Current Status of Data Protection Impact Assessment and AI Human Rights Impact Assessment in Korea[1]

Kim Byung-wook (Attorney, Minbyun Digital Information Committee)

1. Introduction

As the scope of artificial intelligence (AI) expands and its applications deepen, concerns are growing about the negative impacts stemming both from risks inherent in AI systems and from risks derived from their deployment.

AI systems are premised on large-scale data training. However, the process of generating parameters is opaque, the systems possess a degree of autonomy, and there is the potential for bias and errors. Considering the opacity and autonomous nature of AI technology, along with the breadth and ripple effects of its development and use, it is often difficult to provide remedies or impose sanctions after harms have occurred. Therefore, it is necessary to establish a systematic framework that identifies risk factors in advance, before the introduction of AI technologies, and that mitigates or eliminates such risks proactively.

Since AI technologies affect not only personal data protection but also a wide range of fundamental rights, there is an urgent need to introduce a comprehensive impact assessment system—namely, an AI Human Rights Impact Assessment. At the same time, it is important for specialized assessments in areas such as personal data protection to effectively identify and control the potential negative impacts and risks posed by AI technologies, ensuring both independence and expertise.

However, when examining the current status and substance of Korea's human rights impact assessment system for artificial intelligence, as well as the status of its data protection impact assessment system, it becomes evident that they are not effectively performing the functions of identifying the new threats and negative impacts that AI technologies pose to human rights—including personal data protection—nor are they adequately controlling such risks.

In the following sections, we will review the current status of Korea's AI Human Rights Impact Assessment and Data Protection Impact Assessment systems.

2. Current Status of AI Human Rights Impact Assessments in Korea

1) The Need for AI HRIA

AI technologies inherently carry the potential to negatively affect a wide range of fundamental rights. AI systems hold the possibility of infringing upon nearly all rights on the

[1] This presentation paper includes portions of the content from Research on the Introduction of AI Human Rights Impact Assessments by Yoo Seung-ik et al., National Human Rights Commission of Korea (2022).

list of human and fundamental rights.[2] For example, an AI algorithm may learn from biased data and make discriminatory decisions, or an AI system may collect personal data without authorization and use it independently of the data subject's control. There is also the potential for AI technologies to threaten individuals' lives, physical safety, and the protection of privacy.

In response to the risks of AI, many countries around the world are preparing various countermeasures, one of the most significant being impact assessment systems, including human rights impact assessments.

In particular, the European Union, following a risk-based approach, categorizes AI systems into unacceptable risk, high risk, and limited risk. For high-risk AI systems, it imposes strong obligations, including the requirement to conduct fundamental rights impact assessments (Article 27 of the AI Act). In addition, several countries have already implemented or are preparing to introduce AI-related impact assessments or human rights impact assessments—for example, Canada's Algorithmic Impact Assessment, the Netherlands' Fundamental Rights and Algorithms Impact Assessment, and Denmark's Human Rights Impact Assessment.

Meanwhile, United Nations human rights bodies have recommended the implementation of human rights due diligence to identify, prevent, and mitigate the negative impacts of AI and other new technologies on human rights. They have consistently emphasized that human rights impact assessments are a valuable tool within human rights due diligence, helping to identify and address potential adverse effects on human rights.

The UN High Commissioner for Human Rights has emphasized that many inferences and predictions made by artificial intelligence significantly affect the enjoyment of the right to privacy, while also raising serious concerns for other rights such as freedom of thought and opinion, freedom of expression, and the right to a fair trial. Accordingly, the High Commissioner recommended that systematic human rights due diligence be carried out across the entire life cycle of AI systems—from design, development, deployment, sale, and procurement to operation—identifying HRIAs as a core component of such due diligence (Right to Privacy in the Digital Age, 2021).

On May 11, 2022, the National Human Rights Commission of Korea established the "Guidelines on the Development and Use of Artificial Intelligence." The Commission recommended to the Prime Minister and relevant ministers that AI-related policies be formulated and implemented, and relevant laws enacted or revised, on the basis of these guidelines. The guidelines specifically provide for the implementation of AI HRIAs.

According to the guidelines, AI HRIAs should take into account the potential and degree of human rights violations and discrimination, the number of individuals affected, and the volume of data used. The content of such assessments must include principles and elements presented in the human rights guidelines, international human rights standards,

[2] Yoo Seung-ik, The Impact of Artificial Intelligence on Human Rights and Democracy and Issues of the AI Bill, Proceedings of the National Assembly Forum (Why Are the EU and the United States Seeking to Regulate AI?), July 2023.

and obligations prescribed by relevant laws, all while considering the characteristics, context, scope, and purpose of the AI in question. In addition, the assessments are required to identify and analyze risk factors for human rights violations and to propose necessary improvements.

Furthermore, the guidelines stipulate that when the results of an AI HRIA reveal negative effects on human rights, biases, or risks, measures must be developed and applied to prevent or mitigate such impacts, and, in principle, the content of these measures should be made public.

In order to effectively identify the risks of AI and to mitigate its risks and negative impacts in advance, it is necessary to introduce HRIAs for AI technologies and systems as part of a comprehensive normative framework.

2) Current Status of AI HRIAs in Korea

a. Existing Impact Assessment Systems

For impact assessments to be stably implemented, with clear procedures and binding force through the imposition of obligations, a legal basis is required. However, as of now, there is no legislated impact assessment system in Korea that independently designates AI technologies or services as its exclusive subject.

The Framework Act on the Promotion of AI Development and the Establishment of a Trust-Based Foundation (hereinafter, the “AI Framework Act”), which is scheduled to take effect in January 2026, refers to the fundamental rights impact assessment of high-impact AI. Nevertheless, the Act does not mandate its introduction, stipulating only a duty of effort (Article 35 of the AI Framework Act).

That said, among related impact assessment systems, there are certain mechanisms through which services or businesses utilizing AI technologies may, to some extent, be subject to assessment. Yet, even in such cases, it is difficult to conclude that these systems adequately perform—or are capable of performing—the function of identifying negative impacts or risks to fundamental rights, including the right to informational self-determination, or of preventing such risks in advance.

In this regard, particular attention is often drawn to the Social Impact Assessment under the Framework Act on Intelligent Informatization and the Technology Impact Assessment under the Framework Act on Science and Technology, which will be examined below.

a-1. Social Impact Assessment of Intelligent Information Services

Article 56 of the Framework Act on Intelligent Informatization provides for a “social impact assessment,” which may investigate and evaluate “the impacts of the use and dissemination of intelligent information services, etc., that have a significant effect on the lives of the people, on society, economy, culture, and everyday life.”

Under the Act, “intelligent information services, etc.” include intelligent information technologies, defined as technologies that implement learning, reasoning, and decision-making through electronic means (Article 2, subparagraph 4(a) of the Act). This definition is generally understood to encompass services utilizing AI technologies; therefore, the social impact assessment under the Act may be applied to AI technologies.

The evaluation items for the social impact assessment under the Act include “safety and reliability” and “impacts on information protection” (Article 56 of the Act).

> Article 56(1) of the Act
>
> 1. Safety and reliability of intelligent information services, etc.;
> 2. Impacts on the information culture, such as closing the digital divide, protection of privacy and ethics for the intelligent information society;
> 3. Impacts on the society and the economy, such as employment, labor, fair trade, industrial structure, rights and interest of users, etc.;
> 4. Impacts on information protection;
> 5. Other impacts of intelligent information services, etc. on the society, economy, culture and citizens’ daily lives.

However, the social impact assessment of intelligent information services is closer to an ex post evaluation system aimed at securing social acceptability and normative legitimacy of AI services, and it evaluates the impacts of such services on society at a very broad level.

The entities responsible for conducting the assessment are limited to the state (the Minister of Science and ICT) and local governments. Moreover, as there are no detailed provisions in subordinate regulations, such as enforcement decrees, regarding the scope and procedures of the assessment, the process remains fluid, and there are no mechanisms to enforce the assessment.

In addition, the evaluation item of “information protection,” as prescribed in the Act, differs in purpose and aim from “personal data protection,” which is intended to faithfully guarantee the constitutional right to informational self-determination. The Act defines “information protection” as “administrative and technical measures (referred to as ‘information protection systems’) designed to prevent the damage, alteration, or leakage of information that may occur during its collection, processing, storage, retrieval, transmission, or reception.” This definition aligns more closely with the concept of “information security” and is considerably different from “personal data protection,” which centers on reflecting the will of the data subject in controlling and managing data.[3]

a-2. Technology Impact Assessment under the Framework Act on Science and Technology

The social impact assessment system under the Framework Act on Intelligent Informatization stipulates that assessments of intelligent information technologies are to be

[3] Seoul Economic Daily, May 17, 2023. “[Rotary] Information Protection and Personal Data Protection.” https://www.sedaily.com/NewsView/29PM4TS3JV. Accessed September 12, 2025.

substituted by the technology impact assessment under the Framework Act on Science and Technology (Article 56 of the Framework Act on Intelligent Informatization). However, it is difficult to regard the technology impact assessment as a system that can effectively identify and mitigate the negative impacts of AI on human rights, including personal data protection.

According to Article 14 of the Framework Act on Science and Technology and Article 23 of its Enforcement Decree, the subjects of technology impact assessments are defined as "technologies with significant technological, economic, and social impacts and ripple effects in the future, as determined by the Minister of Science and ICT in consultation with the heads of relevant central administrative agencies." The Act and its Enforcement Decree stipulate that the following items must be included in the evaluation criteria (Article 14 of the Act; Article 23 of the Enforcement Decree).

Article 23(2) of the Enforcement Decree
1. Impact of the relevant technology on the enhancement of benefits to citizens and on the development of relevant industries; 2. Impact of the new science and technology on the economy, society, culture, ethics and the environment; 3. Measures to prevent adverse effects of the relevant technology, where the relevant technology has any potential adverse effect; 4. Impact of the nature and ripple effects of the relevant technology on characteristics, such as gender.

Since its introduction around 2003, the technology impact assessment has been conducted almost annually under the leadership of the Ministry of Science and ICT. Artificial intelligence was included among the target technologies in 2015, followed by virtual and augmented reality in 2016, Level 4 and above autonomous vehicles in 2021, and "safe and trustworthy AI technologies" in 2024.

However, the technology impact assessment is conducted only once a year, led by the Ministry of Science and ICT, and is rarely utilized by other ministries. Moreover, rather than evaluating the impacts of emerging technologies such as AI from a human rights perspective, the assessment is designed to broadly examine their effects on the economy, society, culture, ethics, and the environment, with the aim of deriving policy recommendations and guiding desirable policy directions. While it is a positive aspect that the results of the assessment are formally required to be reflected in policy, it has been pointed out that the process risks being reduced to a mere administrative formality without an in-depth process of public deliberation.

b. AI Voluntary Self-Assessment Standards

Although not established in a legislated form, various standards and tools have been proposed domestically as voluntary instruments or guidelines that can be utilized.

These tools are limited in that they lack binding force with respect to implementation, procedures, or the incorporation of results. Nevertheless, they are noteworthy in terms of

their awareness of the risks associated with AI and their frameworks for identifying such risks, and thus may serve as useful references.

b-1. AI Ethics Impact Assessment Framework by the Ministry of Science and ICT

In February 2024, the Ministry of Science and ICT and the Korea Information Society Development Institute announced the AI Ethics Impact Assessment Framework, setting out the following objectives:

- To support companies' voluntary efforts to practice AI ethics and trustworthiness, and to provide standards for users to utilize AI in an ethical and responsible manner.
- To evaluate the ethical impacts of AI products and services in advance, thereby deriving insights for management, institutional, and policy measures aimed at maximizing positive impacts and minimizing negative ones.
- To provide reference materials for companies, civil society, academia, and the public sector (government) to systematically understand the ethical impacts of AI, and to encourage the development, deployment, and utilization of AI products and services in more ethical ways.

The AI Ethics Impact Assessment Framework presents individual measurement items divided into ten areas. These ten areas are named similarly to the ten core requirements of the National AI Ethics Standards announced in 2020 by the Ministry of Science and ICT and the Korea Information Society Development Institute: ① Guarantee of human rights, ② Protection of privacy, ③ Respect for diversity, ④ Prohibition of harm ⑤ Public interest, ⑥ Solidarity, ⑦ Data management, ⑧ Accountability, ⑨ Safety,, ⑩ Transparency

However, the implementation of the AI Ethics Impact Assessment Framework is limited to the government. While it is a positive step that the framework incorporates participation from diverse stakeholders—such as industry, government, academia, research institutes, evaluation committees, and civil society—in order to ensure expertise, fairness, objectivity, reliability, and transparency, there remains concern that this participation could amount to little more than a formal process of collecting opinions.

For the framework to provide a substantive evaluation of the impacts of AI technologies or services, access to information about the development process—such as the operating principles of the service and the sources of training data—is also necessary. Yet, the framework offers no explanation as to whether the companies developing or deploying the AI in question will be involved in the impact assessment.

Even when examining the measurement items in each area, the framework poses abstract questions such as whether "a particular AI service poses a risk of undermining human dignity and value," asking respondents to indicate their degree of agreement or disagreement on a quantitative scale. As a result, the framework's ability to effectively identify the negative impacts of AI or the risks of human rights violations, and to propose improvements, appears to be very limited.

In other words, the meaning of “ethics” within the ethical impact assessment is itself ambiguous. It seems to function merely as a device that is much looser and easier to comply with than laws or policies.[4] Therefore, the Ethical Impact Assessment Framework cannot be regarded as an appropriate tool for effectively identifying impacts on human rights or for improving the risks posed by AI.

b-2. AI Human Rights Impact Assessment Tool of the National Human Rights Commission of Korea

In May 2022, the National Human Rights Commission of Korea (NHRCK) announced the Guidelines on the Development and Use of Artificial Intelligence. Building on this, in May 2024, the Commission introduced an AI Human Rights Impact Assessment Tool, recommending to the Minister of Science and ICT that AI human rights impact assessments be introduced for all AI systems applied in both the public and private sectors.

Although this tool has not yet been institutionalized by law and therefore lacks binding force, it can nonetheless be evaluated as addressing a broad range of considerations regarding the characteristics of AI technologies and their potential impacts on human rights. The specific content of the AI Human Rights Impact Assessment Tool proposed by the NHRCK is as follows.

Scope of Assessment

The AI Human Rights Impact Assessment Tool covers all AI systems adopted by public institutions and high-risk AI systems adopted by the private sector. Since AI developed in the public sector tends to have a broader impact compared to that in the private sector, the framework designates all AI systems in the public sector as subject to assessment, while limiting the scope in the private sector to high-risk AI systems.

This approach reflects a differentiated application of requirements according to the level of risk posed by AI systems. AI deemed “prohibited”—those where risks cannot fundamentally be mitigated or eliminated—are excluded from the scope of assessment. For high-risk AI systems, however, the tool requires the implementation of human rights impact assessments. That said, the tool does not clearly delineate the specific scope of what constitutes “high-risk AI.”

Timing of Assessment

AI human rights impact assessments are, in principle, to be conducted prior to the development of the system as a preliminary assessment. However, risks must also be continuously managed through regular and ex post assessments.

This reflects the understanding that, since the underlying technologies of AI continue to evolve, and since the risks of the same technology may vary depending on the geopolitical,

[4] The concept of “ethics” is being used as a means to avoid the application of codified laws or regulations, while treating the development and dissemination of AI technology as an absolute good (Heo Yuseon et al., Why Ethics?: A View on Contemporary Discussions of AI Ethics, Their Characteristics and Limitations, March 2020)

social, and economic context in which it is deployed, a preliminary assessment alone is insufficient to fully prevent or manage potential human rights violations and risks.

Assessment Entity

The AI Human Rights Impact Assessment Tool states that it is desirable for the assessment to be conducted either by an internal organization independent from the development entity and relevant business department, or by a third-party institution possessing independence and expertise in both human rights and AI technologies. This recommendation is based on the concern that, when assessments are carried out directly by the developers of the AI system or by the implementing business entity, it is difficult to ensure objectivity and neutrality, and the process risks becoming a mere formality.

However, even when an internal organization independent from the business department or an external third-party body with expertise and independence is tasked with the assessment, there remains the possibility that it may lack accurate information or in-depth understanding of the AI technology or service in question. Thus, this issue should not be framed as a simple either-or choice, but rather addressed by considering the overall design of the system and its circumstances, with appropriate complementary measures.

For instance, if the development or business department conducts the assessment directly, objectivity and neutrality may be enhanced by requiring disclosure of the assessment report or by instituting follow-up review procedures. Conversely, if the assessment is carried out by a third-party institution or an independent internal body, the system should be designed to ensure that sufficient information is smoothly provided by the development or business department.

Assessment Procedure

The AI Human Rights Impact Assessment is structured to proceed in the following four stages:
Planning and Preparation Stage – establishing the assessment plan and making preliminary preparations (Stage 1) → Analysis and Evaluation Stage – analyzing and assessing potential negative impacts (Stage 2) → Improvement and Remedy Stage – identifying measures for prevention, mitigation, and remedy of the risks identified (Stage 3) → Disclosure and Review Stage – ensuring transparency and reviewing the overall assessment (Stage 4).

One of the central features of the procedure is the emphasis on stakeholder participation at every stage. The tool designates stakeholder involvement—particularly the participation of those individuals whose human rights are at risk of being infringed—as a core principle, stressing that consultations must be conducted in as diverse and inclusive a manner as possible.

Reflecting this principle, procedures for collecting opinions and consulting stakeholders are incorporated across all stages. For example, in Stage 1, consultations with stakeholders are to be used to gather information and identify the specific human rights that may be at risk. Similarly, in Stage 3 (Improvement and Remedy), stakeholders' input must be sought to

develop measures to prevent and mitigate human rights violations and to secure remedies for affected parties.

Composition of Assessment Items

The assessment items are organized into four stages, reflecting the structure of the assessment procedure.

As shown below, Stage 1 includes questions designed to gather descriptions of the AI system subject to assessment, to identify stakeholders, to understand the temporal and spatial context in which the AI system will be introduced, and to obtain materials related to consultations with stakeholders.

Q1-2-3. From the stakeholders previously identified, have you collected or discussed their opinions regarding the potential human rights impacts of the AI system, and have you documented them?

☐ Yes ☐ Needs supplementation ☐ No ☐ No information available ☐ Not applicable

Explanation: (_____)

Q1-2-4. When collecting or discussing stakeholders' opinions, the following information must be included:

- Name, affiliation, and contact information of the stakeholder(s) consulted
- Date(s) of consultation
- Materials provided to stakeholders regarding the AI system
- Stakeholders' opinions about the AI system

In Stage 2, the assessment includes questions related to the analysis and identification of impacts concerning personal data protection and data management, algorithmic reliability, non-discrimination, explainability and transparency, the degree of automation and human involvement, security, accessibility, and licensing. At this stage, the tool also lists individual human rights that may be at risk of infringement, and includes questions designed to assess the severity of potential impacts using scales that measure their scope and magnitude.

In Stage 3, the assessment examines measures that can prevent or mitigate human rights violations, as well as remedies for infringements. Stage 4 consists of questions regarding the disclosure and review of the assessment report.

c. Summary

As examined above, while some legislated impact assessment systems are technically capable of evaluating the impacts of AI technologies, they have limitations in effectively

identifying negative impacts or risks to human rights, devising improvement measures, or exercising control over such risks.

Similarly, the various voluntary self-assessment tools that have been proposed are limited in that they lack legal grounds and therefore binding force. However, in the case of the AI Human Rights Impact Assessment Tool of the National Human Rights Commission of Korea, if it is actually applied in practice, experiences are accumulated, and shortcomings are addressed, it has the potential to evolve into a substantive tool for identifying, mitigating, and preventing the negative impacts and risks of AI technologies and systems.

3. Current Status of Data Protection Impact Assessments in Korea

1) The Need for Effective DPIAs of AI Systems

While the negative impacts and risks of AI on human rights are being discussed from multiple perspectives, one of the most clearly visible risks is the infringement of the right to informational self-determination.

AI technologies are premised on the collection of large volumes of data, including training and learning datasets. Such data often contain various forms of personal information and do not distinguish between structured and unstructured data, including facial recognition data and sensitive information. The collection of this data should, in principle, take place within a lawful framework, such as through notice or consent from the data subject. However, in the process of AI development, there is often indiscriminate and excessive collection—for example, through the scraping of publicly available information.

De-identification measures applied to large-scale datasets for the purpose of personal data protection may prove meaningless. It can be relatively easy and low-cost to re-identify individuals or to infer sensitive attributes from vast datasets. Moreover, companies utilizing generative AI indiscriminately collect not only the information they have gathered directly but also various data available online, thereby generating new personal information.[5] Inferences based on data regarding gender, race, or age can exacerbate the risks of prejudice and discrimination.

The opacity of data processing makes it impossible for data subjects to exercise their rights. Because identifying individuals within large datasets—including unstructured data—is technically difficult and costly, the rights of data subjects, such as access, rectification, erasure, restriction of processing, and destruction of data, are not substantively guaranteed.

When the results generated through generative AI include personal attributes or personal information—particularly when such information is used in critical areas such as hiring decisions, promotions, or determinations of eligibility for social welfare benefits—the risks of personal data infringement, along with the impacts on individual human rights, can be extremely serious.

---

[5] Jang Jae-young, Threats of Generative AI and Measures for Protecting the Right to Control Personal Data, SW Policy Research Institute, April 2025

Therefore, from the perspective of data protection, there is a need to establish a separate Data Protection Impact Assessment (DPIA) system as part of a comprehensive regulatory framework, in order to effectively identify in advance the potential negative impacts and risks of AI systems and to exercise control over those risks.

While the assessment items or indicators for evaluating the negative impacts or risks of AI systems on data protection may overlap with those of human rights impact assessments, they cannot fully converge, as each domain has its own independent scope. If DPIAs are operated in a complementary manner alongside human rights impact assessments, this should not be considered as imposing redundant regulatory burdens on companies, individuals, or institutions subject to assessment.

The European Union, for example, provides that where the same or similar assessment has been conducted under the General Data Protection Regulation (GDPR) through a DPIA, the fundamental rights impact assessment should function in a supplementary manner. Likewise, Denmark's human rights impact assessment points out that it cannot fully converge with a DPIA.

2) Current Status and Challenges of DPIAs in Korea

The DPIA system was introduced in 2011 alongside the enactment of the Personal Information Protection Act(PIPA). It has been implemented in the public sector in the form of preliminary assessments when operating personal data files above a certain scale. However, it has not been effectively fulfilling a preventive function of identifying risk factors that pose actual threats to personal data protection or of deriving improvement measures.

In response to the new threats posed by artificial intelligence, there is a need to strengthen and substantiate the DPIA system. The following sections will examine in greater detail the current shortcomings of the existing DPIA framework.

a. Mandatory Subjects (Public Institutions) and Quantitative Threshold Issues

Under the current DPIA system, the obligation to conduct assessments is limited to public institutions, while for private entities and other personal data controllers, the implementation of DPIAs is left to their discretion (Article 33(11) of the PIPA). It is understandable, to some extent, that heavier obligations are imposed on public institutions, given the broader impact when they process personal data using AI and the heightened accountability required in the public domain.

However, the scope of mandatory assessments should be based primarily on the risk of personal data infringement. Even in the private sector, there are serious risks to data protection when high-risk AI systems are developed, introduced, or utilized. Considering that the development and deployment of AI systems largely take place in the private sector, there can be little disagreement on the necessity of requiring private entities to conduct mandatory DPIAs in order to identify risk factors in advance and to mitigate or eliminate them.

Furthermore, the PIPA stipulates mandatory DPIAs based on quantitative criteria tied to the number of data subjects. Specifically, DPIAs are required when building, operating, or

modifying personal data files involving the processing of sensitive information or unique identifying information of 50,000 or more individuals; when linking personal data files results in files containing the personal data of 500,000 or more individuals; and when modifying operational systems of personal data files (such as search systems) after a DPIA has already been conducted (Article 35 of the Enforcement Decree of the PIPA).

However, in the case of AI systems, even when the number of personal data processed is relatively small, the severity of negative impacts can be significant depending on the field of application, the type of data being processed, and the way the system operates.

Therefore, AI systems should be included as mandatory subjects of DPIAs under a risk-based approach, particularly in fields such as real-time biometric data collection for identification in public spaces, or use in employment-related decisions such as recruitment and promotion.

b. Issues Concerning Considerations, Evaluation Criteria, and Assessment Items in DPIAs

Under the PIPA, the factors to be considered when conducting a DPIA include: the number of personal data processed, whether personal data is provided to third parties, the likelihood and degree of risk of infringing the rights of data subjects, whether sensitive information or unique identifying information is processed, and the retention period of personal data (Article 33(3) of the PIPA and Article 37 of its Enforcement Decree).

In addition, the Enforcement Decree of the PIPA sets out the following four evaluation criteria (Article 38(1) of the Enforcement Decree of the PIPA):

1. The type and nature of personal information contained in the relevant personal information files, the number of data subjects, and the possibility of subsequent personal information breach;
2. The level of measures to ensure safety taken under Articles 23 (2), 24 (3), 24-2 (2), 25 (6) (including cases applied mutatis mutandis in Article 25-2 (4)), and 29 of the Act, and the subsequent possibility of personal information breach;
3. Countermeasures against risk factors of personal information breach, if any;
4. Other necessary measures subject to the Act or this Decree, or any factor affecting breach of duties.

The Notification on DPIAs further specifies the evaluation areas and domains under the evaluation criteria. However, the previous evaluation framework (areas and domains) did not take into account any of the new or potential negative impacts and risks that may arise in the course of developing or utilizing AI systems.

According to the revised Notification, which entered into force on September 5, 2025, “Artificial Intelligence” has been added as an evaluation domain. This domain is subdivided into “AI system learning and development” and “AI system operation and management,” and

the accompanying guidelines provide more detailed assessment items under these subcategories. Nevertheless, this revision cannot be regarded as sufficient in itself.[6]

Evaluation Domain (Notification): Artificial Intelligence (AI)

- **Subdomain: AI System Learning and Development**
    - Ensure adequacy of personal data processing
    - Prevent collection of unnecessary sensitive information
    - Clarify retention periods for training data
    - Minimize personal data leakage through AI vulnerability attacks
- **Subdomain: AI System Operation and Management**
    - Clarify responsibilities among entities involved in AI development and operation
    - Ensure transparency in personal data processing
    - Provide guidelines on the permissible use of generative AI services
    - Implement safeguards to protect the rights of data subjects

Even with the newly added items, the DPIA framework still fails to take into account the risks of personal data infringements that may arise or be amplified depending on the specific context or field of application of AI systems. It also does not sufficiently address requirements such as explainability and transparency that are integral to AI systems. Moreover, the guarantee of data subjects' rights is treated only at a formal level, limited to asking whether plans and measures have been established and implemented.

In addition, safety measures and remedies specifically tailored to the inherent risks of personal data infringements in AI systems should also be incorporated into the assessment items, but these too are absent, which represents another limitation of the current framework.

c. Lack of Stakeholder (Affected Parties) Participation

One of the most important elements not only in establishing systems to respond to AI-related risks but also in ensuring the substantive effectiveness of impact assessments is stakeholder participation. The AI Human Rights Impact Assessment Tool disseminated by the National Human Rights Commission of Korea designates stakeholder involvement as a core principle throughout the entire assessment procedure. The purpose is to promote objectivity and

[6] Boan News, September 4, 2025, "Public Institutions to Strengthen Personal Data Protection When Using AI". Accessed September 12, 2025.

fairness in the assessment and to prevent results skewed toward the interests of only one party.

However, under the current DPIA framework, there is no consideration of, nor any provisions for, the participation or consultation of affected stakeholders—namely, the data subjects whose personal data may be impacted by AI systems. This represents a significant limitation of the current system.

d. Disclosure and Verification of Assessment Reports

For the credibility of impact assessments, disclosure and verification procedures are essential. The PIPA stipulates that when a public institution conducts a DPIA, the results must be submitted to the Personal Information Protection Commission (PIPC) (Article 33(1) of the PIPA). Furthermore, when registering personal data files, the results of the DPIA must be attached (Article 33(5) of the PIPA). Following the amendment of the law on September 15, 2023, it is now possible to publish a summary version of the DPIA results (Article 38 of the Enforcement Decree of the PIPA and Article 12-2 of the Notification).

However, in the private sector, where DPIAs are conducted, there are no provisions at all regarding the disclosure or verification of assessment results.

As noted earlier, in cases where AI systems are deployed in the private sector—particularly with respect to high-risk AI systems—it is necessary to mandate the conduct of DPIAs. In such cases, obligations should also be imposed to submit the results to the PIPC and to disclose at least a summary of the findings.

e. Timing of Impact Assessments

AI technologies are continuously evolving at the foundational level, and even the same technology may present different risks depending on the geopolitical, social, and economic contexts in which it is deployed. Accordingly, risk analysis and management should not only take place at the stage when an AI system is first introduced into a project but should also be conducted whenever there are significant modifications to the system itself or substantial changes in its scope of application.

However, under the current PIPA, DPIAs are only required prior to the establishment or operation of an AI system involving the processing of personal data, or when an existing AI system is modified (Article 33 of the PIPA). This structure makes continuous risk management impossible.

f. Weak Enforcement of Obligations to Implement Results

For an impact assessment system to be truly effective, it must go beyond the mere preparation of an assessment report. There must be mechanisms in place to ensure the implementation of improvement measures derived from the assessment. Once the negative impacts or risks of an AI system are identified through a DPIA, and measures to mitigate or eliminate them are proposed, these must be incorporated into the design, development, and deployment processes of the AI system.

However, under the current PIPA, the PIPC is only authorized to provide opinions on the results of DPIAs, and even this is left to its discretion (Article 33(4) of the PIPA). Even when opinions are provided, there are no means to ensure their implementation.

The PIPC recently amended its Notification in November 2024, shortening the deadline for submitting implementation results and plans concerning improvement measures identified in DPIAs from one year to two months. Nevertheless, this remains insufficient to guarantee compliance. To address this gap, the obligation to reflect improvement measures derived from DPIAs should be clearly stipulated in law, and specific sanctions for non-compliance should also be introduced.

3) Summary

AI technologies pose significant threats to fundamental rights, including the protection of personal data. However, despite the widespread use of AI systems that involve the processing of personal data, the current DPIA system has fallen far short of fulfilling its intended function of risk prevention and control.

In order to respond effectively to the new threats posed by AI, a comprehensive overhaul of the DPIA framework is necessary so that it can serve a genuine preventive function in identifying and mitigating negative impacts and risks in advance.

4. Conclusion

A systematic framework is essential for managing and controlling the risks and negative impacts of AI. The DPIA system must be strengthened so that, together with the AI Human Rights Impact Assessment, it can serve as an effective component of a comprehensive regulatory framework for controlling and managing the risks posed by AI.

References

Yoo, Seung-ik et al. A Study on the Introduction of AI Human Rights Impact Assessments. National Human Rights Commission of Korea, 2022.

Yoo, Seung-ik. The Impact of Artificial Intelligence on Human Rights and Democracy and Issues of the AI Bill. Proceedings of the National Assembly Forum (Why Are the EU and the United States Seeking to Regulate AI?), July 2023.

Lee, Ho-jung et al. A Study on the Improvement of Personal Information Protection Legislation in Accordance with International Human Rights Standards such as the EU GDPR. 2020.

Heo, Yu-seon et al. Why Ethics?: An Overview of Contemporary AI Ethics Discussions, Their Characteristics and Limitations. March 2020.

Jang, Jae-young. Threats of Generative AI and Measures for Protecting the Right to Control Personal Data. SW Policy Research Institute, April 2025.

Guidebook to the AI Human Rights Impact Assessment Tool. National Human Rights Commission of Korea, December 2024.

AI Ethics Impact Assessment Framework (2024). Ministry of Science and ICT & Korea Information Society Development Institute, February 2024.