IRSHAD

International Journal of Law and Policy |
Volume: 3, Issue: 4
2025

Methods and Tools for Personal Data Protection in Big Data: Analysis of Uzbekistan's Legal Framework

Mamanazarov Sardor Shukhratovich
Tashkent State University of University

Abstract

This study examines methods and tools for protecting personal data in the Big Data context, with a focus on Uzbekistan's legal framework. The research analyzes anonymization, pseudonymization, privacy notices, privacy impact assessments, privacy by design, and ethical approaches to data protection. Through comparative analysis with international standards such as GDPR, the study identifies significant gaps in Uzbekistan's "On Personal Data" law, which lacks specific provisions on modern data protection tools. Research findings reveal that while basic protections exist, Uzbekistan's legislation requires enhancement to address Big Data challenges effectively. This paper proposes legislative amendments to include comprehensive anonymization guidelines, formal pseudonymization processes, and privacy impact assessment requirements. Additional recommendations include establishing personal data repositories, implementing privacy certification mechanisms, and developing national data ethics principles. These measures would strengthen Uzbekistan's data protection framework while enabling innovation in the digital economy, balancing technological advancement with individual privacy rights.

Keywords: Big Data, Personal Data Protection, Anonymization, Pseudonymization, Privacy Notices, Data Protection Law, Uzbekistan

APA Citation:

Mamanazarov, S. (2025). Methods and Tools for Personal Data Protection in Big Data: Analysis of Uzbekistan's Legal Framework. *International Journal of Law and Policy, 3* (4), 20-40. https://doi.org/10.59022/ijlp.322

I. Introduction

The era of Big Data has transformed how organizations collect, process, and analyze information, creating unprecedented opportunities for innovation while simultaneously posing significant challenges to personal data protection. As vast amounts of data are aggregated and analyzed, traditional approaches to privacy may no longer suffice in safeguarding individuals' rights while enabling beneficial data uses. Big Data is characterized by the "three Vs": volume (large quantities of data), velocity (rapid data processing), and variety (diverse data types from multiple sources). These characteristics fundamentally alter the privacy landscape, as data that might appear anonymous in isolation can often be re-identified when combined with other datasets. This reality necessitates robust frameworks for personal data protection that balance technological advancement with individual rights (Ajah & Nweke, 2019).

It is becoming harder to keep personal data truly anonymous in the age of Big Data. Even if names and other details are removed, people can still be identified by combining different pieces of information. For example, just a few details about someone's location and time can be enough to figure out who they are. This shows that old ways of hiding identities may not work well anymore. However, some experts believe that if done carefully with strong rules and checks in place, anonymization can still help protect privacy. Balancing the need to use data and the need to protect people's privacy is now more difficult and important than ever. Privacy notices are meant to help people understand how their personal data is used, but they often don't work well. Most of them are long, complicated, and take too much time to read. In today's digital world, people see so many of these notices that they get tired of reading them and just accept without understanding. This makes it hard for users to truly protect their privacy online (Ortega-Fernandez et al., 2022).

Uzbekistan is going through a digital transformation, with more people using the internet and mobile data every year. As of 2023, about 65% of the population uses the internet, and mobile data use is growing quickly. To protect people's personal information, Uzbekistan passed a law in 2019, but it may not fully cover the new risks that come with modern technologies like Big Data. It is important to design systems that protect privacy from the very beginning, making sure that data is handled safely by default and not just as an afterthought (AllahRakha, 2024). This approach helps build trust and keeps people's information secure as the country becomes more digital.

This paper examines various methods and tools for protecting personal data in the Big Data context, including anonymization, pseudonymization, privacy notices, privacy impact assessments, privacy by design, certification mechanisms, and ethical approaches. By analyzing these tools through the lens of Uzbekistan's legal framework and comparing them with international standards such as the EU's General Data Protection Regulation (GDPR), this research aims to identify gaps and provide recommendations for enhancing data protection in Uzbekistan.

The primary research question guiding this study is: How does Uzbekistan's legal framework for personal data protection compare with international standards in addressing Big Data challenges, and what improvements could be made to enhance data protection while enabling innovation? The significance of this research lies in its potential to inform policy development in Uzbekistan during a critical period of digital transformation. As the country builds its digital economy, establishing robust data protection mechanisms will be essential for fostering trust, ensuring compliance with international standards, and protecting citizens' rights in an increasingly data-driven world.

II. Methodology

This research employs a qualitative approach centered on document analysis and comparative legal research. The methodology is designed to thoroughly examine Uzbekistan's current legal framework for personal data protection while benchmarking it against international standards and best practices. The study utilizes comparative legal analysis to evaluate how Uzbekistan's data protection framework addresses Big Data challenges compared to more established regimes. This approach enables the identification of gaps, strengths, and potential areas for enhancement in the national legislation. The comparative method is particularly well-suited for this study because it allows for the systematic examination of legal frameworks across different jurisdictions while considering their cultural, economic, and social contexts.

Primary data sources include Uzbekistan's Law "On Personal Data" (No. LRU-547), the EU's General Data Protection Regulation (GDPR), legislative acts from countries with advanced data protection frameworks (Japan, South Korea, Singapore), academic literature on data protection in the Big Data context, and international guidelines and standards from organizations such as UNESCO and ISO. Secondary sources include scholarly articles, policy papers, and reports from international organizations on data protection practices in the digital age. These were systematically collected through academic databases including Web of Science, Scopus, and Google Scholar, using search terms related to data protection, privacy, Big Data, and relevant legal frameworks.

The collected data was analyzed through a systematic comparative framework examining how different jurisdictions address key aspects of data protection in Big Data contexts. The analysis focused specifically on: 1) legal definitions and approaches to anonymization and pseudonymization; 2) requirements for privacy notices and transparency; 3) privacy impact assessment frameworks; 4) implementation of privacy by design principles; 5) certification mechanisms for data protection; and 6) ethical frameworks for data governance. For each aspect, Uzbekistan's provisions were compared with international standards to identify alignment, gaps, and potential improvements. The CRAAP test (Currency, Relevance, Authority, Accuracy, and Purpose) was applied to evaluate the quality and reliability of all sources.

This study primarily focuses on the legal and regulatory aspects of data protection and may not fully capture implementation challenges or practical aspects of enforcement. Additionally, as Big Data technologies and regulatory approaches continue to evolve rapidly, the findings represent a snapshot of the current landscape rather than a definitive long-term assessment. The research also does not include primary data collection from stakeholders in Uzbekistan, which could provide additional insights into practical challenges and priorities.

III. Results

A. Current State of Uzbekistan's Legal Framework

Analysis of Uzbekistan's Law "On Personal Data" reveals a basic framework for personal data protection that includes general provisions on data collection, processing, and security. However, the legislation demonstrates significant gaps when compared to more comprehensive frameworks like the GDPR, particularly in addressing Big Data-specific challenges. The law defines personal data and establishes basic principles for its processing, including purpose limitation, data minimization, and security requirements. It also outlines the rights of data subjects and the obligations of data controllers and processors. However, it lacks detailed provisions on several key tools and methods essential for protecting personal data in Big Data environments (Allah Rakha, 2023).

B. Anonymization

Uzbekistan's law addresses the concept of anonymization in Article 16, terming it "depersonalization" (egasizlantirish). "When processing personal data for historical, statistical, sociological, or scientific research, data controllers must depersonalize the data so that it can no longer be associated with specific individuals." However, the law lacks detailed guidance on specific anonymization techniques and standards, criteria for determining when data is sufficiently anonymized, and risk assessment procedures for potential re-identification. This contrasts with more comprehensive frameworks like the GDPR, which provides extensive guidance on anonymization techniques and their implementation.

Anonymization means changing personal data so that it can no longer be used to identify someone. This can be done by removing extra details, reducing specific information, or adding random changes to the data. It is not something that is done just once, but a process that needs to be checked and updated regularly to make sure people stay unidentifiable. Some countries have laws that describe anonymization as a way to make sure no one can be recognized directly or even indirectly from the information. It helps protect people's privacy while still allowing useful analysis. For example, location data from mobile phones can be used to study how groups of people move around, but all personal details are removed first (Rupp & von Grafenstein, 2024). In medical research, information from clinical trials is also carefully anonymized so that researchers can study the data without knowing who the patients are. This way, important insights can be gained without putting anyone's privacy at risk

The question of whether truly effective anonymization is possible in the Big Data era remains disputed. A study at MIT examined records of 1.1 million people's credit card transactions over three months and found that using the dates and locations of just four purchases, 90% of individuals in the database could be identified. While these researchers could identify spending patterns, they didn't actually identify any specific individuals. He also noted that in practice, access to such databases could be controlled, and the anonymization methods applied weren't particularly sophisticated and could be improved.

Organizations using anonymized data must conduct thorough risk assessments for potential re-identification and implement solutions proportionate to the risk. This may include technical measures such as data masking, pseudonymization, and aggregation, as well as legal and organizational safeguards. Anonymization should be viewed not as a means to exempt data processing from regulatory requirements but as a tool to reduce the risk of unlawful disclosure or loss of personal data (Toom & Miller, 2018). It serves as a mechanism that assists Big Data operations and helps organizations conduct research or develop products and services. It also allows organizations to assure individuals that their data will not be used for such analysis, forming an important part of the trust-building process essential for the advancement of Big Data technology.

C. Pseudonymization

The concept of pseudonymization is notably absent from Uzbekistan's data protection law. Unlike the GDPR, which explicitly defines pseudonymization in Article 4(5) and promotes it as a security measure, Uzbekistan's legislation does not recognize or regulate this important data protection tool. This gap is significant because pseudonymization offers a balanced approach to data protection in Big Data contexts, allowing for meaningful analysis while reducing privacy risks. The absence of provisions on pseudonymization limits the legal tools available to organizations in Uzbekistan seeking to implement privacy-enhancing technologies.

Pseudonymization is a method of protecting personal data that involves replacing identifiable information with other data, while still allowing for restoration when necessary (Varanda et al., 2021). For example, an email address or name might be replaced with confidential markers. This de-identifies the data but allows for re-identification using a special key. The U.S. National Institute of Standards and Technology (NIST) has identified the "pseudonym" substitution of data as an important method of data protection. Japan's "Act on Anonymously Processed Information" clearly defines the pseudonymization process and considers it of great importance for developing the digital economy.

GDPR Article 4(5) defines pseudonymization as: "The processing of personal data in such a manner that the personal data can no longer be attributed to a specific data subject without the use of additional information, provided that such additional information is kept separately and is subject to technical and organizational measures to ensure that the personal data are not attributed to an identified or identifiable natural person." Pseudonymization facilitates data processing and reduces the risk of confidential information being disclosed to unauthorized persons. For instance, when sending Excel spreadsheets containing confidential information via email, IT staff also gain access to this data. If the data contains information about executives' bonuses or salaries, the disclosure of this information could be dangerous. Pseudonymized data significantly reduces such risk.

In this context, a pseudonym is an identifier associated with a person. Just as writers use pseudonyms to hide their identity and protect their privacy, pseudonyms are used for the same purpose in data protection. A pseudonym can be a number, letter, special character, or combination thereof, and is associated with specific personal data or a person. This makes data use safer in a business environment. According to GDPR, pseudonymized data is still considered personal data since the process is reversible and the person can be identified with the appropriate key. Article 26 explains: "Personal data that could be attributed to a natural person through the use of additional information, including pseudonymized data, should be considered information about an identifiable person."

Under Singapore's amended Personal Data Protection Act (PDPA), derived personal data is defined as personal data obtained by an organization from other personal data in the course of business, or from other data in the possession or control of the organization. This includes pseudonymized data. Unlike anonymization, which transforms data so that a person cannot be directly or indirectly identified, pseudonymization modifies data but allows for restoration using a special key. Pseudonymization can be used in companies handling confidential information, such as in HR, marketing, or IT departments, to reduce risk and prevent data leakage. It also supports companies' overall GDPR compliance.

Pseudonymization can be implemented through various methods, including data masking, encryption, or tokenization. It’s also recommended for protecting personal data from systems not used for production, testing, or training purposes. Anonymized personal data sets can still be useful for development, statistics, and analysis. South Korea’s Personal Information Protection Act (PIPA) clearly distinguishes between pseudonymization and anonymization. It defines “pseudonymization” as “removing directly identifying information from personal data and replacing it with other data” and anonymization as “modifying personal data in a way that excludes the possibility of directly or indirectly identifying an individual.”

Both methods are widely recommended, but which one to choose depends on many factors, including the use case, risk level, and how data is processed in the company. The best method also depends on the purpose of processing, the type of data being processed, and the risk of data leakage it presents. Pseudonymization is more complex than anonymization because it leaves a key to “unlock” the data (Tinabo et al., 2009). In this method, the data is not directly identifiable and is not anonymized, so it doesn’t lose its original value. According to GDPR Article 28, applying pseudonymization to personal data can reduce risks to the data subjects and help controllers and processors fulfill their data protection obligations.

A common use case for pseudonymization in production systems that process personal data is as a temporary storage of original values during anonymization and as a rollback mechanism against failures. In such cases, pseudonyms may be stored for a short time, just long enough for the business to confirm the successful completion of anonymization. Countries like Japan and South Korea have explicitly recognized pseudonymization in their data protection laws. For example, South Korea’s Personal Information Protection Act (PIPA) clearly distinguishes between pseudonymization and anonymization, providing a legal foundation for implementing these techniques appropriately.

D. Privacy Notices

Regarding transparency and information provision, Uzbekistan’s law includes basic requirements. Article 31 requires data controllers to notify data subjects when their personal data is modified, deleted, restricted, or transferred to third parties. Additionally, Article 22 specifies that operators must provide data subjects with information about the operator’s location, the purpose of personal data processing, recipients of personal data, categories of personal data processed, processing period, and legal consequences. While these provisions establish a foundation for transparency, they lack specific requirements for the format, accessibility, and readability of privacy notices. The law does not address the unique challenges of providing meaningful notice in Big Data contexts, where data collection may be ubiquitous and data uses might evolve over time.

Except in certain specified situations, personal data processing cannot be considered fair unless the data subject is provided with certain basic information, including the identity of the data controller, the purpose of processing, and other information necessary for ensuring fair processing (privacy notice). The GDPR requires data controllers to provide more detailed information, particularly in the context of "automated decision-making, including profiling." In such cases, data controllers must explain the logic involved and the "significance and envisaged consequences" of profiling for the data subject. This applies only to decisions based solely on automated processing, not to decisions made by humans. It is relevant for Big Data processes where algorithms are developed initially and then applied to specific situations, such as determining credit scoring. Data controllers need to find reasonable ways to explain how decisions are made.

The transparency of privacy notices and their compliance with actual practices are crucial in the Big Data context, as they allow data subjects to be informed about how their data is being used and to make informed decisions. The U.S. Department of Commerce's data protection standards establish requirements for transparency and providing privacy notices. Specifically, they state that "organizations should provide clear and easily understandable notice regarding the collection, use, disclosure, and retention of their personal information." Japan's "Act on the Protection of Personal Information" also incorporates the concept of privacy notices. It states that "personal information controllers shall provide, in an easy and understandable manner, the purpose of collecting personal information, how personal information is processed, and other necessary information."

In the Big Data context, these requirements can be problematic, and there's an argument that privacy notices may be impractical for Big Data analysis. This argument is based on several grounds:

- People are unwilling to read lengthy privacy notices.
- The sources from which data is collected (e.g., smartphone apps or IoT devices) may make providing information practically difficult.
- The analyses applied in Big Data can be too complex to explain in terms people can understand.
- Big Data analysis often involves reuse of data, so the data controller may not foresee all possible uses of the data initially.

The Korea Internet and Security Agency (KISA) acknowledges these difficulties, stating that "applying the traditional privacy notice model in the Big Data environment may be difficult.” However, KISA also asserts that “the right of data owners to know how their personal information is processed should be guaranteed through a privacy notice.” According to Uzbekistan’s Law “On Personal Data,” “the owner and (or) operator must notify the subject in writing when personal data has been modified, deleted, restricted, or when personal data has been transferred to a third party” (Article 31, Part 2). The law also requires operators to provide subjects with the following information:

- The operator’s location (postal address);
- The purpose of personal data processing;
- The range of personal data users (recipients);
- The content of the personal data provided;
- The period of personal data processing;
- The legal consequences.

Thus, Uzbekistan’s legislation also imposes an obligation to provide privacy notices to personal data subjects. Checking the box “I have read and agree to the terms and conditions” is described as the “biggest lie on the web.” Undoubtedly, when people want to buy something online or download an app, they check “I agree” without reading the privacy notice. The propensity to read privacy notices is well documented:

- It was found that reading all terms and conditions encountered on the Internet would take one month per year.
- The White House report on Big Data noted the phenomenon of “privacy fatigue” and found that although U.S. advertisers provided information about data use, few people read or understood it.
- According to a WIK Consult report, there are few incentives for people to read privacy policies when using the Internet, as it takes much more time than using the content or application.

However, it would be wrong to conclude that the requirement to provide a privacy notice is inappropriate or inapplicable in the Big Data context just because people are indifferent to how their data is used. It’s understandable that people don’t want to read lengthy privacy notices written in legal terms or designed to protect the organization using the data rather than the data subjects.

Big Data organizations should find innovative ways to deliver privacy notices to subjects in a concise format. Specifically, they can use videos, animations, timely notifications, and standardized icons. Privacy notices should be written in plain language, considering the average reading level of a person. Textual information can be provided along with other methods of delivering information in a user-friendly format. Channel 4's use of a YouTube video is an example of an innovative approach to following the "Viewer Promise." The Guardian and O2 use animations to explain their privacy policies. A combination of different approaches can be used to make the information more accessible for understanding.

Several organizations are currently developing practical ways to make privacy notices understandable. These approaches include promoting the use of plain language, identifying commonly used terms, and creating a database for reusing them in different contexts with standard icons. Just as nutrient content is conveyed on food packaging using standard methods, a similar approach could be applied to privacy notices.

The GDPR states that information intended for the public or the data subject should be "concise, easily accessible and easy to understand, and that clear and plain language and, additionally, where appropriate, visualization be used." It notes that "this is of particular relevance in situations such as online advertising, where the proliferation of actors and the technological complexity of practice makes it difficult for the data subject to know and understand whether, by whom, and for what purpose personal data relating to him or her are being collected." It also mentions the possibility of using standardized icons to explain processing.

The approaches of Asian countries to privacy notices are interesting. In Japan, companies are required to ensure that privacy policies are provided before collecting personal information. In South Korea, the data protection law requires operators to provide privacy notices "in an easy and understandable format." The International Organization for Standardization (ISO) also emphasizes in standard 29184:2020 that privacy notices should be understandable, easy to read, and simple for users. It recommends enriching privacy notices with visual elements and pictograms.

In Uzbekistan, privacy notices in the Big Data field should be developed in a simpler, clearer, and more user-friendly manner. It would be useful to develop national guidelines based on the experience of developed countries and international standards. These guidelines should establish specific requirements for the form, volume, design, and language of privacy notices, and provide sample templates. When Big Data includes information not directly provided by individuals but collected or observed by applications or devices, providing a privacy notice can be more difficult, but there are solutions to this problem.

For data protection related to IoT devices, particularly wearable gadgets like watches, glasses, and home devices like smart thermostats, privacy information can be provided on the device itself or transmitted via Wi-Fi or QR code. It's suggested that explaining processing in a privacy notice can be very difficult because Big Data relies on complex analysis and algorithms. However, this misunderstands the purpose of the privacy notice. Processing cannot be fair if people are deceived or misled about these purposes. Even if it's difficult to explain in simple terms how the analysis works, it's necessary to explain the purposes in a way that doesn't deceive or mislead people.

Organizations analyzing Big Data may initially collect data for one purpose but later wish to use it for other purposes (Acciarini et al., 2023). In such cases, they should immediately identify the purposes for which the data will be used, inform people, and obtain their consent for using the data for new purposes. If one organization buys personal data from another organization and wishes to analyze it, the seller's initial privacy notice must have indicated this possibility. Otherwise, the buyer must provide its own privacy notice to individuals, explaining the new purpose of data processing.

Privacy notices are also important when organizations merge or are acquired by another organization. Such mergers or acquisitions in the technology sector, such as Facebook's acquisition of WhatsApp or Microsoft's acquisition of LinkedIn, are frequent. In such cases, data protection obligations transfer along with the data. Therefore, people need to be informed of the situation and assured that their personal data will only be used within their reasonable expectations. Providing the original privacy notice, informing about the new organization, and explaining what's happening helps fulfill this requirement.

The use of data posted on social networks by other organizations is also becoming increasingly widespread. Social media platforms like Facebook and Twitter provide information posted by subscribers to third parties under certain conditions. Social media companies like Twitter may collect information through a software interface (API) or sometimes independently through third-party "web harvesting." The data may be used to analyze opinions or identify general trends. In some cases, the data is also used for profiling individuals, such as assessing credit risk. When data is transferred to another party, it can be difficult to anonymize, so the third party may need to process personal data. In such cases, it's necessary to consider whether to provide a privacy notice to the relevant individuals (Jain et al., 2021).

The social media company's terms of service may include terms related to third-party use, but in reality, people may not know how their data is being used. Research conducted by Ipsos Mori showed that two out of five adults knew their social media data could be shared for research by companies or the government, and three out of five adults thought it shouldn't be. These challenges are exacerbated in Big Data contexts, where data collection may occur through devices with limited display capabilities (IoT), and where data uses might evolve over time in ways not foreseeable at collection. The "biggest lie on the web" remains the claim "I have read and agree to the terms and conditions" (biggestlie.com), highlighting the ineffectiveness of traditional notice mechanisms.

E. Privacy Impact Assessment

Uzbekistan's legislation does not explicitly require privacy impact assessments. While Article 27 mandates that data controllers implement legal, organizational, and technical measures to protect personal data, it does not establish a formal framework for assessing privacy risks before initiating data processing activities. This contrasts with the GDPR's requirement for Data Protection Impact Assessments (DPIAs) for high-risk processing activities, particularly those involving systematic evaluation based on automated processing or large-scale processing of special categories of data.

An important tool that helps identify and mitigate privacy risks before processing personal data is the Privacy Impact Assessment. According to GDPR rules, when analyzing Big Data that involves processing personal data, a privacy impact assessment a process called "data protection impact assessment" is likely to be required. Some aspects of Big Data analysis may complicate certain stages of the privacy impact assessment, but these difficulties can be overcome. The U.S. Federal Trade Commission recommends privacy impact assessments as a "best practice" for companies before collecting data. The Council of Europe also describes PIAs as a "very useful tool" in Big Data environments.

Some Asian countries, such as Japan and Singapore, have made PIAs mandatory in their legislation. In South Korea, while conducting a PIA is not mandatory, it is supported by the Personal Data Protection Committee. In Uzbekistan's legislation, the concept of PIA has not yet been established, but elements of it are expressed in Article 27 of the Law "On Personal Data." In particular, it states that the owner and (or) operator, as well as third parties, must take legal, organizational, and technical measures to protect personal data to ensure the subject's right to protection from interference in personal life, to observe the integrity and confidentiality of personal data and ensure their preservation, and to prevent illegal processing of personal data.

Big Data analysis may involve new, complex, and sometimes unexpected uses of personal data. To determine if processing is fair, especially before it begins, it's important to assess how it might affect the individuals whose data is being used and identify possible mitigations (Bormida, 2021). This is where a privacy impact assessment (PIA) is applied. It is now considered good practice to conduct a PIA in projects involving new uses of data. However, the GDPR requires a process called "data protection impact assessment" (DPIA) in certain cases, especially when new technologies are used, which may pose a high risk to the rights and freedoms of individuals. Most Big Data applications involving personal data processing are likely to fall into this category. The GDPR includes the "systematic and extensive" assessment of individuals based on automated processing, including profiling, where decisions have a significant effect on individuals.

F. Privacy by Design

The concept of Privacy by Design embedding privacy protections into systems and processes from the earliest stages is not explicitly recognized in Uzbekistan’s data protection law. While Article 27 requires technical and organizational measures for data protection, it does not mandate a proactive, design-based approach to privacy. This represents another significant gap compared to international standards, as Privacy by Design has become a fundamental principle in modern data protection frameworks, including the GDPR (Article 25). The use of Big Data should not come at the expense of privacy. Implementing privacy-enhancing solutions in Big Data analysis helps protect privacy through a range of technical and organizational measures. Under GDPR rules, the concept of privacy by design, termed “data protection by design and by default,” is a legal requirement.

The U.S. National Institute of Standards and Technology (NIST) has developed a “Guide to Privacy Engineering” (SP 800-53) for Big Data developers, which provides recommendations for implementing privacy by design principles. The concept of privacy by design is often associated with applying methods of anonymizing or pseudonymizing personal data. One such method is “differential privacy,” which involves adding “noise” to queries in a database. The noise should be sufficient to provide anonymity at the individual level but not so much as to affect the usefulness of the query response. Differential privacy is becoming a popular privacy-ensuring method among major technology companies like Apple and Google. However, privacy by design solutions includes not only anonymization methods but also a range of other technical and organizational measures, including:

- Security measures to prevent misuse of data,
- Data minimization measures to process only necessary personal data at each stage,
- Purpose limitation to store personal data separately from data intended for identifying general trends and relationships,
- “Sticky policies” that record individuals’ preferences and corporate rules in the metadata that comes with the data.

Asian countries are also applying privacy by design principles. The “Smart City” model developed by the Japanese government includes specific technical solutions for ensuring privacy. South Korea’s national identification system also employs privacy by design, using a distributed rather than centralized storage system. ENISA has published a comprehensive report on using privacy-ensuring methods in Big Data. It provides examples of how the privacy by design approach can be applied in various use cases for "smart cities," such as smart parking applications, smart meters, and civic platforms. It envisages a shift from "Big Data versus privacy" to "Big Data with privacy." They acknowledged that this is not easy to achieve and that more work on Privacy-Enhancing Technologies (PET) is needed, but concluded that "the concept of privacy by design is important in identifying privacy requirements early in the Big Data analysis chain and subsequently implementing the necessary technical and organizational measures."

The privacy by design concept was incorporated into the GDPR rules under the heading "Data Protection by Design and by Default." Thus, it became a legal requirement, as data controllers are obliged to "implement appropriate technical and organizational measures to ensure that, by default, only personal data necessary for each specific purpose of processing are processed." In Uzbekistan's Law "On Personal Data," the concept of privacy by design is not established. However, the requirement to implement technical and organizational measures to ensure privacy is provided for in Article 27. Privacy by Design encompasses various technical and organizational measures, including data minimization strategies, access controls, privacy-enhancing technologies like differential privacy, data segregation approaches, and "sticky policies" that encode privacy preferences.

G. Certification Mechanisms

Uzbekistan's law does not establish certification mechanisms or privacy seals that would allow organizations to demonstrate compliance with data protection standards. Such mechanisms, which are encouraged under the GDPR (Articles 42-43), could help build trust in data processing activities and provide incentives for organizations to implement strong privacy protections. The certification system helps demonstrate compliance with data protection laws in Big Data analysis.

A system for certifying that certain personal data processing complies with data protection requirements, often called a trust mark or privacy seal, is encouraged. It is noted that this could help increase consumer confidence in processing in the Big Data context. For example, Huawei announced that it had received a form of certification from the German company ePrivacy for its Hadoop-based Fusion Insight product. Certification is also provided for in the GDPR. It requires the introduction of "data protection certification mechanisms and data protection marks and seals" to demonstrate that Big Data analysis complies with legal requirements. These are issued by data protection authorities or accredited certification bodies.

Among APEC countries, there is also a Cross-Border Privacy Rules (CBPRs) system, which allows companies to certify their compliance with data protection requirements. In Japan, companies can certify their privacy and information security systems under the "General Rule for Protection of Personal Information" (PIPA). South

Korea also introduced a data protection certification system in 2011. Certification confirms that a particular service, product, or process (not the entire organization) meets the requirements of data protection laws. In Uzbekistan, the issuance of privacy certificates is not yet provided for in legislation.

H. Ethical Approaches

The law does not address ethical dimensions of data protection or establish frameworks for ethical data governance. This contrasts with emerging international practices, where ethics councils and ethical guidelines increasingly complement legal requirements in guiding responsible data use. Ethical approaches are very important in the use of Big Data and are an essential means of complying with legal requirements. Ethics councils at organizational and national levels help assess issues and apply ethical principles. Ethical approaches to using personal data help develop trusting relationships with people. They also play an important role in setting Big Data standards to promote best practices across various sectors.

Ethics, in particular, is based on principles such as respecting human dignity in working with data, a fair and non-discriminatory approach, and ensuring privacy. In developing ethical principles, companies and organizations pay special attention to transparency, fairness, and gaining customer trust. Aspects such as how to collect, use, and share data are reflected in these principles. Giving customers control and choice over their data is considered important. Organizations in both private and public sectors are actively involved in developing their ethical principles. For example, a Privacy Advisory Council has been established in Seattle, USA, to support “smart city” initiatives.

Organizations also consider business interests when developing ethical principles. Adhering to ethical principles ensures customer trust and guarantees the proper use of their personal data. As a result, companies can increase their revenues. The Japanese government adopted “Data Ethics” principles, highlighting issues of respecting human dignity, diversity, and inclusivity in the use of data. Singapore also published guiding principles on data ethics for artificial intelligence system developers in 2019. Organizations such as the International Telecommunication Union and the International Organization for Standardization are actively involved in developing standards to establish best practices and reduce risks for organizations dealing with Big Data processing. In Uzbekistan, bold steps should be taken to implement data ethics. The following is proposed:

- Develop “National Data Ethics Principles” that prioritize humanity, justice, and diversity in the collection, storage, and use of data.
- Study and adapt the data ethics rules of advanced foreign companies and government agencies to the national context.

- Establish advisory councils and committees on data ethics in government bodies and the private sector.
- Establish teaching and research on data ethics in higher education institutions and scientific institutions.

Japan's "Data Ethics" principles (2019) emphasize respect for human dignity, diversity, and inclusivity in data use. Similarly, Singapore has developed ethical guidelines for AI and data use (PDPC Singapore, 2019). These approaches recognize that legal compliance alone may be insufficient to address all concerns related to Big Data processing.

I. Personal Data Repositories

Personal data repositories represent an innovative approach to enhancing individual control over personal information that is not currently addressed in Uzbekistan's legislation. Implementing a framework for personal data repositories could enhance individual control while potentially facilitating innovation in data-driven services. Using personal data repositories can address fairness and transparency issues by giving individuals greater control over their personal information. Personal data repositories can support the concept of data portability (according to GDPR, under certain conditions) of an individual's personal data under their control.

One way to increase individuals' control over the use of their data is usually suggested through a personal data repository or sometimes through personal data management services. These are third-party services that store individuals' data on their behalf and provide it to organizations when individuals wish. Rubinstein, an early advocate of this system, saw it as a way to implement privacy controls by managing organizations' access to personal data and setting "fine-grained" privacy settings. The European Data Protection Supervisor also sees personal data repositories as a way to address problems related to individuals losing control over their data.

The idea that individuals can effectively manage how their personal data is used in a Big Data environment as the "data privacy self-management fallacy." He states that people don't know that their data is being collected or how it's being used, and they don't have time to read privacy notices. Instead, he proposes "delegate data management," a system of intermediaries who manage a person's data on their behalf. It has been proposed to organize personal data repositories on a cooperative basis, where individuals who store their data in the repository can receive financial benefits when their data is used (Obar, 2015).

The GDPR introduced the concept of data portability. If a data controller is processing personal data based on consent or a contract, the data subject has the right to receive the data they have provided in a "structured, widely used and machine-readable format." They also have the right to transmit this data to another data controller. In Uzbekistan's Law "On Personal Data," the concept of a personal data repository is not established. To implement personal data repositories in Uzbekistan, the following steps are recommended:

- Create an online platform for citizens to collect and manage their personal data (My Data Portal).
- Establish a licensing system for personal data repository operators, where operators are required to comply with privacy and security requirements.
- Create a rating system to evaluate how companies and government agencies use data from personal data repositories.

The above analysis shows that various methods and tools for ensuring privacy and protecting personal data are used in the development of the data economy. Only through their correct and effective application can Big Data technologies be used for the benefit of society. In this regard, it is necessary to rely on the experience of developed countries and the recommendations of international organizations, while taking into account the peculiarities of the national legal system and the level of technological development. In Uzbekistan, it is necessary to create an appropriate legal framework and infrastructure in this area, particularly to introduce modern tools such as anonymization, pseudonymization, privacy by design, as well as establish personal data repositories and certification mechanisms, and develop Big Data ethics principles.

Control over data is not only a technological but also a socio-legal tool. Therefore, protecting the rights of data subjects and users and ensuring information security is one of the important conditions for the innovative development of Uzbekistan. Thus, Uzbekistan should recognize data security and privacy as a priority issue in the transition to a digital economy and the development of innovations. To do this, it is necessary to create legal and organizational foundations, develop technological solutions, and improve the literacy of the population. We hope that these analyses and recommendations will contribute to improving state policy on protecting personal data when using Big Data technologies in Uzbekistan. This, in turn, will serve to develop the digital economy and strengthen citizens' trust.

IV. Discussion

The analysis of Uzbekistan's data protection framework reveals several significant gaps when compared with international standards and best practices. These gaps limit the effectiveness of the current framework in addressing the unique challenges posed by Big Data technologies.

A. Anonymization and Pseudonymization

The absence of detailed provisions on anonymization techniques in Uzbekistan’s legislation creates significant uncertainty for organizations processing Big Data. Without clear guidelines on what constitutes effective anonymization, organizations may implement inadequate measures that fail to protect privacy or overly restrictive approaches that unnecessarily limit data utility. The complete absence of pseudonymization provisions represents an even more significant gap. Pseudonymization offers a middle ground between fully identified and fully anonymized data, allowing for continued utility while reducing privacy risks.

It is particularly valuable in Big Data environments where complete anonymization might severely limit analytical possibilities. To address these gaps, Uzbekistan should consider amending its legislation to provide detailed guidance on anonymization techniques and standards, introduce the concept of pseudonymization with clear definitions and requirements, establish criteria for evaluating the effectiveness of both anonymization and pseudonymization, and develop sector-specific guidelines for implementing these techniques in different contexts.

B. Privacy Notices and Transparency

While Uzbekistan’s law establishes basic transparency requirements, it fails to address the challenges of providing meaningful notice in Big Data environments. Traditional notice and consent models face significant limitations in the Big Data context, particularly given the phenomenon of “privacy fatigue” among consumers faced with numerous complex notices. To address these challenges, Uzbekistan’s legal framework should encourage innovative approaches to privacy notices, such as videos, animations, and layered notices. The GDPR’s emphasis on information that is “concise, easily accessible and easy to understand” provides a useful model for enhancing Uzbekistan’s transparency requirements.

C. Privacy Impact Assessment and Privacy by Design

The absence of formal Privacy Impact Assessment (PIA) requirements and Privacy by Design provisions represents critical gaps in Uzbekistan’s legislation. PIAs serve as a proactive tool for identifying and mitigating privacy risks before they materialize, while Privacy by Design ensures that privacy considerations are embedded into systems and processes from the earliest stages. Incorporating these tools into Uzbekistan’s legal framework would strengthen protection while supporting responsible innovation. They would help organizations identify and address privacy risks early in the development process, reducing the likelihood of privacy violations and building trust with data subjects.

D. Certification and Ethical Approaches

Certification mechanisms and ethics councils can complement legal requirements by providing additional assurance and guidance. Implementing certification mechanisms would help build trust in Uzbekistan's digital economy by allowing organizations to demonstrate their commitment to privacy, while ethical frameworks would guide responsible data use beyond legal compliance.

E. Personal Data Repositories

Personal data repositories represent an innovative approach to enhancing individual control over personal information that is not currently addressed in Uzbekistan's legislation. Implementing a framework for personal data repositories could enhance individual control while potentially facilitating innovation in data-driven services. Specific measures could include creating an online platform for citizens to collect and manage personal data, establishing a licensing system for personal data repository operators, and developing a rating system to evaluate how companies and government agencies use data from personal data repositories. Based on this analysis, several recommendations emerge for enhancing Uzbekistan's data protection framework:

- Amend the Law "On Personal Data" to include detailed provisions on anonymization techniques and standards, introduce the concept of pseudonymization, establish requirements for accessible privacy notices, introduce mandatory PIAs for high-risk processing, and incorporate Privacy by Design principles.
- Establish a national privacy certification scheme, develop a national data ethics council, and create a legal framework for personal data repositories.
- Develop educational programs on data protection for professionals, create public awareness campaigns about data rights, and provide technical guidance for implementing data protection tools.
- Participate in international forums on data protection, establish bilateral cooperation with countries having advanced data protection frameworks, and align national standards with international best practices.

Conclusion

This study has examined the methods and tools for protecting personal data in the Big Data context, with a particular focus on Uzbekistan's legal framework. The analysis reveals significant gaps in Uzbekistan's legislation compared to international standards, particularly regarding advanced data protection tools such as anonymization techniques, pseudonymization, privacy impact assessments, and privacy by design. While

Uzbekistan’s Law “On Personal Data” establishes basic protections, it lacks the detailed provisions necessary to address the unique challenges posed by Big Data technologies. The absence of specific guidance on anonymization methods, the complete omission of pseudonymization, the lack of formal privacy impact assessment requirements, and the failure to incorporate privacy by design principles limit the effectiveness of the current framework.

To strengthen data protection while enabling innovation, Uzbekistan should implement comprehensive recommendations including legislative amendments, institutional mechanisms, capacity building initiatives, and international cooperation. By implementing these recommendations, Uzbekistan can create an environment that protects individual privacy while enabling the responsible use of data for innovation and economic development. A robust framework that addresses the complexities of Big Data will build trust in digital systems, facilitate international data flows, and create a foundation for sustainable growth in the digital economy. As Uzbekistan continues its digital transformation journey, strengthening data protection should be recognized as a priority.

Bibliography

Acciarini, C., Cappa, F., Boccardelli, P., & Oriani, R. (2023). How can organizations leverage big data to innovate their business models? A systematic literature review. *Technovation*, *123*, 102713. https://doi.org/10.1016/j.technovation.2023.102713

Ajah, I. A., & Nweke, H. F. (2019). Big Data and Business Analytics: Trends, Platforms, Success Factors and Applications. *Big Data and Cognitive Computing*, *3*(2), 32. https://doi.org/10.3390/bdcc3020032

Allah Rakha, N. (2023). Artificial Intelligence strategy of the Uzbekistan: Policy framework, Preferences, and challenges. *International Journal of Law and Policy*, *1*(1). https://doi.org/10.59022/ijlp.27

AllahRakha, N. (2024). Constitutional Safeguards for Digital Rights and Privacy. *International Journal of Law and Policy*, *2*(4), 31–43. https://doi.org/10.59022/ijlp.172

Bormida, M. Da. (2021). *The Big Data World: Benefits, Threats and Ethical Challenges* (pp. 71–91). https://doi.org/10.1108/S2398-601820210000008007

Jain, A. K., Sahoo, S. R., & Kaubiyal, J. (2021). Online social networks security and privacy: comprehensive review and analysis. *Complex & Intelligent Systems*, *7*(5), 2157–2177. https://doi.org/10.1007/s40747-021-00409-7

Obar, J. A. (2015). Big Data and *The Phantom Public* : Walter Lippmann and the fallacy of data privacy self-management. *Big Data & Society*, *2*(2). https://doi.org/10.1177/2053951715608876

Ortega-Fernandez, I., Martinez, S. E. K., & Orellana, L. A. (2022). Large Scale Data Anonymisation for GDPR Compliance. In *Big Data and Artificial Intelligence in Digital Finance* (pp. 325–335). Springer International Publishing. https://doi.org/10.1007/978-3-030-94590-9_19

Rupp, V., & von Grafenstein, M. (2024). Clarifying "personal data" and the role of anonymisation in data protection law: Including and excluding data from the scope of the GDPR (more clearly) through refining the concept of data protection. *Computer Law & Security Review*, *52*, 105932. https://doi.org/10.1016/j.clsr.2023.105932

Tinabo, R., Mtenzi, F., & O'Shea, B. (2009). Anonymisation vs. Pseudonymisation: Which one is most useful for both privacy protection and usefulness of e-healthcare data. *2009 International Conference for Internet Technology and Secured Transactions, (ICITST)*, 1–6. https://doi.org/10.1109/ICITST.2009.5402501

Toom, K., & Miller, P. F. (2018). Ethics and Integrity. In *The European Research Management Handbook* (pp. 263–287). Elsevier. https://doi.org/10.1016/B978-0-12-805059-0.00013-4

Varanda, A., Santos, L., Costa, R. L. de C., Oliveira, A., & Rabadão, C. (2021). Log pseudonymization: Privacy maintenance in practice. *Journal of Information Security and Applications*, *63*, 103021. https://doi.org/10.1016/j.jisa.2021.103021