Artificial Intelligence and Big Data: Technological Advances, Applications, and Ethical Issues

PROF. DR. ŞEREF SAĞIROĞLU
Gazi University Faculty Member,
Vice President of the Information Security Association,
IEEE Blockchain Turkey President

In this article, I will try to explain the recent studies on big data and artificial intelligence (AI), which have gained importance in recent years, and the points to consider in this field. I will share my PhD in the field of AI 35 years ago and being a faculty member who has been working on big data since 2012, who trained the first MS and PhD students in the field of "big data" at Gazi University, who chaired the first IEEE-supported International "Big Data Analytics and Cyber Terrorism Conference" in 2018, who chaired two national workshops in this field, who opened the Big Data Analytics Security and Privacy English MS Program in 2020, who established the Artificial Intelligence and Big Data Analytics and Security Application and Research Center in 2021, who opened the Artificial Intelligence and Big Data Analytics Laboratory in 2024 with private sector support, who participated as an invited speaker in national and international conferences on these subjects, and who is the most cited faculty member in this field in our country.

AI and big data are the two fundamental technologies shaping today's digital transformation. We know that the most important component underlying AI is data. Those who understand the world of big data are also the ones who develop AI large language models.

Indeed, although the concept of AI was introduced nearly 80 years ago, it has become a technology of today, not the future. A major factor in this development is the enormous amount of data generated in recent years and the increasing computational power that has enabled AI algorithms to train large models. These models have evolved into structures capable of answering millions of questions, producing output, presenting information, and translating. While a total of 5 billion GB of data was generated in the digital world in the 2000s, today, data of even greater magnitude is generated in mere days or even hours. This data has made significant contributions to the advancement of science and scientific advancements, as well as being crucial for training and developing AI models. Ultimately, all these developments have fueled each other, leading to the current leap forward.

The term big data, in its most basic sense, refers to massive amounts of structured or unstructured data. This data: Data is constantly generated in various formats and at high speeds from countless sources, including online transactions, smart devices, social media posts, sensors, generative AI (GENERATIVE AI) model outputs, health data, meteorological data, satellite data, and enterprise systems. To explain the concept of big data, it's essential to understand its V characteristics. These three V characteristics, such as very high volume, high velocity, and great variety, make it difficult to manage big data using traditional methods. While volume, velocity, and variety are the distinguishing characteristics of big data, traditional database software is inadequate to process, store, and analyze such large and complex structures.

The set of methods developed to extract value from big data is called big data analytics. Big data analytics provides significant benefits to institutions, companies, and organizations by uncovering trends and relationships hidden within vast data masses. Indeed, organizations that generate value from data gain significant advantages over their competitors, develop new, different, and advanced solutions, and achieve superiority. On the other hand, AI is a branch of science that aims to enable machines to imitate human intelligence and perform tasks such as learning, reasoning, solving engineering problems, and understanding language. Previously limited to rule-based expert systems and simple algorithms, AI applications have now made significant strides with machine learning, deep learning, and AI techniques. Especially since the 2010s, thanks to the advancements in graphics processors (GPUs) and cloud computing, deep neural networks have been trained on massive datasets and have begun to successfully solve challenging problems such as image processing and recognition, speech recognition, and natural language

processing at human-level levels. In recent years, AIs like AlphaGo defeating the world champion in Go, the AlphaFold algorithm breaking new ground in biology by solving protein structures, and the launch of ChatGPT in November 2022. The remarkable tools offered to end users with the large language models that began in the 1960s are among the striking advances in this field. Systems that understand written, spoken, photographic, and video, answer questions, and engage in real-time conversations are now part of our lives. In short, advancements in big data and AI are progressing together, creating a transformation that feeds on each other.

Application Areas and Opportunities

AI and big data have permeated nearly every sector in today's world. From healthcare and education to finance and manufacturing, from education to innovation, and from theory to practice, we see the applications of these technologies in many areas, and we hear more and more about their successes every day. These include:

- Healthcare: Big data analytics and AI are revolutionizing disease diagnosis and treatment. For example, image processing-based AI systems can detect tumors in medical images like X-rays and MRIs faster and with higher accuracy than human experts. Robotic surgical systems like DaVinci assist surgeons in surgeries, reducing human error. Furthermore, in drug discovery, AI can scan billions of molecules to suggest new drug candidates and identify the most promising options for clinical trials, presenting them to us. The rapid production of over 200 million proteins is a prime example of this.
- Education: AI-powered education platforms can provide personalized content that adapts to students' learning pace and level. Virtual teachers and chatbots can provide 24/7 one-on-one support, answering students' questions, developing software and software applications based on given instructions, transcribing voice to text, converting text to images and videos, interpreting any digital data presented to them, performing language translations—in short, they can easily perform many difficult tasks. Thanks to big data analytics, educational data can be analyzed to identify students' strengths and weaknesses, create personalized study plans, address unclear topics, terms, or content, identify students' levels, learning desires, or capacities, and implement customized educational programs. However, when preparing assignments and reports using AI, ethical principles must be adhered to, citations and sources must be cited, and original content must be produced more carefully than ever.
- Finance: In the banking and finance sector, big data and AI models are used in everything from fraud detection and algorithmic bias, to forecasting and identifying trends, to real-time detection of anomalies in credit card transactions and loss and fraud, and to preventing fraud, privacy, or security breaches before they occur. Systems that analyze customer data can provide personalized banking services and investment recommendations, and can predict potential risks. Furthermore, advanced language models such as ChatGPT, Gemini, DeepSeek, Grok, and Claude are used in finance to summarize reports, graph data, and automatically answer customer questions by voice.
- Manufacturing and Industry 4.0: Big data from sensor-equipped smart factories and Internet of Things (IoT) devices can be analyzed to increase production efficiency. AI-powered quality control systems can be used in many areas, including detecting defects on the production line at levels subtle enough to escape the human eye, reporting driver and vehicle anomalies, developing predictive maintenance applications, optimizing and planning maintenance schedules by analyzing sensor data before machinery malfunctions, creating new graphics and animations from images, proposing new tools, equipment, objects, and designs, and critiquing and improving existing ones. Within the scope of Industry 4.0, robots and autonomous systems collaborate with humans on production lines to provide speed and flexibility. Recently, "artificial intelligence factories" have offered new perspectives in the application of AI to industry.
- Energy: Big data and AI also play a significant role in smart grids and renewable energy management. Data collected from solar panels, geothermal power plants, batteries, and wind turbines can be analyzed with AI to generate energy production forecasts. This makes it possible to balance energy demand with production and increase efficiency.

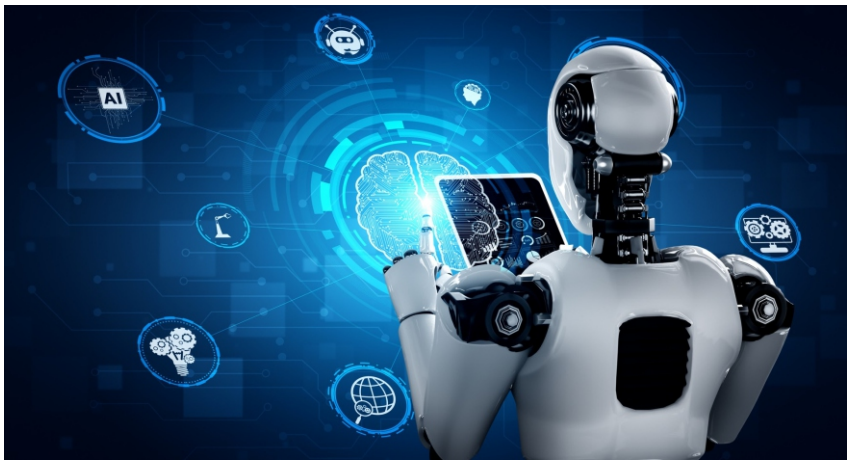
This allows for real-time monitoring of large data streams to detect anomalies in electricity distribution networks, allowing for rapid detection and control of leaks or faults. This allows smart grids to manage energy more reliably and effectively, increase efficiency and reliability, and, most importantly, contribute significantly to balancing energy supply and demand.

- Daily Life: As individuals, we experience AI every day, test and use new models, and encounter new technologies. Many solutions, such as the voice assistants on our smartphones, unlocking screens with facial recognition, voice searches, and making our lives easier with personalized content on social media, are actually applications of AI and big data. E-commerce sites use big data analysis to learn our preferences and habits and offer personalized product recommendations.

  Navigation apps process real-time traffic data and offer us the fastest routes.

  Augmented reality-supported language translation apps can instantly translate text when you point the camera at it. Translation is available to all languages. Even if we don't speak a foreign language, we can easily communicate with citizens of different countries. It's possible to increase this number. All these examples demonstrate that AI and big data offer countless opportunities that can improve our quality of life.

  Thanks to AI, it's now possible to "possess all the world's data, languages, culture, knowledge, technology, behavioral models, teaching methods, and ultimately,



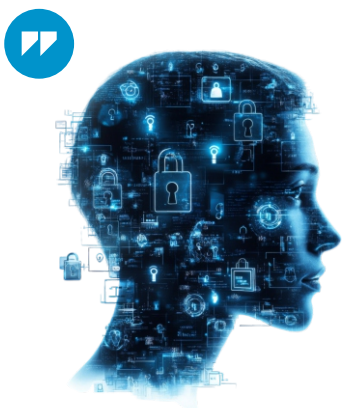
all scientific advancements." These technologies go far beyond just an intelligent robot. These systems can not only understand what's written, photographed, pictured, visualized, coded, spoken, and all the data presented to them, but can also translate and interpret in any language, summarize text, generate code, detect and correct errors, offer systematic solutions, and engage in real-time dialogue. These systems can summarize texts and even produce articles or code. Furthermore, when equipped with medical knowledge, an advanced AI model can pass medical specialty exams and become a good student, teacher, or expert professor in the field. AI-powered robots can perform surgical operations, and driverless cars can navigate roads without human intervention. When used correctly, these technologies transform into technological assistants that make our lives easier, increase productivity and innovation, relieve us of routine and tedious tasks, and offer us great opportunities.

Ethical and Security Issues

The transformation brought about by AI and big data also raises serious ethical, privacy, and security issues. While significant risks include information security breaches, encouraging crime, increasing public fear, deception, data manipulation, personal data disclosure, profiteering, sexism, racism and discrimination, society manipulation, system hacking, and copyright infringement, they also pose societal, national, and global risks. Some of the key issues of these risks can be summarized as follows:

- Data Privacy: Success in big data analytics often depends on collecting large amounts of data related to tasks, personal data, and corporate information assets, and developing the most appropriate solutions for that specific area. For example, when considering personal data, there is a danger of violating individual privacy. Our location information, online habits, phone usage habits, app preferences, health records, eating habits, Sensitive data, such as our habits and walking information, can be misused when collected and analyzed. The Cambridge Analytica Scandal of the 2010s is a prime example of this. This example demonstrated that unauthorized

analysis of social media data can even affect democratic processes, and that these processes concern all countries. As we emphasized in a 2013 study, privacy and security are among the most critical issues in the field of big data, a field whose importance remains unabated today and will only grow in the future. Therefore, legal regulations for the protection of personal data (e.g., GDPR in the European Union, KVKK in our country) are of paramount importance. Companies, institutions, and countries should invest more in data security, prevent all attacks on personal and corporate information assets, close vulnerabilities, eliminate abuses, develop strong encryption approaches, and process data using anonymization techniques.



In order for our country to have a say in the AI transformation, it is vital that it produces its own technologies, strengthens the legal infrastructure based on ethical principles, and increases the AI literacy of society.

- Bias and Discrimination: Artificial intelligence systems can unknowingly learn biases in training data. If the training data contains biases reflecting social inequalities, AI models may reflect these biases in their decisions. For example, it has been revealed that the AI software used by some companies in recruiting has historically discriminated against Black people, tended to systematically exclude women from new candidates, and was more tolerant of white people. Similarly, facial recognition systems have a higher error rate when recognizing dark-skinned individuals and can exhibit racial bias. Such undesirable effects undermine the principles of fairness and equity in AI applications. To overcome this problem, data sets must be diversified to ensure equal and fair representation, the explainability of AI model decisions (the reasoning behind why the model reached a conclusion) must be increased, and human oversight must be retained. Indeed, "explainable AI" is a current research topic and is particularly used to explain critical decisions in areas such as defense, healthcare, or cybersecurity.

- Cyber Threats: AI systems can be subject to cyberattacks or themselves be used as attack tools. For example, AI-assisted deepfake technologies can generate fake images, audio recordings, pictures, or text that are difficult to distinguish from the real thing. These images are misused for disinformation and reputational damage.

  Large language models can help attackers create more convincing phishing content, develop malware code, or, in particular, create personalized persuasion, deception, and deception content. Furthermore, a cyberattack on AI systems managing critical infrastructure can disrupt vital networks such as the power grid or transportation system. Therefore, it is recommended to take extra precautions for the security of AI applications, develop defense mechanisms against algorithmic attacks, establish specialized testing environments for AI security, develop a blockchain-based model security stack, and establish national AI Security Institutes.

- Employment and Economic Impact: It is clear that AI and automation, as with the development of internet technologies, will lead to changes in the labor market as machines take over some routine tasks. This will pose a risk of unemployment, especially for workers in routine jobs. For example, it is clear that with the advancement of driverless vehicle technologies, the need for the driver profession may decrease in the long term, the need for translators and interpreters will gradually disappear, customer representatives will no longer be needed with the increase in automated call center assistants, there will no longer be a need for tourist guides, there will be less need for language tutors, and many other professions will disappear.

  On the other hand, it is clear that new technologies will also create new professions and employment opportunities. According to World Economic Forum reports, AI and robotic automation could eliminate approximately 80 million jobs globally by 2030, but could also create nearly 100 million new ones. While this net increase appears

positive, it also means that the employment structure will radically change, and jobs and ways of doing business will evolve. To adapt to this change, countries need to update their education systems, equip their workforce with new skills, and establish social support mechanisms.

- Legal and Regulatory Challenges: When flawed decisions or unintended consequences of AI systems occur, it is unclear who bears legal responsibility. For example, if an autonomous vehicle or drone crashes, will the manufacturer, the software developer, or the vehicle owner be held responsible? Questions like these highlight the need for new regulations in AI ethics and law. The EU has taken pioneering steps in this regard, drafting a comprehensive Artificial Intelligence Law that classifies risks into four distinct groups. By classifying AI applications according to their risk levels, it aims to strictly regulate high-risk systems and ban AI applications that pose unacceptable risks.

Similarly, countries such as the US, China, the UK, and Germany have established AI ethical principles and are preparing to implement sanctions for violations within a short period. The National Artificial Intelligence Strategy, published for 2021-2025 and revised from time to time, outlines principles for establishing an ethical and regulatory framework. In this area, work continues on the transparency, auditability, and accountability of AI systems, and emphasis is placed on developing new large language models.

Despite all these risks and challenges, abandoning AI seems impossible. The crucial point is to act with a responsible understanding of AI and develop technology for the benefit of humanity. For our country to have a say in this field, we must create our own large language models and make them freely available to the entire country. Additionally, individuals and institutions must increase their AI literacy, use systems with awareness of the risks, expand the use of AI tools in routine tasks, and correctly understand and use the technology with confidence. This will ensure both user awareness and the safe management of the entire process, adhering to ethical principles at the developer level. This requires continuing research, development, and use of AI in our country, recognizing that AI is a tool that will be shaped by the human factor, and expanding these technologies while being aware of the opportunities and threats it presents.

Scientific Contributions and Initiatives in Türkiye

Turkey is striving to take its place in the global race in AI and big data. In the International Artificial Intelligence Index, Turkey ranks 47th out of 160 countries. Significant advances have been made in this field in recent years; Institutions such as the Ministry of Industry and Technology, the Presidential Digital Transformation Office, TÜBİTAK, and the Information and Communication Technologies Authority are conducting various studies to strengthen the AI ecosystem. The National Artificial Intelligence Strategy (2021-2025), published in 2021, outlined Turkey's roadmap in the field of AI and defined the steps to be taken in education, R&D, infrastructure, and legal aspects.

In addition, new undergraduate programs such as AI and Data Engineering are being opened at universities, AI startups are being supported in technoparks, and the private sector is increasing its investments in this field.
In the academic field, Turkish researchers are producing internationally acclaimed publications, events, patents, and studies on AI and big data. A notable initiative in Turkey on AI and big data is the "Artificial Intelligence and Big Data" open-access book series. This series, conceived as part of the Turkish Brain Project, conducted in collaboration with the Presidential Digital Transformation Office, aims to increase Turkish open-source knowledge in the fields of artificial intelligence and big data in Turkey. These comprehensive works compile the experiences of academics and industry experts in Turkey and contribute to the literature, while also providing up-to-date Turkish resources for young researchers.

Open-source approaches are crucial for the scientific development of countries. Our country's progress in Artificial Intelligence and Big Data is driven by publications, awareness campaigns, and the development of ecosystems.

The creation of technology development environments and the content produced on this topic are directly proportional to them. Indeed, there are reports showing that open-access studies have a

multiplier effect on countries' digital transformation. Acting with this perspective, our country is striving to expand the AI ecosystem by encouraging data sharing and collaboration at the national level.

For example, AI Institutes and AI Research Centers at universities, established through partnerships with various institutions, strive to develop local and national AI applications through interdisciplinary projects. Annual academic and sectoral conferences bring together researchers and sector representatives, paving the way for the sharing of knowledge and experience.

Conclusions and Evaluations

AI and big data are the most critical technologies of the 21st century and contribute most significantly to human development. These technologies offer numerous benefits, including early diagnosis in healthcare, comfort and safety in transportation, equal opportunity in education, robotization and increased productivity in industry, automation of routine tasks, ease of design, a revolution in language learning, optimization in solutions, and unlimited communication. These technologies, when used incorrectly or uncontrolled, can lead to a wide range of social problems, including privacy violations, social collapse, unemployment, disinformation, bullying, distrust, oppression, exploitation, and destruction of vulnerabilities. Therefore, both as individuals and as a society, we must be active participants in this transformation, recognize and manage risks, be open to innovation, and, most importantly, develop and use our own technologies and solutions to solve our own problems. Ultimately, as we confidently advance into the future shaped by artificial intelligence and big data, it is essential to maintain the "human-centered perspective" that forms the core philosophy of our national AI strategy. If we learn how to best generate value from data, prioritize the tools used to generate value, and prioritize collaboration and open science, we can build a future powered by artificial intelligence and big data faster and make our future more livable.

Note: This article was prepared based on my presentation titled “Artificial Intelligence and Our Future” that I presented at the Murat Karakaya Academy on May 23, 2025.