Systemy Logistyczne Wojsk
Zeszyt 60 (2024)
ISSN 1508-5430, s. 199-216
DOI:

Instytut Logistyki
Wydział Bezpieczeństwa, Logistyki i Zarządzania
Wojskowa Akademia Techniczna
w Warszawie

Military Logistics Systems
Volume 60 (2024)
ISSN 1508-5430, pp. 199-216
DOI:

Institute of Logistics
Faculty of Security, Logistics and Management
Military University of Technology
in Warsaw

Data is the fuel of generative artificial intelligence

Dane paliwem generatywnej sztucznej inelegancies

Wojciech Sokołowski
w.sokolowski@amw.gdynia.pl; ORCID: 0000-0002-8238-6917
Faculty of Command and Naval Operations

Krzysztof Ficoń
k.ficon@amw.gdynia.pl; ORCID: 0000-0002-9153-474X
Faculty of Command and Naval Operations

Marzena Florczyk-Żółtowska
m.florczyk-zoltowska@amw.gdynia.pl; ORCID: 0000-0002-6248-3373
Faculty of Command and Naval Operations

Abstract. The aim of the article is to present the generative technology of artificial intelligence and a general analysis of its achievements and technological limitations. In turn, the hypothesis is as follows: data is the main factor ensuring the development of generative AI. The identified research niche includes stimulants of this development, the need to clarify the concept of generative AI and the indication of the most frequently used sources and methods of obtaining data. At the beginning, the concept and genesis of artificial intelligence were explained, which was necessary to move on to further considerations regarding the analyzed generative artificial intelligence. The self-learning mechanisms used by this type of artificial intelligence are discussed, consisting mainly in the analysis of specific data sets, thanks to which computers learn things that have not been known before. The advantages of generative artificial intelligence were also indicated, including increased productivity, greater efficiency and high creativity. Part of the attention was also paid to the limitations of this technology, including inaccurate and fuzzy output information, uncritical analyzes and often biased conclusions, copyright infringement and high requirements for memory and computing power of computers. Two factors have been identified that have recently stimulated the development of generative intelligence, i.e. the growing computing power of computers and easy access to large data sets. The main subject of considerations are large data sets, dynamically and uncontrollably self-replicating in Internet servers. Today, data resources are collected very efficiently, including: in Big Data, Cloud Computing and Internet of Things technologies, and their

use is almost cost-free and publicly available. It was enough to improve the technology of artificial neural networks, expand multi-layer machine learning and create deep learning for generative intelligence to gain enormous momentum and become a symbol of the beginning of the 21st century. The whole investigation was summarized with a synthetic ending.
Keywords: data, generative intelligence, artificial intelligence, neural networks, machine learning

Abstrakt. Celem artykułu jest przybliżenie technologii generatywnej sztucznej inteligencji i ogólna analiza jej osiągnięć i ograniczeń technologicznych. Z kolei postawiona hipoteza brzmi następująco: dane są głównym czynnikiem zapewniającym rozwój generatywnej sztucznej inteligencji. Dostrzeżona nisza badawcza obejmuje stymulanty tego rozwoju, potrzebę doprecyzowania samego pojęcia generatywnej sztucznej inteligencji oraz wskazanie najczęściej wykorzystywanych źródeł i metod pozyskiwania danych. Na wstępie wyjaśniono pojęcie i genezę sztucznej inteligencji, co było niezbędne, żeby przejść do dalszych rozważań dotyczących analizowanej generatywnej sztucznej inteligencji. Omówiono mechanizmy samouczenia wykorzystywane przez ten rodzaj sztucznej inteligencji, polegające głównie na analizie określonych zbiorów danych, dzięki którym komputery uczą się od postaw rzeczy dotychczas nieznanych. Wskazano również zalety generatywnej sztucznej inteligencji, wśród których wymienić można zwiększoną produktywność, większą wydajność i dużą kreatywność. Cześć uwagi poświęcono także ograniczeniom tej technologii, są to między innymi niedokładne i rozmyte informacje wyjściowe, bezkrytyczne analizy i często stronnicze wnioski, czy naruszanie praw autorskich oraz wysokie wymagania w zakresie pamięci i mocy obliczeniowej komputerów. Wskazano dwa czynniki stymulujące w ostatnich czasach rozwój inteligencji generatywnej, tj. rosnące moce obliczeniowe komputerów oraz łatwą dostępność do wielkich zbiorów danych. Zasadniczym przedmiotem rozważań są wielkie zbiory danych, dynamicznie w sposób niekontrolowany samoreplikujące się w serwerach internetowych. Zasoby danych są dziś bardzo sprawnie gromadzone m.in. w technologii Big Data, Cloud Computing i Internet of Things, a ich użytkowanie jest prawie bezkosztowe i ogólnodostępne. Wystarczyło poprawić technologię sztucznych sieci neuronowych, rozszerzyć wielowarstwowe nauczanie maszynowe i stworzyć uczenie głębokie, aby inteligencja generatywna nabrała ogromnego tempa i stała się symbolem początków XXI wieku. Całość dociekań podsumowano syntetycznym zakończeniem.
Słowa kluczowe: dane, inteligencja generatywna, sztuczna inteligencja, sieci neuronowe, uczenie maszynowe

Introduction

The phrase "artificial intelligence" (AI) was first introduced in the summer of 1956 by John Mc Carthy during a historic summer symposium in at Dartmouth College, which brought together enthusiasts fascinated by the emerging artificial intelligence. The idea for that conference was conceived by John McCarthy, Claude Shannon and Marvin Minsky, all pioneers of future computer science. The concept of artificial intelligence is commonly defined as imitation of human intelligence by computers with the aim of executing genuine tasks. Notwithstanding the passing of more than 50 years, the idea of artificial intelligence is still hardly definable and ambiguous because of its complex and vivid context. The same goes for the concept and definitions of natural (human) intelligence (Ficoń, 2013, pp. 69-80). As far as computer science is concerned, the concept of intelligence is quite often "personalized" based on the example of an agent, namely a software product equipped with a certain situational awareness, understood as the ability to behave effectively under new situations (Wodecki, 2021, p. 13). The agent's goal is to single-handedly solve tasks assigned to it by its environment.

One theory states that "artificial intelligence is a field of computer science dealing with the study of the rules governing intelligent human behavior, the creation of formal models of this behavior and, consequently, computer software that simulates this behavior" (Thagard, 1993, p. 11). Yet another definition identifies artificial intelligence as "a science that incorporates algorithmic, fuzzy logic, evolutionary computing, neural networks, artificial life and robotics" (Russell, Norvig, 2003, p. 16). Artificial intelligence refers to computer hardware and intelligent software capable of handling tasks that typically require human intelligence. What is distinctive about artificial intelligence is that it does not need to be pre-programmed, since it has the capacity for self-learning derived from data and its own programmatic reflection. All it requires is rules and instructions on how to learn and from what data. This is how, in a nutshell, artificial intelligence learns new things "spontaneously".

The concept of artificial intelligence has been gaining ever-increasing importance in the world of science, technology and practice for more than 20 years, although it has actually been in the public eye for more than half a century. The sudden explosion of artificial intelligence is associated with the emergence of very user-friendly and useful generative intelligence technology, namely relying on computer analysis (purposeful scrambling) of vast data sets collected on the Internet. Formally generative artificial intelligence implies no qualitative or theoretical progress in this area, rather a progress in terms of quantity, tools and applications. The fascination with generative intelligence applications stems from its great practicability and the great user-friendliness of its software interfaces. "Gossip" chat technology GPT-3 is a virtually new social network, generating personalized and astonishing results for all its users. "With this kind of computer I can chat on almost all topics", but as long as I am an active player in the network (smartphone) life.

The core research problem has been phrased around the following question: how has generative artificial intelligence technology developed in recent years? The core research problem thus defined has been decomposed into specific problems, taking the shape of the following questions:

1. What is generative artificial intelligence?
2. What technologies have influenced the rapid development of generative artificial intelligence?
3. How is the data underlying the development of generative artificial intelligence collected?

On the other hand, the research tasks has featured the following:

- analysis of technological achievements and constraints in the development of generative artificial intelligence,
- identification of factors that stimulate the development of generative intelligence,
- identification and analysis of technologies used to collect big data.

Meanwhile, the group of key research methods used in the research process can be identified as follows: the method of document study, the method of analysis and logical construction, the method of analysis and criticism of the written materials, including the subject-matter, domestic as well as foreign professional and scientific literature, covering the key items for the research.

The concept and substance of generative artificial intelligence

Conventional artificial intelligence applies to computer systems (programs) capable of executing particular sets of tasks under pre-defined rules or algorithms. These consist primarily of systems built on logical rules and fixed algorithms, which – sadly enough – cannot learn from their experience or improve their functionality over time (SAP, 2024). The most popular classical computational models of artificial intelligence are: artificial neural networks (Tadeusiewicz, 1993), evolutionary and genetic algorithms, logic and fuzzy sets, as well as expert systems and cognitive systems. Save for cognitive science, all of these models rely on fixed logical computational algorithms processed by digital computers. A groundbreaking development in artificial intelligence has been deep learning technology, that is, the self-learning of computers based on big data. The concept of deep learning was known as far back as the 20th century; however, due to the insufficient computing power and training data, computers did not produce the expected results (Lee, Qiufan, 2023, p. 10).

Generative artificial intelligence is an advanced type of artificial intelligence that can create (generate) new content from available patterns detectable in existing data that formally did not exist before. It refers to artificial intelligence models engineered to generate new content of any form whatsoever, such as audio, text, images or videos. Not only does generative artificial intelligence synthesize existing data, it also creates new value, original material that was not there before. Its primary purpose is to create or produce – on the basis of statistical quantity – a new quality (content) in any editorial form, utilizing specific lexical models and self-learning-based algorithms. Advanced self-learning, or in other words, machine learning is guaranteed by today's high-speed computers, running on the basis of upgraded multilayer artificial neural network technology. Machine learning generally necessitates a great deal of corrections before the computer comes up with the best method that yields the best result. However, a human does absolutely not care about this problem; the computer perfects its results on its own with no user input (Tadeusiewicz, 2021, p. 15).

The cornerstone of generative artificial intelligence is deep learning, representing an advanced form of machine learning. Deep Learning relies on the mechanisms employed by the human brain in data processing, pattern formation and decision-making (SAP, 2024). Deep learning models leverage sophisticated symbolic architectures, otherwise known as multilayer artificial neural networks. The said networks comprise a single input and output layer as well as a number of so-called hidden layers, all of which are interconnected to process and transmit information, imitating the action of neurons in the human brain. Deep learning entails a systematic approach to training a multilayer neural network. The input data runs through multiple hidden layers of the network, and at each stage the data from the previous stage becomes the input data for the next stage (Błażewicz, 2023, p. 150).

The mechanism of machine learning involves utilizing the apparatus and methods of statistics to learn models from data (Hurbans, 2021, p. 13). In other words, machine learning might be defined as a set of techniques that rely, among other things, on mathematical statistics and optimization theory as a means of accomplishing tasks such as classification, linear regression, decision trees, Bayesian inference, or Boltzmann machines, to name a few (Hurbans, 2021, pp. 27-39). Notably, when statistics embarked upon its first steps as a science, it aroused a great deal of objection since human individuals were forfeiting their individuality. All of a sudden they would become a number inside a larger set of numbers, be given numerical labels, and be lumped into a bag called a population (Szreder, 2015).

In contrast, the narrower definition of deep learning refers to a set of useful research methods and software tools employed in solving machine learning problems through algorithms relying on multilayer artificial neural networks. The most sophisticated learning through interacting with the environment, categorized as amplified learning, opens up brand new cognitive and application horizons guiding to so-called autonomous intelligence, yielding practical artifacts in the areas of automation, robotics and active mobile objects (Wodecki, 2021, p. 8).

Generative artificial intelligence systems provide greater flexibility than classical computer software; this is due to a machine learning mechanism that does not necessitate the explicit coding of any functions. There is absolutely no need to develop any software to review and synthesize big data sets in a meaningful and purposeful fashion. Instead, a human gives a computer access to a vast amount of diverse data, which ultimately appear in digital form at all times. Based on this data, computers (computer software) train themselves to recognize patterns in this data and, most importantly, to draw conclusions from it using a machine learning mechanism.

Regrettably, traditional machine learning can only train machines to execute one specific task. The size and quality of the data set play an immense role in this respect, for artificial intelligence is only as good as the data upon which it is trained (Adobe, 2024). Ultimately, generative artificial intelligence is evolving into a personal virtual expert (helper, collaborator), which, in fact, can enhance efficiency and productivity across a number of industries as well as the quality of our daily lives (e.g. Alexa, Siri, Cortana). The omnipresence of generative systems has now become a commonly accepted social fact, whose popularity can only be compared to that of the Internet or the smartphone.

In the past, computer software was incapable of handling tasks other than those commanded by the developer. It was up to humans to articulate precise instructions for performing a given task, meaning to program the computer. Although some software did perform highly complex tasks yielding impressive results, it never went beyond the scope of the instructions prescribed by a human (Ficoń, Krasnodębski, 2018). Deterministic programming formed the cornerstone of computers prior to the AI era.

Present-day generative artificial intelligence systems feature greater flexibility due to their intelligence, namely a self-learning mechanism that does not necessitate explicit coding of intended functions. Instead of developing software for a computer in the form of a line of source code, it is enough to provide it with access to a huge amount of diverse data. By analyzing these data sets, computers undergo self-training, meaning they learn from scratch things hitherto unknown to them (Tadeusiewicz, 2021, pp. 10-16). Computer machine learning has brought a new dimension to statistical data analysis, in which the statistical size of the data plays a huge role. For this reason, there is a saying that artificial intelligence is only as good as the data on which it has been trained and taught. The notion of data excellence refers primarily to a quantitative attribute, since high-speed computers must find the qualitative features themselves in these big data sets.

Generative artificial intelligence analyzes vast sets of diverse (digital) data in a much faster fashion than natural human intelligence. Hence arise the following potential quantitative advantages of generative artificial intelligence: increased productivity, higher and better efficiency as well as virtually endless creativity. The first two features relate to quantitative aspects associated with the enormous processing speed of modern computers. Creativity, by contrast featuring a qualitative nature, stems from the concept of machine learning that relies on artificial neural networks and the related ability to make decisions, for example, improvising. At the same time, generative artificial intelligence renders it possible to analyze all potential data sources, detect connections between them, perform generalizations as well as answer a variety of questions.

Generative artificial intelligence can, among other things, develop machine code for software, produce fantastic images, write poetry or produce music, or even modify protein structures based on DNA code. The automatic production of original source software can be conducive to expanding the development horizons of machine learning in the future as well. The constraints of generative artificial intelligence include inaccurate and fuzzy output information, indiscriminate analysis and often biased conclusions, copyright infringement and high requirements for computer memory and processing power, as well as high energy consumption. Arguably copyright compliance appears to pose the greatest concern at the moment. For this reason, it is advocated that a “Do Not Train” label should be placed next to proprietary data, requiring individual author approval (Figure 1).

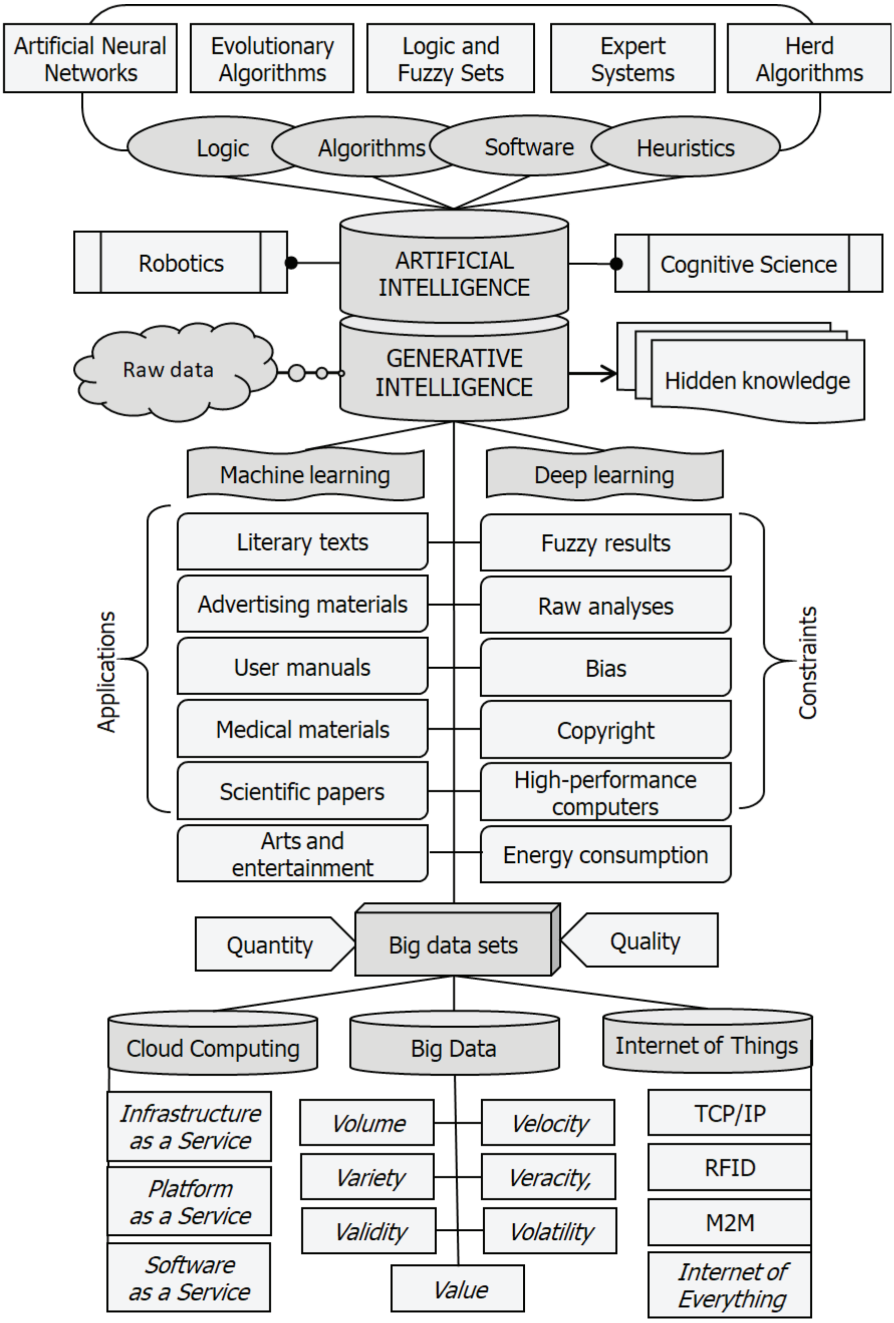

Fig. 1. Generative intelligence vs. artificial intelligence
Source: Own study

The above considerations lead to the fundamental conclusion that, in keeping with the motto of this paper, the fuel, or active resource, of generative artificial intelligence is context-free data, preferably raw and unprocessed. Given this context, data can be defined as: "Data means the representation of facts or concepts in a formalized form that allows their transmission and transformation" (Pasieczny, 1981, p. 86). "Data refers to facts, figures, events, signs that have informational value" (Penc, 1997, p. 72). "Data is the raw input from which information is produced" (Lucey, 1991, p. 14).

The enormous load of potential yet hidden, useful knowledge, as is widely known, is embodied in boundless datasets in web servers that need to be mined, e.g. using generative artificial intelligence technologies. In order to utilize computers for automated processing and analysis of Big Data sets containing static and behavioral data, a new scientific discipline of Data Science was conceived (Szeliga, 2017, p. XVIII), whose underlying task is data-driven machine learning. The following sections will outline the essential sets of data sources and technologies, which are running modern generative systems based, needless to say, on the unrestricted and self-replicating resources of the universal Internet.

Big Data

Big Data means extremely vast data sets with an unstructured, heterogeneous form, the analysis of which requires specialized technologies that allow the mining of new knowledge hidden in such data. Big Data sets are too extensive, diverse and unstructured to be processed using classic database systems. That is why the original Data Mining methods of scrambling and extracting source data are in place (Berry, Linoff, 2000).

Back at the turn of the centuries, when Big Data technology began to emerge, this phenomenon was defined by the so-called 3 V's: *Volume, Velocity* and *Variety.* As the attractiveness of Big Data grew stronger, its description began to incorporate other attributes. Ultimately, a set of 7 V's is in use today, additionally including; *Veracity, Validity, Volatility,* and *Value* (Błażewicz, 2023, pp. 91-96):

- the word *Volume* refers to the size of the collected data. In view of the fact that Big Data includes data originating from various technical and social, public and commercial, random and targeted, structured and unstructured sources, it is disorganized, often chaotic and, above all, unformatted. Data going into Big Data takes the form of text, images, photographs, video, as well as social behavior, road traffic, commercial transactions, satellite images, social surveys, medical materials, court reports, computer software, robotic decisions, etc. Their nature is unstructured, thus making it necessary to employ specialized technologies for selective formatting, analysis and sharing;

- *Velocity* refers to the tremendous increase in the size of data stored on IT media. Nowadays, the principal source of data for Big Data is the Internet and its more than 3 billion vibrant users. It is also worth remembering the Internet of Things (IoT) is developing just as rapidly. The enormous Velocity of the stream of data flowing into Big Data continues to expand its resources, namely Volume. Due to advanced IT solutions, the size of Big Data and the speed of data growth are virtually unlimited. These two extremely dynamic factors have substantially complicated the ability to process a stream of endless data in real time, in terms of extracting useful information and useful knowledge;
- *Variety* stands for the qualitative peculiarity of Big Data, namely the huge variety and complexity of data derived from a wide array of sources, which can be static or mobile, graphical or textual, encoded and encrypted, secure and insecure, or public and private. The Big Data resources include YouTube videos, advertising and promotional materials, scientific and religious texts, legal regulations and administrative acts, as well as millions of photographs from individuals, and more. The majority of this data is collected in an automatic manner by devices and sensors requiring no human intervention, although there is also a huge stream of targeted and dedicated data. Such a huge variety of sources of forms, contents and shapes of data collected in Big Data necessitates its standardization, integration and unification with respect to the capabilities of IT technology on the one hand, and the needs of individual users on the other;
- *Veracity* relates to the reliability of the (input) data, which determines the veracity of the final outcomes. In view of the great variety of sources feeding Big Data, it is a challenging task to assess the veracity of the data flowing in. The data portrays certain facts and phenomena that the recipients may interpret in different ways. As regards large data sets, any potential deviations can be statistically smoothened, hence mitigating their harmfulness;
- *Validity* refers to the up-to-dateness and relevance of data, which, under certain circumstances, must be subjected to prior validation, that is, valuation in terms of situational needs. The fact that data is useful at one stage does not necessarily make it useful in another situational context;
- *Volatility* pertains to the variability of data over time. Data is always characterized by certain discrete conditions, which are subject to natural changes at particular moments. The truth of yesterday does not automatically mean the truth of today. The evaluative criterion of data veracity is subject to sequential changes in the timeline. Ahistoricism of data is widely eliminated by data that is more relevant at a specific point in time;

- *Value* pertains to data quality and is mainly applicable to outcome data derived from Big Data. The Big Data systems make it quite clear that the results generated are reliable and of high quality, e.g. in terms of content. This attribute is also applicable to the costs incurred in the processing of data with the aim of obtaining the desired result, which should not exceed the benefits provided by the outcome data (Hand, Mannila et al., 2005).

Big Data technology refers to large (huge) data sets, also known as gigadata, having high volume, variety and great variability. The dynamics of gigadata changes are enormous; they change extremely rapidly, whether in terms of value, structure, topics, or volume. They appear as quickly as they disappear, often growing rapidly, changing location and access clauses. They may comprise information of considerable value, as long as one knows how to capitalize on them. Furthermore, they make it possible to uncover new knowledge, often concealed among the data. Conventional SQL-type database methods prove worthless in this regard, thus making it necessary to employ separate data scrambling and extrapolation technologies (Ficoń, Sokolowski, 2023).

If properly analyzed, Big Data sets can reveal new correlations between data, expose unseen relationships and interdependencies, and contribute to the generation of new knowledge. The analytical integration of data from multiple sources enables streamlined decision-making processes, better forecasting of development trends and the provision of genuinely needed knowledge. As far as managing big data sets is concerned, it does require specialized technical, social and analytical-synthetic skills (Moczydłowska, 2013, p. 43). Not only does the successful processing of vast data sets affect the realm of practical information management, both at the individual and institutional levels, but it also affects the realm of scientific research methodology and trends in scientific and technological progress. The huge sets of Big Data – acquired and collected at almost no cost – provide an invaluable information resource for state-of-the-art artificial intelligence systems driven by machine learning and multi-layered artificial neural networks.

Cloud Computing

According to the European Commission, the "cloud computing" model can be simply defined as „the storage, processing and use of data accessed via the Internet on computers in another location" (EC, 2012). In the view of the Ministry of Administration and Digitization, cloud computing is: "a computing model based on the rental of services provided by a service provider with regard to: technical infrastructure (e.g., computers, memory, network hardware, networks); software platforms (e.g., databases, identification systems, AI building tools, runtime modules); utility applications (e.g., financial, HR, statistical systems)" (Gaza, 2013). Having generalized the various definitions, one might say that cloud computing refers to applications and services that run on a distributed network, leveraging virtual resources accessible via network communications that use standard Internet protocols. The fact that the information resources of cloud computing are unlimited, as well as that the hardware hosting the cloud is separate from the user, is also consistently emphasized (Oktawave, 2024).

In other words, cloud computing can be further defined as a service for the remote provision of dynamically configured IT resources, both hardware (servers, disks, networks) and software (system, tool, and utility software) used for data storage and processing (Ambrust, Fox et al., 2010, pp. 50-51). Originally, Cloud Computing technology was utilized as a base for software not installed on computers of individual users. As time went on, cloud databases became popular and are now dominant. The evolution of the cloud to the present model started with the classic Client/Server architecture, through Grid Computing and Utility Computing systems.

The concept of cloud computing represents a natural course of evolution of modern ICT technologies; this is because it combines two previously used data processing methods: Grid Computing and Utility Computing. The Grid Computing model relies on the use of extremely high computing power, by processing data in parallel on multiple computer resources employing workstations on a local or wide area network, such as the Internet. The Utility Computing concept, on the other hand, involves the virtualization of existing IT resources and making them available to individual units.

Cloud computing services are broken down into 3 basic models: IaaS (Infrastructure as a Service), PaaS (Platform as a Service) and Saas (Software as a service). IaaS – is primarily about providing technical infrastructure – hardware and networks. PaaS – in addition to hardware, includes hardware-specific operating systems. The most advanced service – SaaS – involves the provision of hardware infrastructure, the respective operating systems and the entire application environment. All installation activities – related to hardware or software in case of SaaS, accordingly to the users' requirements – are performed by the cloud computing provider, while users themselves can focus only on carrying out their own activities, e.g. business.

Virtualization, characterized by simulated availability of any hardware and software resources, as well as information resources, has heavily stimulated the expansion of cloud technologies (Dziembek, 2016). By way of example, so-called virtual machines installed on other computers were used to simulate various operating systems. A single physical computer would run multiple virtual machines that could be attached at any time, depending on users' needs. All attach/detach activities of individual cloud resources take place automatically, thus leveraging such features as cloud scalability and flexibility. Scalability means that cloud computing, subject to the load (number and needs of users), can reduce or increase its capacity. Scalability is directly correlated with the flexibility of a given cloud. The development of cloud services is closely related to the use of the Internet, making it highly popular and generally accessible.

The paramount benefit for the user of cloud computing services lies in minimized costs of operating the server room and the entire IT infrastructure. The cost of cloud fees only applies to the actual resources used, with no extra investment, maintenance, or emergency costs, etc. The user does not have to have any concern about service availability, data transmission speed, software maintenance, data security, equipment reliability, etc. The provider is expected to guarantee continuous availability, the required speed, stable transmission and, above all, data security and robustness of the service provided. Cloud Computing has forged a new paradigm, where the boundary of computing services will be set by economic justification rather than technological constraints (Szpor, Grochowski, 2021, p. 82).

In terms of the organizational side, the basic models of cloud operation include public, private, group (community) and hybrid models. Private cloud (Cloud Internal) is owned or leased by a single entity, enterprise or organization. It is a customer-specific service or infrastructure, unavailable to other users. Public cloud (Cloud External) is available to a potentially unlimited group of recipients (large-scale services). In a group cloud (Cloud community), the infrastructure is shared by a group of recipients having certain common characteristics (e.g., policy, joint venture); hybrid – combines two or more of the previously mentioned types of clouds, although each of the clouds included in the hybrid retains its own characteristic features.

Cloud computing, as a widely available technology in the age of cyber threats, generates the greatest concerns about the integrity and security of data stored on shared drives hosted off-site, which the company has no control over. Notwithstanding the intensive advancement of security systems, this means a lack of absolute privacy and confidentiality of data, which not all users may find acceptable. In addition, there is a risk of becoming heavily dependent on a single cloud service provider and of monopolizing the relationship. Not without significance is also the lack of uniform formal and legal regulations in the relationship between the provider and recipient of cloud services, especially with regard to the continuity and reliability of the services provided.

Internet of Things

The Internet, understood as a global network of interconnected computers, servers and devices, has virtually changed our world, across all its dimensions. For at least 50 years, it has been developing in its classic version of being a universal communication tool for those using advanced computer technology. The popularization of RFID radio codes at the end of the 20th century laid the groundwork for the integration of tagged items into the traditional Internet computer network as well, mainly in the logistics supply chain stream. Thus, in parallel, a new actor emerged in cyberspace resources at the turn of the century – Internet of Things (IoT). As one of multiple definitions puts it, the Internet of Things is "an ecosystem in which sensor-equipped objects communicate with a computer" (Malucha, 2018). In this context, the Internet of Things is a system that enforces omnipresent and continuous communication between Smart objects residing in a specific space. In other words, the Internet of Things can be defined as the combination of physical objects (things, objects) along with their virtual representation on the Internet – global computer network.

The backbone of the Internet of Things IoT is M2M (Machine to Machine) communication, featuring active mediation by the Internet and, ever-increasingly, Cloud Computing. The Internet of Things is already playing a fundamental role in the operation of the TSL sector and driving its modernity and robustness, but above all, its high profitability and market competitiveness. At the same time, it must be said that the Internet of Things began its global career by streamlining logistics, which has always been keen to exploit all scientific and technological innovations.

One could hypothesize that the predecessor of the modern Internet of Things in the 1990s was barcode technology, EAN-128. As early as the 1970s, the EAN graphic bar code was a static and passive carrier of business information, mainly of goods, while dedicated readers and scanners were active information processing elements. Barcode readers were connected by radio to servers operating upon the collected source information.

The early 21st century saw the proliferation of active RFID chips that, when placed on various goods, cargo, or packaging, would automatically emit their own unique identifier in the form of encoded radio emissions. RFID technology has significantly automated business processes – primarily transactional and commercial processes – thus essentially contributing to the vast expansion of global supply chains. Information retrieved by mobile RFID terminals would be the first to go to the database servers of various cloud computing facilities, enhancing the efficiency and flexibility of managing global supply chains. On the basis of RFID technology and modern cloud computing, the global TFL (Transportation Forwarding Logistics) sector was computerized in the first place, amplifying the dynamics of global freight flows.

The technical definition of IoT comprises smart devices with their own IP addresses, connected to other similar devices via the Internet's TCP/IP protocol. To put it another way, a lot of small and big things are connected to a lot of other small and big things via the Internet. According to M. Miller, the Internet of Things is a technological masterpiece. It leverages a variety of existing and in-development technologies and protocols, combining them into cutting-edge models, ultimately creating genuine utility applications. The Internet of Things integrates a number of relatively simple technologies to create something far superior and greater than just the aggregate of the constituent elements (Miller, 2016, p. 34); combines various technologies into a partially autonomous network, each element of which has a mutual, often direct, connection to one another.

Technically speaking, the IoT is a collection of devices comprising embedded sensors that record or generate data and wireless active transmitters/receivers that connect to a network. These devices run in so-called Internet backbone networks using various data transmission protocols. Software that analyzes and processes the collected data and initiates appropriate external communications forms an important piece of the IoT. Automatic identification and communication of physical objects is achieved via wireless RFID technology. By means of radio detectors and sensors, there is automatic detection of the device's status, current reading of technical parameters, exchange of messages with other devices in this network as well as continuous wireless transmission of information (Kamiński, 2018, p. 118).

In keeping with the concept of the Internet of Things, ever-improving RFID radio chips are placed on an ever-increasing array of various mobile and static objects and things; given sufficiently increased range and power, they are employed to automatically identify these objects in time and space. Thanks to telecommunications technologies such as GIS, GPS (Global Positioning System) or GSM (Global System Mobiles), source information about the current location of the tagged objects and their status is continuously collected on cloud computing servers and can be effectively utilized to manage these objects. Cloud Computing may also serve as an intermediary for automatic communication between various Internet of Things objects, sharing up-to-date information with relevant decision-makers – both personal and impersonal.

To make the IoT ecosystem work holistically, it is necessary to have a so-called infrastructure layer relying on both conventional and modern technologies of the classic Internet, nowadays augmented through Big Data technology, Cloud Computing, Smart mobile devices, mobile telephony, and cyberspace. It must bed pointed out that there is no such thing as a one overarching, universal Internet of Things; in fact, each of its applications is individually configured to meet the needs of a specific user. Upon the foundation of the Internet of Things, the concept of the Internet of Everything (IoE), a network that connects people, processes, data and objects, has already germinated, in effect producing entirely new opportunities for action and new application horizons.

The area of applications of the IoT technology is virtually limitless and is used in the public, private and commercial sectors, as well as in scientific research (Moczydłowska, 2013, pp. 47-50). The origins of business IoT applications lie in commercial fields of use, such as telecommunications, transportation, logistics, trade, agriculture, or manufacturing. The IoT technology finds widespread use in healthcare or education, and most importantly in the areas of defense, public safety or public administration. Originated from the TFL sector, the Internet of Things is boosting global commodity flows across the globe to the highest extent and amplifying the intensity of economic processes across all dimensions.

Conclusions

The ever-escalating popularity of the Internet has yielded a tremendous influx of mass data and information across all technical and social sensors in the late 20th century. The Internet has produced – almost spontaneously – a vibrantly expanding huge data sets, generating technologies such as Big Data, Cloud Computing and the Internet of Things. Originally, these huge collections of spontaneously flowing data were a major problem for potential users – conventional OLTP (OnLine Transaction Processing) transactional systems were not likely to reach selective information that was valuable to a specific user. For this reason, OLAP (Online Analytical Processing), namely analytical systems relying on large Internet data sets like Big Data, Cloud Computing or Internet of Things, and especially on specialized data mining and scrambling applications like Data Mining, soon appeared (Berry, G.S. Linoff, 2000). Meanwhile, the business domain yielded attractive business intelligence applications integrating OLTP transactional and OLAP analytical system technologies, running on dedicated data warehouses.

The dynamically growing information resources of the Internet, in addition to the adverse phenomena of the information avalanche and information smog, appeared to be a great opportunity for the resurgence, thanks to deep machine learning, of a new installment of generative artificial intelligence. As far as generative intelligence is concerned, the fascination with large data sets makes perfect sense, for it utilizes the achievements of mathematical and statistical sciences to the highest degree. It is not a secret that statistics most readily operates with large numerical populations, inasmuch as only then are the results relatively reliable (statistically averaged). The seemingly disorderly and context-free large volumes of data collected in Big Data, Cloud Computing or the Internet of Things appeared to be the perfect fuel for such statistical methods as linear regression theory, covariance and correlation, Bayesian inference, cluster analysis, decision trees, to name but a few. Generative intelligence also draws on the tool output of other scientific disciplines sharing a broader theoretical background. More specifically, these include forecasting theory and methods, probability calculus, graph and network theory, game theory or a variety of heuristic methods, such as, search methods (Albrzykowski, 2023).

The deterministic and discrete nature of the data streamlines its digitization and, consequently, its efficient processing in line with heuristic utility criteria. Admittedly, advanced deep learning algorithms also operate through multi-criteria optimization methods, thus forcing researchers to use continuous functions. The rediscovered great flexibility and utility of multilayer neural networks (Ficoń, 2013, pp. 133-174) represents the top performing engine for accelerating generative intelligence. Generative intelligence's tool apparatus, so far as can be seen, is very extensive and diverse, while requiring the powerful computing power of today's computers. The ever-more-real prospect of quantum computers and the exponential growth of the Internet's information resources have created simply limitless development opportunities for generative artificial intelligence. The title fuel is unlikely to run out, and the computing power of computers supporting multilayer neural networks keeps growing stronger.

Nowadays, the philosophy of generative intelligence is a being of versatile importance, penetrating all spheres of our lives. First, it is the *spiritus movens* of the modern Internet; second, it has proven indispensable in the world of economics, finance and business; third, it is the driving force behind many economic industries and scientific disciplines; and last but not least, it is the great hope of robotics, automation and autonomous vehicles (Lee, Qiufan, 2021, pp. 277-299). Artificial generative intelligence is no longer an abstract and academic concept, rather a tangible engineering technology for hands-on activities to streamline our lives across a wide range of its social, technical, research areas.

Networked accessibility to huge computer farms and large database server rooms has fostered the development of generative intelligence and boosted its rapid growth over the past 20 years. Given this kind of environment, training experiments with multi-layered artificial neural networks, accompanied by the vast information resources of the Internet, were a genuine passion for some researchers, hence the rash of so numerous excellent applications. Generative artificial intelligence, powered by infinite and cost-free online database fuel, is likely to grow tremendously. One further driving force in this regard is the huge public interest in the increasingly user-friendly generative intelligence applications, which today successfully compete with the iconic smartphone. The growing computing power of computers and the real vision of quantum computers facultatively amplifies the "intellectual potential" of partner generative intelligence.

BIBLIOGRAPHY

[1] Albrzykowski, L., 2023. Uczenie maszynowe. Elementy matematyki w analizie danych.

[2] Gliwice: Helion.

[3] Ambrust, M., Fox, A., et al., 2010. A view of cloud computing. Communications of the AMC, 53(4).

[4] Berry, M., Linoff, G.S., 2000. Mastering Data Mining. New York: Wiley.

[5] Błażewicz, G., 2023. Marketing Automation. W kierunku sztucznej inteligencji i hieperpersonalizacji. Warsaw: PWN.

[6] Oktawave 2024. Available at: https://oktawave.com/pl/blog/chmura-obliczeniowa/ [16 March 2024].

[7] Adobe 2024. Available at: https://www.adobe.com/pl/products/firefly/discover/how-generative--ai-work.html [22 February 2024].

[8] SAP 2024. Available at: https://www.sap.com/poland/products/artificial-intelligence/what-is--generative-ai.html [22 February 2024].

[9] Dziembek, D., 2016. Cloud Computing – charakterystyka i obszary zastosowań w przedsiębiorstwach, [in:] Knosala, R., 2016. Innowacje w zarządzaniu i inżynierii produkcji. T.2., Opole: Oficyna Wydawnicza Polskiego Towarzystwa Zarządzania Produkcją (PTZP).

[10] Ficoń, K., 2013. Sztuczna inteligencja nie tylko dla humanistów. Warsaw: BEL Studio.

[11] [10] Ficoń, K., Krasnodębski, G., 2018. Zastosowanie algorytmów sztucznej inteligencji do rozwiązania problemu komiwojażera. Gospodarka Materiałowa i Logistyka, 12/2018.

[12] Ficoń, K., Sokołowski, W., 2023. Wykorzystanie komputerowych metod eksploracji danych i odkrywania wiedzy w zarządzaniu kryzysowym. Rocznik Bezpieczeństwa Morskiego, No. XVII/2023.

[13] Gaza, M., 2013. Wykorzystanie potencjału chmury obliczeniowej w Europie. Warsaw: Ministerstwo Administracji i Cyfryzacji, Departament Informatyzacji.

[14] Hand, D., Mannila, H., Smith, P., 2005. Eksploracja danych. Warsaw: WNT.

[15] Hurbans, R., 2021. Algorytmy sztucznej inteligencji. Gliwice: Helion.

[16] Kamiński, A., 2018. Inteligentna fabryka – nowe trendy w rozwoju systemów informatycznych dla przemysłu. Zarządzanie i Finanse, R. 16, 3/2.

[17] European Commission, 2012. Wykorzystanie potencjału chmury obliczeniowej w Europie. Bruksela, dated 27.9.2012 COM(2012) 529 Final.

[18] Lee, K.F., Qiufan, Ch., 2021. Sztuczna inteligencja 2041, Dziesięć wizji przyszłości. Poznań: Media Rodzina.

[19] Lucey, T., 1991. Management Information System. London: DP Publications London Ltd.

[20] Malucha, M., 2018. Internet rzeczy – kontekst technologiczny i obszary zastosowań. Studia i Prace, 54/2, 2018.

[21] Miller, D., 2016. Internet rzeczy. Warsaw: PWN.

[22] Moczydłowska, J. M., 2023. Przemysł 4.0 (?). Ludzie i technologie. Warsaw: Difin.

[23] Moczydłowska, J.M., 2013. Przemysł 4.0. ludzie i technologie. Warsaw: Difin.

[24] Pasieczny, L., 1981. Encyklopedia organizacji i zarządzania. Warsaw: PWE.

[25] Penc, J., 1997. Leksykon biznesu. Warsaw: AW Placet.

[26] Russell, S., Norvig, P., 2003. Artificial Intelligence: A Modern Approach, 2nd edition. Hoboken, New Jersey: Prentice Hall.

[27] Szeliga, M., 2017. Data Science i uczenie maszynowe. Warsaw: PWN.

[28] Szpor, G., Grochowski, L., 2021. Wielka Encyklopedia Prawa XXII Prawo informatyczne. Warsaw: Fundacja USII.

[29] Szreder, M., 2015. Big Data wyzwaniem dla człowieka i statystyki. Wiadomości Statystyczne, 8.

[30] Tadeusiewicz, R., 1993. Sieci neuronowe. Warsaw: AOW EXIT.

[31] Tadeusiewicz, R., 2021. Archipelag sztucznej inteligencji. Warsaw: AOW EXIT.

[32] Thagard, P., 1993. Computational Philosophy of Science. Cambridge, Massachusetts: The MIT Press.

[33] Wodecki, A., 2021. Sztuczna inteligencja we współczesnych organizacjach. Warsaw: PWN.