IDFI

Institute for Development
of Freedom of Information

ARTIFICIAL INTELLIGENCE:

MEANING, INTERNATIONAL STANDARDS, ETHICAL NORMS, RECOMMENDATIONS AND CONCLUSIONS

POLICY PAPER

Tbilisi, 2021

Author: Nata Goderdzishvili

The study was financed by the International Center for Not-for-Profit Law (ICNL). The Institute for Development of Freedom of Information would like to thank ICNL for the support in the Study's development. ICNL does not necessarily share the opinions here within expressed. IDFI bears the sole responsibility for the content.

Artificial Intelligence: Meaning, International Standards, Ethical Norms, Recommendations and Conclusions, 2021

TABLE OF CONTENTS

INTRODUCTION

Information and communications technology (ICT) is a rapidly growing and developing area of the fourth industrial revolution – digital transformation era. It plays a special role in the formation of the global agenda, global economic developments, societal development, public administration and governenance. For the modern person, regardless of their occupation, interests or work style, adaptation to the digital world and the acquisition and improvement of the corresponding skillset is vitally important.

Is it necessary to use these technologies in the process of public governance? The answer to this question is unambigiously clear – modern challenges require modern solutions that utilize any available tools. The public sector of any developed country in the process of dgital transformation should be actively researching and learning new technologies and modern innovative solutions, and through the use of these technological innovations it should be trying to increase the effectiveness, accountability, and openness of the public sector and provide user-oriented, trustworthy and reliable services to the information society.

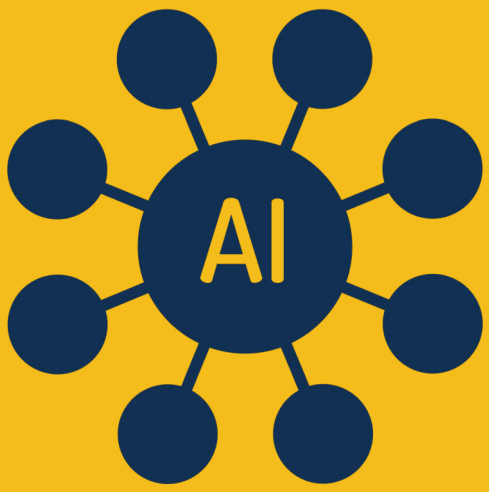
"Digital" is not merely a term denoting a specific technology; rather, it defines the capabilities of ICT in terms of the technological transformation of socio-economic processes and various other spheres. "Digital transformation" implies new solutions to problems, the creation of unique experiences and the improvement of business processes.

Technological solutions that have the capacity to change our lives already exist today! Which technologies? Across the world, 8 technological developments have been recognized, the first of which being artificial intelligence (henceforth – "AI").

In the modern world, the interest towards artificial intelligence is growing rapidly. International news media pay a staggering amount of attention to artificial intelligence, while leading politicians and scientists bring up AI and its role in the future of humanity with increasing frequency in their debates. The topic of artificial intelligence also dominated the studies and political and judicial agreements of various international players on the topics of the future of the public sector and its digital transformation. In the last five years, scientists, researchers, activitists and public policy makers are becoming increasingly interested in the potential of AI technology.

The following study aims to provide the readers with information about the main aspects of artificial intelligence, not only in its capacity as a technological solution to various problems, but also as a completely new methodology for the development of business and public administration processes, and a modern instrument of decision-making, governance, and supervision.

The study was developed within the framework of the project - "Promoting Transparency and Ethical Standards in the Use of Artificial Intelligence (AI)", implemented by the Institute for Development of Freedom of Information (IDFI) with the support of the International Center for Not-for-Profit Law (ICNL). This initiative aims to promote more transparency in the use of artificial intelligence, development of ethical standards, and its better implementation by public institutions.



ARTIFICIAL INTELLIGENCE

2.1 MEANING, MAIN CONCEPTS

Artificial intelligence is the achievement of digital transformation, the fourth industrial revolution. Artificial intelligence is a field in computer science – a special type of ICT that carries out intellectual actions that are generally only typical to humans. These include automatic decision-making processes/systems that imitiate, anticipate and/or replace human activities.

The European Council defines artificial intelligence as: "a set of sciences, theories and techniques whose purpose is to reproduce by a machine the cognitive abilities of a human being. Current developments aim, for instance, to be able to entrust a machine with complex tasks previously delegated to a human."[1]

According to the definition developed by the High Commission of Experts (AI HLEG), on the other hand: ""Artificial intelligence (AI) refers to systems that display intelligent behaviour by analysing their environment and taking actions –with some degree of autonomy – to achieve specific goals. AI-based systems can be purely software-based, acting in the virtual world (e.g. voice assistants, image analysis software, search engines, speech and face recognition systems) or AI can be embedded in hardware devices (e.g. advanced robots, autonomous cars, drones or Internet of Things applications)."[2]

The humanlike skills characteristic of artificial intelligence include: perception, argumentation, reasoning/evaluation, activity with the elimination of or small interaction with the human factor.

Contemporary scientific literature (Gutwirth and Friedwald) describes three principal directions within the framework of the categorization of information and communications technologies in terms of function: 1) internet-technologies; 2) interface-technologies, and 3) artificial intelligence technologies.

Three concepts are noted when difining artificial intelligence: narrow, general, and artificial superintelligence. Examples of **narrow artificial intelligence** are the most widespread nowadays – these are meant for singular tasks and for providing technological solutions to specific problems. At this time, the development and successful future of artificial intelligene without human effort and intervention is unthinkable. When it comes to **general artificial intelligence**, the capacities of human and technology are equilized. Unmanned flight systems are an example of such. **Artificial superintelligence**, meanwhile, is only a theoretical product, comprising the technologies where artificial intelligence exceeds human capabilities and becomes able to act/make decisions better and faster.

The questions of whether or not computer capacity has matches those of humans, whether it has exceeded the thinking of mankind, and to what extend can technology self-study and self-develop without the interference of human factors is still the subject of more future research and evaluation. Artificial intelligence should not become the rival of human intellence. Instead, ideally, mankind should be able to use artificial intelligence to tackle difficult, complicated issues with maximum efficiency and precision, in addition to the ability to fully delegate simple tasks.

It is also important to separate artificial intelligence and machine learning, as the latter is only a component of artificial intelligence, comprising a subcategory, and the two concepts should not in fact be treated interchangeably. In general, machine learning is always related to statistical work and the processing of more and more data. Such systems have the capability to improve results with the experience built on this data.

Deep learning is a type of artificial intelligence and comprises the unsystematic, unstructured processing of data by neiral networks.

Artificial intelligence is of course strongly interlinked with data science, as it is built on the processing of large datasets and uses data as its primary resource.

1 *https://www.coe.int/en/web/human-rights-rule-of-law/artificial-intelligence/glossary*

2 *https://ec.europa.eu/futurium/en/ai-alliance-consultation*

2.2 MAIN CHARACTERISTICS

Artificial intelligence allows us to find desired information promptly and without exerting human effort based on the collection, processing and analyzis of large quantities of digital information, as well as allowing us to make statistically most rational decisions based on objective criteria and analysis of similar data.

Artificial intelligence is defined by two main characteristics:

- **Autonomy**: carrying out tasks in complex environments with many different variables without the constance control/ supervision of the user.
- **Ability to adapt:** the capacity to (self-)learn and improve based on experience.

Analyze tremendous volumes of data to make much more accurate and complete predictions by humans and/or to support decision-making in an automated (e.g., algorithm-based) form, such as for crime prevention or to study economic and societal trends.

The following are the most commonly recognized examples of AI: remote control systems, unmanned flying vehicles, self-driving cars. The opportunities for the use of artificial intelligence for transportation purposes are growing in parallel to the continuing development of traffic signs and related infrastructure. The same mechanism is at work in the cases of self-propelled vessels, delivery robots, and drones. Artificial intelligence selects optimal routes, avoids obstacles encountered and makes difficult decisions with the consideration of a changing, dynamic, and complex environmental data.

While in print media every customer encounters the exact same information, the use of online media outlets, advertising offers, and search engines presents a contrasting environment where the system offers personalized information based on out interests and preferences as determined by specific algorithms. This is made possible for companies like Google, Facebook, Netflix or Instagram precisely thanks to the algorithms generated through the use of artificial intelligence.

Artificial intelligence is used for receiving specific directions and recommendations on individual topics with the consideration of existing knowledge and experience on these topics. This is especially the case in the field of healthcare.

Some other general examples of artificial intelligence include: voice recognition systems, tex and visual data decoding, face recognition, voice assistants, Apple's SIRI, Amazon's ALEXA. **Virtual assistants** found on smart mobile devices and computer systems. In Customer Relationships Management (CRM), the first-line Q&A mode **autoresponder**; digital assistants significantly simplify our everyday life, as they enable us to quickly search information; partially automated communication with service users through chatbots and virtual assistants, interpretation of textual data supported by Natural Language Processing (NLP), perception of audio and textual material.

An instance of artificial intelligence known to the Georgian public is the humanoid Sophia, who attended the global summit held in Tbilisi in 2018 within the framework of the Open Government Partnership (OGP). Its use is envisioned in the direction of service delivery.

2.3 ENABLING FRAMEWORKS

There are several prerequisites and conditions that need to be met before artificial intelligence can be implemented in a given country. First of all, it is important to have a developed ICT infrastructure base and a practice of the widespread use of digital governance, proper functioning of interoperable databases and information systems, unconditional assurance of data quality, integrity and accessibility, and Use of correct and documented business processes in both departmental and inter-organizational transactions. Databases should be digitized and accessible, the information/data contained in these databases should be complete, correct and updated.

Along with infrastructural and governance factors, the existence of qualified human resources and, in general, the

presence of a developed information society with proper digital skills is an equally important prerequisite for the initiation of artificial intelligence projects.

It is essential that countries establish an environment conducive to artificial intelligence in order to promote AI research and development and achieve a competitive advantage with artificial intelligence. Ethical standards, as well as legislative, strategic, and regulatory framework, and fiscal policy – all of these factors can be considered supporting mechanisms.

Of course, without regulatory norms, technological solutions oriented toward processing tremendous quantities of data can become a source of great risk. The topic of developing a proper regulatory framework for questions such as technology deciding whether a specific citizen will receive credit from a bank, or if an applicant's resume will be included in the list of selected candidates, or moreover, whether a person will be subject to release on bail instead of incarceration as a preventive measure, is a particularly complex issue. On one hand, regulations should take into account any subtle nuances when guaranteeing the balance of rights and responsibilities on all sides. On the other hand, legislation should not become an obstacle to the new industrial revolution or a barrier to technological advantages.

Without a proper strategic framework, the development of artificial intelligence may encounter a large number of significant challenges. The strategy for artificial intelligence should reflect the direction the country wishs to take in terms of its ICT, innovations, economic development, and the field of digital transformation. The strategy should be in agreement with the combination of long-term goals, priorities, framework, and infrastructural projects defined on the national level in compliance with the general public policy of the state. The strategy, as a national-level guide in the field of artificial intelligence, should be able to be executed with the involvement of all interested parties in the process.

The artificial intelligene realized with the use of existing conditions and the creation of supporting instruments and framework and the successful initiatives implemented due to its success will bring economic growth, social wellbeing, better national services, and effective public administration.

ARTIFICIAL INTELLIGENCE IN THE PUBLIC SECTOR

3.1 THE USE OF ARTIFICIAL INTELLIGENCE ACROSS THE WORLD, FIELDS, SERVICES, CASES

We can observe a growing trend of the use of artificial intelligence in the public sector. Multiple countries have already created special public institutions aimed at the development of this field. For instance, Great Britain and United Arab Emirates have developed a Office for Artificial Intelligence and a Ministry of Artificial Intelligence, respectively.

The use of artificial intelligence furthers the development of the process of providing public services and improves the quality of services (governmental services for government structures (G2G), governmental services for business (G2B), and governmental services for citizens (G2C). The technological solutions provided by artificial intelligence greatly improve the provision of governmental services and enable the growing engagement of citizens in the decision-making process, resulting in a more responsible, accountable, transparent, and effective public sector. [3]

Developed countries actively use artificial intelligence in the following directions:

- Improvement of governance processes and public sector activities through the use of algorithms; modernization of public policy;
- Automatization of the business processes of public institutions, updating document management and case management systems;
- Stimulating innovative projects in the public sector;
- Provision of public electronic services and the improvement of citizen engagement, including the provision of proactive and personalized services.

Artificial intelligence is most commonly used to fulfill the following public tasks:

- **Enforcement**: enforcement, monitoring, and inspection of legislative and regulatory requirements.
- **Research, analysis, and monitoring of public policy**: support evidence-based decision-making process. Analyses/ argumentation (Case base Reasoning) based on existing cases, making new decisions based on existing data and examples of old cases.
- **Public services and participation**: improvement of public services aimed at citizens and business subjects, service personalization, perceiving the importance of information based on logical-deductive methods and simplifying the search for such information; identifying and offering public services personalized for a given citizen based on their preferences and requests.
- **Subsidies, granting social benefits**: Identify specific target groups for social assistance.
- **Internal document workflow systems**: Support for organizational management systems (personnel, procurement, ICT systems).

Processing of audio materials - hotline systems, including emergency medical assistance or other emergency situation management systems (911, 112), are in the process of actively implementing artificial intelligence with the aim of identifying and processing audio data, which increases the likelihood of saving the life of the caller through constructing appropriate plans for critical situations and providing immediate aid. For example, in Denmark, the implementation of artificial intelligence in emergency medical services increased the identification rate of cardiac arrest cases among patients from 73% to 93% during emergency calls. [4] In Spain, false police reports are identified precisely through the use of artificial intelligence processing audio data.

Opportunities for the identification/recognition of visual data and individuals – artificial intelligence enables the

3 S. Ghosh, "Application of Natural Language Processing (NLP) Techniques in E-Governance", in E-Government Development and Diffusion: Inhibitors and Facilitators of Digital Democracy. PA: Hershey, 2009, pp.122-132.

4 *https://sifted.eu/articles/ai-that-eavesdrops-on-a-and-e/*

creation of face recognition, visual data processing, and identification systems. For instance, in Estonia, the Government issues certain subsidies to citizens in the regions for the cultivation of agrarian lands, but recognizing how these funds were spent by agrarian workers was oftentimes either beyond control or would incur disproportionate administrative costs with the necessity to sent inspectors to the locations and conduct monitoring activity. Currently, the Estonian government controls agrarian activity on agricultural lands through the use of satellite data/images[5].

Management of knowledge rooted in artificial intelligence - search for related cases, information, and decisions pertaining to a specific case for expert analysis. As an example, the government of Slovakia uses artificial intelligence to collect semantic data. Meanwhile, Portugal uses artificial intelligence in seeking alternative methods for the resolution of judicial proceedings and litigation.

Cognitive management, process automatization, and interconnected and automated machines - automation of processes and the joint use of data-based artificial intelligence systems enables the development of a number of service areas and the optimization of back-offices of organizations. For instance, the hiring process in a public agency of a Swedish municipality is based on an interview with an AI robot[6].

Threat study, analytics, intelligence - instances of using artificial intelligence to identify vulnerabilities, weak points, and defense solutions in networks are on the rise around the world. As an example, the Norwegian Security Service uses machine learning to identify malicious code to ensure cybersecurity.

3.2 THE ROLE OF THE PUBLIC SECTOR IN THE PROCESS OF DEVELOPMENT OF ARTIFICIAL INTELLIGENCE

The public sector plays a dual role in the direction of the development of artificial intelligence, and thereby faces two kinds of challenges. Firstly, the public sector must support the process of creating an ecosystem suitable for AI for the formation of national startups and industry, attracting investors and donors, as well as promoting the use of AI applications by various sectors and the pursuit of greater socio-economic growth and prosperity through artificial intelligence. Simultanously, for the development of artificial intelligence, the government is responsible for creating a regulatory-legislative framework that will balance the risks, threats, challenges, and legislative and ethical questions related to artificial intelligence and for establishing a mechanism for its effective enforcement.

THE FIRST – THE FACILITATOR OF THE INTRODUCTION OF ARTIFICIAL INTELLIGENCE, IMPLIES:

- Formation/development of a enabling frameworks
- Establishment of a governance model
- Access to information
- Communication with interested parties and promoting awareness

THE SECOND - THE ROLE OF CREATING, IMPLEMENTING, USING ARTIFICIAL INTELLIGENCE, WHICH IMPLIES:

- Initiating the widespread use/introduction in public services
- Initiating innovative projects
- Conducting research and analysis

5 *https://forbes.kz/process/technologies/how_estonia_uses_artificial_intelligence_in_the_healthcare_legal_industry_and_agriculture/*

6 *https://www.smartcitiesworld.net/news/news/swedish-municipality-deploys-robots-for-safer-recruitment-5251*

3.3 THE INFLUENCE OF ARTIFICIAL INTELLIGENCE ON THE PROCESS OF DEVELOPMENT OF THE PUBLIC SECTOR

In the modern worls, as both developed countries with strong economies and those that are still developing and possess relatively scarce resources recognize the potential of contemporary data-based technologies, these same countries actively use artificial intelligence to automate public tasks, make better decisions, and improve the experience of the citizens' engagement with public institutions. Artificial intelligence provides a number of important opportunities to the public sector; among these are:

- Effectiveness: doing more and increasingly important work: new opportunities, increased opportunities, concentration of human resources from bureaucracy to highly intellectual work, reduction of the risk of human error.
- Optimization: the following are done better: reduction in costs, maximized savings of human and time resources, speed, quality, flexibility, exercise of public authority at a qualitatively high level, quickly and cheaply. Repetitive and time-consuming tasks that can be monotonous and tiring for humans can be instead delegated to technology.
- Increased Value: updated business models: participation, creativity, innovation:

In order to introduce AI to the public sector and to create new opportunities through it, the main task of government structures is to update the relevant business processes, involve qualified staff, review approaches among the stakeholders, and update the data and technical infrastructure needed to implement AI[7]. Strategy, human resources, processes, data, technologies and infrastructure, legislative and ethical framework – all these constitute the fields that the public sector should be working to develop better in the process of inroducting artificial intelligence.

7 *https://www2.deloitte.com/us/en/insights/industry/public-sector/ai-readiness-in-government.html*

INTERNATIONAL REGULATION OF ARTIFICIAL INTELLIGENCE

4.1 INTERNATIONAL INSTRUMENTS AND ENDEAVOURS

The growing consumption of computer products created by private entities, in the absence of a regulatory framework, naturally poses many risks and dangers. Public awareness of the colossal volume of personal data provided to machine learning technologies as behavioral data is virtually absent or extremely low.

In the course of conducting online transactions, users frequently reveal their personal information (health, family, or political and societal data), oftentimes without a direct and obvious notification that this might be the case. The majority of users are also not aware that their personal data may be observed and monitored, and that the data may then be sold. The data revealed by us knowingly or unknowingly subsequently enables computer systems to anticipate and estimate individual preferences, a process that is far more precise when done by technology rather than in human hands. These goods that can be predicten are sold for solid sums in the "new type" marketspace, giving rise to the "surveillance capitalism" of the 21st century. This practice, on the one hand, presents significant challenges to the control of individuals over their own personal data and the guarantees of the right to privacy, while simultaneously giving rise to the risk of the monopolization of data by specific multinational companies.

Moreover, behavioral data and "predictive" technologies are also used to influence the personal preferences and subsconscious desires of individuals through the control of the informational resources that are or can be provided to said individual. Persons might essentially become the subjects of behavioral experiments. Algorithms with a narrow purpose that have the capacity for subsconscious influence or persuasion have an impact on the formation of an individual's thoughts and their independent decision-making capabilities.

All of this effectively forms an environment that promotes manipulating individuals in an easy, effective and undetectable manner. Artificial intelligence that is used with wrongfully and with malicious intent will harm all aspects of public life and completely change daily lives. Technological power makes other democratic mechanisms of legal control and oversight by the state all the more relevant. An example that can be used here is the case of Cambridge Analytica and the public scandal related to ethical questions. This is the case of one of the most high-profile incidents in the history of social media, when a political data analysis company involved in the 2016 Donal Trump election campaign – Cambridge Analytica – acquired and processed the personal data of 50 million Facebook users (American electorate) without their explicit consent. This colossal volume of data was subsequently used by the company to execute their personalized political advertising campaign. [8]

Apart from the abovementioned hazards, one should also pay attention to the methods used for processing personal data. For instance, what criteria are used to categorize individuals into various groups. Should the method used in this kind of process not be constructed according to law, it creates a risk of segregation and discimnation when users of certain profiles are given an advantage over others without basis. This has the potential to fundamentally alter the life and social environment of an individual. Algorithms created based on incorrect preferences can set a precedent for making unfair and unlawful decisions in healthcare and social subsidies, employment, or litigation. Another risk arises when it comes to the freedom of speech and expression. The openness of the filtering methodology used by internet mediators in their algorithms is also very noteworthy, as well as the protection of the rights to privacy, freedom of assembly and the right to move freely when using facial recognition technologies.

During the introduction of algorithms rooted in artificial intelligence, special care should be taken to clearly identify the influence this technology can wield and the outcomes its use can lead to for individuals as well as various fields of every life and the democratic development of different countries on the global scale. The positive impact of artificial intelligence as well as the risks and hazards associated with its implementation have become the subject of research by a multitude of international organizations. Currently, no single legislative framework, set of regulatory standards, or internationally recognized agenda that has been a subject of universal consensus exists. Accordingly, this study will present the important decisions, legal frameworks and strategic approaches on artificial intelligence developed by various international actors.

EUROPEAN UNION

The world community and its global leaders recognize that a country developed with artificial intelligence has a fairly

8 *https://www.cnbc.com/2018/03/21/facebook-cambridge-analytica-scandal-everything-you-need-to-know.html*

strong competitive advantage in the international arena. The European Union declares that, in terms of the use of human-centered artificial intelligence in implementing human rights and freedoms, the rule of law and the principles of democracy, it currently has no valid competitors.

The use of artificial intelligence is subject to the requirements of the international legal regime, and in this regard, the EU, among other international stakeholders, plays a leading role. Various EU institutions, including the European Commission, have unequivocally supported the use of artificial intelligence, promoted and funded its large-scale introduction. "All the key ingredients that make it a leader in the artificial intelligence revolution, upholding its own approaches and values, are present in the EU" - with this 2018 announcement, the European Commission began developing an artificial intelligence strategy and implementing large-scale knowledge, development, and research projects.

In analyzing the legal framework for the regulation of artificial intelligence, the EU relies primarily on the existing legal environment. There is no doubt that the Data Protection Framework Regulation (GDPR) applies to the scope of operations of artificial intelligence, and this regulation has a truly great potential for ensuring a high level of transparency in the field of artificial intelligence.

The EU, on the one hand, seeks to attract talent and investment in the field of artificial intelligence, and, on the other hand, attaches no less importance to the protection of human rights and freedoms and to the preservation of European values in the process of using technological innovations. Executing a balanced policy in this direction is a significant challenge for the EU.

The seven basic principles of GDPR regulation (on the processing of personal data) from the process of creating artificial intelligence are important for consideration, namely: Transparency, fairness and legality; Purpose limitation of processing; Data minimization; Accuracy; Accountability; Integrity and confidentiality; Storage limitation.

There are three basic rights guarantees when applying the GDPR:

- Right of Access - Each individual has the right to gain access to the personal data about him/her that is subject to processing;
- Right to be Forgotten - The subject of the data has the right to request the deletion of the processed data, and the processor is obligated to ensure the implementation of this right, unless he has an obligation to further process this data;
- Right to Clarification - The data processor should answer the questions regarding the data processing that may arise.

Despite the consensus on the use of the GDPR, many issues remain unclear, e.g.: the entire process of artificial intelligence is built from the ground up on assumptions and probabilities, while the GDPR requirement for the accuracy of data is unambiguous; tt is also difficult to fully comply with the requirements for transparency and accountability due to the limited ability of technical explanation for some artificial intelligence processes (e.g., deep learning, neural networks); practical implementation of the principles of minimization, goal limitation, and anonymization under the conditions of the large scale of artificial intelligence applications and the technological environment of massive data analysis. Additionally, processing data with fully automated processes (in the case of AI), which has a significant impact on humans, if the main purpose of artificial intelligence is to scan huge streams of data to detect trends, find patterns, then a problem arises in the conflict with the key principles of GDPR (e.g., data collection / processing only for a specifically identified purpose).

There are a number of guidance documents on artificial intelligence adopted/published within the EU that offer member states good practicse in the process of the introduction of artificial intelligence. [9]

9 European Commission, "White Paper on Artificial Intelligence – A European Approach to excellence and trust", 2020 *(https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf)* ; 2) European Commission, „Report on the safety and liability implications of Artificial Intelligence, the Internet of Things and robotics" , 2020 *(https://eur-lex.europa.eu/legal-content/en/TXT/?qid=1593079180383&uri=CELEX%3A52020DC0064)*; 3) Council of the European Union, „Governance of the Artificial Intelligence Strategy for Europe" *(https://data.consilium.europa.eu/doc/document/ST-11352-2018-INIT/en/pdf);* 4) European Economic and Social Committee, „Opinion about AI" *(https://www.eesc.europa.eu/en/our-work/opinions-information-reports/opinions/artificial-intelligence-consequences-artificial-intelligence-digital-single-market-production-consumption-employment-and)*

The European Commission considers it important to update the European ePrivacy Regulation and the Cybersecurity Act with standards governing artificial intelligence, as well as to continue research in this area and increase funding. The relevant research on the application of GDPR requirements to artificial intelligence is continued by Article 29 Working Party and the EU Data Security Council, which issues guidelines and recommendations for AI developers and users. Some of the notable guidelines include:

- Informing subjects of the data that their data is being used for profiling;
- Maximize information regarding the importance, goals, outcomes, and potential impact of data processing;
- Reasoned explanation of the decision made to the subject of the data and informing them about the possibility of appealing said decision.

The guidelines also explicitly stipulate that data processing should be kept to a minimum, data storage time should be limited in accordance with the principle of necessity, the data processor/authorized person should bear the responsibility for the data processing process, and the developer of the artificial intelligence application should answer the question: how necessary it is to process given data for the functioning of the system.

THE COUNCIL OF EUROPE

Risk - Technology should not have a negative impact on the core European values that form the bedrock for the unification of European countries in the form of the European Union, and which the Council of Europe is designed to uphold.

The Council of Europe is an important international player when it comes to protecting the requirements and guarantees of human rights, democracy, and the rule of law in the technological space. For example, the Council of Europe has developed an Internet Governance Strategy 2016-2019[10], as well as its updated project - "Digital Governance Strategy 2020-2023".

The Council of Europe Committee on Artificial Intelligence (CAHAI) is an ad-hoc committee[11] whose main purpose is to study and analyze the legal framework, regulatory norms, standards, and ethical principles related to artificial intelligence. Naturally, under the umbrella of the Council of Europe, this institution places particular emphasis on the issues of the protection of human rights standards, democracy and the rule of law at various stages in the creation, development and application of artificial intelligence. The Committee was established in 2019 and brings together representatives from 47 countries, including Georgia (Office of the State Inspector) as well as observer countries (e.g., Canada) and international organizations interested in the field - the United Nations, the European Union, and the OSCE. CAHAI's work environment is quite large and inclusive, with many stakeholders actively involved, including private companies, academic and research sectors, and NGOs.

CAHAI is an important authority of the Council of Europe in the field of establishing the legal norms and ethical standards around artificial intelligence. It consists of three working groups operating in specific directions:

- Artificial Intelligence Policy Development Group
- Artificial Intelligence Legal Research Group;
- Communications and Consulting Group.

CAHAI is currently in active consultation and operations mode and is obligated to submit its final research reports by 2021. The feasibility study of the CAHAI Working Group will cover international legal instruments governing artificial intelligence, ethical guidelines, a review of national instruments, policies and strategies on artificial intelligence, and the advantages, disadvantages, and limitations of existing international and national instruments.

10 *https://edoc.coe.int/fr/internet/7128-internet-governance-council-of-europe-strategy-2016-2019.html*

11 *https://www.coe.int/en/web/artificial-intelligence/cahai*

Significant legal tools, standards, reports, and initiatives have been developed in recent years within the Council of Europe that partly address/influence the development of artificial intelligence[12]. Relevant Council of Europe research and reports on current trends in artificial intelligence[13] are also informative and instructive.

ORGANIZATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT (OECD) ARTIFICIAL INTELLIGENCE WORKING GROUP

The OECD defines artificial intelligence as "a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy"[14].

In 2019, 42 countries signed the OECD Artificial Intelligence Principles, which aim to agree on the safe, equitable, and reliable use/development of artificial intelligence. The OECD's principles of artificial intelligence are[15]: 1) Inclusive growth, sustainable development and prosperity; 2) Human-centered values and justice; 3) Transparency and clarity; 4) Strength, safety and reliability; 5) Accountability.

The OECD recommends the signatory states[16]: 1) invest in research and development of AI; 2) development of a digital ecosystem conducive to AI; 3) formation of regulatory framework conducive to AI; 4) development of human capital and transformation of the labor market; 5) international cooperation for trustworthy AI future.

UNITED NATIONS (UN)

More than 35 special institutions or bodies within the UN system use artificial intelligence to deal with various important challenges in the world, including humanitarian crises or climate change. The United Nations Specialized Organization - International Telecommunication Union holds an annual global summit - "AI for Good Global Summit", which is a platform for global and inclusive dialogue with AI innovators and other stakeholders and aims to develop a common vision for the development of the capabilities of artificial intelligence and its use.

---

12 1)"Unboxing Artificial Intelligence: 10 steps to protect Human Rights" Recommendations of the Council of Europe Commissioner for Human Rights, 2019 *(https://rm.coe.int/unboxing-artificial-intelligence-10-steps-to-protect-human-rights-reco/1680946e64);* 2) Declaration by the Committee of Ministers of the Council of Europe "On the manipulative capabilities of algorithmic processes", 2019 *(https://search.coe.int/cm/pages/result_details.aspx?objectid=090000168092dd4b)*; 3) Council of Europe Directorate-General for Human Rights and Rule of Law "Guidelines on Artificial Intelligence and Data Protection", 2019 *(https://rm.coe.int/guidelines-on-artificial-intelligence-and-data-protection/168091f9d8)*; 4) The European Commission for the Efficiency of Justice (CEPEJ) „European Ethical Charter on the Use of Artificial Intelligence in Judicial Systems and their environment" , 2018 5) Recommendation of the Parliamentary Assembly of the Council of Europe „Technological convergence, artificial intelligence and human rights", 2017 *https://pace.coe.int/pdf/eb957373c02acc48df23359a09d52367e46172953326667a8259ffe25682ae848428feba12/recommendation%202102.pdf*

13 1) Council of Europe study „A study of the implications of advanced digital technologies (including AI systems) for the concept of responsibility within a human rights framework", 2019 *(https://pace.coe.int/pdf/eb957373c02acc48df23359a09d52367e46172953326667a8259ffe25682ae848428feba12/recommendation%202102.pdf);* 2) „Governing the Game Changer – Impacts of artificial intelligence development on human rights, democracy and the rule of law" , Conclusions of the Council of Europe Conference, Helsinki, 2019 *(https://rm.coe.int/conclusions-from-the-conference/168093368c)*; 3) Directorate-General for Human Rights and the Rule of Law of the Council of Europe „Artificial Intelligence and Data Protection: Challenges and Possible Remedies", 2018 *(https://rm.coe.int/artificial-intelligence-and-data-protection-challenges-and-possible-re/168091f8a6)*; 4) Academic Survey by the Council of Europe Anti-Discrimination Department on "Discrimination, Artificial Intelligence and Algorithmic Decision Making", 2018 *(https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73);* 5) Council of Europe study "STUDY ON THE HUMAN RIGHTS DIMENSIONS OF AUTOMATED DATA PROCESSING TECHNIQUES (IN PARTICULAR ALGORITHMS) AND POSSIBLE REGULATORY IMPLICATIONS", 2018 *(https://rm.coe.int/algorithms-and-human-rights-study-on-the-human-rights-dimension-of-aut/1680796d10)*

14 *https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449*

15 OECD Artificial Intelligence Principles *(https://www.oecd.org/going-digital/ai/principles/)*

16 *https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449*

UN Special Rapporteurs, through their research and reporting on artificial intelligence and human rights[17], provide important recommendations and guidance to countries that want to implement artificial intelligence in adherence with the high standards of human rights.

The UN Special Rapporteur on the Privacy of Privacy, Joseph Cannataci, and the working groups under his leadership are developing "Guidelines for the Protection of Personal Data in Artificial Intelligence Solutions" in accordance with the requirements of the UN Universal Declaration of Human Rights and the International Covenant on Civil and Political Rights. At this stage, a draft research paper has been submitted for feedback from UN member states.

Article 7 (prohibition of discrimination) and Article 12 (inviolability of privacy) of the UN Universal Declaration of Human Rights, Articles 2, 3 and 17 of the International Covenant on Civil and Political Rights are binding for all signatory member states. The guidelines indicate the need for the transparency of algorithms, as well as the transparency of information on decision-making procedures based on them, weighted scores of the criteria used therein, and other steps. The decisions to be made by artificial intelligence and possible outcomes must be foreseeable, and it is necessary to identify all cases where the decision/observed outcome reached by artificial intelligence is not explicable, predictable, and does not fit into established patterns. Decisions made by artificial intelligence should be monitored by humans, not "machines". The volume and quality used in artificial intelligence should carry special importance. Therefore, it is critical to have publicly available information about the databases used and the basis for their processing. The development of artificial intelligence must take place in an environment controlled by ethical, moral and legal mechanisms.

The UN Special Rapporteur on the Inviolability of Private Life calls for adherence to seven key principles in the use of artificial intelligence:

1. **Jurisdiction.** It is important that in the process of creating and using artificial intelligence, the legal framework applied to it be clearly defined, thus achieving foresight of legal consequences and sustainability of civil turnover;

2. **Legislative Bases and Limitation of Aims.** A decision rooted in artificial intelligence necessarily needs a solid legal basis for the processing of personal data;

3. **Accountability** - Artificial intelligence solutions should be subject to accountability and responsibility mechanisms, and it should be possible to evaluate, control, and audit algorithms and the process of creating and using them.

4. **Control** - AI solutions must be under the complete control of the data processor. From the initial design of the system to its final implementation and shutdown, oversight should be exercised over the data processing process – namely, what data is processed by the AI, what parameters and data measurements are the basis for its decision-making, and how they should be weighted. The results should be subject to constant monitoring, issues of possible bias and discrimination should be checked and corrected both before the system is launched and at regular intervals in all phases of its operation.

5. **Transparency and Explainability** - Stakeholders and users of the artificial intelligence system will not be able to build trust unless the system's capabilities, tasks performed, and causal relationships between these and observed results are clearly and unequivocally demonstrated. The decision of artificial intelligence must be unambiguously perceptible and be able to be predicted in terms of processes and outcomes.

6. **Rights of the subject of the data**, which involves acquiring the consent of the subject, which may not lead to negative consequences for him. The subject of the data should have the opportunity to request a decision from an authorized person if there is doubt that the decision by the AI is erroneous or based on incorrect data.

7. **Guarantees.** The introduction of control mechanisms in the organizational context is a necessary prerequisite for assessing and dealing with risks and threats. ISO 27000 (Information Security Management System) 27701 (in the context of personal data protection) is one of the proposals for data protection, risk identification, and response measures.

17 Report of the UN Special Rapporteur on freedom of thought and expression „Implications of artificial intelligence technologies for human rights in the information environment" *(https://www.undocs.org/A/73/348)*, 2018 2) United Nations – Secretary-General's High-level Panel on Digital Cooperation *(https://www.un.org/en/digital-cooperation-panel/)* 3) United Nations – Special Rapporteur on the promotion and protection of the right to freedom of opinion and expression *(https://undocs.org/A/73/348)*

UNESCO

According to the United Nations Educational, Scientific and Cultural Organization (UNESCO), the world must positively utilize the significant potential of artificial intelligence to achieve the goals of sustainable development, the development of the knowledge society, and socio-economic progress[18]. The principal direction taken by UNESCO focuses on the development of human-centered artificial intelligence, in particular the organization explores the potential uses of artificial intelligence for the advancement of humanity, better protection of human rights, promoting knowledge and experience, and enhancement of inclusive digital media literacy.

G20

During the G20 summit of 2019 in Japan, the Trade and Digital Economy Ministers agreed on a commitment to develop human-centered artificial intelligence based on the OECD artificial intelligence recommendations. The ministers recognize that the free and secure exchange of data between countries is essential for the development of the digital economy, and in this process the observance of the norms of international law, as well as respect for data security, intellectual property, and privacy are absolutely essential. The G20 countries consider it important to continue discussions around artificial intelligence in relation to issues such as inclusive growth, sustainable development and prosperity, the reflection of human values, fairness, transparency, security and safety, and accountability in artificial intelligence systems. [19]

GLOBAL INITIATIVES OF NON-GOVERNMENTAL AND RESEARCH ORGANIZATIONS

There are a multitude of private initiatives and platforms centered around the topic of artificial intelligence across the globe. Some of the notable onec include the European Expertise and Expert Institute (EEEI), Institute of Electrical and Electronics Engineers (IEEE), AI Transparency Institute, AccessNow, The Future Society, Homo Digitals, Article 19, Algorithm Watch

The Institute of Electrical and Electronics Engineers (IEEE), the world's largest technology organization, has launched a global initiative tackling the ethical standards for autonomous and intelligent systems. The initiative aims to protect and guarantee human rights and wellbeing in the course of the exploitation of new technologies, including artificial intelligence, in accordance with the global standards for human rights, including those set out by the European Convention on Human Rights (ECHR) and the 108+ Modernized Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data. The IEEE has also developed a program for the ethics certification of autonomous and intelligent systems and a standardization guide.

As a human rights organization, **Human Rights Watch** investigates human rights abuses in the use of artificial intelligence in any corner of the world, studies the use of lethal autonomous weapons, the introduction of face recognition tools for minority monitoring and profiling, and the cases of misallocation and reduction of pension savings.

The Toronto Declaration on the Right to Equality and Non-Discrimination in Learning Systems was initiated in 2018 by Amnesty International, Human Rights Watch, the Wikimedia Foundation, and other key players.

18 "ON A POSSIBLE STANDARD-SETTING INSTRUMENT ON THE ETHICS OF ARTIFICAL INTELLIGENCE", 2019 *https://unesdoc.unesco.org/ark:/48223/pf0000369455*

19 Ministerial Statement on Trade and Digital Economy – Human Centered Artificial Intelligence *https://www.mofa.go.jp/files/000486596.pdf*

4.2. INTERNATIONAL EXPERIENCE AND EXAMPLES

Assessing the readiness of countries in terms of the level of development of artificial intelligence is important if one wishes to reveal international leaders, global trends, and regional competitiveness in this field. In this regard, the activities of the UK-based international consulting group Oxfod Insight is of particular interest. On the one hand, Oxfod Insight advises different countries on how to develop artificial intelligence-based solutions, and, on the other hand, it processes data pertaining to artificial intelligence in 172 countries based on corresponding methodology and criteria and subsequently published an artificial intelligence readiness index of the countries.

The Artificial Intelligence Readiness Index consists of three components and ten corresponding directions:

- Governance component: goals, governance framework and ethical norms, digital skillsets and ability to adapt;
- Technological sector: human capital, capability to innovate, scale of the technological sector
- Data and infrastructure: infrastructure, data accessibility, data quality

According to the Oxford Insight AI Readiness Index 2020, the top ten in terms of AI readiness among the 172 countries looks like this: USA, UK, Finland, Germany, France, Singapore, South Korea, Denmark, the Netherlands, and Norway.

However, it should be noted that the results of a survey conducted by the same organization in terms of "Responsible AI" show that most of the rules and principles, legal and ethical requirements are developed by EU member states, and the top five in this regard are: Estonia, Norway, Finland. Sweden, Denmark. The US and Britain lag behind Scandinavian-Baltic countries in terms of responsible use of artificial intelligence, while Russia, China and India have the lowest scores in this regard.

The United States became the first country in 2016 to adopt an AI national strategy. In the footsteps of the United States, strategic frameworks for AI were soon developed by China, the United Kingdom, South Korea, France, and Finland. Overall, the US, with its Silicon Valley high-tech hub and private sector innovation, has an undisputed advantage, although the EU stands out due to the high level of development of its strategic and legal-ethical frameworks.

THE U.S.: The largest-scale, most extensive, and influential AI research to date has been conducted in the US, with 3,000 students studying and researching artificial intelligence every year, up to 1,400 startups, and 7 of the 10 largest AI companies in the world working in this field. All of this, with years of collaborative experience accumulated by the aforementioned research centers, universities, companies, and public structures, uniquely establishes the United States as a global leader in the field of AI. The US has established a joint artificial intelligence department with the Department of Defense that aims to develop US defense capabilities and implement AI in military solutions. The US is also actively working on the commercialization of artificial intelligence and is similarly searching for experts in this field in order to import intellectual resources into the country.

CHINA: China presents serious competition to the US for world leadership in artificial intelligence. China recognizes artificial intelligence as a strategic advantage in the knowledge economy and plans to gain world dominance in this field by 2030 with its strategic plan. Its strategy, adopted in 2017, is rooted in three main issues: a) to catch up with the leading countries by 2020 and create an AI economy worth €19 billion and related areas worth € 126 billion. b) To become a leader in the world by 2025 with a worth of 51 billion euros for AI and 635 billion for its related fields. c) To become the sole superpower in 2030 with a €130 billion AI economy and related fields bing worth €1.2 trillion. China is the only country that measures the development of artificial intelligence on the basis of macroeconomic indicators.

SOUTH KOREA: South Korea is focused on finding and training talent in the field of artificial intelligence, strengthening the research component, and collaborative on projects between the science and economics sectors.

MEMBERS OF THE EUROPEAN UNION AND GREAT BRITAN:

The most strategic plans aimed at the growth, regulation and management of artificial intelligence have been developed among the EU member states, as almost all EU member states have elaborated a strategic framework and action plan for the development of artificial intelligence. Bulgaria, Croatia, Greece, Romania, Slovenia all plan to approve relevant

strategies by the end of 2020. The majority of member states approved their strategies in 2019-2020[20].

FRANCE'S AI Strategy: The main emphasis is placed on research and the creation of target groups that will study the issues under consideration and prepare corresponding recommendations. The title of the 2018 strategy is "Artificial Intelligence for Humanity", and the strategy highlights several important issues: 1) Development of an open data policy: Giving "second life" to unused data in the public sector and making it public in order to promote public good. 2) Creating regulatory frameworks and an environment conducive to innovative AI products (e.g., self-driving vehicles); 3) Publicity of all algorithms used in the public sector, ensuring the maximum level of transparency principle for AI in general. To this purpose, France creates expert panels and observer groups. Algorithms should be subject to audit that will evaluate them in terms of predictability and outcomes. 4) France names healthcare, transport, environment, security, and defense as the fields where artificial intelligence should be implemented most actively. 5) As an important principle, France unequivocally states that the organizations using such systems are fully liable and accountable for the decisions made by artificial intelligence. 6) Prior to launching the AI, it is necessary to conduct impact assessment research on discrimination, bias, and exclusion.

GERMANY: Germany has established a fairly comprehensive legal framework on responsibilities, rights, and accountability for the use of artificial intelligence (especially in the case of self-driving cars). For instance, discrimination against citizens during an accident is prohibited. Germany has an ethics committee for self-operating vehicles, which issues relevant guidelines. Additionally, a national commission working on the study of the social consequences of the decision-making by algorithms has been established.

GREAT BRITAIN: The British House of Lords and the Office of the Information Commissioner are actively working on the issues surrounding artificial intelligence. The House of Lords puts an emphasis on the perceptibility of artificial intelligence algorithms, while the Data Protection Commissioner conducts research and makes recommendations on regulatory and legal frameworks[21]. Britain intends to invest heavily in artificial intelligence in higher education. Several structures have been set up in the UK for artificial intelligence, including the Ministerial Working Group, the Artificial Intelligence Council, the Center for Data Ethics and Innovation, the Government Digital Services and Artificial Intelligence Office[22].

FINLAND: Finland has similarly set up a group of high-level artificial intelligence experts to study the circumstances for the development of an innovative artificial intelligence market within the country. The Nordic-Baltic group believes that excessive regulation should be avoided and all efforts should be made to standardize artificial intelligence. Infrastructure, software, hardware, and data used to create artificial intelligence should be based on the standards of interoperability, privacy, security, reliability, conscientious use, and portability.

DENMARK: Denmark views artificial intelligence in the context of the country's "digital growth and development". Promoting data openness, easing regulations and creating a flexible legal environment, strengthening cybersecurity – these constitute the country's approach to the process of development of artificial intelligence.

NORWAY: Introducing artificial intelligence ethically, responsibly, and with a high level of credibility is the cornerstone of Norway's artificial intelligence strategy. The strategy notes that: Artificial intelligence creates a competitive advantage; Introduction of ethical AI from theory to practice taking into account sectoral standards and requirements; The Norwegian Data Protection Inspector has the authority to inspect and audit algorithms; New talent equipped with the necessary skills for artificial intelligence are provided with the appropriate education; Norway will seek to concentrate Scandinavian capabilities in NLP (Natural Language Processing) development projects.

SWEDEN: Sweden's strategy on artificial intelligence focuses on four main areas: education and training; research; innovation and application; frameworks and infrastructure.

ITALY: The main public policy document on artificial intelligence, "Artificial Intelligence in the Service of the People," promotes the use of artificial intelligence in the public sector. It is important to ensure that the public procurement process as it related to AI is transparent, with a legal framework governing liability for damages caused by AI robots. Currently, Italy uses the EU-funded machine-learning data counter-terrorism tool - DANTE (Detecting and Analyzing

20 *https://ec.europa.eu/knowledge4policy/ai-watch/national-strategies-artificial-intelligence_en*

21 *https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf*

22 *https://www.gov.uk/government/collections/a-guide-to-using-artificial-intelligence-in-the-public-sector*

Terrorists). This system raises a number of questions in regard to potential profiling and lack of compliance with procedural norms by law enforcement agencies.

SPAIN: The Spanish government is investigating the economic impact of the use of artificial intelligence, the possible social impact on the labor market and the empowerment of people through this technology.

ETHICAL STANDARDS IN THE USE OF ARTIFICIAL INTELLIGENCE

As it is known, ethical principles and norms are perceived as a "soft" legal mechanism that is not binding, but its observance promotes good governance practices and the formation of a strong corporate culture, signalling that the organization takes highly reputable and responsible approach in its activities.

Over the past decade, a number of papers have been published on the ethical use of artificial intelligence by research organizations, from among the public and private sectors, within the academia, and by various companies. Despite their diversity, all stakeholders develop one specific approach in addressing the central topic - only by using artificial intelligence "ethically" is it possible to attain public welfare and socio-economic progress.

It is also interesting to outline the intersection of mandatory and voluntary requirements at each stage of the realization of artificial intelligence, namely the correlation of ethical principles, technical standards, best practices, and other similar non-binding governance instruments with the fundamental principles of the rule of law, democracy, and human rights.

Artificial intelligence is not created in a legal-ethical vacuum. On the contrary, the technological decision is accompanied by the process of forming a corresponding governance model and self-regulatory framework. The soft regulatory power of artificial intelligence and self-regulation in the form of ethical principles is most actively developed and used among the European and North American states, especially by the member states of the Council of Europe. Numerous initiatives for the development of artificial intelligence have been implemented in the United States, the United Kingdom, Germany, Finland, Belgium, Japan, the Netherlands, Singapore, Spain, China, France, India, South Korea, and Norway.

One of the most important initiatives within the European Union is the Ethical Guidelines for Trustworthy Atificial Intelligence developed by the EU High Level Expert Group on Artificial Intelligence.[23] The group worked to define three main characteristics, the combination of which forms an example of reliable artificial intelligence. The following are the three requirements presented therein: (1) artificial intelligence must be compliant with the law; (2) it must be ethical and respond to ethical and value principles; (3) it must create a strong and resilient environment, taking into account both technological and social contexts. Artificial intelligence should not cause harm or create the threat of such harm. The Ethical Guidelines on Reliable Artificial Intelligence set out the following basic principles for artificial intelligence: (1) human oversight of technology; (2) technological strength and security; (3) the inviolability of the right to privacy and data management; (4) transparency; (5) diversity, non-discrimination, fairness; (6) social and environmental well-being; (7) accountability.

The ethical guidelines developed by public and non-governmental organizations, as well as private companies and academia, should certainly not be considered as mechanisms that can effectively replace legislative requirements, although their positive contribution is not insignificant. Ethical norms can have a practical impact on the decision-making processes on artificial intelligence and contribute to the development of AI systems for social good in observance of ethical values and legal norms.

The most important ethical principles regarding artificial intelligence that are commonly recognized include transparency, lawfulness, responsibility, inviolability of privacy, and protection from harm - these five ethical principles are named and most widely used in almost all documents developed so far, although based on the analysis[24] of the documents from the European Union, the Council of Europe and other countries, the following 11 ethical principles can be distinguished:

1. **TRANSPARENCY** (openness, clarity, interpretation, communication, demonstration, disclosure): Transparency has the strongest support among ethical standards, some experts consider it not merely a principle, but a prerequisite of all other ethical norms, as the enforcement of other norms is only possible due to the existence of transparent processes.

- TRANSPARENCY OF ALGORITHMS AND DATA PROCESSING METHODS - artificial intelligence must be explicable, its actions and results should be possible to be understood and analyzed by people, thereby allowing for auditing or other testing. Artificial intelligence should not be a "black box", but a clearly interpretable technological solution.

- TRANSPARENCY OF HUMAN ACTIONS IN THE PROCESS OF CREATION, DEVELOPMENT, AND USE OF ARTIFICIAL INTELLIGENCE SYSTEMS. In this case, the whole focus is on the public nature of the practice of using artificial

23 *https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai*

24 "AI Ethics Guidelines: European and Global Perspectives Provisional report by Marcello Ienca* and Effy Vayena" *https://rm.coe.int/cahai-2020-07-fin-en-report-ienca-vayena/16809eccac*

intelligence and AI strategies, on disclosing information about the processing of personal data to the subjects of said data, and on ensuring that oversight bodies are fully involved in the process.

2. **LAWFULNESS** (consistency, inclusion, equality, impartiality, diversity, accessibility, reversibility, challenge, accessibility): Fairness, equality, impartiality primarily imply the creation and use of algorithms in such a way as to avoid preferential affiliations that may lead to discrimination as much as possible. Taking into account all possibilities for inclusion, diversity, and equality when compiling databases and "learning" them. The implementation of the principle of justice also implies the possibility of appealing a decision and the opportunity of potentially receiving compensation.

3. **NON-MALEFICIENCE** (safety, caution, integrity): It is unambigious that artificial intelligence should not cause intentional or negligent harm. Nevertheless, it is also clear that technologies are a potential target of hacking and cyber attacks, and consequently the risk of harm through technology is on the increase. When implementing the principle of safety, it is possible to establish sectoral technical standards and assess the quality of data used in technologies in the process of creating a product or service while taking into account the standards and framework of data protection (security and privacy by design). It is also important to involve stakeholders, to take into account their interests, to enforce the requirements of the law in a consistent manner, and to meet the need to establish oversight processes, reviews, monitoring and auditing practices by internal departments, third parties, independent institutions, and specific users.

4. **RESPONSIBILITY** (accountability): This principle is still the subject of much debate. Specifically, the international community has yet to agree on who takes responsibility for the actions and consequences of using artificial intelligence - technology, its creators, specific individuals, or user organizations. There is a clear trend, however, according to which it is not the technology but the authorized person/body that should be responsible for technological decisions and its outcomes.

5. **INVIOLABILITY OF PERSONAL LIFE** (protection of personal data, confidentiality): The implementation of the right to privacy and personal data security and protections are often challenged by the introduction of data-intensive technologies. Inviolability of privacy in the course of using artificial intelligence can be achieved in several ways: 1) Technological solutions such as differential privacy (aggregating data while excluding data that can be used for identification), privacy-preserving computation, homomorphic encryption (data is processed by machine learning and artificial intelligence systems in an encrypted format, thereby reducing the risk of its disclosure); 2) Increasing the involvement of the public, data subjects by enhancing their level of awareness; 3) Defining clear requirements within the legal framework governing artificial intelligence.

6. **BENEFICIENCE** (benefits, common welfare, social good): Strategies designed to encourage the use of artificial intelligence typically recognize that this technology should not benefit any target group or user, but should be aimed at protecting the public welfare, universal advancement, and fundamental human values, while minimizing the concentration of power.

7. **FREEDOM AND AUTONOMY** (self-determination, choice, consent, empowerment): Artificial intelligence should facilitate the realization of a number of rights and freedoms, including the freedom of expression, information self-determination, the inviolability of privacy, and the implementation of the right to personal autonomy. These represent a kind of positive requirements for artificial intelligence, although in order to avoid negative consequences it is also important for AI to protect people from technological experiments. Protection from manipulation and surveillance, for which, among other mechanisms, it is important to use transparent and explicable AI, increase the level of awareness about AI, and adhere to informed consent policies.

8. **TRUSTWORTHINESS.** Maintaining a rational level of credibility has been the subject of much international discussion. Specifically, how morally justified is the trust in technology, will increased trust lead to neglecting the issue of reviewing technological decision-making? If, on the contrary, it is necessary to strengthen trust through education and by raising public awareness. It is important that AI enjoy balanced trust from users as well as developers. In this regard, education, accountability, reliability of AI systems, integrity monitoring and evaluation in the proper timeframes and in compliance with standards carry a vital importance.

9. **SUSTAINABILITY**. Not actively applied in relation to AI, although the impact of artificial intelligence on ecosystems, biodiversity, labor markets and unemployment, and sustainable data processing is still highlighted.

10. **Dignity**. A high standard of protection of human dignity by artificial intelligence will be achieved if this principle is embedded in the architecture of its creation from the very beginning and if the legislative requirements, new governing initiatives or technical and methodological guidelines issued by the government are observed.

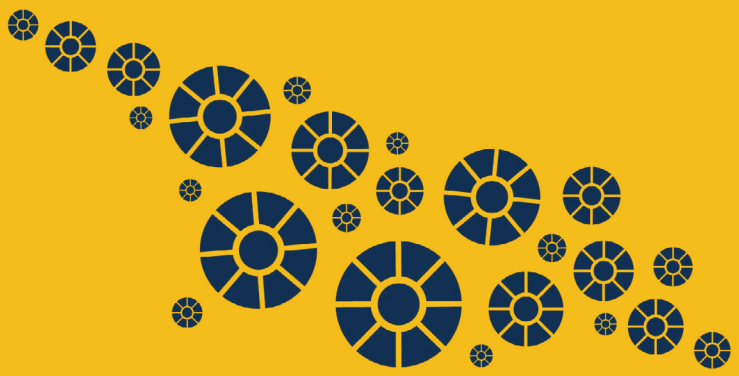
11. **Solidarity**. Solidarity is mainly related to the protection of human rights in the labor market, as the need to redistribute the benefits of AI arises in order prevent a threat to social sustainability and to better take into account the interests of potentially vulnerable individuals and groups.

Ethical self-regulation of artificial intelligence is being pursued by both the public and non-governmental sectors of a number of countries, indicating that this issue is of great importance in all areas, requiring the involvement of many stakeholders and an overall inclusive approach.

There is no internationally agreed definition of ethical artificial intelligence, and countries operate with heterogeneous, in some cases incompatible approaches. However, even at the national level, identifying ethical principles is unambigiously a significant step forward. This is only the beginning stage that requires further development. At this stage, it is important to translate the identified ethical norms into clearly defined actions, to establish appropriate rules and mechanisms for implementation, to develop guidelines for action in the presence of conflicting interests, to bring into agreement the main and auxiliary factors involved in the implementation process and the supervisory institutions.

The document on the ethical side of artificial intelligence developed by the European Data Protection Supervisor (EDPS) is also noteworthy. According to the document:

- The dignity of an individual should not be violated in the digital age;
- The individual and personal data are not separate from each other;
- Digital technologies threaten the foundations of democratic governance;
- Digital data processing gives rise to the possibility of new forms of discrimination;
- The treatment of data as a “commodity” leads to the transfer of value/importance from an individual to his or her personal data.



ARTIFICIAL INTELLIGENCE AND GEORGIA

Georgia has yet to develop a legislative framework, national development concept, strategy and action plan for dealing with the question of artificial intelligence. Instances of the use of artificial intelligence can be found in both the public and private sectors, although issues of their functioning and development are dealt with in the narrow organization framework, as the country does not have a unified national vision with which compatibility would be required for any of these specific cases.

Some examples of the use of artificial intelligence in the public sector are publicly available:

- Risk management system to be used through the automated customs data system "ASYCUDA" when carrying out customs control over declared goods on the customs territory of Georgia. [25] Risk management process in the customs clearance process, for cases when the data of the declared goods meets risk criteria, the relevant corridor and customs clearance procedure are defined for the declared goods registered in the system. The Revenue Service intends to use chatbots in the tax-customs consultation process. The system of automatic refund of VAT excess is also noteworthy.

- "Emotions are Georgia" - a campaign of the National Tourism Administration of the Ministry of Economy and Sustainable Development aimed at analyzing the opinions expressed by tourists about Georgia on social media through the use of artificial intelligence. [26]

- Analytical software of the Ministry of Internal Affairs of Georgia used for face recognition and the identification of a vehicle's state number.

In terms of legislative regulation of the use of artificial intelligence, we should note the order "On the approval of the state of regulations for managing the risks of data-based statistical, artificial intelligence and machine learning models" issued by the President of the National Bank of Georgia in 2020. This statute sets out the principles of risk management at the design, outsourcing, and start-up stages of the models and administrative-organizational control mechanisms for the oversight subjects of the National Bank (e.g., commercial banks, microfinance organizations) in the implementation of statistical, artificial intelligence and machine learning systems.

The statute sets out the requirements for the establishment of processes and mechanisms, the provision of appropriate documentation, system/model control and audit forms, and the requirements to be observed when using data. The following standard is established for data quality: 1) data completeness; 2) data compatibility; 3) timeliness of data; 4) access to data; 5) audit footprint. It is also important that the entities using the model develop ethical principles. Additionally, in the case of commercial banks, they are required to apply the relevant principles set out in the National Bank's Code of Ethics and Standards of Professional Conduct. This statute will come into force in 2021 and aims to guarantee the transparency of artificial intelligence-based models in financial institutions through process planning-control and documentation, as well as to establish a high standard of artificial intelligence risk management.

According to a study conducted by DG NEAR and DG CONNECT for the Eastern Partnership countries in 2018 ("Innovations and Startup Ecosystems"), in terms of artificial intelligence development, Georgia lags far behind the Eastern Partnership countries such as Armenia and Moldova and only slightly exceeds the average EaP indicator. The regional leader in this regard is Armenia, where multiple startups working on AI are currently in operation. The results of the research show that the biggest gap in Georgia among the fields of ICT and innovation training lies in artificial intelligence, e-logistics, open API and open data, industry 4.0, smart contracts. [27]

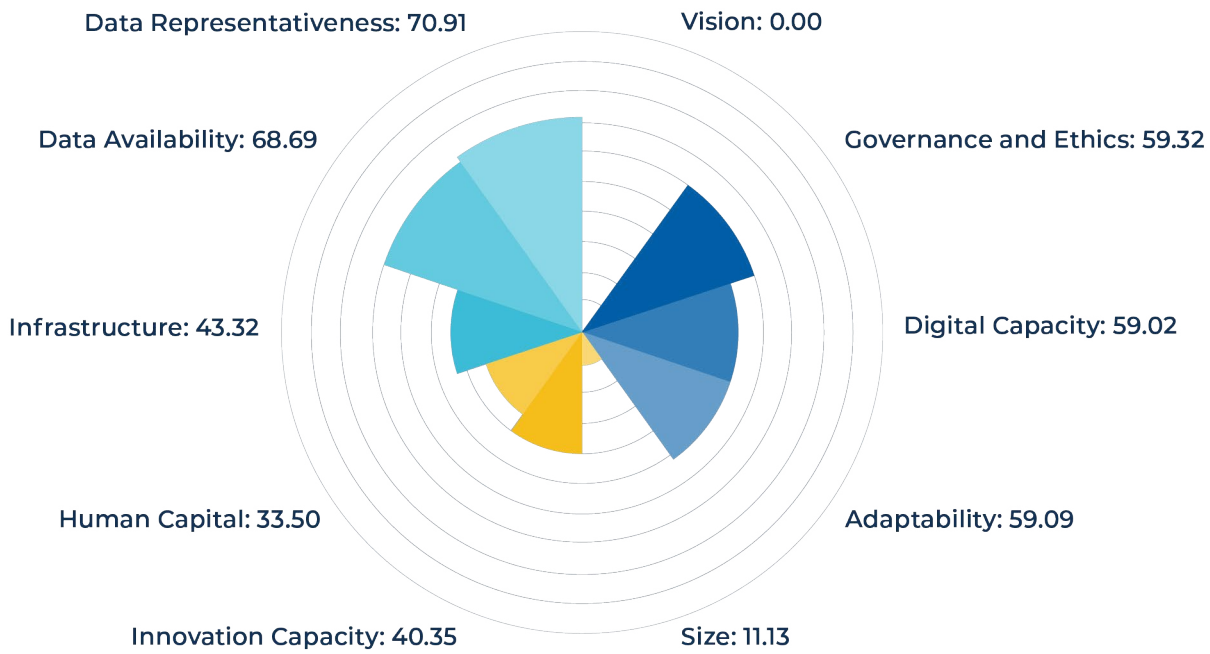
Another study of interest is the research conducted by OXFORD INSIGHT - AI Readiness Index 2020. Within its framework, 172 countries were evaluated based on 10 directions and 33 indicators characteristic to artificial intelligence. Georgia ranks 72nd in this index in the global context, and 5th out of 16 countries in the region (South and Central Asia), lagging behind countries of the region such as Turkey, Azerbaijan, and Kazakhstan.

The results of Georgia according to indicators are distributed as follows: Vision: 0.00; Governance and ethics: 59.32; Digital capabilities: 52.02; Adaptability: 52.09; Share in the economy: 11.13; Innovative opportunities: 40.35; Human capital: 33.50; Infrastructure: 43.32; Data availability: 68.69; Data representation: 70.91.

25 *https://matsne.gov.ge/ka/document/view/81454?publication=3*

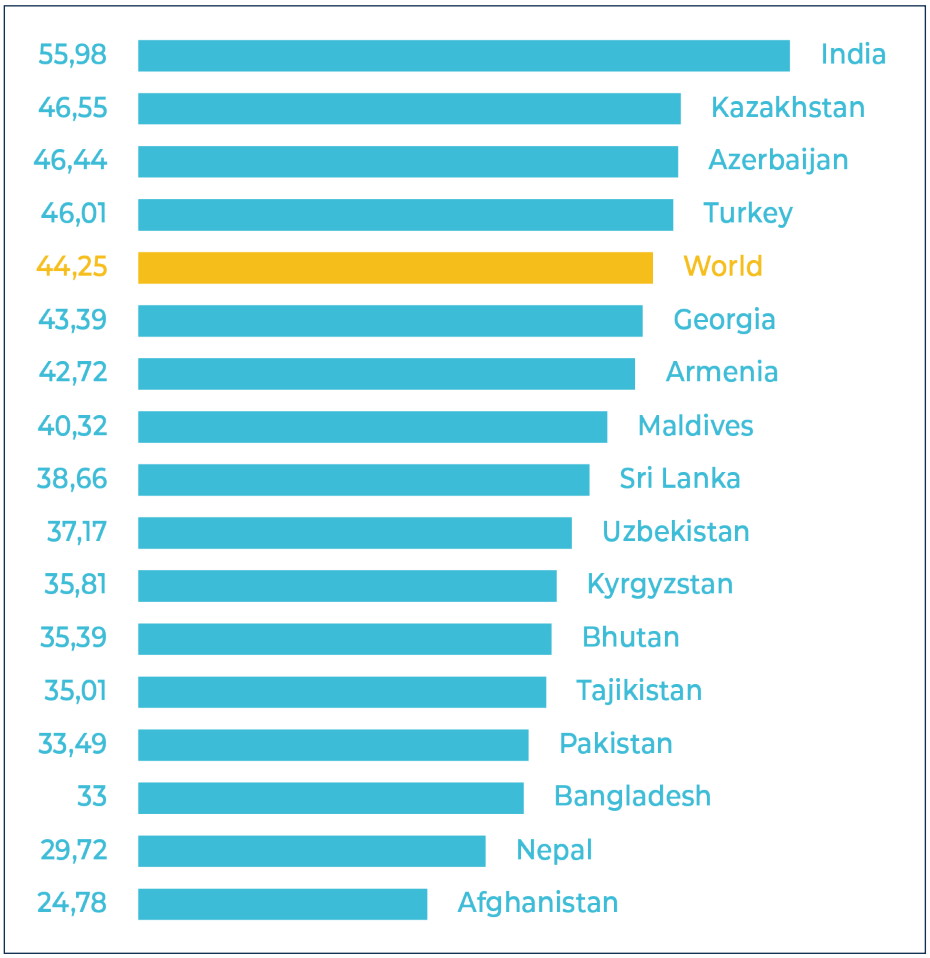
26 *https://gnta.ge/ge/emotion-are-georgia/*

27 *https://eufordigital.eu/wp-content/uploads/2019/10/3.ICT-innovation-and-start-up-ecosystems.pdf*



The region of South and Central Asia as a whole, with its 16 countries, lags behind the world average (South and Central Asia - 39.03, world average - 44.25). Among the countries in the region, only India enters the top 50 category, being ranked at 40. The countries of the region do not have an AI strategy, while precisely this issue has been named as the main barrier to the development of the field. Only India has developed such a strategy, and in our immediate neighborhood - the South Caucasus and the Black Sea countries – only Turkey and Azerbaijan have initiated the process of the development of this kind of strategy.

Regarding Georgia, the 2020 Index emphasizes that despite a supportive business environment (World Bank Doing Business Index - 2019 data, Georgia ranks 7th in ease of doing business), there is no systematic approach to ICT ecosystem development in Georgia, including private sector investment policies and sector development programs. This is further compounded by underdeveloped infrastructure.



RECOMMENDATIONS

THE FOLLOWING **CHALLENGES** WITH REGARD TO THE USE OF ARTIFICIAL INTELLIGENCE CAN BE DISTINGUISHED:

- Adherence to the principles of human rights and freedoms, the rule of law, justice and democratic, good governance;
- Reliability, accountability, security, transparency, and accountability of artificial intelligence-based algorithms;
- The public sector "partially" uses artificial intelligence purposefully and still uses human resources in the decision-making process and/or defers to the will of the management, the problem of delegation;
- Certain imbalance between the potential capabilities of artificial intelligence and the outcomes of existing examples;
- Less transformative, more productivity-oriented, faster, more efficient service delivery examples, and less concern for openness and transparency.

AS A RESPONSE TO THESE CHALLENGES, THE FOLLOWING **RECOMMENDATIONS** WERE ELABORATED:

- Defining the role of the public sector as a facilitator and initiator of the introduction of artificial intelligence;
- Research and analysis of the frameworks promoting artificial intelligence;
- Georgia needs to develop a legal framework, national concept, strategy and action plan regulating artificial intelligence;
- At the national level, establishment of thematic working groups with the involvement of stakeholders, academic research centers, business organizations, public agencies, and donor partners;
- Identification of potential services/areas in the public sector;
- Maximum limitation of bias and discriminatory approaches by algorithms;
- Use of artificial intelligence against fraud, corruption and other common "harmful practices";
- Introduction of a high standard of privacy and personal data protection during the implementation of artificial intelligence systems;
- Formation of the most equitable environment for the use of artificial intelligence with the aim of preventing economic inequality and social discrimination;
- In order for a society to choose public governance based on artificial intelligence and to be able to make critical and argumentative judgments about its pros and cons, it must be given the opportunity to acquire knowledge about this technology.

CONCLUSIONS

Digital transformation is not an optional choice, but an inevitable way for modern developed countries to move forward and succeed.

Modern technologies, including artificial intelligence, cannot have a positive impact and become a prerequisite for public welfare unless the state guarantees the reforms for digital transformation in adherence to high ethical principles, by balancing conflicting interests, with the involvement of all stakeholders, while taking into account best practices and international regulatory standards, and implementing transparent procedures and governance systems.

Artificially supported services can bring a lot of benefits to humanity. The widespread use of this technology in both the public and private sectors of Georgia will lead to a positive indicator in terms of effectiveness, efficiency, resource conservation, and success. Technological readiness alone will not suffice to this end. It is also critical to establish supportive frameworks and gain public trust, which can only be achieved through open, transparent, and accountable processes.

The Georgian public sector should actively take action to stimulate the use of modern technologies, including artificial intelligence, both in public institutions and on a large scale – across the entire country. Georgia must utilize all of the advantages that accompany the introduction of artificial intelligence in a timely and effective manner, so as not to lag far behind the economic-social transformation inspired by artificial intelligence and not miss out on the benefits provided by this process.

20 T. SHEVCHENKO STREET, 0180, TBILISI

+ 995 32 292 15 14 INFO@IDFI.GE WWW.IDFI.GE