

EMERGING LEGAL ISSUES AROUND AI, SYNTHETIC DATA, MACHINE LEARNING

Prepared by Intizar Aliyeva
State Statistical Committee of the Republic of Azerbaijan

Abstract

The paper discusses how to adjust the legislation of the Republic of Azerbaijan in accordance with the requirements of the new and dynamic era and the future projected world, what steps can be taken in this field by government.

I. MAKING USE OF THE POWER OF SYNTHETIC DATA FOR STATISTICAL PURPOSES

1. During the 1990s, statistics offices faced demands to decrease expenses and manpower while boosting the efficiency of the statistical production process. Statisticians were compelled to consider alternatives to traditional data collection techniques in this area. Consequently, the question of combining data gathering through statistical surveys with administrative data sources arose. Beginning in the 1990s, all industrialized nations started using administrative data, which is the foundation for many statistical operations. We are dealing with something similar right now. Technology advancements, particularly the rise in big data processing operations, have raised questions about the need to use a third technique of producing statistical data in addition to administrative and survey resources. Concerns regarding privacy and secrecy have grown as a result of the issue of using big data for statistical analysis. These worries also demonstrated how quickly data privacy laws are evolving. Regulations pertaining to data protection encouraged data scientists to work with other data rather than personal data, which gave rise to the notion of synthetic data. Interest in synthetic data has grown, particularly with the pace of advancements in areas like artificial intelligence and machine learning.

2. Any machine or system that exhibits a human behavioral pattern is said to possess artificial intelligence. In the most basic kind of artificial intelligence, computers are taught to mimic a pattern of behavior that a human could exhibit in a given situation by utilizing enormous volumes of data from prior occurrences of similar patterns of behavior. In general, machine learning and synthetic data are regarded as subcategories of artificial intelligence.

3. Data scientists commonly employ synthetic data, which is a sort of data that is created using different tools and algorithms in place of real data. Data that is meticulously designed to mimic authentic data but does not genuinely belong to an identifiable individual is known as synthetic data.

4. A data controller who uses synthetic data will respect the privacy of real people, whose personal information is processed less. Since synthetic data does not intrude on the personal spaces of the parties involved and differs from data that could be considered personal¹.

5. Organizations can benefit from practically limitless scalability and extremely low-cost synthetic data generation in a number of ways. Numerous copyright, privacy, and compliance laws are in place to

¹ <https://tr.linkedin.com/pulse/6698-say%C4%B1l%C4%B1-ki%C5%9Fisel-verilerin-korunmas%C4%B1-kanunu-kapsam%C4%B1nda-burak-memi%C5%9F>

safeguard sensitive information in industries like healthcare, finance, and law. Nonetheless, businesses in these industries must use data for analysis and research, and they frequently have to provide data to outside parties in order to make the most use of it. To accomplish the same goal as these private data sets, synthetic data—that is, comparable data that displays the same statistically significant information without disclosing sensitive or private data—can be produced in place of personal data. It is feasible to generate "fake" information—that is, information lacking a name, address, or other personally identifiable information—by synthesizing data from a real data set.

6. Machine learning techniques are used to create synthetic data. A branch of artificial intelligence called "machine learning" focuses on building systems that can learn from and perform better as a result of the data they process. Artificial intelligence and machine learning are frequently combined in discussions. Although they have different meanings, they might be used interchangeably in some situations. The fact that not all AI solutions are machine-learning, even though all machine-learning solutions are, is a crucial distinction. Statistics and machine learning research are closely related fields of study. Data is processed, summarized, and guided by statistics for researchers. It supports our ability to forecast outcomes, look at the connections between factors, and make wise choices. We can use statistics primarily for three purposes: decision-making, prediction, and description. In machine learning, statistics are crucial since they aid in the selection, assessment, and interpretation of predictive models.

7. In summary, machine learning can be defined as a methodological paradigm that uses statistical and mathematical techniques to process available data in order to predict unknowns. In essence, machine learning works similarly to human learning—that is, learning from experiences or observations and using that knowledge to inform future judgments. The sole difference is that machine learning is derived from data, whereas human learning is derived from actual observations.

8. By 2030, artificial intelligence models are expected to use synthetic data to the exclusion of real data, according to Gartner. Even though artificial intelligence (AI) has many advantages, it's crucial to keep in mind that machine learning models are built on real data.

II. GAPS, CHALLENGES, AND VULNERABILITIES IN THE LEGISLATION OF THE REPUBLIC OF AZERBAIJAN IN THE FIELD OF THE APPLICATION OF ARTIFICIAL INTELLIGENCE

9. Even if the use of artificial intelligence has produced many benefits in the current period, a number of factors continue to have a detrimental effect on the advancement of this sector. The issue brought about by legal regulation is one of them. As a result, the Republic of Azerbaijan lacks a distinct normative legal act governing the field of artificial intelligence. The use of artificial intelligence systems in our nation is fraught with legal issues, despite the fact that several acts represent distinct rules on the subject. The absence of a legislative framework is one of them. Therefore, despite the fact that artificial intelligence is being used in practice today, neither the Republic of Azerbaijan's primary field law "On Information, Informatization, and Information Protection" nor any other legislative act have established the idea of artificial intelligence. While the term "information system" is used in the Act, it is not synonymous with the term "artificial intelligence system". As a result, the information system is a mechanism that makes available the data we input into it as required. A system with artificial intelligence is one that can evaluate the data that we input and produce output. The fact that the data stored in information systems is arranged in a way that is predetermined by law is another problem. Later on, laws will establish the requirements for acquiring this data. It is debatable, therefore, if the new information gleaned by the

artificial intelligence system's application breaches human rights and freedoms because the matters surrounding it are not governed².

10. The inconsistent usage of terminology is another significant problem in the realm of artificial intelligence applications. The aforementioned law classifies information as either limited or open, depending on the type of access.

11. Subsequently, confidential and secret information are separated from the limited information. Information deemed to be a state secret is referred to as confidential information. The Republic of Azerbaijan's Law "On State Secret" defines a state secret as any information that is protected by the state and whose disclosure could jeopardize the security of the Republic of Azerbaijan. This includes information pertaining to military, foreign-political, economic, intelligence, and counterintelligence operations.

12. Information that is classified as secret includes commercial, legal, investigative, and professional (medical, lawyer, notary) secrets, to which access is prohibited in order to safeguard the rights of persons and legal entities.

13. As is evident, information that qualifies as confidential under Republic of Azerbaijani law also includes information that is considered a state secret. All information must be subject to the confidentiality laws if we are to uphold the "confidentiality" principle. This will blur the line between secret and confidential information and lead to misunderstandings regarding the law. Consequently, we must use the word "confidentiality" in response to the word "privacy". The way the terms were approached by the legislator needs to be altered in order to get around such legal issues.

14. The primary factors that demonstrate how the application process of artificial intelligence is negatively impacted are the high level of population lack of digital literacy, the total elimination of documentation procedures in paper carriers, the population's low recognition of electronic services, the inadequate teaching of digital knowledge and skills in educational institutions, the taught knowledge remaining mostly in theory, and the unsatisfactory access to financial resources.

15. Despite the obstacles in the way, a lot has happened in the past few years, especially in terms of integrating artificial intelligence and modern technology innovations into the public and private spheres in our country. Consequently, the Republic of Azerbaijan will rank 73 out of 193 countries in the global artificial intelligence index released by Oxford Insight magazine at the end of 2023, third behind Turkey (47) and Russia (38) in terms of its neighbors and the South Caucasus³.

16. In line with international norms, the Republic of Azerbaijan's Electronic Security Service works to combat cyber security issues. Established in 2013, this service functions as a coordinating organization under the Republic of Azerbaijan's Ministry of Digital Development and Transport, ensuring the detection of actions intended to violate cyber security and the implementation of preventive measures for prevention⁴.

17. The establishment of Electronic Services Register was carried out by State Agency for Public Service and Social Innovations under the President of the Republic of Azerbaijan, in accordance with Decree No. 262 of the President of the Republic of Azerbaijan dated September 11, 2014, "On measures related to the creation of the Electronic Register of State Services". The register offers all of the state bodies' services through a single interface. The electronic services offered by different government departments are included below, along with links to their websites, contact details, and other relevant information. The

² Aytikin Ibrahimova and Gulnaz Rzayeva: Artificial Intelligence, Human Rights, and Security of Personal Data, April 26, 2021.

³ <https://oxfordinsights.com/ai-readiness/ai-readiness-index/>

⁴ <https://cert.az/>

primary objectives of creating and maintaining the register are to guarantee the integration of the utilized information systems and resources, gather and organize service-related data from a single source, and give citizens access to comprehensive information⁵.

18. Approved by Decree No. 4060 of the President of the Republic of Azerbaijan dated August 28, 2023, the "Strategy of the Republic of Azerbaijan on Information Security and Cybersecurity for 2023-2027" is the primary state policy and the first information security and cyber security strategy adopted in the nation. It decides on the objectives, guiding principles, paths, and most important tasks for the activity. The strategy is to enhance the institutional and regulatory framework, foster a culture of cyber and information security, foster domestic and international collaboration, recognize risks and manage them, and, by the end of 2027, seek to gradually apply such measures.

19. We can also provide instances of artificial intelligence in action, such as the "smart city" and "smart village" initiatives, which are being implemented as test projects in the Karabakh region. Agdam is home to the first "smart city" project, while Agali village in the Zangilan district is home to the first "smart village" initiative. As a result, plans are in place to fully digitize the infrastructure in these locations, build a sizable database, and enhance both the population's quality of life and the efficiency of the city's expenses.

III. ISSUES RELATED TO THE APPLICATION OF SOFT LAW IN THE NEW DRAFT LAW OF THE REPUBLIC OF AZERBAIJAN "ON OFFICIAL STATISTICS"

20. Artificial intelligence makes it necessary to regulate the creation, gathering, storing, and use of data, which combines these two concepts. The Republic of Azerbaijan pays attention to identifying the needs in legal regulation and keeping up with the times by tracking the changes in this sector, even if it is a state that consumes rather than produces technology and thus lags somewhat behind these technical developments. Since evolving technologies are starting to surpass the scenarios envisioned by rules, it appears that numerous adjustments will inevitably occur in the future. Despite the fact that the law is not keeping up with technology, with the help of lawyers who are interested in this area, a legal framework for this quickly evolving technology can be developed. For instance, soft law is the favored model for the regulatory requirements imposed by technology. It differs from hard law in that it lacks binding legal authority.

21. There isn't a widely acknowledged definition of a soft law norm in practice. Even though there is substantial disagreement over the definition of soft law and some academics believe it is unneeded, this idea can be applied to a variety of current international legal instruments that include non-binding normative manifestations. Soft law encompasses many forms of legal provisions, such as guidelines, suggestions, non-binding decisions, principles, proposals, guidelines, notes, standards, and so on. Hard law, on the other hand, refers to legally enforceable texts in international law, frequently in the form of treaties⁶.

22. Soft law, which refers to non-binding standards, unquestionably has a big influence on how the international legal system is established and grows. As a result, soft law tools may eventually develop into enforceable rules. Subsequently, this recommendatory norm's change may take the form of a legal custom or a mandate.

⁵ <https://e-gov.az/az/services>

⁶ <https://opiniojuris.org/2021/11/30/overcoming-the-soft-vs-hard-law-debate-in-the-development-of-new-global-health-instruments/>

23. The three primary characteristics of soft law instruments are summed up by Cassezi as follows:
- (a) Soft law reflects the new concerns of the international community;
 - (b) It illustrates contemporary trends arising in global society;
 - (c) It aids in the convergence of legal norms.
24. States frequently use soft law instruments and are more ready to participate in non-binding agreements than in agreements with legally enforceable provisions. This is primarily because soft law instruments are non-binding, do not necessitate internal procedures, and offer a suitable and adaptable setting for reaching global agreement and meeting the changing requirements of emerging nations.
25. The following categories apply to how soft law affects contract law:
- (a) The result of the initial phase in the contract-finalization procedure:
Shelton highlighted the proactive role of soft law in contract law with regard to this matter, noting that the process of deliberation and the creation of non-binding papers can significantly aid in achieving the consensus required for the ratification of a legally binding multilateral agreement;
 - (b) Closing the gaps in the currently in effect agreements:
Shelton points out that non-binding soft law instruments can be added to binding contracts in this way;
 - (c) Acceptable interpretation of the agreements' terms:
When a contract's concerns are contested and not settled in the document itself, soft law can also be utilized to "authoritatively interpret" or "enforce" its terms.
26. The Economic Commission for Europe of the United Nations created a model law, which served as the basis for the new draft law "On Official Statistics" created by the State Statistical Committee of the Republic of Azerbaijan. The project defines official statistics as those activities that fall under the purview of the statistical works program and are carried out in compliance with this law's provisions as well as international standards and recommendations. These activities include the production and distribution of statistical data that characterizes demographic, social, economic, and environmental events and processes in the Republic of Azerbaijan. The new project also adheres to the model law's guiding principles, which include relevance, impartiality and objectivity, accuracy and reliability, coherence and comparability, clarity and transparency, and statistical confidentiality. In other words, the State Statistical Committee of the Republic of Azerbaijan should not fall behind the current era's rapid pace of technological advancement when it comes to implementing new solutions and safeguarding human rights, particularly those pertaining to privacy and personal information. Therefore, it believes it is suitable to apply soft law and seeks to implement the legal regulation of issues in accordance with international norms and recommendations—the fundamentals of official statistics.
27. In summary, soft law plays a major role in fostering a climate of understanding between nations that might otherwise prefer not to join legally enforceable international accords as well as in generating novel responses to emerging issues facing the global community. Soft law will become more important in international law in the near future because of the regulation of artificial intelligence, which is still developing. Strict legal norms will completely control how far it develops, how far it can be used, and how much it can expand. This will cause legal regulation to lag far behind the development of technology.