tuned for accuracy, they are powerful tools in computer science and artificial intelligence that can be used to classify and cluster data at high speed. Tasks in speech or image recognition often take minutes instead of hours compared to manual recognition by human experts (IBM 2024).

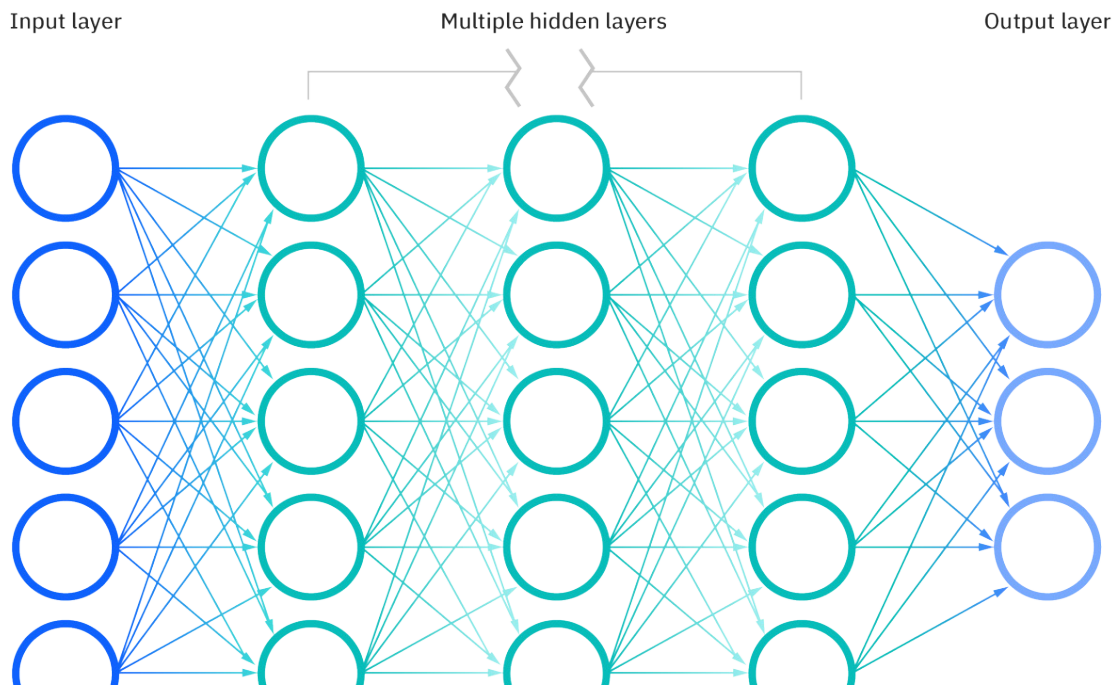


Fig. 2. Schematic representation of the structure of a neural network

Deep Learning and neural networks are often used interchangeably in conversation, which can cause confusion. Therefore, it is worth noting that the “deep” in Deep Learning refers only to the depth of layers in a neural network. A neural network that consists of more than three layers-including both the input and output layers can be considered a Deep Learning algorithm. A neural network that consists of only two or three layers is just a simple neural network.

What is generative AI?

At its core, it is the combined application of NLP with neural networks and deep learning. These models were fed with extremely large amounts of data. According to a McKinsey publication, the data input and thus the use of computing power is exorbitantly high:

“When you’re asking a model to train using nearly the entire internet, it’s going to cost you. OpenAI hasn’t released exact costs, but estimates indicate that GPT-3 was trained on around 45 terabytes of text data – that’s about one million feet of bookshelf space, or a quarter of the entire Library of Congress – at an estimated cost of several million dollars. These aren’t resources your garden-variety start-up can access” (McKinsey 2024).

As we can see above, an “artificial intelligence” has to be trained based on training data. I.e. with a simple image recognition, the program is trained to distinguish for example the image of a cat from the image of a dog. If the image of a horse is presented to this image recognition, the image recognition will initially fail in the task or make a false statement. This, admittedly very simple example, shows the dependency of artificial intelligence on data and its quality and quantity as well as the training of the algorithms. This leads us to the question of security.

The outputs generative AI models produce may often sound extremely convincing. This is by design. But sometimes the information they generate is just plain wrong. This is called a hallucination and even worse, sometimes it’s biased (because it’s built on the gender, racial, and myriad other biases of the internet and society more generally) and can be manipulated to enable unethical or criminal activity.

To avoid these negative effects the provider of Gen AI services need to invest heavily in training and data quality (McKinsey 2024).

CONCLUSION

The strong and the weak AI hypothesis

In the context of Artificial Intelligence, philosophy distinguishes the weak AI hypothesis and the strong AI hypothesis. The weak AI hypothesis is the assumption that machines can act as if they were intelligent. In contrast, the strong AI hypothesis assumes that machines can really think, i.e., they do not just simulate thinking. Or as Searle says originally:

What psychological and philosophical significance should we attach to recent efforts at computer simulations of human cognitive capacities? In answering this question, I find it useful to distinguish what I will call “strong” AI from “weak” or “cautious” AI (Artificial Intelligence). According to weak AI, the principal value of the computer in the study of the mind is that it gives us a very powerful tool. For example, it enables us to formulate and test hypotheses in a more rigorous and precise fashion. However according to strong AI, the computer is not merely a tool in the study of the mind; rather, the appropriately programmed computer is a mind, in the sense that computers given the right programs can be said to understand and have other cognitive states. In strong AI, because the programmed computer has cognitive states, the programs are not mere tools that enable us to test psychological explanations; rather, the programs are themselves the explanations (Searle 1980).

The weak AI hypothesis is taken for granted by most. and the example of image recognition and its possible problems presented above supports this thesis. The strong AI hypothesis is not generally at the forefront of AI research, but rather the pragmatic view that it is irrelevant whether the “thinking” is merely demonstrative or actually happens – the main thing is that the system works.

It is to this first the Strong AI hypothesis that the warnings against AI mentioned at the beginning, have referred to, presumably under the impression of the performance of Chat GPT.

Do we now have to be afraid that in a not-too-distant future mankind will be wiped out by AI-controlled robots, like in the Hollywood movies Terminator with Arnold Schwarzenegger? This question can be answered quite quickly with NO.

However, damage can still occur if the existing tools are used for unethical and criminal purposes.

The simplest are so-called deep fakes, i.e. the artificial generation of (moving) images including the sound (voice) of celebrities, e.g. politicians, who spread fake messages. In addition to these applications, which are created by criminal energy, the danger also lurks in the unclean/unprofessional use of artificial intelligence. As we have seen above, the system’s function depends on data and training.

Let’s imagine a system for automated cancer detection is poorly trained or works with qualitatively and/or quantitatively insufficient data. A false negative diagnosis that ends in a non-treatment of cancer would be an extremely dangerous and harmful result.

The conclusion therefore is to use the possibilities of artificial intelligence BUT always with the premise that a human expert can check the results again and correct them where necessary. This implies, that the user of the artificial intelligence service will need an understanding of how the system came to its conclusion.

REFERENCES

- Bosse, S.** (2020). *Script Einführung in die Künstliche Intelligenz*. Bremen: Universität Bremen; Fakultät Informatik.
- Ertel, W.** (2009). *Grundkurs Künstliche Intelligenz – Eine praxisorientierte Einführung*. Wiesbaden: Vieweg+Teubner Verlag.
- Finance, Y.** (2024). *Yahoo! Finance*. [Online]. Available at: <https://finance.yahoo.com/quote/MSFT/>
- Gillenkirch, R.** (2018). *Gabler Wirtschaftslexikon*. [Online]. Available at: <https://wirtschaftslexikon.gabler.de/definition/fuzzy-logic-34198/version-257704>. [Accessed April 2024].
- IBM** (2024). *Was ist ein neuronales Netz*. [Online]. Available at: <https://www.ibm.com/de-de/topics/neural-networks>
- IBM** (2024). *Was ist ein Chatbot*. [Online]. Available at: <https://www.ibm.com/de-de/topics/chatbots#:~:text=den%20nächsten%20Schritt-,Was%20ist%20ein%20Chatbot%3F,indem%20es%20menschlichen%20Datenaustausch%20simuliert>. [Accessed April 2024].
- IBM** (2024). *Was verbirgt sich hinter der Verarbeitung natürlicher Sprache (NLP)?* [Online]. Available at: <https://www.ibm.com/de-de/topics/natural-language-processing> [Accessed April 2024].
- Kaplan, J.** (2017). *Künstliche Intelligenz*. Frechen: mitp Verlags GmgH & Co. KG.
- Kelly, S.** (2023). *CNN*. [Online]. Available at: <https://edition.cnn.com/2023/10/31/tech/sam-altman-ai-risk-taker/index.html#:~:text=Two%20weeks%20after%20the%20hearing,as%20pandemics%20and%20nuclear%20war>. [Accessed April

2014].

Liddy, E. D. (2001). Natural Language Processing. In: *Encyclopedia of Library and Information Science, 2nd Ed.* New York: Marcel Decker, Inc.

Lindner, J. (2023). *Gitnux*. [Online]. Available at: <https://gitnux.org/chat-gpt-statistics/>. [Accessed April 2024].

Lourdusamy, R. & J. Gnanaprakasam (2023). Expert Systems in AI Components, Applications and Characteristics Focusing on Chatbot. In: *Data Science with Semantic Technologies 1st Edition*. Boca Raton FL: CRC Press, p. 15.

McKinsey (2024). *What is generative AI?*. [Online]. Available at: <https://www.mckinsey.com/featured-insights/mckinsey-explainers/what-is-generative-ai>. [Accessed April 2024].

Searle, J. R. (1980). Minds, brains, and programs. In: *Behavioral and Brain Sciences*. Cambridge: Cambridge University Press, pp. 417–457.

Stecking, R. (2000). *Marktsegmentierung mit Neuronalen Netzen*. Wiesbaden: Springer Fachmedien.

ИЗКУСТВЕН ИНТЕЛЕКТ: ИСТОРИЯ, ВИДОВЕ И ГРАНИЦИ

Резюме: В тази статия се разглежда въпросът какво е изкуствен интелект. Вдъхновена от големия успех на Chat GPT на организацията Open AI и медийните дискусии за предизвикателствата и рисковете на изкуствения интелект, тази статия прави преглед на появата, видовете и границите на изкуствения интелект. За разработването на изследването са използвани източници от академичната среда и източници от ИТ индустрията въз основа на литературно проучване. След преглед на хронологията на възникването на изкуствения интелект читателят ще получи обща представа за историята на изкуствения интелект, която датира десетилетия назад. В статията са обяснени най-важните технологични тенденции и тяхното въздействие. Ще бъдат обяснени различните потоци на развитие, както и различните видове, като обработка на естествен език или невронни мрежи. Накрая статията дава оценка на възможностите и рисковете на новата технология. Читателят ще получи обща представа за историята на изкуствения интелект и за това, че всъщност историята му датира десетилетия назад.

Ключови думи: генеративен ИИ, обработка на естествен език (NLP), невронни мрежи – дълбоко обучение, хипотеза за силен и слаб AI

Кристиан Функ, докторант

Университет по библиотекознание и информационни технологии

E-mail: chfunk@hotmail.com