

Principles-Code for Protection of Intellectual Property Protection and Transparency for the Appropriate Use of Generative AI

1. Overview

(1) Basic Concepts (Purpose)

This document establishes principles for ensuring transparency and protecting intellectual property rights that should be implemented by generative AI businesses, drawing upon the spirit of the "Act on the Promotion of Research and Development and Practical Application of Artificial Intelligence-Related Technologies" (Act No. 53 of Reiwa 7) while referencing initiatives such as the Stewardship Codes in the field of corporate governance (comply-or-explain approach), thereby aiming to create a balance between the promotion of advancements in generative AI technology with the appropriate protection of intellectual property rights, in order to creating a safe and secure environment for rights holders and users alike.

(2) Applicable Scope of This Document

a. Definitions

While recognizing that it should not be discouraged that parties involved in generative AI may voluntarily and independently negotiate terms, this document applies to "generative AI developers" and "generative AI providers" (collectively referred to as "generative AI businesses") as follows.

- "Generative AI" refers to any AI system based on AI models capable of generating text, images, programs, or other content.
- "Generative AI developer" means an individual or entity responsible for constructing generative AI systems, who has substantially provided these systems to the public (whether to unspecified individuals or specific groups; the same applies below), regardless of purpose or whether it is a corporation or individual.
 - The following individuals are in principle excluded from the definition of "generative AI developer," unless it can be reasonably determined that they substantially provided these systems to the public:
 - Individuals who, upon commission from a generative AI developer, solely handle partial aspects of the system development process—such as developing generative AI

models/algorithms, data collection (including purchase), preprocessing, generative AI model training, and validation—before delivering these components to the developer or providing related services (regardless of purpose or whether it is a corporation or individual).

- "Generative AI provider" means an individual or entity that has provided services incorporating generative AI systems into applications, products, existing systems, or business processes (hereinafter referred to as "generative AI services") to the public.
 - The following individuals are in principle excluded from the definition of "generative AI provider," unless it can be reasonably determined that they substantially provided these generative AI services to the public:
 - Individuals who, upon commission from a generative AI businesses, solely handle partial aspects of the service—such as generative AI system verification, integration with other systems, provision of generative AI systems or generative AI services, or user-side operational support for normal system operation—before delivering these components to the businesses or providing related services (regardless of purpose or whether it is a corporation or individual).
- Individuals who conduct research and development but do not provide services to the public do not qualify as generative AI businesses.

b. Cases Where Generative AI Systems or Services Are Developed or Provided Exclusively to Data Providers or Their Designated Limited Circles

- For clarity, it should be noted that an individual or corporation that develops specialized generative AI systems using only their own data and exclusively provides these systems to said data providers or their designated limited circles does not fall under the definition of "generative AI developer" in this document. Similarly, an individual or corporation that provides generative AI systems or services equipped with specialized generative AI systems developed using only data provided by a specific corporate group or organization—data exclusively obtained from corporations or individuals belonging to that group or organization—does not qualify as a "generative AI developer" in this document (provided that the corporations or individuals providing the data have granted permission for its use).
- Additionally, individuals or corporations that provide generative AI services equipped with specialized generative AI systems developed using only their own data exclusively to said data providers or their designated limited circles do not qualify as "generative AI providers" in this

Note: This is a provisional translation, and the Cabinet Office is not responsible for its content. For official information, please refer to the Japanese Version.

document. Similarly, individuals or corporations that provide generative AI services equipped with specialized generative AI systems developed using only data provided by a specific corporate group or organization exclusively to said group or its designated limited circles do not qualify as "generative AI providers" in this document (provided that the corporations or individuals providing the data have granted permission for its use).

[Examples]

For instance, the following cases typically do not qualify as "generative AI developers" or "generative AI providers" under this document:

- ① In cases where Corporation B, licensed from Corporation A that provides general-purpose generative AI systems or services to the public, develops specialized generative AI systems or services tailored to Corporation C's internal operations using data provided by Corporation C, and provides these to Corporation C ("Corporation B") (provided that the generative AI system or service is configured to exclusively generate content based on data provided by Corporation C).
- ② In cases where Corporation C, in the same scenario as ①, exclusively provides said generative AI system or service (e.g., chatbot) to Corporation C's customers ("Corporation C").

c. Cases Where Generative AI Systems or Services Are Developed or Provided Specializing in Specific Industries

- Unless falling under case b. above (cases involving generative AI systems or services developed or provided exclusively to data providers or their designated limited circles), individuals or corporations that develop or provide generative AI systems or services specialized for specific industries (defined as social groups comprising entities engaged in the same industry; the same applies below) may also qualify as generative AI businesses.

[Examples]

- When Corporation E, licensed by Corporation D (which provides generic generative AI systems or services to the public), develops specialized generative AI systems or services for industry X based on these generic systems or services, and then provides them to corporations F, G, and H within industry X ("Corporation E").
- However, with regard to each principle, only the additional modifications made by Corporation E to the generic generative AI systems or services are subject to compliance requirements as generative AI businesses. The generic systems or services themselves are assumed to maintain intellectual property protection and transparency through Corporation D's compliance to these principles.

- However, this document does not apply to entities as for the scope of generative AI businesses that develop or provide generative AI systems or services where the risk of generating content that infringes third-party rights is extremely low.

[Examples]

- Those who develop or provide generative AI systems or services that produce the following types of content at an extremely high frequency shall not be considered "generative AI service businesses" under this document:
 - Content that cannot be directly perceived as the original creative expression of existing works
 - Content that consists solely of statistical data, inference results, or other information useful for decision-making
- When developing industry-specific generative AI systems or services using only data provided by corporations I, J, K, and L within industry X, and then providing them to corporations I-L, such entities shall not be considered "generative AI service businesses" (provided that the generative AI system or service is specifically configured to generate content based solely on the data provided by corporations I-L).

D. Application to entities without headquarter or principal office in Japan

- Even generative AI service businesses without headquarter or main office in Japan are subject to this document if their generative AI systems or services are provided to Japan (including cases where these services are available to Japanese citizens).

(3) Methodology Adopted by This Document

This document has been formulated while considering the circumstances of generative AI service businesses, users, and rights holders, as well as their respective intentions. It does not require mandatory disclosure of sensitive information belonging to generative AI service businesses (including trade secrets and information related to safety and security), nor does it constitute legally binding norms. We expect that generative AI service businesses who understand and accept the spirit of this document will implement measures accordingly.

Furthermore, this document requires compliance with or explanation regarding the following principles through the "comply or explain" approach.

The "comply or explain" method requires either implementing the principles or providing reasons for not implementing them. In other words, if a generative AI service provider considers any principle below to be inappropriate to implement in light of their specific circumstances, they are permitted to refrain from implementing that principle by providing a thorough explanation of the reasons for noncompliance. However, of course, it is essential that generative AI service businesses carefully structure their explanations to ensure full understanding by both generative AI users and rights holders.

Additionally, actively explaining one's specific initiatives while complying with principles may also be beneficial for gaining full understanding from generative AI users and rights holders.

(4) Visualization of Acceptance Status of this document

To visualize the acceptance status of this document, we expect the following from generative AI businesses who accept the principles outlined below:

- Publicize the following information on their corporate websites (which serve as official channels for disseminating information about the businesses' overview, business content, product information, etc.) or other websites with equivalent functionality (collectively referred to as "corporate websites, etc."), as well as submit such facts using the prescribed format by the Cabinet Office Intellectual Property Strategy Promotion Office:
 - Statement accepting each principle specified in this document (acceptance declaration)
 - The following details regarding each principle specified in this document:
 - Implementation details of each principle

Note: This is a provisional translation, and the Cabinet Office is not responsible for its content. For official information, please refer to the Japanese Version.

- Explanation of reasons for any principles not implemented
- Regularly review and update each item annually as a general rule (also announce any updates made).

The Cabinet Office Intellectual Property Strategy Promotion Office will create reference forms for the comply or explain approach in accordance with this document, publish a list of generative AI businesses that have submitted their declarations along with links to their corporate websites, etc., and actively encourage submissions from relevant industries in cooperation with related ministries and organizations. However, the Cabinet Office Intellectual Property Strategy Promotion Office will not review the content of these submissions and will not respond to inquiries from third parties.

2. Principles and Exceptions Governed by This Document

(1) Principles that all Generation AI service businesses subject to this document should adhere to

[Principle 1]

Generative AI service providers shall disclose an outline of each of the matters set out in (1) and (2) (hereinafter collectively referred to as the "Outline Disclosure Subject Matter") on their corporate websites (websites that officially publish information about the businesses' overview, business operations, product information, etc., and are accessible to all users) or other equivalent functional websites, making it accessible to all, including rights holders and users. If partial disclosure of the Outline Disclosure Subject Matter is impossible, the service provider must provide an explanation indicating the reasons for such non-disclosure. Furthermore, as stated in Section 1(3), this document does not require mandatory disclosure of sensitive information belonging to the Generation AI businesses, including trade secrets and information related to safety/security.

(1) Measures for Ensuring Transparency

The service provider shall disclose the following matters:

a) Model-related information

- ☐ Name (identifier, version, etc.)
- ☐ Provenance including publication date (past versions, revision history, etc.)
- ☐ Architecture and design specifications (status of licenses contacted with third-party for model development, hardware/software and licenses required for use, etc.)
- ☐ Terms of use (clarification of intended use, restricted/prohibited uses, etc.)
- ☐ Details of model training process (training methods, etc.)

b) Training data-related information

- ☐ Matters related to data used for training and validation (data types including those used in RAG (Retrieval-Augmented Generation), matters related to private datasets obtained through web crawling or third parties, matters related to public datasets, matters related to data collected through other means, and whether synthetic data is used and its purpose, etc.)
- ☐ Crawler (purpose, data collection period, names/identifiers, whether third-party crawlers are used and their names/identifiers, etc.)

c) Accountability-related information

- ☐ Content that can be made traceable to the extent technically feasible and reasonable regarding decisions made during development, provision, or use of Generation AI systems or services (improved traceability, clear identification of responsible parties, allocation of responsibilities among stakeholders, specific responses to stakeholders, documentation, etc.)

(2) Measures for Protecting Intellectual Property Rights

The service provider shall disclose the status of response to the following matters:

- ☐ To ensure proper compliance with intellectual property rights, establish principles for protecting IP rights, clarify the system responsibility, and review these at least annually, publishing the summary of these reviews externally.
- ☐ Regarding data utilization including Generation AI development and training, ensure that these businesses shall not infringe on the intellectual property rights of others.
- ☐ Implement measures such as access restrictions via paywalls or adoption of crawlers that comply with machine-readable instructions like robots.txt. For proper measures by rights holders, publicly disclose these measures by user agent and notify them of any changes.
- ☐ Maintain training logs for a certain period of time¹.
- ☐ Work to avoid crawling so-called pirate sites.
- ☐ Implement technological measures to prevent generation of content that infringes intellectual property rights to the greatest extent possible.
- ☐ Implement technological measures such as digital watermarking, C2PA, and other methods that prove content origin and provenance to the greatest extent possible.
- ☐ Inform Generation AI users that they should refrain from using generated content if they believe it may infringe others' intellectual property rights.
- ☐ Establish appropriate contact channels, including leveraging existing systems when possible, to ensure timely and appropriate relief for rights holders, clarify submission requirements to the greatest extent possible, and preserve records of responses.

¹ Regarding log retention, the Ministry of Internal Affairs and Communications and Ministry of Economy, Trade and Industry's "AI Business Guidelines (1.2 Edition)" (page 19) states: "① Ensuring Verifiability" - "To ensure verifiability of the AI's decision-making process, logs should be recorded and preserved within reasonable limits relative to data volume or content, including: development process of the AI system/service, input/output during operation, learning process, reasoning process, and decision-making rationale." The guidelines further specify: "When recording and preserving logs, consideration should be given to the characteristics of the technology being used and its intended applications, taking into account the importance of the recording methods, frequency, and retention period for purposes such as identifying accident causes, developing preventive measures, and establishing liability for damages."

(Detailed Provisions)

- It is expected that information disclosed through each Generation AI businesses' efforts under Principle 1 will become standardized, deepen mutual understanding among Generation AI businesses, users, and rights holders, and serve as a foundation for proper understanding and voluntary judgment by all parties, thereby contributing to building trustworthy Generation AI. For this reason, when disclosing the summary of information under Principle 1, it is essential to proceed while considering the principles outlined in "1. Overview," identifying the presence and scope of information applicable to each Outline Disclosure Subject Matter, and ensuring compliance with other laws and regulations. Furthermore, this does not preclude Generation AI businesses from proactively disclosing detailed information about Outline Disclosure Subject Matters.
- Regarding the degree of detail to be disclosed for each Outline Disclosure Subject Matter, please refer to the specific examples of Outline Disclosure Subject Matter separately published by the Cabinet Office Intellectual Property Strategy Promotion Office.
- To realize an ecosystem where AI technology advances while properly protecting intellectual property rights, all parties should strive to facilitate orderly dialogue and mutual understanding between rights holders, users, and Generation AI businesses, and should refrain from immediately criticizing a particular Generation AI businesses solely based on their failure to disclose some Outline Disclosure Subject Matter.
- While Principle 1 will be implemented for the time being, to achieve similar objectives, it is expected that stakeholders will make progress on initiatives that take into account international alignment from the perspective of ensuring transparency regarding training data, taking into account future international developments.

【Principle 2】

When a generative AI businesses receives a request for disclosure from a party currently engaged in or preparing to initiate litigation, mediation, alternative dispute resolution (ADR), or other legal proceedings to protect their rights or legally protected interests in relation to films, music, theater, literature, photography, manga, animation, computer games, or other works consisting of text, graphics, colors, sound, movement, or images—or programs designed to provide such information through computer systems, or a lawyer entrusted by such party or an agent authorized to act in court under applicable laws (hereinafter referred to as "Principle 2 Disclosure Requester"), and the request satisfies all the following requirements ("Items that Disclose Requests must satisfy"), the generative AI businesses shall respond to the request for disclosure. As noted in 1(3), this does not require mandatory disclosure of sensitive information belonging to the generative AI businesses (including trade secrets and information related to safety/security) to the generative AI businesses.

【Items that may be requested for disclosure】

- Whether the data used for training or validation purposes—including but not limited to web crawled data, private datasets obtained from third parties, public datasets, synthetic datasets, and other collection methods—contains URLs or other information that the requester is querying (limited to information that the generative AI provider can easily access and verify²; hereinafter referred to as "Principle 2 Reference Information")
- If the requesting party is a generative AI businesses and cannot provide the requested information, the name of the entity that developed the generative AI model incorporated into the generative AI service

【Items that Disclose Requests must satisfy】

- ① The request include sufficient reasons demonstrating that the requester falls under the category of Principle 2 Disclosure Requester
- ② The intended purpose of using the response from the disclosure request be explicitly stated, and the Principle 2 Disclosure Requester pledge not to use the information for any purposes other than those specified
- ③ The Principle 2 Reference Information be provided, and the specific grounds for requesting disclosure to the generative AI businesses in relation to such reference information be clearly identified

² As stated in typical terms of service examples, this provision does not generally permit users who create and post original works on their own websites to inquire beyond the scope of Principle 2 Reference Information regarding whether their original works themselves are included in crawling targets when they discover AI-generated content identical or similar to their works.

(Detailed Provisions)

- The typical examples of disclosure requests under Principle 2 are as follows:
 - A creator who has published original works on the website A discovers AI-generated content that is identical or similar to their works, and requests the generative AI businesses to provide the URL of the works page on the website A, seeking disclosure regarding whether the domain was included in the crawling scope of web crawlers or sourced from training data provided by third parties, etc.
- Regarding the "sufficient reasons demonstrating that the requester falls under the category of Principle 2 Disclosure Requester" specified in ① of the "Items that Disclose Requests must satisfy," the requesting generative AI businesses should provide reasonable grounds for believing that the requester meets the criteria of "a party currently engaged in or preparing to initiate litigation, mediation, ADR, or other legal proceedings to protect their rights or legally protected interests, or a lawyer entrusted by such party or an agent authorized to act in court under applicable laws."
- To ensure no hindrance to the realization of the rights or legally protected interests of the Principle 2 Disclosure Requester, the generative AI businesses should make every effort to provide disclosure in a manner that is as detailed and understandable as possible. Even in cases where certain items related to the disclosure request may constitute trade secrets, it is expected that the generative AI provider will first engage in sincere consideration and consultation.
- Generative AI businesses are encouraged to clearly establish and publicly announce their response policy to disclosure requests from Principle 2 Disclosure Requesters, based on reasonable judgment taking into account technical challenges, costs, intended purposes, and details of legal proceedings (Note: Regarding responses to legal proceedings where the generative AI businesses becomes the opposing party, considerations should appropriately be examined by the generative AI businesses in light of their overall legal proceedings strategy, etc.).

Through the accumulation of such diverse compliance or explanation practices, it is expected that market principles will appropriately evaluate businesses demonstrating excellent practices, while also fostering mutual understanding between generative AI businesses/users and rights holders.
- In implementing Principle 2, measures may be considered to prevent abusive requests, such as setting certain fees or limiting the number of requests to a reasonably appropriate number within a specified period. However, care should be taken to avoid measures—such as fees or request limits—that would discourage, make difficult, or cause the requester to give up on

making disclosure requests.

- While no specific rules are established regarding the timing of responses, it is expected that efforts will be made to provide disclosure promptly within a reasonable period, without hindering the realization of the rights or legally protected interests of the Principle 2 Disclosure Requester.
- From the perspective of fulfilling accountability obligations as providers of generative AI systems or services to the public, merely stating that the implementation framework for Principle 2 has not been established would be considered insufficient as an "explanation," and the businesses should appropriately explain the expected timeline for completing such implementation while considering factors such as their business scale.
- This set of principles is not complete as of this writing, and while acknowledging that there may be matters that cannot be addressed due to each business's services or other circumstances, it will be continued to pursue a balance between promoting advances in AI technology and appropriately protecting intellectual property rights with more significant means.
- If a generative AI businesses describes with the necessary grounds that it is technically impossible to make disclosure with this principle, this shall be considered an "explanation."
- Methods for collecting information other than through disclosure requests under Principle 2 include party interrogatories (Article 163 of the Civil Procedure Code) and motions to compel document production (Article 221 of the same code).

【Principle 3】

When a generative AI businesses receives a request for disclosure from an individual who has generated content using either a generative AI system or generative AI service offered by the provider—such content being in the form of films, music, theater, literature, photography, manga, animation, computer games, or other works consisting of text, graphics, colors, sound, movement, or images—or programs designed to provide such information through computer systems (hereinafter referred to as "Principle 3 Disclosure Requester"), the businesses shall respond to the request if all the following requirements specified in the "Items that Disclosure Requests must satisfy" are met. As noted in Section 1(3), this does not involve mandatory disclosure of sensitive information belonging to the generative AI businesses, such as trade secrets or information concerning safety/security.

【Items that may be requested for disclosure】

- Whether the data used for training, validation, or other purposes—including but not limited to web crawling, non-public datasets obtained from third parties, public datasets, synthetic datasets, and other collection methods—contains URLs or other information (limited to data that can be easily accessed and verified by the generative AI businesses³) that include content identical or similar to that generated by the Principle 3 Disclosure Requester (hereinafter referred to as "Principle 3 Reference Information").
- If the requesting party is a generative AI Provider and cannot provide the requested information, the name of the entity that developed the generative AI model incorporated into the Generative AI Service.

【Items that Disclosure Requests must satisfy】

- ① The request clearly demonstrate the requester's status as a Principle 3 Disclosure Requester by providing both the generated content created by the requester and the prompt used to produce that content.
- ② The purpose of the requested disclosure be explicitly stated, and the Principle 3 Disclosure Requester pledge not to use the information for any purposes other than those specified, nor for litigation, mediation, or alternative dispute resolution proceedings.
- ③ The request include Principle 3 Reference Information and specify the circumstances that justify the disclosure request to the generative AI businesses in relation to that reference information.

³ As stated in typical examples of detailed rules, Principle 3 does not intend to inquire whether content identical to or similar to the product generated by the Disclosure Requester—beyond the scope of Principle 3's Reference Information—is included in the crawling target.

Note: This is a provisional translation, and the Cabinet Office is not responsible for its content. For official information, please refer to the Japanese Version.

[Detailed Provisions]

- The typical examples for disclosure requests under Principle 3 are as follows:
 - When an individual generates content using a generative AI Service A capable of image generation and discovers that identical or similar images appear on website B, the requester may submit a disclosure request to the generative AI Provider offering Service A, providing the following: the generated content, the prompt used to produce it, the intended use of the content, and the URL of website B. The request should seek disclosure regarding whether the domain portion of that URL was included in the training data crawled during the development of the generative AI system for Service A, whether it was sourced from third-party providers, or if the generative AI Provider cannot provide this information, the name of the entity that developed the generative AI model incorporated into the Service.

Generative AI businesses should make every reasonable effort to provide detailed yet understandable disclosures to avoid any disruption to the requester's utilization of the generative AI-generated content. Even in cases where the requested information may constitute trade secrets or other confidential information, it is expected that the generative AI provider will first engage in sincere consideration and consultation.
- Generative AI businesses are encouraged to clearly establish and publicly announce their response policy to disclosure requests from Principle 3 Disclosure Requesters, based on reasonable judgment taking into account technical challenges, costs, intended purposes, and details of legal proceedings (Note: Regarding responses to legal proceedings where the generative AI businesses becomes the opposing party, considerations should appropriately be examined by the generative AI businesses in light of their overall legal proceedings strategy, etc.).

Through the accumulation of such diverse compliance or explanation practices, it is expected that market principles will appropriately evaluate businesses demonstrating excellent practices, while also fostering mutual understanding between generative AI businesses/users and rights holders.
- When implementing Principle 3, it may be appropriate to take measures to prevent abusive requests, such as imposing reasonable fees or limiting the number of disclosure requests to a certain maximum within a specified period. However, care should be taken to avoid implementing measures that would discourage, complicate, or discourage individuals from making disclosure requests.
- While no specific rules are established regarding the timing of disclosures, efforts are required to

ensure prompt disclosure within a reasonable timeframe.

- From the perspective of fulfilling accountability obligations as providers of generative AI systems or services to the public, merely stating that the implementation framework for Principle 2 has not been established would be considered insufficient as an "explanation," and the businesses should appropriately explain the expected timeline for completing such implementation while considering factors such as their business scale.
- This set of principles is not complete as of this writing, and while acknowledging that there may be matters that cannot be addressed due to each business's services or other circumstances, it will be continued to pursue a balance between promoting advances in AI technology and appropriately protecting intellectual property rights with more significant means.
- If generative AI businesses describe with the necessary grounds that it is technically impossible to make disclosure with this principle, this shall be considered an "explanation."
- Alternative methods for collecting information other than through disclosure requests under Principle 3 include party interrogation (Civil Procedure Act Article 163) and petitions for document production orders (Article 221 of the same act).

(2) Considerations when choosing the "explanation" approach

These principles in this document do not preclude generation AI businesses from choosing not to implement all or part of any principle due to reasons such as it being inconsistent with technological realities. In such cases, while striving to achieve the document's purpose of ensuring a safe and secure environment for users by promoting advancements in generation AI technology and appropriately protecting intellectual property rights, providers should make sufficient efforts to secure understanding from both rights holders and users regarding their approach to principles they have not implemented.

Furthermore, even if an organization has previously expressed its intention to implement each principle in this document (i.e., made an acceptance statement), if its actual implementation is effectively non-existent—such as by limiting disclosure subjects based on terms of service or other provisions—an additional "explanation" would still be required. For clarity, simply indicating the existence of provisions in contracts or terms of service that override or restrict the application of the principles outlined in this document would be insufficient as an "explanation"; rather, an explanation of why such provisions have been established and implemented is required.

Note: This is a provisional translation, and the Cabinet Office is not responsible for its content. For official information, please refer to the Japanese Version.

(3) Other matters

This document does not in any way restrict the possibility of alternative legal procedures. For example, when harmful information infringing rights is disseminated on social media, exercising the right to claim damages against the sender by obtaining disclosed sender's information from a social media provider through a disclosure request under the Platform Service Act contributes to protecting intellectual property for the proper utilization of generation AI technology.

Based on the state of response by generation AI businesses, operational effectiveness, and trends in international initiatives, when deemed necessary to promote both advancements in generation AI technology and appropriate protection of intellectual property rights, appropriate measures—including revisions to this document—will be taken in accordance with current circumstances.

(End)