

# Outline of “AI Guidelines for Business Appendix Ver1.2”

---

Ministry of Internal Affairs and Communications  
Ministry of Economy, Trade and Industry  
(March 31, 2026)

# Positioning of main part and appendix of AI Guidelines for Business

- The main part covers “the efforts to be made regarding AI (guiding principles = what)” based on “the ideal society while considering stakeholders’ expectations (basic philosophies = why)” that are important for using AI safely and securely to maximize the benefits of AI.
- The appendix covers “the specific approach to be adopted (implementation = how)” to lead AI business actors to take actual implementation of the principles.

## Main part (why, what)

## Appendix (how)



The ideal society while considering stakeholders’ expectations  
(basic philosophies = why)



The efforts to be made regarding AI  
(guiding principles = what)



The specific approach to be adopted  
(implementation = how)

# Structure of “AI Guidelines for Business”

- This guideline outlines the fundamental concepts necessary for the development, provision, and use of AI.
  - The main part : “the efforts to be made regarding AI (guiding principles = what)” based on “the ideal society while considering stakeholders’ expectations (basic philosophies = why)”
  - The appendix : “the specific approach to be adopted (implementation = how)”

|                            | Main part (why, what) |  | Appendix (how)  |   |
|----------------------------|-----------------------|--|---|---|
| For all AI business actors | Part 1                | Definitions  | 1. Relevant to Part 1 [About AI]  | A. Preconditions for AI<br>B. AI’s benefits and risks   |
|                            | Part 2                | Society to aim for with AI, and matters each AI business actor works on<br>A. Basic philosophies<br>B. Principles<br>C. Common Guiding Principles<br>D. Hiroshima AI Process “Hiroshima Process International Guiding Principles for All AI Actors”<br>E. Building AI governance | 2. Relevant to Part 2 [E. Building AI Governance]   | A. Building of AI governance and monitoring by management<br>B. Examples of business operator’s efforts at AI governance  |
| For each AI business actor | Part 3                | Matters Related to AI developer<br>* Includes Hiroshima AI Process “Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems” as well  | 3. Relevant to Part 3 [For AI developer]  | A. Descriptions of Part 3 “Matters Related to AI developer”<br>B. Descriptions of “Common Guiding Principles” in Part 2<br>C. Hiroshima AI Process “Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems” |
|                            | Part 4                | Matters Related to AI provider   | 4. Relevant to Part 4 [For AI provider]   | A. Descriptions of Part 4 “Matters Related to AI provider”<br>B. Descriptions of “Common guiding principles” in Part 2  |
|                            | Part 5                | Matters Related to AI Business user  | 5. Relevant to Part 5 [For AI Business user]  | A. Descriptions of Part 5 “Matters Related to AI Business user”<br>B. Descriptions of “Common Guiding Principles” in Part 2   |
| Other references           |                       |  | 6. Major precautions for referring to “Contract Guidelines on Utilization of AI and Data”<br>7. Checklist, Worksheet<br>8. Cross-actor virtual cases<br>9. References for overseas guidelines, etc. | The appendixes 7, 8, and 9 are <u>Japanese only.</u>  |

# Appendix 1. Relevant to Part 1

- Appendix 1 provides a detailed explanation of the premises related to AI and AI's benefits and risks
- Appendix 1 leads to a deeper understanding of the content of this guideline through these explanations.

## A. Preconditions for AI

- Flow of training and use of AI
  - The general flow of training and use of AI
- Overview of AI system
  - Scope of AI system
- Value chain of AI from development to use
  - The general flow of applying AI and responses of each AI business actor
- Examples of AI systems and services
  - Specific examples of typical AI systems and services and the actors involved
- Patterns of AI companies
  - Specific examples of AI value chain at the time of business utilization
- About data providers
  - Definitions of data providers and the like, that are not covered by the guidelines

## B. AI's benefits and risks

- AI's benefits
  - Focus on end-users who will be primarily benefited
- AI's risks
  - Typical examples (including assumptions)
  - Systematic Classification of Risk Examples (tentative ver.) (Technical risks and social risks)
  - Key common guiding principles corresponding to each risk example, and examples of countermeasures taken by AI business actors.

# Appendix 1 A. Preconditions for AI

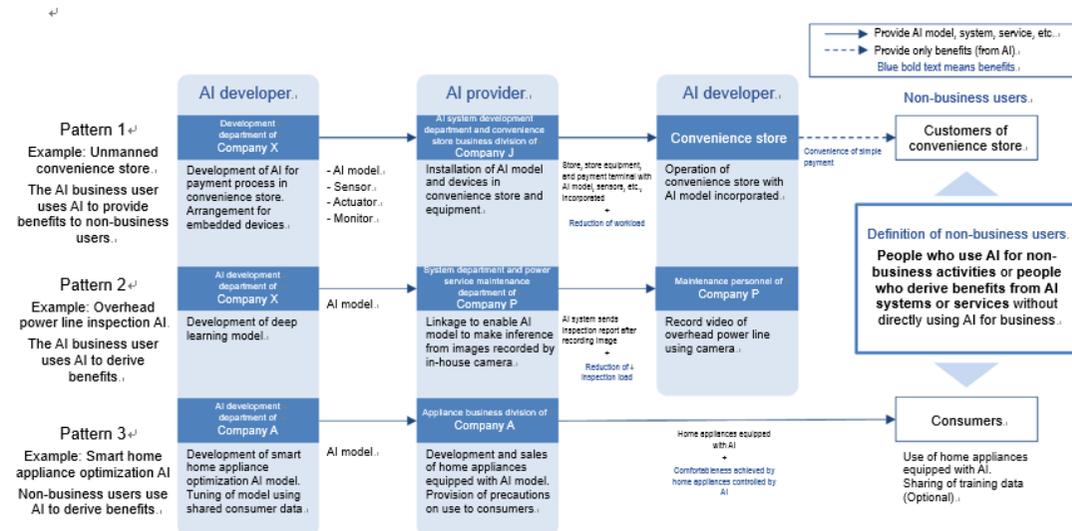
## Appendix 1

- To deepen the understanding of AI itself, the contexts in which it is utilized, and the roles of each AI business actor, terms such as “AI systems,” “AI business actors,” and “data providers” are explained with specific examples (the following is an excerpt of the content).

### Examples of AI systems and services (excerpt)

| Case name                  | Used AI                 | Overview  | AI developer                          | AI provider  | AI business user   | Non-business user               |
|----------------------------|-------------------------|---|---------------------------------------|--|--|---------------------------------|
| Recruitment AI             | Text analysis           | The recruit department of each foreign subsidiary of Company A Group uses an AI service that provides reference information for conducting the applicant screening process based on applicants' applications. The AI development department of Company A has created an AI model that receives past application data and decision of acceptance (judgment on whether to employ each applicant) from the recruit department of Company A (AI business user; including the recruit departments of overseas group companies) and processes them through machine learning (classification model) for supporting in making acceptance decisions. | Company A (Development department)    | Company A (System department and human resource development department)              | Company A Group (Recruit department)                       | Applicants for recruitment      |
| Unmanned convenience store | Image analysis          | Company J, which holds convenience store franchises across Japan, operates unmanned convenience stores in which image recognition AI is used. In the unmanned convenience stores, AI calculates the price for items taken by each customer and carries out the payment process for all the items through digital money, etc., when the customer leaves the store. An AI system for unmanned convenience stores developed by Company X is incorporated into the AI service.  | Company X                             | Company J (AI system development department and convenience store business division) | Convenience stores   | Customers of convenience stores |
| Cancer diagnosis AI        | Text and image analyses | Using the multimodal learning, this system imports “information of the medical history, genes, etc., of a patient (data 1)” and “endoscopic image (data 2)” to highlight areas that are highly possibly affected by cancer in real time during an endoscopic examination. It enables physicians to observe output images and diagnose potential cancer. Company A has developed AI and provides the cancer diagnosis AI system to health facilities.  | Company A (AI development department) | Company A (Healthcare IT service department)   | Health facilities (System department and gastroenterology) | Patients examined               |

### Patterns of AI business actors



# Appendix 1 B. AI's benefits and risks (AI's benefits)

## Appendix 1

- Focusing on the end-users who will benefit from AI, the benefits of AI are organized by industry and business operation.

Examples of benefits available conventionally

Improved by generative AI ( Improved by partly AI agents, and physical AI )

|  | Development   | Marketing  | Sales  | Logistics/distribution  | Customer support   | Legal   | Finance  | HR   |
|--|---|--|--|---|--|---|--|--|
|  | Automation of code verification and documentation             | Automatic distribution of ad emails  | Support after order intake<br>Automatic transmission of emails, etc. | Optimization of production and inventory based on demand prediction | Automatic response through chat bot                                | Translation   | Automatic creation of financial statements   | Automation of payroll calculation, etc.                |
|  | Extraction and verification of similar code and data          | Personalized ad based on data  | Sales prediction for each channel and need                           | Optimization of delivery route                                      | Prediction of customer conversion and churn rates                  | Review of legal text  | Future prediction based on past records, and detection of malpractices                               | Human resources demand matching based on résumés, etc. |
|  | Generation of training data                                   | Automatic creation of sales promotion (marketing materials, sales                      | Automatic creation of sales talk script                              | Assistant for negotiation for logistics conditions                  | Automatic generation and summarization of transcription of support | Automatic generation of draft of contract based on stipulations   | Expense settlement assistant (automatic journalizing, application review, voucher acquisition, etc.) | AI recruitment assistant (interviews and evaluations)  |
|  | Coding assistant (code generation, automatic bug fixes, etc.) | SNS management agent that autonomously performs tasks from posting through to analysis | Automated customer service and sales by in-store robots              | Automated delivery by autonomous transport robots and drones        | Automatic FAQ creation based on past inquiry records               | Similar case search and automatic summarization of key precedents | Automatic generation of investment reports and market analyses                                       | Personalized career plan proposals                     |

Examples of benefits unique to generative AI, AI agents, and physical AI

# Appendix 1 B. AI's benefits and risks (AI's risks)

## Appendix 1

- AI's risks have been systematically organized to help AI business actors comprehensively understand and consider countermeasures for these risks.
  - The table below does not cover all AI risks and includes hypothetical cases, and it is expected to be recognized as just one example.
  - The table below also includes risks that require responses and discussions from society as a whole, including governmental and public institutions.

| Major categories   | Subcategories                                    | Risk examples   |
|--|--|---|
| <b>Technical Risks</b><br>(=risks primarily associated with AI systems)                    | Risks during the learning and input stages of AI | Attacks on AI systems such as data poisoning attacks  |
|  | Risks during the output stage of AI              | Biased outputs and inconsistent outputs<br>Incorrect outputs due to Hallucinations and similar issues |
|  | Risks during the post-response stage             | Black-boxing and inadequate explanations of decisions   |
| <b>Societal Risks</b><br>(=existing risks that may also arise in AI or be amplified by AI) | Risks related to ethics and law                  | Inappropriate use of personal information   |
|  |  | Occurrence of accidents related to lives, etc.  |
|  |  | Discriminatory outputs<br>Excessive dependence  |
|  | Risks related to economic activities             | Misuse  |
|  |  | Infringement of intellectual property rights, etc.  |
|  |  | Financial loss  |
|  |  | Leak of confidential information<br>Unemployment of workers   |
|  |  | Concentration of data and profits<br>Infringement of qualifications, etc.                             |
|  | Risks related to the information space           | Distribution and diffusion of disinformation  |
|  |  | Negative influence on democracy   |
|  |  | Filter bubble and echo chamber phenomena  |
|  |  | Loss of diversity and inclusion<br>Reproduction of biases   |
|  | Risks related to the environment                 | Energy consumption and environmental load   |

# Appendix 1 B. AI's benefits and risks (AI's risks)

## Appendix 1

- To further connect the identified risks to the consideration of countermeasures by AI business actors, the key common guiding principles corresponding to each risk and examples of countermeasures by AI business actors have been described\*1.
  - The table below does not cover all AI risks and includes hypothetical cases, and it is expected to be recognized as just one example.
  - The table below also includes risks that require responses and discussions from society as a whole, including governmental and public institutions.

| Risk Examples  | Key common guiding principles corresponding to each risk example  | Important matters for each AI business actor in addition to common guiding principles   |  |   |
|--|---|---|--|---|
|  |   | Part 3. AI Developer  | Part 4. AI Provider  | Part 5. AI Business User                            |
| Attacks on AI systems such as data poisoning attacks       | 5) Ensuring security  | i. Deployment of mechanisms for security measures<br>ii. Consideration for the latest trends  | i. Deployment of mechanisms for security measures<br>ii. Handling of vulnerabilities   | i. Implementation of security measures              |
| Biased outputs and inconsistent outputs                    | 1) Human-centric<br>(1) Human dignity and autonomy of individuals<br>(3) Countermeasures against disinformation |   |  |   |
| Incorrect outputs due to Hallucinations and similar issues | 2) Safety   | i. Proper data training<br>ii. Development that takes into consideration the lives, bodies, properties and minds of humans and the environment<br>iii. Development contributing to proper use (of AI) | i. Actions against risks that consider the lives, bodies, properties, and minds of human and the environment<br>ii. Provision contributing to proper use (of AI) | i. Proper use (of AI) that considers safety         |
| Technical Risks  | 3) Fairness   | i. Consideration for bias in data<br>ii. Consideration for bias in algorithms, etc., of AI models   | i. Consideration for bias in configurations and data of AI systems and services  | i. Consideration for bias in input data or prompt   |
|  | 8) Education/literacy   |   |  |   |
| Black-boxing and inadequate explanations of decisions      | 6) Transparency   | i. Ensuring verifiability<br>ii. Providing relevant stakeholders with information   | i. Documentation of system architectures and the like<br>ii. Providing relevant stakeholders with information  | i. Providing relevant stakeholders with information |

\*1 : It is advisable to refer to the relevant sections of Parts 3 to 5 of the main text and Appendices 3 to 5 for measures and specific methods for each entity.

# Appendix 2. “Section 2. E. Building AI Governance”

- Appendix 2 includes “Behavioral Goals” and “Practice Guidelines” for establishing “AI Governance, as well as corresponding hypothetical “Practical Examples” and examples of actual company initiatives.
- Specific examples will enable each business to consider how to establish its own AI governance.

## A. Creating AI governance and monitoring by management

- Behavioral Goals
  - The general and objective goals
- Practice Guidelines
  - Based on the “Governance Guidelines for Implementation of AI Principles,” incorporating elements from national and international guidelines, ISO, etc.
- Practical Examples
  - Includes examples based on hypothetical cases
  - Incorporates examples of responses to the latest trends in generative AI, etc.

## B. Examples of business operators’ AI governance initiatives

- Examples of business operators
  - Columns of 11 entities’ (10 companies and 1 local government) AI governance initiatives based on the “Governance Guidelines for Implementation of AI Principles”
  - Describes perspectives that many companies encounter when promoting AI governance

List of entities Featured in the Column (in order of appearance)

- |                   |                       |
|-------------------|-----------------------|
| • ABEJA           | • NTT DATA            |
| • NEC Group       | • Ubie, Inc.          |
| • Toshiba Group   | • Kobe City           |
| • Panasonic Group | • IBM                 |
| • Fujitsu Group   | • Amazon Web Services |
| • Softbank        |                       |

# Appendix 2 A. Building of AI governance and monitoring by management (Content)

## Appendix 2

- Support the establishment of AI governance by ensuring that each AI business understands the significance of and utilizes the governance behavioral goals.

| Category                           | Behavioral Goals ※ Some are further subdivided like 「3-1-1」   |
|------------------------------------|---|
| 1. Environmental and risk analysis | 1-1 Understanding benefits/risks<br>1-2 Understanding social acceptance of AI<br>1-3 Understanding company's AI know-how  |
| 2. Goal setting                    | 2-1 Setting AI governance goals   |
| 3. System design                   | 3-1 Requiring evaluation of goal deviation and measures to minimize it<br>3-2 Improving literacy of those in charge of the AI management system<br>3-3 Enhancing AI management through cooperation between AI business actors and divisions<br>3-4 Reducing burden related to incidents involving AI Business Users and non-business users through preventive and prompt action |
| 4. Operation                       | 4-1 Ensuring that the operation of AI management system is explainable<br>4-2 Ensuring that the operation of each AI system is explainable<br>4-3 Considering proactive disclosure of AI governance practices   |
| 5. Evaluation                      | 5-1 Verifying AI management system functions<br>5-2 Considering opinions of outside stakeholders  |
| 6. Environment and risk reanalysis | 6-1 Reimplementing Behavioral Goals 1-1 to 1-3 at an appropriate time   |

# Appendix 2 A. Building of AI governance and monitoring by management (Structure)

## Appendix 2

- Organize “Practice Guidelines” and “Practical Examples” for each “Behavioral Goal.”
- Enables each AI business to consider AI governance by referring to them according to the type of AI to be employed and the degree of risk.

### Appendix Content

### Descriptions

#### Behavioral Goal 1-1 [Understanding benefits/risks]

AI business actors will, under the leadership of the management team, clarify the purpose of development, provision and use of AI, specifically understand that there are not only benefits offered by AI, but also unintended risks, report them to the management, share with the management and update the understanding at an appropriate time.

#### Behavioral Goal

- Provides general and objective goals that are important for AI businesses to work toward
- Provides material for each AI business to consider when reviewing its policies

#### [Practice Guidelines]

AI business actors will, under the leadership of the management team, take the following measures:

- Specifically understand the benefits and risks including unintended risks in a way that is relevant to a company’s own business.
- Establish a system for promptly reporting to / sharing with the management.

“Risks” include the following examples:

- Common risks of AI in general: Output of biased results or discriminative results, filter bubbles, echo chambers, disinformation, handling of inappropriate personal data, data pollution attacks, obscuring, leak of confidential information, abuse of AI system services, energy consumption and environmental burden, reproduction of bias, etc.

#### Practice Guidelines

- Summarizes important matters and points to keep in mind to implement the above action goals
- Provides material for each AI business to consider specific actions

#### [Practical Examples]

##### [Practical Example i: Understanding benefits and risks]

It is important that AI business actors examine not only benefits, but also risks under the leadership of the management (including the examination the management itself implements by taking the lead instead of leaving it up to an officer in charge or staff in charge), share the results of examination and update the understanding at an appropriate time.

#### Practical Examples

- Provides examples of hypothetical cases in practice
- Makes it easier for each business operator to take action by having a specific example

# Appendix 3-5. For each AI business actor (Content)

## Appendix 3-5

- Attachments 3 through 5 provide detailed explanations in the main part for each AI business actor.
- It can be used in conjunction with the main part to refer to and discuss specific approaches related to the main part's content.

### A. Descriptions of Part 3-5

### B. Descriptions of “Common Guiding Principles” in Part 2

### C. Hiroshima AI Process “Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems” ※Only in Appendix3. For AI developers

#### • Points

- Supplements important points in addition to those described in the main part

#### • Specific Methods

- Specific explanations with reference to other guidelines, etc.

#### • References

- List other references

(e.g.)

- Digital Agency “Data Quality Guidebook (β Edition)” (June 2021)
- National Institute of Advanced Industrial Science and Technology “Guidelines for Machine Learning Quality Management, 4th Edition” (December 2023)
- Consortium of Quality Assurance for Artificial-Intelligence-based Products and Services “AI Product Quality Assurance Guidelines” (April 2024)
- NIST, “AI Risk Management Framework Playbook” (January 2023)
- Council of Europe, “Risk and Impact Assessment Method for AI Systems from the Perspective of Human Rights, Democracy, and the Rule of Law (HUDERIA)” (November 2024)

# Appendix 3-5. For each AI business actor (Structure)

- For each important matter of each AI business actor, “Points,” “Specific methods,” and “References” are provided in the following structure:

## Appendix Content

## Descriptions

### A. Descriptions of Part 3 “Matters Related to AI Developers”

[Contents of the main part (repeat)]

During data preprocessing

#### D-2) i. Proper data training

- Properly collect training data through privacy-by-design, etc., and if it contains third-parties’ personal data, data requiring attention to intellectual property rights, etc., ensure that such data is properly handled in compliance with laws and regulations throughout the lifecycle of AI (“2) Safety,” “4) Privacy protection,” “5) Ensuring security”).
- Implement proper protective measures before and across training by, for example, considering the deployment of any data management and restriction function that controls access to data (“2) Safety,” “5) Ensuring security”).

[Points]

In order to improve the quality of AI models, it is important for AI Developers to pay close attention to the quality of data used for AI training, etc.

- Pay close attention to the quality (accuracy, integrity, etc.) of the data used for AI training, etc., in view of the characteristics and applications of the AI to be used.

[Specific methods]

- Verify that the data does not contain personal data, confidential information, rights including copyrights or legally protected interests.
  - Extraction of unique expressions

[References]

- National Institute of Advanced Industrial Science and Technology “Guidelines for Machine Learning Quality Management, 4th Edition” (December 2023)
- NIST, “AI Risk Management Framework Playbook” (January 2023)

### B. Descriptions of “Common guiding principles” in Part 2

Although not mentioned in the Main Part, “Part 3 Matters Related to AI Developers,” specific methods for the Main Part, “Part 2” “Common guiding principles,” which are especially important for AI Developers, are explained here.

Repeat of main part content

#### Points

- Supplements important points in addition to those described in the main part

#### Specific methods

- Specific explanations with reference to other guidelines, etc.

#### References

- Provide references to “Points” and “Specific methods”

Descriptions corresponding to Part 3-5 in the main part

Descriptions corresponding to Part 2

Specific approaches of each AI business actor regarding the Common Guiding Principles

# Appendix 6. Major precautions for referring to “Contract Guidelines on Utilization of AI and Data”

- Appendix 6 contains the precautions for referring to the “Contract Guidelines on Utilization of AI and Data” as of March 2026, the first edition of which was formulated and published in June 2018, due to changes in the situation regarding the development and use of AI and new technological advancements.
- By clearly defining the rights and obligations of the parties through contracts, transactions related to AI can be facilitated and unnecessary disputes can be prevented.

## Contract Guidelines on Utilization of AI and Data



Trends in international society

Emergence of new technology

Contract Checklist for AI Utilization and Development



## Appendix 6. Major precautions for referring to “Contract Guidelines on Utilization of AI and Data”

The following matters should be considered in light of changes in the situation since the release of the Contract Guidelines:

- (1) Diversification of contract models
  - Contracts related to the use and development of AI can be broadly classified into the following three types, with each type having different contractual considerations and negotiation points.
  - Type 1: General-purpose AI Service Usage
  - Type 2: Customization
  - Type 3: New Development
- (2) Risk distribution under a complex value chain
  - Responsibilities based on diversity or complexity of value chain needs to be considered.
- (3) Development, provision and use of AI, and accountability
  - The following are beneficial points to consider contractually in relation to accident risk; Organizing new types of risks, Implementation of reasonable explanations, Presentation of objective evidence.

# Appendix 7. Checklist, Worksheet (Content)

- Appendix 7 provides a “checklist” and “worksheet for specific approaches” to ensure the planning and implementation of initiatives to reduce risks and reap benefits from AI.

Main part·  
Appendix 1~5



Checklist



Worksheet for specific  
approaches



Understand the importance of AI governance and what is expected of each AI business actor by reading the main part and the Appendix.

Confirm the approaches (“What”) of each AI business actor by using the checklist.

Use the “worksheet for specific approaches” to discuss the specific approaches (“how”) of each entity.

※ It is also important for each business operator to create and effectively utilize its own checklist and worksheet suited to its business details, AI policy, AI regulations, and other relevant circumstances. (As for the worksheet, it is not necessarily required to adopt all of the listed items; it is effective to identify and utilize only those items that are necessary for the company.)

# Appendix 7. Checklist Utilization

## Appendix 7

- All AI business actors confirm an overview of their own efforts by using “Appendix 7. Checklist A [for all AI business actors].”
- In addition, please also utilize, as necessary, “Appendix 7 B Checklist [Part 2 D. Hiroshima AI Process 'Hiroshima Process International Guiding Principles for All AI Actors’].”

別添 7 A チェックリスト 令和8年3月31日  
[全主体向け]

本チェックリストは、AI事業者ガイドライン「第2部C.共通の指針」を要約したものです。事業者に求められる重要な取組事項のチェックにご活用ください

別添 7 B チェックリスト 令和8年3月31日  
第2部D. 広島AIプロセス「全ての AI 関係者向けの広島プロセス国際指針」

本チェックリストは、AI事業者ガイドライン「第2部D.広島AIプロセス「全ての AI 関係者向けの広島プロセス国際指針」」の項目です。取組事項のチェックにご活用ください  
※①～⑩については、適時適切に、適切な範囲で、適用されるべきである。また、⑫については、従うべきである。

Japanese only

### チェック項目

- 人間中心の考え方を基に、憲法が保障する又は国際的に認められた人権を侵すことがないようにしているか？
- AIに関わる全ての者の生命・身体・財産、精神及び環境に危害を及ぼすことがないように安全性を確保しているか？
- 潜在的なバイアスをなくすよう留意し、それでも回避できないバイアスがあることを認識しつつ、回避できないバイアスが人権及び多様な文化を尊重する公平性の観点から許容可能か評価しているか？
- プライバシーを尊重・保護し、関係法令を遵守しているか？
- 不正操作によってAIの振る舞いに意図せぬ変更又は停止が生じることのないように、セキュリティを確保しているか？
- 透明性を確保するために、AI自体やAIシステム・サービスの情報をステークホルダーに対し合理的で技術的に可能な範囲で提供しているか？
- データの出所、AIの意思決定等のトレーサビリティに関する情報やリスクへの対応状況等について、関連するステークホルダーに対して合理的な範囲でアカウンタビリティを果たしているか？
- AIガバナンスやプライバシーに関するポリシー等を策定しているか？
- 上記の実現のため、各事業者の状況に応じた具体的なアプローチは検討しているか？

検討には「具体的なアプローチ検討のためのワークシート」をご活用ください

### チェック項目

- ① AI ライフサイクル全体にわたるリスクを特定、評価、軽減するために、高度なAIシステムの開発全体を通じて、その導入前及び市場投入前も含め、適切な措置を講じているか？
- ② 市場投入を含む導入後、脆弱性、及び必要に応じて悪用されたインシデントやバグを特定し、緩和しているか？
- ③ 高度な AI システムの能力、限界、適切・不適切な使用領域を公表し、十分な透明性の確保を支援することで、アカウンタビリティの向上に貢献しているか？
- ④ 産業界、政府、市民社会、学界を含む、高度なAI システムを開発する組織間での責任ある情報共有とインシデントの報告に向けて取り組んでいるか？
- ⑤ 特に高度な AI システム開発者に向けた、個人情報保護方針及び緩和策を含む、リスクベースのアプローチに基づく AI ガバナンス及びリスク管理方針を策定し、実施し、開示しているか？
- ⑥ AI のライフサイクル全体にわたり、物理的セキュリティ、サイバーセキュリティ、内部脅威に対する安全対策を含む、強固なセキュリティ管理に投資し、実施する
- ⑦ 技術的に可能な場合は、電子透かしやその他の技術等、ユーザーが AI が生成したコンテンツを識別できるようにするための、信頼できるコンテンツ認証及び来歴のメカニズムを開発し、導入しているか？
- ⑧ 社会的、安全、セキュリティ上のリスクを軽減するための研究を優先し、効果的な軽減策への投資を優先しているか？
- ⑨ 世界の最大の課題、特に気候危機、世界保健、教育等（ただしこれらに限定されない）に対処するため、高度な AI システムの開発を優先しているか？
- ⑩ 国際的な技術規格の開発を推進し、適切な場合にはその採用を推進しているか？
- ⑪ 適切なデータインプット対策を実施し、個人データ及び知的財産を保護しているか？
- ⑫ 高度な AI システムの信頼でき責任ある利用を促進し、貢献しているか？

検討には「具体的なアプローチ検討のためのワークシート」をご活用ください

- The checklist is a summary of the main part
- With the checklist, an overview of the status of their own initiatives can be obtained

- Utilize the “worksheet for specific approaches” to examine specific practices (see next page for how to utilize the worksheet).

# Appendix 7. Worksheet for specific approaches Utilization

## Appendix 7

- Describes important matters when considering specific approaches to the guideline content.
- Presumes that each business will customize and utilize this worksheet according to its own business activities and circumstances.

### 別添 7 C. 具体的なアプローチ検討のためのワークシート (共通の指針関連)

利用上の留意点

ガイドラインに記載した内容に関して取り組むべき事項は、各事業者の事業内容や置かれた状況等により、個々に異なります。このため、本ワークシートは、あくまで各事業者が取り組むべき事項が何かを検討する際の材料をご提供するものであり、各事業者それぞれの状況に応じ、カスタマイズして必要に応じて活用いただくことを前提としたものです。したがって、必ずしも、全ての事項について、検討が必要となるものではないので、活用の要否、各自の事情に応じた修正や取捨選択を検討ください。D列、E列の記載内容を基に、E列以降を各事業者にてご検討いただき、各自が取り組むこと（あるいは内容）を具体化の上、活用ください。

各自用カスタマイズし運用するには、α. 取組内容を作成する者（当該ワークシートを基に、各自の取組内容のカスタマイズを行う者）、β. 実施状況の確認を行う者（現場において実際の確認を行う者）γ. 責任者（確認内容に対して責任を負う者）を特定ください（各主体の規模によって、α～γが重複する場合もある）。

α. ワークシートの作成者

β. 実施状況の確認を行う者

γ. 責任者

| 対応箇所    | 分類                      | 検討にあたって重要な事項   | 各自の事業において検討対象とする事項<br>(該当しない場合はその理由) | 他の主体との関係についての事項 | ✓ | 具体的なアプローチ | 最終検討日<br>(見直し日) |
|---------|-------------------------|--|--------------------------------------|-----------------|---|-----------|-----------------|
| 1) 人間中心 | 1) 人間の尊厳及び個人の自律         | a. AIが活用される際の社会的文脈を踏まえ、人間の尊厳及び個人の自律を尊重しているか？<br>b. 特に、AIを人間の脳・身体と連携させる場合には、その周辺技術に関する情報を踏まえつつ、諸外国及び研究機関における生命倫理の議論等を参照しているか？<br>c. 個人の権利・利益に重要な影響を及ぼす可能性のある分野においてAIを利用したプロファイリングを行う場合、個人の尊厳を尊重し、アウトプットの正確性を可能な限り維持させつつ、AIの予測、推察、判断等の限界を理解して利用し、かつ生じうる不利益等を慎重に検討した上で、不適切な目的に利用していないか？ |                                      |                 |   |           |                 |
|         | 1) AIによる意思決定・感情の操作等への留意 | a. 人間の意思決定、認知等、感情を不当に操作することを目的とした、又は意識的に知覚できないレベルでの操作を前提としたAIシステム・サービスの開発・提供・利用は行っていないか？<br>b. AIシステム・サービスの開発・提供・利用において、自動化/バイパス等のAIに過度に依存するリスクに注意を払い、必要な対策を講じているか？<br>c. フィルター/バブルに代表されるような情報又は価値観の偏斜を助長し、AI利用者を含む人間が本来得られるべき選択   |                                      |                 |   |           |                 |

Japanese only

- Before using the worksheet, **decide who will be responsible for its implementation.**

- Each business will **customize its own approach** according to the nature of its business, its circumstances, etc.

- **Periodically update** by considering and writing the “review date”.

- The worksheet contains important information when considering specific approaches and can be used as a **reference for businesses to customize their own approaches.**

# Appendix 8. Attachment 8. Hypothetical case study across AI business actors (Overview)

## Appendix 8

- Attachment 8. provides a “hypothetical case study across AI business actors” in which AI developers, AI providers, and AI business users consider important matters in accordance with the guideline.
- Enables each actor to recall specific cases in which they actually incorporated the content of the guideline and clarify the points where cooperation is essential.

Japanese only

| Case 採用AI   |    | AI開発部門  | 人材採用部門（採用AIチーム）  | 人材採用担当者  |
|---|----|---|--|--|
| 機械学習モデル：XGBoost（エントリーシートの文章で、応募者に対して可否を判断する）<br><a href="https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2022/10/RCModel_Case01_Recruitment-AI_JP.pdf">https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2022/10/RCModel_Case01_Recruitment-AI_JP.pdf</a> |    | AI開発者   | AI提供者  | AI利用者  |
| No  | 分類 | 共通の指針/各主体に関する事項   | 本UCにおいて主体が実施している活動   | 本UCにおいて主体が実施している活動   |
| <b>1) 人間中心</b>  |    |   |  |  |
| 各主体は、AIシステム・サービスの開発・提供・利用において、後述する各事項を含む全ての取り組むべき事項が導出される土台として、少なくとも憲法が保障する又は国際的に認められた人権を侵すことがないようにすべきである。また、AIが人々の能力を拡張し、多様  |    |   |  |  |
| <b>①人間の尊厳及び個人の自律</b>  |    |   |  |  |
| 1   | 共通 | AIが活用される際の社会的文脈を踏まえ、人間の尊厳及び個人の自律を尊重する   | AIシステムの開発において、学習データの収集やラベリング、モデルの性能評価等は、AI開発者だけで完結せず、AI提供者側で   | AIサービスの提供において、AI利用者が最終判断(応募者の可否)を行えるようになっている(Human-in-the-loop)  |
| 2   | 共通 | 特に、AIを人間の脳・身体と連携させる場合には、その周辺技術に関する情報を踏まえつつ、諸外国及び研究機関における生命  | 脳・身体と連携するケースではないため対象外  | 脳・身体と連携するケースではないため対象外  |
| 3   | 共通 | 個人の権利・利益に重要な影響を及ぼす可能性のある分野においてAIを利用したプロファイリングを行う場合、個人の尊厳を尊重し、アウトプットの正確性を可能な限り維持させつつ、AIの予測、推奨、判断等の限界を理解して利用し、かつ生じうる不利益等を慎重に検討した上で、不適切な目的に利用しない | AIシステムの開発において、実際の予測結果を学習データに用いる際には個人情報の取扱いに関わる誓約書の締結やアクセス権管理等を実施している。<br>※公平性とプライバシーについては、「3）公平性」「4）プライバシー保護」を参照 | AIシステムの開発において、実際の予測結果を学習データに用いる際には個人情報の取扱いに関わる誓約書の締結やアクセス権管理等を実施している。<br>※公平性とプライバシーについては、「3）公平性」「4）プライバシー保護」を参照 |
| <b>②AIによる意思決定・感情の操作等への留意</b>  |    |   |  |  |
| 1   | 共通 | 人間の意思決定、認知等、感情を不当に操作することを目的とした、又は意識的に知覚できないレベルでの操作を前提としたAIシステム・サービスの開発・提供・利用は行わない   | 本ケースに関しては、2)①-3と同じ論点になる  | 本ケースに関しては、2)①-3と同じ論点になる  |
| 2   | 共通 | AIシステムの開発・提供・利用において、自動化バイアス等のAIに過度に依存するリスクに注意を払い、必要な対策を講じる  | 本ケースに関しては、2)①-3と同じ論点になる  | 本ケースに関しては、2)①-3と同じ論点になる  |

# Appendix 9. Comparison with international guidelines (Overview)

## Appendix 9

- Appendix 9. summarizes the correspondence between the elements of the guideline and other guidelines, etc.
- This is to be used to clarify the correspondence between the guideline and other (domestic and international) guidelines, etc., when the AI business actor is considering more detailed content.

### Primary guidelines, etc. used as references

- Advancing accountability in AI (February 2023, OECD)
- Hiroshima AI Process Comprehensive Policy Framework (December 2023, G7)
- Artificial Intelligence Risk Management Framework (AI RMF 1.0) (January 2023, NIST)
- CYBERSECURITY FRAMEWORK (CSF) (April 2018, NIST)
- Blueprint for an AI Bill of Rights (October 2022, THE WHITE HOUSE)
- Artificial Intelligence Act (EU AI Act) (August 2024, EU)
- Guidelines for secure AI system development (November 2023, NCSC)
- ETHICS GUIDELINES FOR TRUSTWORTHY AI (April 2019, EU)
- Guidelines for privacy impact assessment (ISO)
- Recommendation on the Ethics of Artificial Intelligence (November 2021, UNESCO)

| AI事業者ガイドライン本編                |                                     | AI事業者ガイドライン別添           |   | ①OECD<br>AIにおけるアカウンタビリティの高度化                      | ②G7<br>広島AIプロセス  |
|------------------------------|-------------------------------------|-------------------------|---|---|--|
| はじめに                         | -                                   | 別添.はじめに                 | -   |   |  |
| 第1部 AIとは                     | -                                   | 別添1.第1部関連               | A.AIに関する前提  | AI Terms & concepts (OECD AI Principles overview) | 高度なAIシステムを開発する組織向けの広島AIプロセス国際行動規範(2023年10月, 広島AIプロセスに関するG7首脳声明)                            |
|                              | -                                   |                         | B.AIによる便益/リスク   |   |  |
| 第2部 AIにより目指すべき社会及び各主体が取り組む事項 | A.基本理念                              | -                       | -   |   |  |
|                              | B.原則                                | -                       | -   |   |  |
|                              | C.共通の指針                             | -                       | -   |   | 別添版「各主体に共通の指針」にて広島AIプロセス包括的政策枠組み(2023年12月, 広島AIプロセスG7デジタル・技術閣僚声明)を引用<br>別添版「各主体に共通の指針」にて引用 |
|                              | D.広島AIプロセス「全てのAI関係者向けの広島AIプロセス国際指針」 | -                       | -   |   | 別添版「各主体に共通の指針」にて引用   |
|                              | E.ガバナンスの構築                          | 別添2.「第2部E.AIガバナンスの構築」関連 | A.経営層によるAIガバナンスの構築及びモニタリング<br>B.AIガバナンスの構築に関する実際の取組事例 |   | 別添版「各主体に共通の指針」にて引用<br>別添版「各主体に共通の指針」にて引用<br>別添版「各主体に共通の指針」にて引用<br>別添版「各主体に共通の指針」にて引用       |

Japanese only