Outline of
"AI Guidelines for Business Appendix Ver1.0"

**Ministry of Internal Affairs and Communications
Ministry of Economy, Trade and Industry
(April 19,2024)**

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"

- The descriptions in the Appendix correspond to those in the main part and serve as a supporting document for the reading of the main part and considerations and actions based on the main part.

Main part (why, what) ▶ **Appendix (how)**

For all AI business actors

	Main part (why, what)		Appendix (how)	
Part 1	Definitions		1. Relevant to Part 1 [About AI]	A. Preconditions for AI B. AI's benefits and risks
Part 2	Society to aim for with AI, and matters each AI business actor works on	A. Basic philosophies B. Principles C. Common Guiding Principles D. Common Guiding Principles for AI business actors involved in advanced AI systems E. Building AI governance	2. Relevant to Part 2 [E.Building AI Governance]	A. Building of AI governance and monitoring by management B. Examples of business operator's efforts at AI governance

For each AI business actor

Part 3	Matters Related to AI Developers	* Includes additional matters described in "Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems" as well	3. Relevant to Part 3 [For AI Developers]	A. Descriptions of Part 3 "Matters Related to AI Developers" B. Descriptions of "Common Guiding Principles" in Part 2 C. Matters to be observed in developing advanced AI systems
Part 4	Matters Related to AI Providers		4. Relevant to Part 4 [For AI Providers]	A. Descriptions of Part 4 "Matters Related to AI Providers" B. Descriptions of "Common guiding principles" in Part 2
Part 5	Matters Related to AI Business Users		5. Relevant to Part 5 [For AI Business users]	A. Descriptions of Part 5 "Matters Related to AI Business Users" 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"
7. Checklist
8. Cross-actor virtual cases
9. References for overseas guidelines, etc.

The appendixes 7, 8, and 9 are Japanese only.

Appendix 1. Relevant to Part 1

- Appendix 1 presents examples of AI systems and specific use examples, examples of patterns of AI business actors, examples of benefits of AI for each industry / business operation, and risks taking actual cases as examples.
- Appendix 1 leads to a deeper understanding of the content of this guideline through these.

A. Preconditions for AI	B. AI's benefits and risks

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)
  (Clarify the relationship with the main part by tying the description to the "common guiding principles")
  (Include risks when using a generative AI)

Appendix 1 A. Preconditions for AI

- Further the understanding of AI itself, its usage scenarios, and the roles of each AI business actor, which encompass each actor, definition, etc. in the main part.
  - Specify relevant AI business actors with examples of major AI systems and services.
  - Organize the process and the roles and patterns of AI business actors for each use case.

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

5
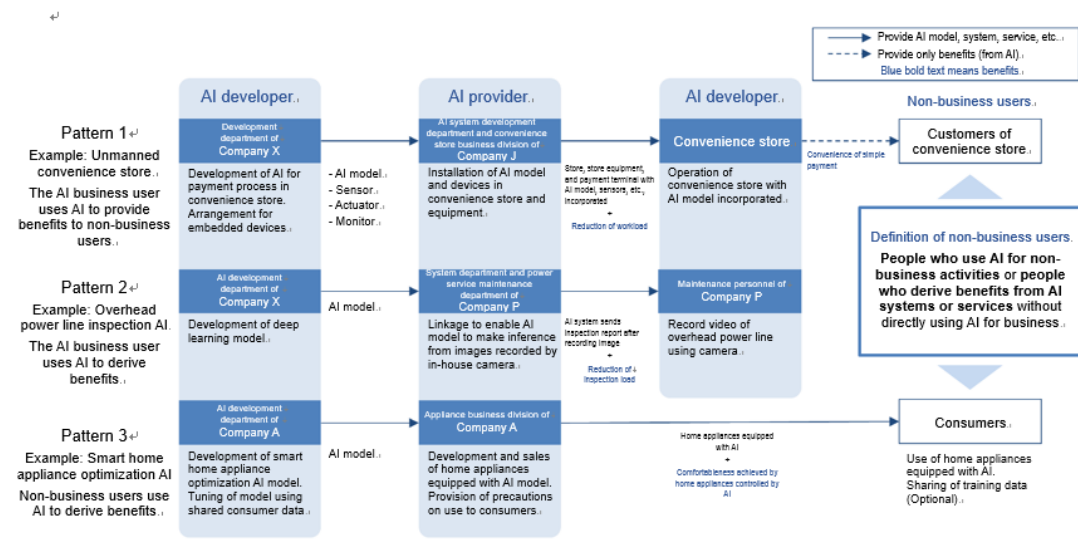
Appendix 1 B. AI's benefits and risks (AI's benefits)

- Consider the benefits of AI with a focus on the end-users who will benefit to promote its understanding.

	Development	Marketing	Sales	Logistics/distribution	Customer support	Legal	Finance	HR
Examples of benefits available conventionally↓ (Improved by generative AI)←	Automation of code verification and documentation←	Automatic distribution of ad emails←	Support after order intake← 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←	Creation of FAQ based on past inquiries←	Review of legal text←	Future prediction based on past records, and detection of malpractices←	Human resources demand matching based on résumés, etc.←
Examples of benefits unique to generative AI←	Generation of training data, coding assistant, brainstorming for new products←	Automatic creation of sales promotion (marketing materials, sales copy, etc.)←	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←	Response to in-house inquiry according to context←	Holding human resources interview according to context←

6

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

- Consider the major risks from traditional AI and the risks that have become especially apparent with generative AI.
- Deepen understanding of the necessity of AI governance by learning about the risks of AI.

	Risks	Examples	"Common Guiding Principles"
Risks from traditional AI	Output of result that includes bias or discrimination	• AI human resources recruitment system developed by an IT company had a defect in machine learning that discriminated against women.	1) Human-centric 3) Fairness
	Filter bubble and echo chamber phenomena	• The social division is caused by recommendations given by SNS, etc.	1) Human-centric
	Loss of diversity	• If the whole society uses the same model in the same way, the derived opinions and replies might converge through LLM, losing diversity.	1) Human-centric
	Inappropriate use of personal data	• The nontransparent use of personal data and the political use of personal data are problematic.	1) Human-centric 4) Privacy protection
	Infringement on lives, bodies, and properties	• During AI training, there is a risk of intrusion of invalid data into learning data, causing performance degradation and misclassification. • In medical settings, if AI has an ethical bias for determining prioritization, fairness might be lost.	2) Safety 3) Fairness
	Data poisoning attack	• During AI training and service operation, there is a risk of intrusion of invalid data into learning data and cyberattacks aimed at the application itself.	5) Ensuring security
	Black-box AI, and requirements for explanation about judgment	• Black-box AI's judgments caused a problem as well. • There is also a rising demand for transparency regarding AI's judgments.	6) Transparency 7) Accountability
	Energy consumption and environmental load	• As the use of AI spreads, the demands for calculation resources also increase. As a result, data centers are enhanced, and some people are concerned about the increase in energy consumption.	1) Human-centric
Risks that have become apparent with generative AI	Misuse	• The use of AI for fraud is also problematic.	2) Safety 8) Education/literacy
	Leak of confidential information	• In using AI, there is a risk that personal data or confidential information is entered as a prompt becomes leaked through output.	5) Ensuring security 8) Education/literacy
	Factual errors	• Response represented by generative AI as facts contained dis/misinformation, and a lawsuit was filed against an AI developer and provider.	2) Safety 8) Education/literacy
	Blindly trusting disinformation and misinformation	• Blindly trusting misinformation produced by generative AI can be a risk. • Misuse of deepfakes has occurred in various countries.	1) Human-centric 8) Education/literacy
	Relationship with copyright	• The handling of intellectual property rights is an issue that needs discussed.	2) Safety
	Relationship with qualifications, etc.	• There might be risks of infringement of legally prescribed licenses and qualifications caused by using generative AI.	2) Safety
	Reproduction of bias	• Because generative AI creates answers based on existing information, biases contained in existing information might be amplified, continuing and enhancing unjust output containing discrimination.	3) Fairness

7

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 5 companies' AI governance initiatives based on the "Governance Guidelines for Implementation of AI Principles"
  – Describes perspectives that many companies encounter when promoting AI governance

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

- 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」

Category	Behavioral Goals
1. Environmental and risk analysis	1-1 Understanding benefits/risks 1-2 Understanding social acceptance of AI 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 3-2 Improving literacy of those in charge of the AI management system 3-3 Enhancing AI management through cooperation between AI business actors and divisions 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 4-2 Ensuring that the operation of each AI system is explainable 4-3 Considering proactive disclosure of AI governance practices
5. Evaluation	5-1 Verifying AI management system functions 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)

- 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

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.

[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.
.
.
.

[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.
.
.

Descriptions

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

10

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

- 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. Matters to be observed in 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" (June 2023)
    - NIST, "AI Risk Management Framework Playbook" (January 2023)

- Hiroshima Process International Code of Conduct for Organizations Developing Advanced AI Systems

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

A. Descriptions of Part 3 "Matters Related to AI Developers"

[Contents of the main part (repeat)]
During data preprocessing and training
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.

Descriptions

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

12

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 2024, 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

<

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) Concepts of development and use of AI
  - Paying attention to the emergence of transactions that do not fit into the category of development and utilization.

(2) Development, provision and use of AI, and distribution of responsibilities
  - Responsibilities based on diversity or complexity of value chain needs to be considered.

(3) Development, provision and use of AI, and accountability
  - It is expected to sort out policies before and after execution of a contract to ensure the appropriateness of explanations, as certain explanations may be required when risks arise.

Appendix 7. Checklist
(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.
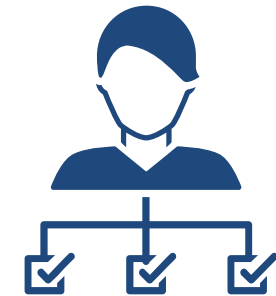
Main part・Appendix 1〜5	»	Checklist	»	Worksheet for specific approaches

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.

14

Appendix 7. Checklist Utilization

- All AI business actors confirm an overview of their own efforts by using " Appendix 7. Checklist A [for all AI business actors]."
- If the business actor is concerned with advanced AI systems, refer also to "Appendix 7. B Checklist [for AI business actors concerned with advanced AI systems]."

Japanese only

別添7 A　チェックリスト　令和6年4月19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
※高度なAIシステムに関係する事業者は、
「チェックリスト[別添7 B. 高度なAIシステムに関係する事業者向け]」も実施ください

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**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

別添7 B　チェックリスト
[高度なAIシステムに関係する事業者向け]

本チェックリストは、AI事業者ガイドライン「第2部D. 高度なAIシステムに共通の指針」を要約したものです。
高度なAIシステムに関係する事業者に該当する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(高度なAIシステムを開発するAI開発者にのみ適用される内容もあるため、AI提供者及び...

チェック項目

□ 高度なAIシステムの市場導入前及び開発全体を通じて、AIライフサイクル全体にわたる**リスクを特定、評価、軽減するための適切な措置**を講じているか？

□ 市場投入後に**脆弱性、インシデント、悪用パターンを特定、緩和**しているか？

□ 十分な透明性の確保やアカウンタビリティの向上のため高度なAIシステムの**能力、限界、適切・不適切な使用領域を公表**しているか？

□ 産業界、政府、市民社会、学界を含む関係組織間で、**責任ある情報共有とインシデントの報告**に努めているか？

□ **リスクベースのアプローチにもとづくAIガバナンス及びリスク管理方針を策定、実施、開示**しているか？

□ 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、内部脅威を含む、強固なセキュリティ管理に投資**し、実施しているか？

□ 技術的に可能な場合は、AIが生成したコンテンツを識別できるように、**電子透かしやその他の技術等、信頼性の高いコンテンツ認証及び来歴のメカニズムを開発、導入**しているか？

□ 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、効果的な軽減策に優先的に投資**しているか？

□ **気候変動、健康・教育などの世界の最大の課題**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？

□ **国際的な技術規格の開発を推進**しているか？

□ **適切なデータ入力措置**と、個人データ及び知的財産の保護を実施しているか？

□ **誤情報の拡散等のAI固有リスクに関するデジタルリテラシー**の向上や脆弱性の検知への協力と情報共有等、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？

□ 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

- 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

- 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.

Japanese only

別添７C．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ございませんので、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、F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。

各自用に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（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（見直し日）
1) 人間中心	1)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？ 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？ 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しつつ、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ていないか？				
	1)②	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？ 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、必要な対策を講じているか？ c.フィルターバブルに代表されるような情報は価値観の傾斜を助長し、AI利用者を含む人間が本来得られるべき選択				

- 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)

- 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.
- Currently, the case study features a business that handles "recruitment AI," but other cases will be added in the future.

Japanese only

Case	採用AI

機械学習モデル：XGBoost（エントリーシートの文章で、応募者に対して合否を判断する）
https://ifi.u-tokyo.ac.jp/wp/wp-content/uploads/2022/10/RCModel_Case01_Recruitment-AI_JP.pdf

No	分類	共通の指針／各主体に関する事項	AI開発部門 / AI開発者 / 本UCにおいて主体が実施している活動	人材採用部門（採用AIチーム）/ AI提供者 / 本UCにおいて主体が実施している活動	人材採用担当者 / AI利用者 / 本UCにおいて主体が実施している活動
1）人間中心					
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…					
①人間の尊厳及び個人の自律					
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(応募者の合否)を行えるようになっている(Human-in-the-loop)	AIサービスの提供において、AI利用者が最終判断(応募者の合否)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	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。※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。※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。※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
②AIによる意思決定・感情の操作等への留意					
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	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	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. 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

Japanese only

- Ethics guidelines for trustworthy AI（April 2019, EU）
- Recommendation of the Council on Artificial　Intelligence（May 2019, OECD Cabinet Meeting）
- Recommendation on the Ethics of Artificial Intelligence（November 2021, UNESCO）
- Blueprint for an AI Bill of Rights（October 2022, THE WHITE HOUSE）
- Artificial Intelligence Risk Management Framework（AI RMF 1.0）　（January 2023, NIST）
- Advancing accountability in AI（February 2023, OECD）
- Proposal for a Regulation laying down harmonized rules on artificial intelligence（AI Act）　（June 2023, EU）
- AI Product Quality Assurance Guidelines
  （June 2023, Consortium of Quality Assurance for Artificial-Intelligence-based Products and Services）
- Guidelines for Machine Learning Quality Management, 4th Edition
  （December 2023, National Institute of Advanced Industrial Science and Technology）

Guidelines, etc. cited or integrated into the guideline

- Human-centered AI social principles（March 2019, Integrated Innovation Strategy）
- The Draft AI R&D GUIDELINES for International Discussions（July 2017, The Conference toward AI Network Society）
- AI Utilization Guidelines -Practical Reference for AI utilization-（August 2019, The Conference toward AI Network Society）
- Governance Guidelines for Implementation of AI Principles Ver. 1.1
  （January 2022, Expert Group on How AI Principles Should be Implemented）
- Contract Guidance On Utilization Of AI And Data（March 2018, Ministry of Economy, Trade and Industry）
- The Hiroshima AI Process Comprehensive Policy Framework（December 2023, G7）