

## 事務局説明資料

- 資料3-1 国際的な議論及び国内の議論の動向
- 資料3-2 利活用原則案の検討の方向性及び各論点に対する考え方（案）

事 務 局  
平 成 3 1 年 3 月

# 国際的な議論及び国内の議論の動向

平成 31 年 3 月  
平 事 務 局

## 富士通「富士通グループAIコミットメント」を策定、公表【2019年（平成31年）3月13日】

- 3月13日、富士通グループは、AIの安心・安全な利用に向けて、AI関連の技術やソリューション・サービスの研究開発・提供に際し、AI倫理を含む価値観をまとめた、「富士通グループAIコミットメント」を策定。
- 外部の有識者等で構成される委員会を設置し、その客観的な評価を取締役会と共有することで、AI倫理に関するコーポレート・ガバナンスの充実を図る。
- 同コミットメントの骨子となる5つの原則は以下の通り：
  - AIによってお客様と社会に価値を提供します
  - 人を中心に考えたAIを目指します
  - AIで持続可能な社会を目指します
  - 人の意思決定を尊重し支援するAIを目指します
  - 企業の社会的責任として、AIの透明性と説明責任を重視します

## OECD デジタル経済政策委員会(CDEP)アドホック会合【2019年（平成31年）3月14-15日】

- OECDは、3月14～15日にデジタル経済政策委員会（CDEP）アドホック会合を開催し、AIの取り組みについて議論。日本からは、中大・実積教授（CDEP副議長）、総務省が参加。
- AIに関する理事会勧告案の承認：
  - AI専門家会合（AIGO）からの報告を踏まえ、OECD事務局は理事会勧告案の素案を作成し、会合に先立ち各国に意見照会。会合では、各国からの意見を踏まえて修正された理事会勧告案について議論が行われ、その結果を反映し、同勧告案が承認された。（原則の内容については概括的な事項のみを記載し、具体的に講じるべき措置等については勧告案策定後のCDEP会合（7月、11月）で別途検討。）
  - CDEPにおける勧告案に関する議論は本会合をもって終結し、同年5月の閣僚理事会で公表予定。

# 利活用原則案の検討の方向性及び各論点に対する考え方 (案)

平成 31 年 3 月  
平 事 務 局

- 利活用原則案の体系
- 目的・基本理念（候補）及び  
「人間中心のAI社会原則（案）」との関係
- 最終利用者の分類
- 各論点に対する考え方

### (1) 利活用原則案の位置づけ（「報告書2018」より抜粋）

- A I は、利活用の過程でデータの学習等により自らの出力やプログラムを継続的に変化させる可能性があることから、A I の利活用においても留意することが期待される事項があるものと想定される。また、A I の利活用においても留意することが期待される事項を整理することは、開発者と利用者、データ提供者などステークホルダ間の関係を考える上でも重要であるものと考えられる。  
（「報告書2018」p48）
- そこで、A I の利活用の促進やA I ネットワーク化の健全な進展に向けて、A I の便益の増進とリスクの抑制を図り、A I に対する信頼を醸成するため、利用者及びデータ提供者がA I の利活用において留意することが期待される事項を「A I 利活用原則案」として取りまとめた。（「報告書2018」p54）
- A I 利活用原則案は、A I 開発ガイドライン案と同様に、A I の便益及びリスクが国境を越えて広く波及することが見込まれることから、国際的な議論のためのものとするとともに、規制の導入を目指すものとするは適当ではないことから、非規制的かつ非拘束的なもの（いわゆるソフトロー）としている。（「報告書2018」p54）
- また、A I 関連技術やA I の利活用が今後とも飛躍的に発展することが期待されることから、国際的な議論を踏まえ、A I 利活用原則案を不断に見直し、必要に応じて柔軟に改定することとする。（「報告書2018」p54）

### (2) アウトプットの用途

- これまでもG 7やO E C D 専門家会合等で紹介してきており、今後もO E C D 理事会勧告等に向けた議論に貢献していく。
- 内閣府「人間中心のA I 社会原則」の「4.2. A I 開発利用原則」において、「我々は、開発者及び事業者において、基本理念及びAI社会原則を踏まえたA I 開発利用原則を定め、遵守するべきと考える」との記載あり。（内閣府資料より）
- A I が適用される用途ごとに規範等を議論する際の参照となるべき分野共通のもの。（「報告書2018」p66）

### (3) アウトプットの形式

- A I 開発ガイドライン案と同様に、目的や基本理念を整理し、原則およびその解説をとりまとめた「ガイドライン」の形とする。（「報告書2018」p65）
- 国際的な議論のために資するもの、また、民間が定める原則の参考となるものとして「ガイドライン（案）」とする。

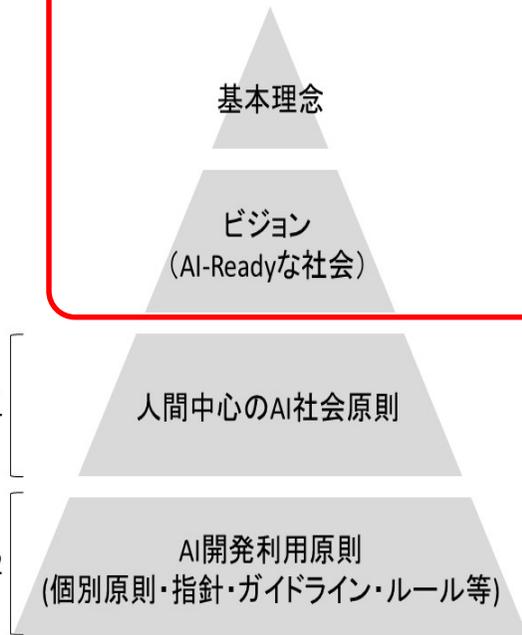
# **目的・基本理念（候補）及び 「人間中心のAI社会原則（案）」との関係**

- (原則/ガイドラインをまとめる) 目的：
  - A I ネットワーク化の健全な進展を通じて、A I の便益の増進とリスクの抑制を図り、A I に対する信頼を醸成することにより、A I の利活用や社会実装を促進する。
  
- 基本理念 (候補)：
  - 人間がA I ネットワークと共生することにより、その恵沢がすべての人によってあまねく享受され、人間の尊厳と個人の自律が尊重される**人間中心の社会を実現**すること
  - **A I の利活用において利用者の多様性を尊重し、多様な背景と価値観、考え方を持つ人々を包摂**すること
  - A I ネットワーク化の便益を増進するとともに、**民主主義社会の価値を最大限尊重**しつつ、権利利益侵害のリスクを抑制するため、**便益とリスクの適正なバランスを確保**すること
  - **利用者それぞれが** A I に関して有していると期待される**知識・能力相応の役割分担**に応じ、**適当と考えられる役割を果たす**こと
  - 非拘束的なソフトローたる**指針やベストプラクティスを国際的に共有**すること
  - A I ネットワーク化の進展等を踏まえ、国際的な議論を通じて、本ガイドライン案を**不断に見直し、必要に応じて柔軟に改定**すること

第2章

第3章

第4章



1. はじめに

2. **基本理念**（3つの価値）

(1) 人間の尊厳が尊重される社会（Dignity）

(2) 多様な背景を持つ人々が多様な幸せを追求できる社会（Diversity & Inclusion）

(3) 持続性ある社会（Sustainability）

3. Society 5.0実現に必要な社会変革「AI-Readyな社会」（ビジョン）

- 「何のためにAIを用いるのか」に答えられるような(1)人、(2)社会システム、(3)産業構造、(4)イノベーションシステム、(5)ガバナンスの在り方について提示

4. AI原則

1. **AI社会原則**（7原則→次頁に記載）：

- AI-Readyな社会において、国や自治体、さらには多国間の枠組みで留意すべき原則

2. **AI開発利用原則**：

- AIの研究開発と社会実装に従事する開発・事業者側が留意すべき原則
- オープンな議論を通じて国際的なコンセンサスを醸成し、非規制的で非拘束的な枠組として国際的に共有する必要（総務省をはじめ、各国等で検討が行われていることから、具体的な原則の内容に言及せず）

5. おわりに

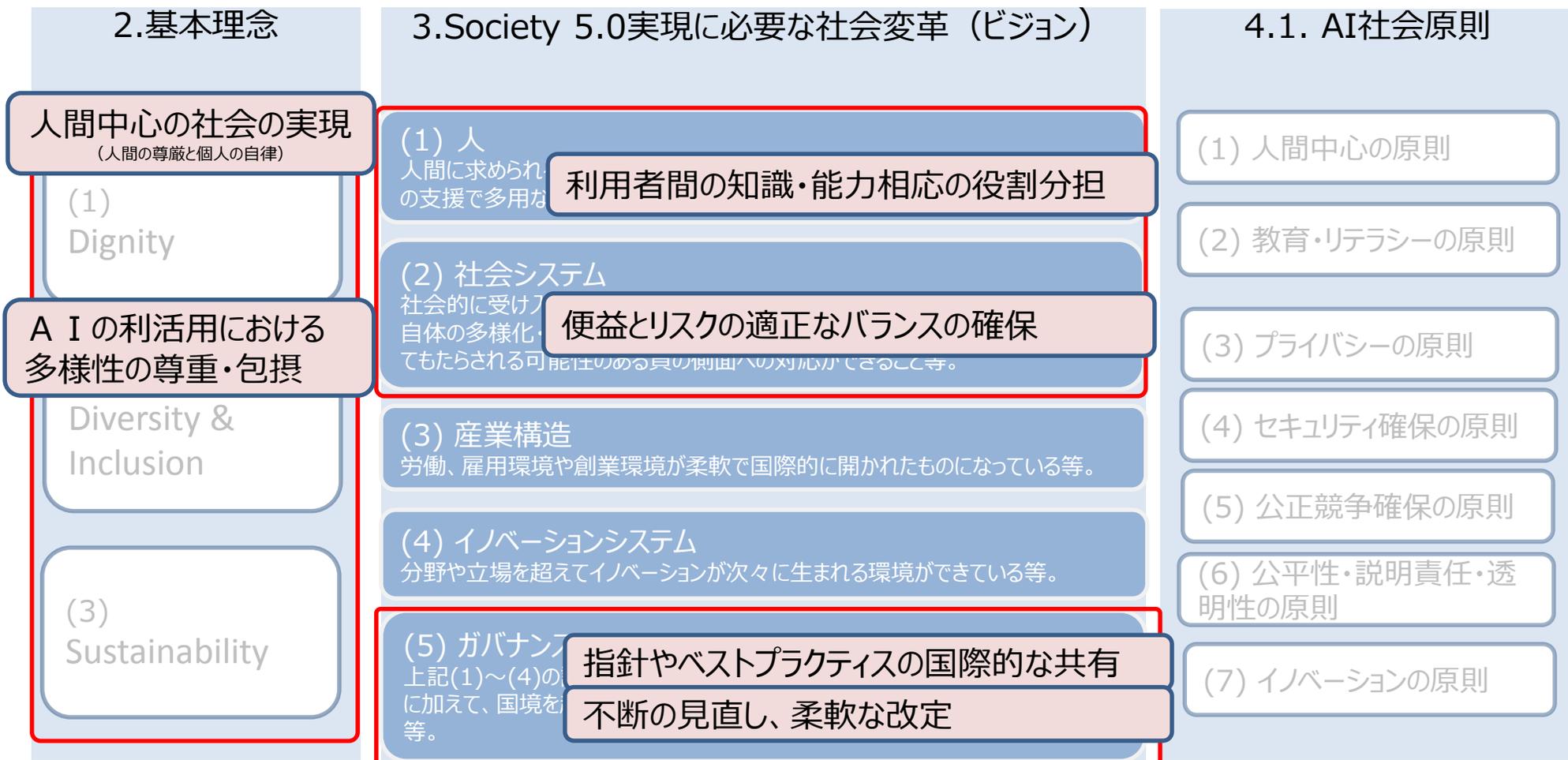
社会が留意すべき事項をとりまとめた「人間中心のAI社会原則（案）」に対し、利活用原則案では、利用者が留意すべき事項をとりまとめている。そこで、**2.基本理念に加え、3.ビジョンにおいては、利活用に関連する、利活用する「人」、利活用における「社会システム」、及びそれらに対する「ガバナンス」**を中心に、利活用原則案に対する基本理念（候補）と対応づける

## 「人間中心のAI社会原則（案）」における基本理念、ビジョン、社会原則

2.基本理念	3.Society 5.0実現に必要な社会変革（ビジョン）	4.1. AI社会原則
(1) Dignity	(1) 人 人間に求められる能力及び役割として、AIの長所・短所の理解、多様な人がAIの支援で多様な夢を実現できること等。	(1) 人間中心の原則
(2) Diversity & Inclusion	(2) 社会システム 社会的に受け入れられた既存の目的に照らした単純な効率化だけでなく、目的自体の多様化・流動化によって生まれる新たな価値の実現や、AIの進化によってもたらされる可能性のある負の側面への対応ができること等。	(2) 教育・リテラシーの原則
(3) Sustainability	(3) 産業構造 労働、雇用環境や創業環境が柔軟で国際的に開かれたものになっている等。	(3) プライバシーの原則
	(4) イノベーションシステム 分野や立場を超えてイノベーションが次々に生まれる環境ができている等。	(4) セキュリティ確保の原則
	(5) ガバナンス 上記(1)～(4)の議論・目的は常に更新されていること。国内におけるガバナンスに加えて、国境を越える問題に対処するための国際協力体制が整っていること等。	(5) 公正競争確保の原則
		(6) 公平性・説明責任・透明性の原則
		(7) イノベーションの原則

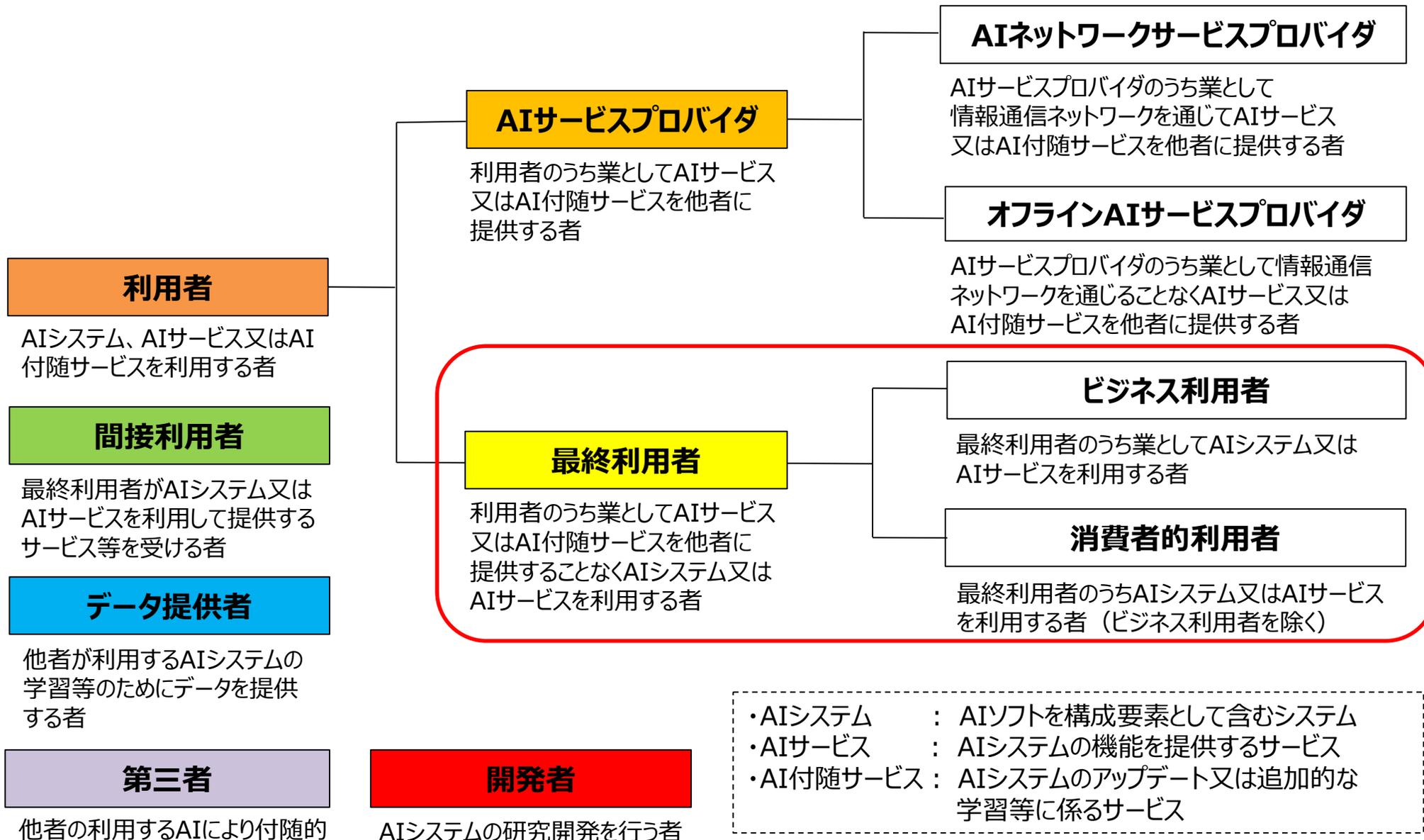
利活用原則案における基本理念（候補）は「人間中心のAI社会原則（案）」の2.基本理念、及び3.ビジョンのそれぞれに以下のように（階層的に）対応づけられる

## 「人間中心のAI社会原則（案）」における基本理念、ビジョン、社会原則



基本理念（候補）	「人間中心のAI社会原則（案）」における言及
人間中心の社会の実現 （人間の尊厳と個人の自律）	<ul style="list-style-type: none"> <li>2(1) <b>人間の尊厳が尊重される社会</b>（Dignity）</li> </ul>
A I の利活用における 多様性の尊重・包摂	<ul style="list-style-type: none"> <li>2(2) <b>多様な背景を持つ人々が多様な幸せを追求</b>できる社会（Diversity &amp; Inclusion）</li> </ul>
利用者間の知識・能力相応の 役割分担	<ul style="list-style-type: none"> <li>4.1(1) 人間中心の原則 AIの開発・提供・利用に関わった種々のステークホルダーが<b>適切に分担して責任を負うべき</b>である。</li> </ul>
便益とリスクの適正なバランスの 確保	<ul style="list-style-type: none"> <li>4.1(3) プライバシー確保の原則 その（パーソナルデータの）<b>利活用と保護のバランス</b>については、文化的背景や社会の共通理解をもとにきめ細やかに検討される必要</li> <li>4.1(4) セキュリティ確保の原則 社会は、常に<b>ベネフィットとリスクのバランス</b>に留意</li> </ul>
指針やベストプラクティスの 国際的な共有	<ul style="list-style-type: none"> <li>3(5)ガバナンス 国境を越える問題に対処するための<b>国際協力体制が整っている</b>ことが求められる。</li> <li>5. おわりに <b>国際的な議論の場において、我が国は、本原則を世界各国と共有した上で、国際的な議論のリーダーシップをとり、コンセンサスの形成を目指す</b>べき</li> </ul>
不断の見直し、柔軟な改定	<ul style="list-style-type: none"> <li>3(5)ガバナンス 社会情勢の変化や技術の進展に伴い（中略）<b>常に更新し続ける必要</b>がある。</li> <li>5. おわりに 本原則は、今後、AI関連技術の進展、社会の変化、世界の情勢等に応じて、<b>今後柔軟に進化・発展させるもの</b>である。</li> </ul>

# 最終利用者の分類



(注) それぞれの名称は仮称。  
(注) 同一の個人・事業者が複数の主体に該当する場合がある。

## (最終) 利用者の分類の修正の方向性

### ■ 本検討会第4、5回にてご提示した分類

#### ① 「ビジネス利用者」(能動的利用者) :

A Iシステム又はA Iサービス等において、A Iの学習等を含む運用等を自ら行う者。

#### ② 「消費者的利用者」(受動的利用者) :

A Iシステム又はA Iサービスを自ら運用等を行うことなく利用するのみの者。ただし、利用形態が②の場合にも、業として利用する場合には、①に記載の内容を踏まえる必要があることが期待される場合があることに留意する。

### 構成員の皆様より頂戴したご意見

- AIに関する十分な知見を有していないビジネス利用者は、社会生活上の地位に基づき継続して行為(業務)を行っており、消費者と同様に消費者的利用者と位置づけるのは適切ではないのではないか。
- 留意すべき事項を整理した上で、それに基づき利用者を分類していく必要があるのではないか。



#### 「ビジネス利用者」:

「報告書2018」のとおり、「最終利用者のうち業としてA Iシステム又はA Iサービスを利用する者」とした上で、以下の注記を付してはどうか。

「ビジネス利用者であっても、A Iシステム又はA Iサービスを自ら運用等を行うことなく利用するのみの者は、サービスプロバイダや他のビジネス利用者と同等の留意を期待することが困難であることも想定される。その場合には、開発者やサービスプロバイダに対し、適切な対応を依頼することが期待される。」

#### 「消費者的利用者」:

「報告書2018」のとおり、「最終利用者のうちA Iシステム又はA Iサービスを利用する者(ビジネス利用者を除く)」とした上で、以下の注記を付してはどうか。

「消費者的利用者であっても、A Iシステム又はA Iサービス等において、A Iの学習等を含む運用等を自身で行う者にあたる場合には、開発者やサービスプロバイダに対してと同等の留意が求められる場合がある。」

# 各論点に対する考え方

原案	修正案
<p data-bbox="147 679 654 815">「報告書2018」 に記載の内容</p>	<p data-bbox="969 291 1908 579">「サービスプロバイダ、ビジネス利用者」 （A I サービス、システム等を提供、 運用（学習等を含む）する者）の 視点で記載した内容</p>
	<p data-bbox="934 943 1960 1158">「消費者的利用者」 （A I サービス、システムを利用するのみ の者）の視点で記載した内容</p>

## ①適正利用の原則（全体構成）

原案	修正案
<p>利用者は、人間とA Iシステムとの間及び利用者間における適切な役割分担のもと、適正な範囲及び方法でA Iシステム又はA Iサービスを利用するよう努める。</p>	<p>利用者は、人間とA Iシステムとの間及び利用者間における適切な役割分担のもと、適正な範囲及び方法でA Iシステム又はA Iサービスを利用するよう努める。</p>
<p>論点 ア) 適正な範囲・方法での利用</p> <p>イ) A Iの便益とリスクの適正なバランス</p> <p>ウ) A Iソフトのアップデート及びA Iの点検・修理等</p> <p>エ) 人間の判断の介在</p> <p>オ) 利用者間の役割分担</p> <p>カ) 関係者間の協力</p>	<p>論点 ア) 適正な範囲・方法での利用</p> <p>イ) 人間の判断の介在</p> <p>ウ) 関係者間の協力</p> <p>基本理念にしてはどうか</p> <p>ア) に統合してはどうか</p> <p>基本理念にしてはどうか</p>

# ①ーア) 適正な範囲・方法での利用

## 原案

①ーア)  
 利用者は、開発者等からの情報提供や説明を踏まえ、社会的文脈や状況にも配慮して、A Iを適正な範囲・方法で利用することが期待されるのではないか。(a)

また、利用者は、A Iの性質、利用の態様等に応じて、利用する前に、便益及びリスクを認識し、適正な用途を理解するとともに、必要な知識・技能を習得すること等が期待されるのではないか。(b)

さらに、利用者は、自らのA Iの利活用が適正な範囲・方法で行われているか定期的に確認することが期待されるのではないか。(c)

①ーウ)  
 利用者は、利活用の過程を通じて、A Iの機能を向上させ、リスクを抑制するため、A Iソフトのアップデート及びA Iの点検・修理等を行うよう努めることが期待されるのではないか。(d)

## 修正案

A Iサービスプロバイダ及びビジネス利用者は、A Iを消費者的利用者等に提供し、又は、自ら運用するに当たり、開発者等からの情報提供や説明を踏まえ、以下に関する情報を適時適切に提供することが期待される。(a)

[提供すべき情報]

- 提供するA Iの利用に関する適正な用途・方法。(a)
- A Iの性質、利用の態様等に応じた、便益及びリスクに関する情報。(b)
- 提供するA Iの利活用の範囲・方法に関する定期的な確認方法（特に、A Iが自律的に更新される場合の観測、確認方法）、及び確認の重要性、頻度、未確認によるリスク等。(c+)
- 利活用の過程を通じて、A Iの機能を向上させ、リスクを抑制するため、A Iソフトのアップデート及びA Iの点検・修理等を行うための情報。(d)

[情報を提供すべきタイミング]

- A Iの利用前に当該情報を提供できることが望ましい。(b)
- 事前に当該情報を提供できない場合に備え、A Iの性質、利用の態様等に基づき考えられるリスクに応じ、消費者的利用者等からのフィードバックに対応する体制が整備されていることが望ましい。(+)

また、利活用の過程を通じて、A Iの機能を向上させ、リスクを抑制するため、A Iソフトのアップデート及びA Iの点検・修理等を提供することが期待される。特に、アップデートのための機能を提供する際に、他のA Iとの関係でリスク<sup>1</sup>が想定される場合は、当該リスク情報を提示した上で提供することが望ましい。

また、提供されるA Iの性質、利用の態様等によっては、提供対象となる利用者が当該A Iを提供するにふさわしい者であるか（信頼性）について事前に考慮することが期待される場合も想定される。

1)アップデートを適用するA Iの動作が周辺A Iに影響を及ぼす場合等に留意が必要となる。例えば、家庭内の家電に含まれるA Iソフトの動作がアップデートにより変わった場合、全体を統括する家庭内執事のようなロボットや周辺のA Iを含む家電がアップデートの内容を知らないと、（家電同士、および家電とロボットの）相互の判断に齟齬が生じる場合がある（「報告書2018」p.138の事例）。

消費者的利用者は、開発者、A Iサービスプロバイダ、ビジネス利用者等からの情報提供や説明を踏まえ、社会的文脈や状況にも配慮して、A Iを適正な範囲・方法で利用することが期待される。具体的には以下に留意することが期待される。(a)

[実施が期待される内容]

- 利用前
  - A Iの性質、利用の態様等に応じて、便益及びリスクを認識し、適正な用途を理解するとともに、必要な知識・技能を習得すること等。(b)
- 利用中
  - 自らのA Iの利活用が適正な範囲・方法で行われているか定期的に確認すること。(c)
  - 利活用の過程を通じて、A Iの機能を向上させ、リスクを抑制するため、A Iソフトのアップデート及びA Iの点検・修理等を行うよう努めること。ただし、他のA Iとの関係でのリスクが存在しうることを<sup>1</sup>理解した上で、アップデートを行うこと。(d)
  - 何らかの問題が発生した場合、問題が起こる予兆があった場合、もしくは、問題に対するフィードバックを要求された場合、開発者及びサービスプロバイダ等に対し、当該情報をフィードバックすること。(+)

# ①ーイ) 人間の判断の介在

原案	修正案
<p>A I によりなされた判断について、必要かつ可能な場合（例：A I を用いた医療等）には、その判断を用いるか否か、あるいは、どのように用いるか等に関し、人間の判断を介在させることが期待されるのではないか。その場合、人間の判断の介在の要否の基準をどのように考えるか。(a)</p> <p>また、アクチュエータ等を通じて稼働するA I の利活用において、一定の条件に該当することにより人間による稼働に移行することが予定されている場合、どのような事項に留意することを期待することが適切か。(b)</p> <p>[基準として考えられる観点(例)]</p> <ul style="list-style-type: none"> <li>• A I の判断に影響を受ける間接利用者等の権利・利益の性質及び意向</li> <li>• A I の判断の信頼性の程度（人間の判断の信頼性との優劣）</li> <li>• 人間の判断に必要な時間的猶予</li> <li>• 利用者に期待される能力 など</li> </ul>	<p>A I によりなされた判断について、必要かつ可能な場合には、その判断を用いるか否か、あるいは、どのように用いるか等に関し、人間の判断を介在させることが期待される。その場合、人間の判断の介在の要否については、例えば以下の基準の下、利用分野、用途等に応じて検討されることが期待される。(a)</p> <p>[人間の判断の介在の要否について、基準として考えられる観点(例)]</p> <ul style="list-style-type: none"> <li>• A I の判断に影響を受ける最終利用者等の権利・利益の性質及び意向（例えば、人生を左右しかねない意思決定に係わるA I の利用時等）</li> <li>• A I の判断の信頼性の程度（人間による判断の信頼性との優劣）</li> <li>• 人間の判断に必要な時間的猶予</li> <li>• 利用者に期待される能力</li> <li>• 判断対象の要保護性（例えば、A I による大量申請への対応等）</li> </ul> <p>また、A I の判断に対し、人間が最終判断をすることが適当とされている場合に、人間がA I と異なる判断をすることが期待できなくなることも想定される。こうした場合に人間が行うべき判断についてその項目、手段などを明確化することにより、人間の判断の実効性を確保することが考えられる。(+) </p> <p>[実効性を確保するための手段の例]</p> <ul style="list-style-type: none"> <li>• 説明可能性を有するA I から得られる説明を前提として、最終判断に必要なチェック項目の作成</li> <li>• A I の判断の適正性を確認するためのチェック項目の作成（能動的な判断の実施（他のA I を利用したダブルチェック、A I への入力を振動させることによるA I 動作の確認など））</li> </ul> <p>また、アクチュエータ等を通じて稼働するA I の利活用において、一定の条件に該当することにより人間による稼働に移行することが予定されている場合、移行前、移行中、移行後等の各状態に伴い、予め責任の所在が明確になっている必要がある。また、前述の移行条件、移行方法等を利用者に事前に告知し、必要な訓練などを前もって実施するなど、人間による稼働に移行した場合に問題が起こらないよう、注意喚起をしておくことが期待される。(b)</p>
	<p>消費者的利用者は、A I の判断に対し、人間が最終判断をすることが適当とされている場合に、適切に判断ができるよう能力を習得しておくことが期待される。加えて、こうした場合に人間が行うべき判断について開発者、AIサービスプロバイダ、ビジネス利用者等により人間が行うべき判断についての項目、手段などが整理されている場合は、当該情報を入手し、これに基づき対応することが期待される。(+) </p> <p>また、アクチュエータ等を通じて稼働するA I の利活用において、一定の条件に該当することにより人間による稼働に移行することが予定されている場合、移行前、移行中、移行後等の各状態に伴う責任の所在を予め認識しておく必要がある。また、前述の移行条件、移行方法等についての説明をサービスプロバイダ等から受け、必要な訓練などを受けておくことが期待される。(b)</p>

## ①ーウ) 関係者間の協力

原案	修正案
<p>利用者及びデータ提供者は、A I の利活用により生じ得る又は生じた事故、セキュリティ侵害、プライバシー侵害等による被害の性質・態様等に応じて、関係者と協力して予防措置及び事後対応（情報共有、停止・復旧、原因解明、再発防止措置等）に取り組むことが期待されるのではないか。(a)</p> <p>また、その実効性を確保するために利用者にとどのようなことを期待することが適当か。(b)</p>	<p>A I サービスプロバイダ、ビジネス利用者及びデータ提供者は、A I を提供または利用するに当たり、A I の利活用により生じ得る又は生じた事故、セキュリティ侵害、プライバシー侵害等による被害の性質・態様等に応じて、関係者と協力して予防措置及び事後対応（情報共有、停止・復旧、原因解明、再発防止措置等）に取り組むことが期待される。(a)</p> <p>また、具体的には、例えば以下に記載の内容等に留意することが期待される。(b)</p> <p>[関係者間で協力して行う予防措置（例）]</p> <ul style="list-style-type: none"> <li>• ①適正利用の原則（本原則） 論点ア：適正な範囲・方法での利用（適正な範囲・方法による利用のための情報の相互提供等）</li> <li>• ④安全の原則（A I がアクチュエータ等を通じて人の生命・身体・財産に危害を及ぼした場合に講ずるべき措置等）</li> <li>• ⑤セキュリティの原則（セキュリティが侵害された場合に講ずるべき措置等）</li> <li>• ⑥プライバシーの原則（他者のプライバシーを侵害した場合に講ずるべき措置等）等</li> </ul> <p>消費者的利用者は、A I を利用するに当たり、A I の利活用により生じ得る又は生じた事故、セキュリティ侵害、プライバシー侵害等による被害の性質・態様等に応じて、関係者と協力して予防措置及び事後対応（情報共有、停止・復旧、原因解明、再発防止措置等）に取り組むことが期待される。(a)</p> <p>また、その実効性を確保するため、開発者、AIサービスプロバイダ及びビジネス利用者等が提供する情報に基づき、関係者と協力して適切に対応することが期待される。(＋)</p>

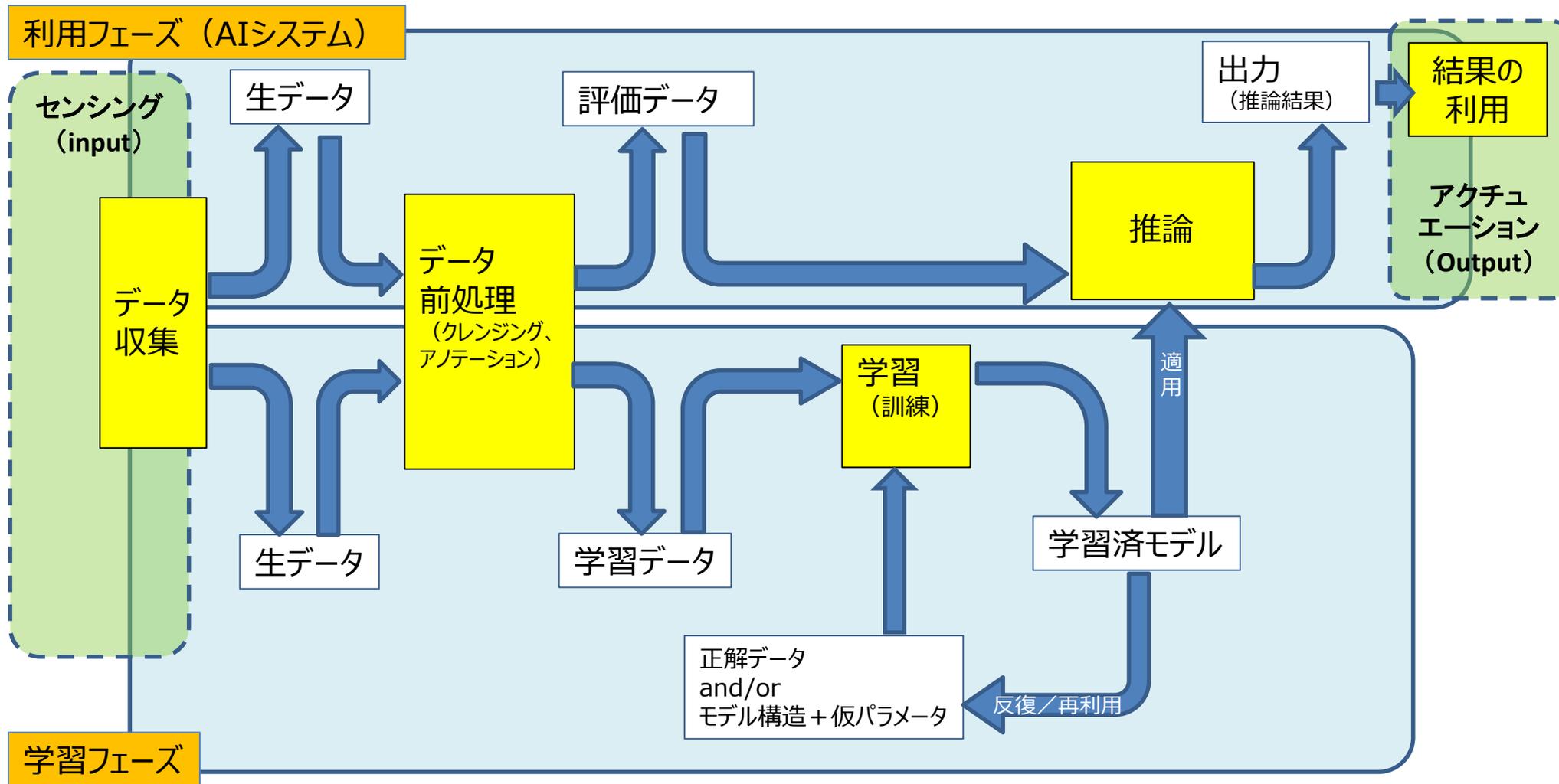
原案	修正案
<p>イ) A I の便益とリスクの適正なバランス A I サービスプロバイダ及びビジネス利用者は、A I のリスクを適切に評価した上で、生産性の向上や業務の効率化のためにA I の積極的な利活用を検討するなど、A I の便益とリスクの適正なバランスに配慮することが期待されるのではないか。</p>	
<p>オ) 利用者間の役割分担 利用者は、それぞれがA I に関して有していると期待される能力及び知識の多寡並びに必要な措置を実施する容易性を踏まえ、適当と考えられる役割を果たすとともに、責任を負担することが期待されるのではないか</p>	<p>それぞれ 「基本理念」の1つとして 位置づけてはどうか。</p>

原案	修正案
<p data-bbox="43 205 982 311">利用者及びデータ提供者は、A I システムの学習等に用いるデータの質に留意する。</p> <p data-bbox="43 376 132 418">論点</p> <p data-bbox="43 429 872 471">ア) A I の学習等に用いるデータの質への留意</p> <p data-bbox="43 529 996 621">イ) 不正確又は不適切なデータの学習等によるA I のセキュリティ脆弱性への留意</p>	<p data-bbox="1050 205 1989 311">利用者及びデータ提供者は、A I システムの学習等に用いるデータの質に留意する。</p> <p data-bbox="1050 376 1147 418">論点</p> <p data-bbox="1050 429 1879 471">ア) A I の学習等に用いるデータの質への留意</p> <p data-bbox="1050 529 1960 621">イ) 不正確又は不適切なデータの学習等によるA I のセキュリティ脆弱性への留意</p>

構成変更無し

## ②ーア) A I の学習等に用いるデータの質への留意

原案	修正案
<p>利用者及びデータ提供者は、利用するA Iの特性及び用途を踏まえ、A Iの学習等に用いるデータの質（正確性や完全性など）に留意することが期待されるのではない。また、A Iによりなされた判断の精度が損なわれたり、低下した場合には、データの質に留意して改めて学習させることが期待されるのではない。</p> <p>どのような場合に、どの程度、学習等に用いるデータの正確性に留意することを期待することが適当か。</p>	<p>A I サービスプロバイダ、ビジネス利用者及びデータ提供者は、提供もしくは利用するA Iの特性及び用途を踏まえ、A Iの学習等に用いるデータの質（正確性や完全性など）に留意することが期待される。なお、特に機械学習においてデータの質を担保するための方法として、例えば以下の対策が考えられる。</p> <p>また、A Iによりなされた判断の精度が損なわれたり、低下することが想定されるため、<b>想定される権利侵害の規模・権利侵害の生じる頻度、実装コスト、及び、技術水準<sup>1</sup>を踏まえ</b>、精度に関する基準を予め定めておくことが期待される。他方、精度が当該基準を下回った場合には、データセットの質に留意して改めて学習させることが期待される。</p> <p><b>加えて、消費者的利用者から提供されるデータを用いることが予定されている場合には、提供もしくは利用するA Iの特性及び用途を踏まえ、データ提供の手段、形式等について、消費者的利用者に情報を提供することが期待される。</b></p> <p>[データ収集時の対策（例）]</p> <ul style="list-style-type: none"> <li>• 前提として、過去のデータ収集時の経験を活かし、収集するデータが目的に適ったものかを確認する。</li> <li>• 社会的に信用の高い者が公開するデータなど<b>データ作成来歴を確認した上で収集する。</b></li> <li>• 自らデータを収集する際には、データに付随する権利にも注意する。特に個人の情報については同意の取得など細心の注意が必要となる。</li> </ul> <p>[データ前処理時の対策（例）]</p> <ul style="list-style-type: none"> <li>• 人間でも判定が困難と考えられるデータは、学習等の対象から除外する<sup>2</sup>。</li> <li>• 一方、機械（学習器）から見て間違われやすいと考えられるデータは積極的に学習の対象とする<sup>3</sup>。</li> <li>• （特に教師あり学習等で）アノテーション（ラベル付与）を行う際には、誤って行わないよう留意する。</li> <li>• 利用時に利用（入力）されるデータの形式を意識してデータセットを作成する。</li> <li>• <b>前処理をどのように行ったのか（データ前処理に関する来歴）について、ログを取得するなど明確にしておく。</b></li> </ul> <p>[学習時の対策（例）]</p> <ul style="list-style-type: none"> <li>• 既存の学習モデルを利用して転移学習<sup>4</sup>等を行う。</li> <li>• 学習の精度を上げるため、特定のデータを拡張<sup>5</sup>した上で学習を行う。</li> <li>• 特に一過性のある時系列データを学習する場合などは、どの範囲のデータを学習対象とすべきかを見極める。</li> </ul> <p><small>1)例えば、機械学習を中心としたA Iが帰納的であることから、当該A I単体で原理的に100%の精度を担保できないこと等が挙げられる。 2)例えば、画像認識などで、対象となるオブジェクトが人間の目で見て同定できない場合など。 3)例えば、画像認識などで、対象となるオブジェクトが端にあるなど。 4)転移学習(Transfer Learning)とは、深層学習を含む機械学習で用いられる技術の1つで、特定の領域（ドメイン）で学習させたモデルを別の領域に適用する技術である。少ないデータで精度の高い学習結果を得ることが出来る可能性がある点がメリットである。 5)「データの拡張」(Data Augmentation)とは、データの正確性を高めるに当たって、特定の学習データが少ない際に、汎化性能（未知のデータに対する性能）を高めるためにとられる手段の1つである。学習に用いる当該データを拡張し（例えば画像データであれば、反転、拡大、縮小を適用し）それぞれを別のソースとして用いることにより、汎化性能が改善されることがある。</small></p> <p>消費者的利用者は、利用するA I等の学習に用いるデータを自ら収集することが<b>予定されている場合には</b>、（データの提供に関する）手段、形式等について、開発者、A I サービスプロバイダ、ビジネス利用者等からの情報を踏まえた上でデータの収集、保存を行うことが望ましい。</p>



## ②ーイ) 不正確又は不適切なデータの学習等によるA Iのセキュリティ脆弱性への留意

原案	修正案
<p>利用者及びデータ提供者は、A Iが不正確又は不適切なデータを学習することにより、A Iのセキュリティに脆弱性が生ずるリスクに留意することが期待されるのではないか。</p>	<p>A Iサービスプロバイダ、ビジネス利用者及びデータ提供者は、A Iが不正確又は不適切なデータを学習することにより、AIのセキュリティに脆弱性が生じるリスクが存在することに留意することが期待される。また、消費者的利用者に対し、そのようなリスクが存在することを予め周知することが期待される。</p> <p>[リスクの例]</p> <ul style="list-style-type: none"> <li>• 学習モデルが、一部学習データからの微少な変動により悪影響を受ける利用されたデータが不均衡な場合に、悪意あるインプット等により当該学習モデルが誤った結果を導出してしまいう攻撃が存在する可能性（例：Adversarial example攻撃）</li> <li>• 学習において不正確なラベリング等がなされたデータを混在させることで、間違った学習が行われる可能性</li> </ul> <p>消費者的利用者は、サービスプロバイダ、ビジネス利用者及びデータ提供者等からの情報を踏まえ、A Iが不正確又は不適切なデータを学習することにより、A Iのセキュリティに脆弱性が生じるリスクが存在することを認識しておくことが期待される。</p> <p>また、A Iを利用するに当たり、セキュリティ上の疑問を感じた場合は、A Iサービスプロバイダ、ビジネス利用者、データ提供者等にその旨を報告することが期待される。</p>

原案	修正案
<p>A I サービスプロバイダ、ビジネス利用者及びデータ提供者は、A I システム又はA I サービス相互間の連携に留意する。また、利用者は、A I システムがネットワーク化することによってリスクが惹起・増幅される可能性があることに留意する。</p> <p>論点</p> <p>ア) 相互接続性と相互運用性への留意</p> <p>イ) データ形式やプロトコル等の標準化への対応</p> <p>ウ) A I ネットワーク化により惹起・増幅される課題への留意</p>	<p>A I サービスプロバイダ、ビジネス利用者及びデータ提供者は、A I システム又はA I サービス相互間の連携に留意する。また、利用者は、A I システムがネットワーク化することによってリスクが惹起・増幅される可能性があることに留意する。</p> <p>論点</p> <p>ア) 相互接続性と相互運用性への留意</p> <p>イ) データ形式やプロトコル等の標準化への対応</p> <p>ウ) A I ネットワーク化により惹起・増幅される課題への留意</p>

構成変更無し

原案	修正案
<p>A I ネットワークサービスプロバイダは、利用するA I の特性及び用途を踏まえ、A I ネットワーク化の健全な進展を通じて、A I の便益を増進するため、A I の相互接続性と相互運用性に留意することが期待されるのではないか。</p>	<p>A I ネットワークサービスプロバイダは、利用するA I の特性及び用途を踏まえ、A I ネットワーク化の健全な進展を通じて、A I の便益を増進するため、A I の相互接続性と相互運用性に留意することが期待される。</p>
	N/A

### ③ーイ) データ形式やプロトコル等の標準化への対応

原案	修正案
<p>A I サービスプロバイダ及びビジネス利用者は、A I 相互間及びA I と他のシステム等との連携を促進するため、データ形式やプロトコル等の標準化に対応することが期待されるのではないか。また、データ提供者についても、データ形式の標準化に対応することが期待されるのではないか。</p>	<p>A I サービスプロバイダ及びビジネス利用者は、A I 相互間及びA I と他のシステム等との連携を促進するため、以下A I の入出力等におけるデータ形式（構文（syntax）及び意味（semantics）<sup>1)</sup>）や、連携のための接続方式、特にネットワークを介す場合は各レイヤにおけるプロトコル等の標準に準拠することが期待される。</p> <p>また、データ提供者についても、A I 相互間及びA I と他のシステム等との連携を促進するため、データ形式（構文（syntax）及び意味（semantics）<sup>1)</sup>）の標準に準拠することが期待される。</p> <p><small>1)データの構文だけが示されていても、意味が示されていないと連携は正しく動作しない。</small></p> <p>消費者的利用者は、利用するA I 等の学習に用いるデータを自ら収集することが予定されている場合には、データの形式について、開発者、A I サービスプロバイダ、ビジネス利用者等から提供された情報を踏まえた上で収集、保存を行うことが期待される。</p>

原案	修正案
<p>A I が連携することによって便益が増進することが期待されるが、利用者は、自ら利用するA I がインターネット等を通じて他のA I 等と接続・連携することにより制御不能となる等、A I がネットワーク化することによってリスクが惹起・増幅される可能性があることに留意することが期待されるのではないか。</p> <p>[A I がネットワーク化することによってリスクが惹起・増幅される可能性の例]</p> <ul style="list-style-type: none"> <li>• 個別の事業者のトラブル等がシステム全体に波及するおそれ</li> <li>• A I システム間の連携・調整が成立しないなどのおそれ</li> <li>• A I の判断・意思決定を検証できないおそれ（システム間の相互作用が複雑となり解析が困難になるおそれ）</li> <li>• 少数のA I の影響力が強くなりすぎるなどのおそれ（少数のA I の判断によって企業や個人が不利な立場になるなどのおそれ）</li> <li>• 領域横断での情報の共有と特定の基盤的なA I への情報の集中によるプライバシー侵害のおそれ</li> <li>• A I が想定外の動作を行うなどのおそれ</li> </ul>	<p>A I が連携することによって便益が増進することが期待されるが、A I サービスプロバイダ及びビジネス利用者は、他者に提供し、または、自ら利用するA I がインターネット等を通じて他のA I 等と接続・連携することにより制御不能となる等、A I がネットワーク化することによってリスクが惹起・増幅される可能性があることに留意した上で、提供・利用するA I の設計を行うことが期待される。また、開発者等からの情報を基に考えられるリスクを分析し、当該情報を連携の相手方と共有するとともに、問題が生じた場合の対応策等を作成の上、消費者的利用者に情報提供することが期待される。</p> <p>[A I がネットワーク化することによってリスクが惹起・増幅される可能性の例]</p> <ul style="list-style-type: none"> <li>• 個別の事業者のトラブル等がシステム全体に波及するおそれ</li> <li>• A I システム間の連携・調整が成立しないなどのおそれ</li> <li>• A I の判断・意思決定を検証できないおそれ（システム間の相互作用が複雑となり解析が困難になるおそれ）</li> <li>• 少数のA I の影響力が強くなりすぎるなどのおそれ（少数のA I の判断によって企業や個人が不利な立場になるなどのおそれ）</li> <li>• <b>多数のA I が同一の判断をし、又は行動をとることにより、市場における競争が機能しなくなるおそれ</b></li> <li>• 領域横断での情報の共有と特定の基盤的なA I への情報の集中によるプライバシー侵害のおそれ</li> <li>• A I が想定外の動作を行うなどのおそれ</li> </ul> <p>A I が連携することによって便益が増進することが期待されるが、消費者的利用者は、自ら利用するA I がインターネット等を通じて他のA I 等と接続・連携することにより制御不能となる等、A I がネットワーク化することによってリスクが惹起・増幅される可能性があることに留意することが期待される。また、問題が生じた場合の対応策等について、開発者、A I サービスプロバイダ及びビジネス利用者等から情報提供があった場合には、利用にあたり留意することが期待される。</p> <p>[A I がネットワーク化することによってリスクが惹起・増幅される可能性と対応策の例]</p> <ul style="list-style-type: none"> <li>• 個別の事業者のトラブル等がシステム全体に波及する可能性がある場合には、開発者、A I サービスプロバイダ及びビジネス利用者等からの情報を踏まえ、当該事業者のシステムを停止させる、もしくは、システム全体を停止させるなどの処理を行う。</li> <li>• A I が想定外の動作を行う場合は、開発者、A I サービスプロバイダ及びビジネス利用者等からの情報を踏まえ、他A I または他システムとの連携を解消する、もしくは当該A I システム全体を停止させる等の処理を行う。</li> <li>• 領域横断での情報の共有と特定の基盤的なA I への情報の集中によるプライバシー侵害があることに留意し、問題の予兆が見られる場合には、当該事業者にその旨報告する。</li> </ul>

原案	修正案
<p data-bbox="43 205 1002 422">利用者は、A I システム又はA I サービスの利活用により、アクチュエータ等を通じて、利用者等及び第三者の生命・身体・財産に危害を及ぼすことがないように配慮する。</p> <p data-bbox="43 482 756 582">論点 ア) 人の生命・身体・財産への配慮</p>	<p data-bbox="1048 205 1987 422">利用者は、A I システム又はA I サービスの利活用により、アクチュエータ等を通じて、利用者等及び第三者の生命・身体・財産に危害を及ぼすことがないように配慮する。</p> <p data-bbox="1048 482 1761 582">論点 ア) 人の生命・身体・財産への配慮</p>

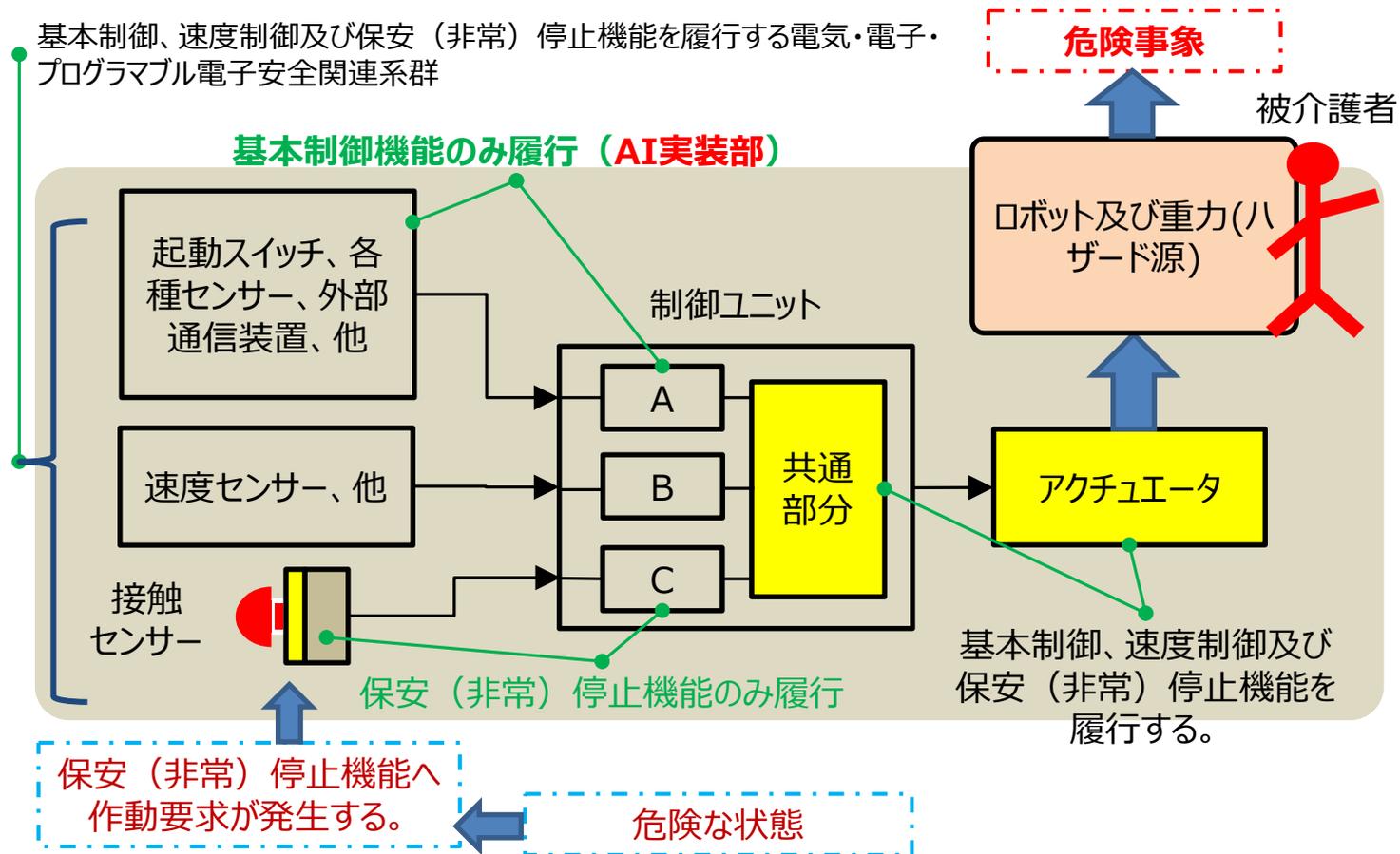
構成変更無し

## ④ーア 人の生命・身体・財産への配慮

原案	修正案
<p>医療や自動運転など人の生命・身体・財産に危害を及ぼし得る分野で A I を利活用する場合には、利用者は、想定される被害の性質・態様等を踏まえ、必要に応じて A I の点検・修理及び A I ソフトのアップデートを行うことなどにより、A I がアクチュエータ等を通じて人の生命・身体・財産に危害を及ぼすことのないよう配慮することが期待されるのではないか。</p> <p>また、利用者は、A I がアクチュエータ等を通じて人の生命・身体・財産に危害を及ぼした場合に講ずるべき措置について、あらかじめ整理しておくことが期待されるのではないか。</p>	<p>人の生命・身体・財産に危害を及ぼし得る分野で A I を利活用する場合には、A I サービスプロバイダ及びビジネス利用者は、想定される被害の性質・態様等を踏まえ、開発者等からの情報を基に、必要に応じて下記の対応策を講じることにより、A I がアクチュエータ等を通じて人の生命・身体・財産に危害を及ぼすことのないよう配慮することが期待される。</p> <p>[対処方法の例]</p> <ul style="list-style-type: none"> <li>• A I の点検・修理及び A I ソフトのアップデートを行う、また消費者的利用者にこれらの実施を促す</li> <li>• A I にはリスクがあるとの前提で、当該 A I の外部で安全を確保できる仕組み<sup>1</sup>を構築する</li> <li>• <b>被害者の救済のために保険を利用する</b></li> </ul> <p>また、A I サービスプロバイダ及びビジネス利用者は、A I がアクチュエータ等を通じて人の生命・身体・財産に危害を及ぼした場合に講ずるべき措置について、あらかじめ整理しておくことが期待される。加えて、当該措置について、消費者的利用者に対し、必要な情報提供を行うことが期待される。</p> <p>[危害時の措置の例]</p> <ul style="list-style-type: none"> <li>• <b>初動措置（当該AIを含むシステムの急用度等の文脈に応じ、必要な手順にて実施）</b> <ul style="list-style-type: none"> <li>➢ 当該システムのロールバック、代替システムの利用などによる復旧</li> <li>➢ システムの停止（キルスイッチ）：可能な場合</li> <li>➢ ネットワークからの遮断：可能な場合</li> <li>➢ 危害の内容の確認</li> <li>➢ 関係者への報告</li> </ul> </li> <li>• <b>補償・賠償等（補償・賠償等を円滑に行うための保険の利用）</b></li> <li>• <b>重大な損害が生じた場合等は、第三者機関の設置とその機関による原因調査・分析・提言など</b></li> </ul> <p><small>1) A I 単体で技術的に安全性を保證することが困難な状況では、A I 以外のシステムにより A I 実装システムの安全確保を実施し、当該システムの運用経験により AI の安全性を実証していくことも可能である。</small></p>
	<p>人の生命・身体・財産に危害を及ぼし得る分野で A I を利活用する場合には、消費者的利用者は、想定される被害の性質・態様等を踏まえ、開発者、A I サービスプロバイダ及びビジネス利用者からの情報を基に、必要に応じて A I の点検・修理及び A I ソフトのアップデートを行うことなどにより、A I がアクチュエータ等を通じて人の生命・身体・財産に危害を及ぼすことのないよう配慮することが期待される。</p> <p>また、消費者的利用者は、A I がアクチュエータ等を通じて人の生命・身体・財産に危害を及ぼした場合に講ずるべき措置について、開発者、A I サービスプロバイダ及びビジネス利用者から提供された情報に留意することが期待される。</p>

## 【参考】 AIの外部での安全確保の仕組みの例

衝突ハザード及び転倒ハザード等をもつ介護ロボットの電気・電子・プログラマブル電子安全関連系群（多重防護層）



引用：A I ガバナンス検討会（第4回）佐藤様講演資料

→制御ユニットにおいて、AI実装部（A）のみでなく、他のシステム（B、C）を含めた安全確保

原案	修正案
<p>利用者及びデータ提供者は、A I システム又は A I サービスのセキュリティに留意する。</p> <p>論点</p> <p>ア) セキュリティ対策の実施</p> <p>イ) セキュリティ対策のためのサービス提供等</p> <p>ウ) 不正確又は不適切なデータの学習による A I のセキュリティ脆弱性への留意</p>	<p>利用者及びデータ提供者は、A I システム又は A I サービスのセキュリティに留意する。</p> <p>論点</p> <p>ア) セキュリティ対策の実施</p> <p>イ) セキュリティ対策のためのサービス提供等</p> <p>ウ) 不正確又は不適切なデータの学習による A I のセキュリティ脆弱性への留意</p>
<p>構成変更無し</p>	

## ⑤ーア) セキュリティ対策の実施

原案	修正案
<p>利用者は、A I のセキュリティに留意し、その時点での技術水準に照らして合理的な対策を講ずることが期待されるのではないかと。</p> <p>また、利用者は、セキュリティが侵害された場合に講ずるべき措置について、あらかじめ整理しておくことが期待されるのではないかと。</p>	<p>A I サービスプロバイダ及びビジネス利用者は、A I はfragileであるという認識を持つと同時にA I のセキュリティに留意し、その時点での技術水準に照らして合理的な対策を講ずることが期待される。</p> <p>また、セキュリティが侵害された場合に講ずるべき措置について、あらかじめ整理しておくことが期待される。また、その措置の内容について、開発者等からの情報を踏まえ、当該AIの用途や特性、侵害の影響の大きさ等社会的文脈に応じたものとするが期待される。</p> <p>[セキュリティ侵害時の措置の例]</p> <ul style="list-style-type: none"> <li>● 初動措置（当該AIを含むシステムの急用度等の文脈に応じ、必要な手順にて実施） <ul style="list-style-type: none"> <li>➤ 当該システムのロールバック、代替システムの利用などによる復旧</li> <li>➤ システムの停止（キルスイッチ）：可能な場合</li> <li>➤ ネットワークからの遮断：可能な場合</li> <li>➤ セキュリティ侵害の内容確認</li> <li>➤ 関係者への報告</li> </ul> </li> <li>● 補償・賠償等（補償・賠償等を円滑に行うための保険の利用）</li> <li>● 重大な損害が生じた場合等は、第三者機関の設置とその機関による原因調査・分析・提言など</li> </ul> <p>1) 例えば、学習モデルから学習データをリバースエンジニアリングできる可能性があることが報告されている。</p> <p>消費者的利用者は、（消費者的利用者側で）セキュリティ対策を実施することが想定されている場合には、A I のセキュリティに留意し、必要な対策を講ずることが期待される。</p>

## ⑤ーイ) セキュリティ対策のためのサービス提供等

原案	修正案
<p>A I サービスプロバイダは、自ら提供するA I サービスについて、最終利用者にセキュリティ対策のためのサービスを提供するとともに、インシデント情報の共有を図ることが期待されるのではないか。</p>	<p>A I サービスプロバイダは、自ら提供するA I サービスについて、最終利用者にセキュリティ対策のためのサービスを提供するとともに、過去のアクシデントやインシデント情報の共有を図ることが期待される。</p> <p>また、AIサービスプロバイダ及びビジネス利用者はセキュリティが侵害された場合の措置について、消費者的利用者に対し必要な情報提供を行うことが期待される。</p> <p>消費者的利用者は、セキュリティが侵害された場合に講ずるべき措置について、開発者、A I サービスプロバイダ及びビジネス利用者から提供された情報に留意することが期待される。</p>

原案	修正案
<p>利用者及びデータ提供者は、A Iが不正確又は不適切なデータを学習することにより、A Iのセキュリティに脆弱性が生ずるリスクに留意することが期待されるのではないか。[再掲：②適正学習の原則 主な論点イ]。</p>	<p>A Iサービスプロバイダ、ビジネス利用者及びデータ提供者は、A Iが不正確又は不適切なデータを学習することにより、AIのセキュリティに脆弱性が生じるリスクが存在することに留意することが期待される。また、消費者的利用者に対し、そのようなリスクが存在することを予め周知することが期待される。[②ーイ 再掲]</p> <p>[リスクの可能性の例]</p> <ul style="list-style-type: none"> <li>学習モデルが、一部学習データからの微少な変動により悪影響を受ける利用されたデータが不均衡な場合に、悪意ある入力等により当該学習モデルが誤った結果を導出してしまいう攻撃が存在する可能性（例：Adversarial example攻撃）</li> <li>学習において不正確なラベリング等がなされたデータを混在させることで、間違っただ学習が行われる可能性</li> </ul> <p>消費者的利用者は、サービスプロバイダ、ビジネス利用者及びデータ提供者等からの情報を踏まえ、A Iが不正確又は不適切なデータを学習することにより、A Iのセキュリティに脆弱性が生じるリスクが存在することを認識しておくことが期待される。</p> <p>また、A Iを利用するに当たり、セキュリティ上の疑問を感じた場合は、A Iサービスプロバイダ、ビジネス利用者、データ提供者等にその旨を報告することが期待される。</p> <p>[②ーイ 再掲]</p>

## ⑥プライバシーの原則（全体構成）

原案	修正案
<p>利用者及びデータ提供者は、A I システム又はA I サービスの利活用において、他者又は自己のプライバシーが侵害されないよう配慮する。</p>	<p>利用者及びデータ提供者は、A I システム又はA I サービスの利活用において、他者又は自己のプライバシーが侵害されないよう配慮する。</p>
<p>論点</p>	<p>論点</p>
<p>ア) 他者のプライバシーの尊重</p>	<p>ア) 他者のプライバシーの尊重</p>
<p>イ) パーソナルデータの収集・分析・提供等におけるプライバシーの尊重</p>	<p>イ) パーソナルデータの収集・<b>前処理</b>・提供等におけるプライバシーの尊重</p>
<p>ウ) A I を利用したプロファイリングを行う場合におけるプライバシー等の配慮</p>	<p>「⑦尊厳・自律の原則」に移動</p>
<p>エ) 自己等のプライバシー侵害への留意</p>	<p>ウ) 自己等のプライバシー侵害への留意及びパーソナルデータ流出の防止</p>
<p>オ) パーソナルデータの流出の防止</p> <p>エ)は消費者的利用者を名宛人として留意すべき事項を記載したものであり、オ) はエ) の結果、情報を取得した者が留意すべき事項をまとめたものとなっているため、今回の整理においては統合してはどうか。</p>	

## ⑥ーア) 他者のプライバシーの尊重

原案	修正案
<p>利用者は、A I を利活用する際の社会的文脈や人々の合理的な期待を踏まえ、A I の利活用において他者のプライバシーを尊重することが期待されるのではないかと。</p> <p>また、他者のプライバシーを侵害した場合に講ずるべき措置について、あらかじめ整理しておくことが期待されるのではないかと。</p>	<p>A I サービスプロバイダ及びビジネス利用者は、A I を利活用する際の社会的文脈や人々の合理的な期待を踏まえ、A I の利活用において他者のプライバシーを尊重する。</p> <p>また、他者のプライバシーを侵害した場合に講ずるべき措置について、あらかじめ整理しておくことが期待される。</p> <p>加えて、当該措置について、消費者的利用者に対し、必要な情報提供を行うことが期待される。</p> <p>[プライバシー侵害時に講ずるべき措置の例]</p> <ul style="list-style-type: none"> <li>● 他者のプライバシーを侵害する情報を誤って取得した場合における、当該情報の消去、A I のアルゴリズムの更新等</li> <li>● 他者のプライバシーを侵害する情報を拡散した場合における、<del>拡散ルートの特定</del>、保存先への消去の依頼、A I のアルゴリズムの更新等</li> <li>● <del>保険の利用</del></li> </ul>
	<p>消費者的利用者は、A I を利活用する際の社会的文脈や人々の合理的な期待を踏まえ、A I の利活用において他者のプライバシーを尊重する。</p> <p>加えて、他者のプライバシーを侵害した場合に講ずるべき措置について、開発者、A I サービスプロバイダ及びビジネス利用者等から提供された情報に留意することが期待される。</p>

原案	修正案
<p>利用者及びデータ提供者は、A I の学習等に用いられるパーソナルデータの収集・分析・提供等において、他者のプライバシーを尊重することが期待されるのではないか。</p>	<p>A I サービスプロバイダ、ビジネス利用者は、及びデータ提供者は、A I の学習等に用いられるパーソナルデータの収集・前処理・提供等<sup>1</sup>及び、それらを通じて生成された学習モデルの提供等において、他者のプライバシーを尊重する。</p> <p>また、データ提供者は、提供したデータに個人情報が含まれる場合、A I を利用する際の社会的文脈に応じて、当該データを誰がどのように利用しているのかを把握しておくことが期待される。</p> <p>1)他者に提供後のパーソナルデータの取り扱いについても留意が必要である。</p>
	<p>消費者的利用者は、利用するA I 等の学習に用いるデータを自ら収集することが予定されている場合には、収集等において他者のプライバシーを尊重する。</p>

原案	修正案
<p>エ) 消費者的利用者は、ペットロボットなどのA Iに過度に感情移入すること等により、特に秘匿性の高い情報（自己の情報のみならず他者の情報を含む。）をむやみにA Iに与えることのないよう留意することが期待されるのではないか。</p>	<p>A Iサービスプロバイダ、ビジネス利用者及びデータ提供者は、A Iの判断により本人同意なくパーソナルデータが第三者に提供されないよう、同意がないデータはシステム上提供不可とするなど適切な措置を講ずる。</p>
<p>オ) A Iサービスプロバイダ、ビジネス利用者及びデータ提供者は、A Iの判断により本人同意なくパーソナルデータが第三者に提供されないよう適切な措置を講ずることが期待されるのではないか。</p>	<p>消費者的利用者は、ペットロボットなどA Iに過度に感情移入すること等により、特に秘匿性の高い情報（自己の情報のみならず他者の情報を含む。）をむやみにA Iに与えることのないよう留意することが期待される。</p>

原案	修正案
<p>利用者は、A I システム又はA I サービスの利活用において、人間の尊厳と個人の自律を尊重する。</p> <p>論点</p> <p>ア) 人間の尊厳と個人の自律の尊重</p> <p>イ) A I による意思決定・感情の操作等への留意</p> <p>ウ) A I と人間の脳・身体を連携する際の生命倫理等の議論の参照</p>	<p>利用者は、A I システム又はA I サービスの利活用において、人間の尊厳と個人の自律を尊重する。</p> <p>論点</p> <p>ア) 人間の尊厳と個人の自律の尊重</p> <p>イ) A I による意思決定・感情の操作等への留意</p> <p>ウ) A I と人間の脳・身体を連携する際の生命倫理等の議論の参照</p> <p>エ) A I を利用したプロファイリングを行う場合におけるプライバシー等の不利益への配慮 (⑥プライバシーの原則より移行)</p>

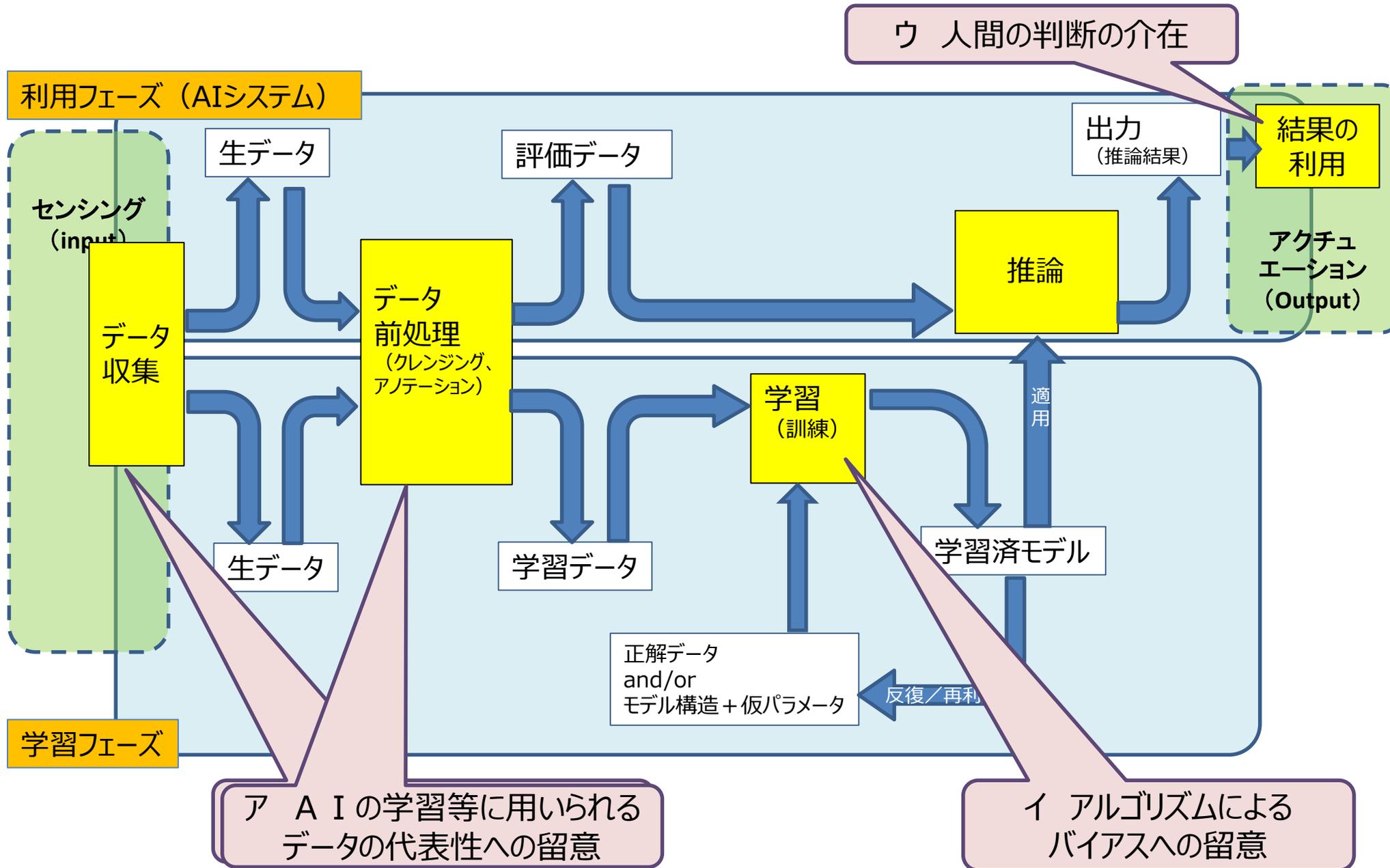
原案	修正案
<p>利用者は、A I を利活用する際の社会的文脈を踏まえ、人間の尊厳と個人の自律を尊重することが期待されるのではないか。</p>	<p>サービスプロバイダ及びビジネス利用者は、A I を利活用する際の社会的文脈を踏まえ、人間の尊厳と個人の自律を尊重することが期待される。その際には、人間とA I の異質性<sup>1</sup>を前提とするとともに、人間の活動を支援するものであるとの認識が重要となる。</p> <p>1)人間とA I が異なる性質を持つことを言い、この前提が成り立つことにより、人間をA I 同様に扱うべきではない（すなわち、人間の尊厳が必要）と考えることが可能となる。</p> <p>消費者的利用者は、A I を利活用する際の社会的文脈を踏まえ、人間の尊厳と個人の自律を尊重することが期待される。その際には、人間とA I の異質性<sup>1</sup>を前提とするとともに、人間の活動を支援するものであるとの認識が重要となる。</p> <p>1)人間とA I が異なる性質を持つことを言い、この前提が成り立つことにより、人間をA I 同様に扱うべきではない（すなわち、人間の尊厳が必要）と考えることが可能となる。</p>

原案	修正案
<p>利用者には、A I により意思決定や感情が操作されるリスクや、A I に過度に依存するリスクに留意することを期待することは適当か否か。</p> <p>このようなリスクについて誰がいかなる役割を果たすべきか。</p>	<p>A I サービスプロバイダ及びビジネス利用者は、消費者的利用者にはA I により意思決定や感情が操作される可能性や、A I に過度に依存するリスクが存在することを踏まえ、例えば以下のような対策を講じる必要がある。</p> <p>[意思決定・感情操作に対する対策例]</p> <ul style="list-style-type: none"><li>• 教育現場等における上記リスクがあることの共有の支援</li><li>• A I システムを含んだコンピュータシステムの開発側による対応</li><li>• 利用者の自覚を促すこと（注意喚起）</li></ul> <p>消費者的利用者は、A I サービスプロバイダ及びビジネス利用者からの情報等<sup>1</sup>を踏まえ、A I により意思決定や感情が操作される可能性や、A I に過度に依存するリスクがあることを自覚することが期待される。</p> <p>1)例えば、教育現場等において、上記リスクがあることの情報を取得するなど</p>

原案	修正案
<p>利用者は、A I を人間の脳・身体と連携させる場合には、生命倫理の議論等を参照し、人間の尊厳と自律が侵害されないよう特に慎重な配慮が求められるのではないか。</p>	<p>A I サービスプロバイダ及びビジネス利用者は、A I を人間の脳・身体と連携させる場合、特に、<b>エンハンスメント（健康の維持や回復を超えた人間の能力の増進の追及）</b>を行う場合には、その周辺技術に関する開発者等からの情報を踏まえつつ、生命倫理の議論等を参照し、人間の尊厳と自律が侵害されないよう特に慎重な配慮が求められる。</p> <p>また、上記に関する情報を消費者的利用者に提供することが期待される。</p>
	<p>消費者的利用者は、A I を人間の脳・身体と連携させたA I を用いる場合には、当該機能及びその周辺技術に関する開発者、A I サービスプロバイダ及びビジネス利用者からの情報を基に、自らの自律性が侵害されないよう留意して利用することが求められる。</p>

原案	修正案
<p>A I サービスプロバイダ及びビジネス利用者は、人事評価・採用や融資など個人の権利・利益に重要な影響を及ぼす可能性のある分野において A I を利用したプロファイリングを行う場合には、対象者のプライバシー等に慎重に配慮することが期待されるのではないか。</p>	<p>A I サービスプロバイダ及びビジネス利用者は、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 A I を利用したプロファイリングを行う場合には、下記により対象者に生じうる不利益等のプライバシー等に慎重に配慮する。</p> <p>[プロファイリングにおいて不利益を生じさせることとなる例]</p> <ul style="list-style-type: none"> <li>• プロファイリング結果が事実と異なることにより誤った判断が下されること</li> <li>• 対象者のプロファイリング結果の一部が特定の集団の特徴と共通であることにより、当該集団にネガティブな判断が下された場合、対象者も同様にネガティブな判断が下されること</li> <li>• 対象者の特定の特徴のみがプロファイリングで用いられることにより、過小評価されてしまうこと</li> <li>• プロファイリング結果から不確実な未来を予測（外挿）することにより、ネガティブな判断が下されること</li> <li>• 匿名データのプロファイリング結果と（匿名でない）特定個人のプロファイリング結果の突合により、匿名データから個人が特定されてしまうこと</li> <li>• プロファイリングの結果、人の権利・利益を損なう取り扱いがなされること</li> </ul> <p>消費者的利用者は、A I によるプロファイリングが行われている可能性があることを踏まえ、自らの情報が正しく利用されているかを意識し、確認することが期待される。</p>

原案	修正案
<p style="text-align: center;">A I サービスプロバイダ、ビジネス利用者及びデータ提供者は、A I システム又はA I サービスの判断によって個人が不当に差別されないよう配慮する。</p> <p>論点 ア) A I の学習等に用いられるデータの代表性への留意 イ) アルゴリズムによる不当な差別への留意 ウ) 人間の判断の介在</p>	<p style="text-align: center;">A I サービスプロバイダ、ビジネス利用者及びデータ提供者は、A I システム又はA I サービスの判断によって個人が不当に差別されないよう配慮する。その原因として、A I システム又はA I サービスの判断にバイアスが含まれる可能性があることに配慮する。 なお、「公平性」には複数の定義・基準がある<sup>1</sup>ことを踏まえ、どのような定義により「公平」と判断しているのかに留意する必要がある。</p> <p>論点 ア) A I の学習等に用いられるデータの代表性への留意 イ) アルゴリズムによる<b>バイアス</b>への留意 ウ) 人間の判断の介在（公平性の確保）</p> <p style="font-size: small; margin-top: 20px;">1) 論点イに公平性基準の例が記載されている。</p>



原案	修正案
<p>A I サービスプロバイダ、ビジネス利用者及びデータ提供者は、A I の判断結果により、人種・信条・性別等によって個人が不当に差別されないよう、A I の学習等に用いられるデータの代表性やデータに内在する社会的なバイアスに留意することが期待されるのではないかと。</p> <p>利用する技術の特性及び用途に照らして、どのような場合に、どの程度、学習等に用いるデータの代表性やデータに内在する社会的なバイアスに留意することを期待することが適切か。</p> <p>(注) データの代表性とはサンプルされて利活用に供されているデータが母集団の性質を歪めないという性質のことをいう。</p>	<p>A I サービスプロバイダ、ビジネス利用者及びデータ提供者は、A I の判断が学習時のデータによって決定づけられる可能性があることを踏まえ、利用時の社会的文脈及び用途に照らして、例えば以下のような観点で、A I の学習等に用いられるデータの代表性<sup>1</sup>やデータに内在する社会的なバイアスに留意することが期待される。</p> <p>[データの代表性やデータに内在する社会的なバイアスに留意する際に考慮する点(例)]</p> <ul style="list-style-type: none"> <li>• 不公平な判断がアルゴリズム上の処理として記述されていなくても（データの代表性やデータに内在する社会的バイアスによっても）不当な差別は起こりうる点に留意する必要がある。</li> <li>• 公平／不公平の如何に関与する属性（センシティブ属性）を持つデータを直接使用しなくても、差別が起こりうる点<sup>2</sup>に注意が必要である。</li> <li>• データ前処理の際、学習データのラベルは多くの場合人間が作成・付与するため、ラベル付与を行う際に、（意図的・無意識どちらのケースもあり得るが）それを行う人のバイアスが反映されることに留意する必要がある<sup>3</sup>。</li> <li>• データの代表性を満足するために大量のデータを集めようとすると、特に個人のデータについてはプライバシーの問題に対処する必要があることに留意する必要がある。</li> </ul> <p>1) データの「代表性」とは、サンプルされて利活用に供されているデータが母集団の性質を歪めないという性質のことをいう。                  2) 例えば、性別に依存しない与信審査を行いたいとする。収入の高い人の割合が男女間で数倍異なった場合、収入の高低を一属性に加えて与信審査を行うアルゴリズムを作ると、結果として性別による差別が生ずることになる。                  3) 本件の対策として、対象や用途に合わせたラベリングの統一基準を作ることも考えられる。</p>
	<p>N/A</p>

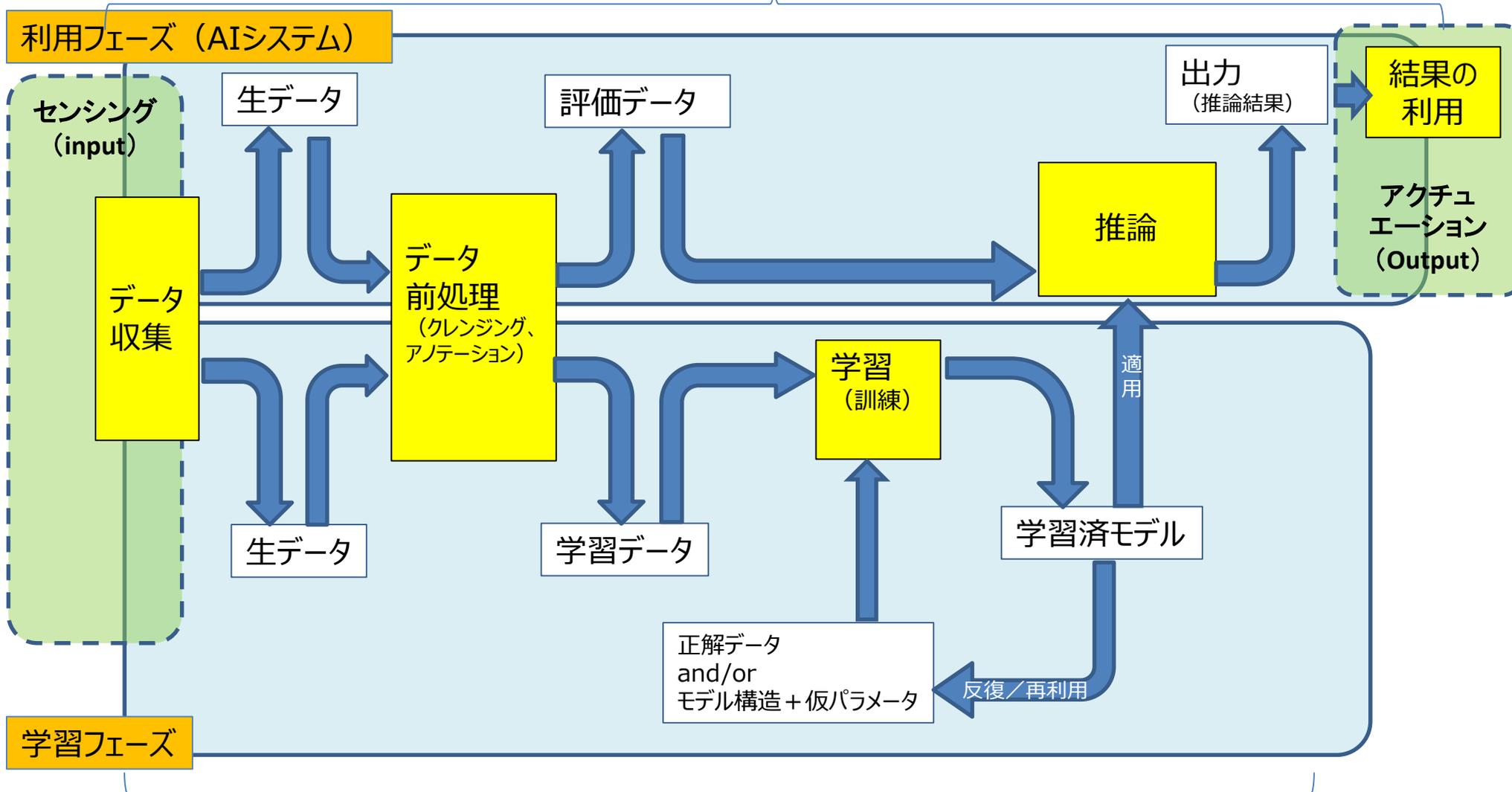
## ⑧ーイ) アルゴリズムによるバイアスへの留意

原案	修正案
<p>A I サービスプロバイダ及びビジネス利用者は、A I に用いられるアルゴリズムにより、人種・信条・性別等によって、個人が不当に差別される可能性があることに留意することが期待されるのではないか。</p>	<p>A I サービスプロバイダ及びビジネス利用者は、A I に用いられるアルゴリズムにより、それによる判断にバイアスが生じる可能性があることに留意することが期待される。特に、機械学習においては、一般的に、多数派がより尊重され、少数派が反映されにくい傾向にある（バンドワゴン効果）。この課題を回避するため、例えば以下の方法が考えられる。</p> <p>[機械学習アルゴリズムにより不当な差別を生み出さないための対策（例）]</p> <ul style="list-style-type: none"> <li>• （統計的）機械学習の枠組みを踏襲しつつ、「公平性基準(definition of fairness)」を守るためのアルゴリズムを活用する。例えば、センシティブ属性（前項参照）を定義し、それに対する中立性の確保（バイアスの除去）をもって「公平」であるとする方法がある。 <ul style="list-style-type: none"> <li>➢ センシティブ属性は、用途に基づく社会の要請<sup>1</sup>に基づき決定すればよい。</li> <li>➢ 公平性基準には、単純にセンシティブ属性を利用しない方法（unawareness）に加え、特定のグループ間の公平性を担保するための人口学的等価性（demographic parity）<sup>2</sup>等があり、センシティブ属性に係わる特定方向のバイアスを除去するものである。</li> <li>➢ ただし、（アルゴリズムにもよるが）公平性について上記の制約を課すことにより機械学習の精度に影響を及ぼす可能性があることに留意すべきである。</li> </ul> </li> </ul> <p>なお、公平性基準を定義すれば、それを満足する判断は可能となるため、公平性基準の決定に資する、意思決定ルールに関する制度設計（メカニズムデザイン）が重要となる。メカニズムデザインにあたっては、どのようなルールがよいのか（特定の者が不利益を被らないルールなど）、また、どのようにルールを確保するのか（契約、不文律、世論など）を検討する。</p> <p>1)例えば、入社試験で性別が問題視されている場合、性別をセンシティブ属性とし、性別に対し中立に（独立に）学習を行うなどが考えられる。</p> <p>2)上記1の例で、男女の平均評価点を同一にする方法のこと（アファーマティブ・アクションの考え方）。</p>
	N/A

原案	修正案
<p>A I サービスプロバイダ及びビジネス利用者は、A I によりなされた判断結果について、人種・信条・性別等によって個人が不当に差別されないよう、A I を利活用する際の社会的文脈や人々の合理的な期待を踏まえ、その判断を用いるか否か、あるいは、どのように用いるか等に関し、人間の判断を介在させることが期待されるのではないか。</p> <p>利用する技術の特性及び用途に照らして、どのような場合に、どの程度、人間の判断を介在させることを期待することが適当か。</p>	<p>A I サービスプロバイダ及びビジネス利用者は、A I によりなされた判断結果の公平性を保つため、A I を利活用する際の社会的文脈や人々の合理的な期待を踏まえ、その判断を用いるか否か、あるいは、どのように用いるか等に関し、人間の判断を介在させることが期待される。</p> <p>人間の判断の介在の要否については、[①ーイ] に掲げる内容を基に、利用する技術の特性及び用途に照らして検討することが期待される。特に、公平性の観点からは以下の基準を踏まえ検討することが期待される。</p> <p>[人間の判断の介在の要否について、基準として考えられる観点 (例) ]</p> <ul style="list-style-type: none"> <li>• 統計的な将来予測が (不確定性が高く) 難しい場合。<sup>1</sup></li> <li>• 意思決定 (判断) に対し納得ある理由を必要とする場合。<sup>2</sup></li> <li>• マイノリティなどに対する人種・信条・性別に基づく差別が想定される場合。</li> </ul> <p>1) 例えば、人事評価・採用等に当たっては、能力や生産性といった変数が重視されるが、これらは時間と共に変わるため、将来予測が難しい。</p> <p>2) 例えば、人事評価・採用等に当たっては、社員・受験者等からの請求があれば意思決定 (判断) の理由を説明できることが期待される。</p> <p>N/A</p>

原案	修正案
<p>A I サービスプロバイダ及びビジネス利用者は、A I システム又はA I サービスの入出力の検証可能性及び判断結果の説明可能性に留意する。</p> <p>（注）本原則は、アルゴリズム、ソースコード、学習データの開示を想定するものではない。また、本原則の解釈に当たっては、プライバシーや営業秘密への配慮も求められる。</p> <p>論点</p> <p>ア) A I の入出力の記録・保存</p> <p>イ) 説明可能性の確保</p>	<p>A I サービスプロバイダ及びビジネス利用者は、A I システム又はA I サービスの入出力等の検証可能性及び判断結果の説明可能性に留意する。</p> <p>（注）本原則は、アルゴリズム、ソースコード、学習データの開示を想定するものではない。また、本原則の解釈に当たっては、プライバシーや営業秘密への配慮も求められる。</p> <p>論点</p> <p>ア) A I の入出力等のログの記録・保存</p> <p>イ) 説明可能性の確保</p>

ア AIの入出力等のログの記録・保存



イ 説明可能性の確保

## ⑨ーア) A I の入出力等のログの記録・保存

原案	修正案
<p>A I サービスプロバイダ及びビジネス利用者は、A I の入出力の検証可能性を確保するため、入出力を記録・保存することが期待されるのではないかと期待されるのではないかと。</p> <p>利用する技術の特性及び用途に照らして、どのような場合に、どの程度、入出力を記録・保存することを期待することが適当か。例えば、自動運転車など人の生命・身体・財産に危害を及ぼし得る分野で利活用する場合には、事故の原因究明や再発防止に必要な範囲において、A I の入出力を記録・保存することが期待されるのではないかと。</p>	<p>A I サービスプロバイダ及びビジネス利用者は、A I の提供にあたり、事故が起こった場合の原因究明（トレース）等を目的として、入出力等のログを記録・保存することが期待される。</p> <p>ただし、ログの記録・保存に当たっては、利用するA I のアルゴリズムや用途に基づき、以下を考慮することが望ましい。</p> <p>[ログの記録・保存に当たり、考慮すべき事項]<sup>1</sup></p> <ul style="list-style-type: none"> <li>• ログ取得・記録の頻度</li> <li>• <b>ログの精度</b></li> <li>• ログの保存期間</li> <li>• ログの保護（機密性、完全性等）</li> <li>• ログ保存場所の容量</li> <li>• ログの時刻の記録</li> </ul> <p>例えば、人の生命・身体・財産に危害を及ぼし得る分野では、事故時の原因究明が可能となるよう、リアルタイムでのログの取得や時刻の記録等をできる限り行う必要があることが想定される。</p> <p>1) 掲げられた考慮すべき事項のいくつかは、それぞれが相互にトレードオフの関係にあるため、社会的文脈及び用途に応じてバランスを考慮することが必要である。例えば、ログ取得頻度や保存期間等はログの機密性、完全性を高めることとトレードオフの関係にある。</p>
	N/A

## ⑨ーイ) 説明可能性の確保

原案	修正案
<p>AI サービスプロバイダ及びビジネス利用者は、AI の判断結果の説明可能性を確保することが期待されるのではないかと。</p> <p>利用する技術の特性及び用途に照らして、どのような場合に、どの程度、説明可能性を確保することを期待することが適当か。特に、医療、人事評価・採用、融資など個人の権利・利益に重大な影響を及ぼす可能性のある分野において利活用する場合には、AI の判断結果の説明可能性を確保することが期待されるのではないかと。</p> <p>(例えば、深層学習は予測精度が高いが、判断結果の説明が困難であるという現状等にも留意することが必要である。)</p>	<p>AI サービスプロバイダ及びビジネス利用者は、利用者の納得感や安心感の獲得、及びそのためのAI の動作に対する証拠の提示等を目的として、AI の判断結果の説明可能性を確保することが期待される(特に、個人の権利・利益に重大な影響を及ぼす可能性のある分野において提供・利用する場合)。</p> <p>ただし、説明可能性の確保に当たっては、その目的に鑑み、どのような説明が必要かを分析・把握し、それに対し、以下の手段を総合的に用いて対処することが期待される。</p> <p>[説明可能性確保のための手段(例)]</p> <p>(データに関する対策)</p> <ul style="list-style-type: none"> <li>• 学習された学習済モデルが具備されたAIソフトならびにAIシステムを提供するに当たり、学習等に利用されたデータが、いつ、どこで、どういう目的で集められたデータなのかを管理(data provenance)しておく。</li> <li>(学習モデルに関する対策)</li> <li>• AI に対する複数の入力と出力の組合せをもとに、その傾向を分析する(例えば、入力パターンを少しずつ変化させたときに出力がどうなるかを観測するなど)。</li> <li>• 下記に示す技術的な手法を用いる(ただし、技術的な「説明可能性」の確保には実装や検証等が必要となるため、一般的にコストとトレードオフの関係にあることにも十分留意する必要がある。) <ul style="list-style-type: none"> <li>➢ 予め可読性の高い解釈可能なモデルを作るための手法<sup>1</sup></li> <li>➢ ブラックボックスモデルに対する説明を可能とする手法 <ul style="list-style-type: none"> <li>✓ 「AIの予測・認識プロセスの可読化」など、解釈可能なモデルに置き換えて説明を行う大域的な説明法</li> <li>✓ 「重要な特徴の提示」、「重要な学習データの提示」およびその「自然言語による表現」など、特定の入力に対する予測の根拠を提示する局所的な説明法</li> </ul> </li> </ul> </li> </ul> <p>(総合的な対策)</p> <ul style="list-style-type: none"> <li>• 説明が不足している部分について、どのような説明が必要なのかを明確にし、開発者とも連携して解決策を模索する。</li> </ul> <p>1) 一般的に、学習モデルを解釈可能な形とすることは、(学習)精度の確保とトレードオフの関係にあることに注意が必要である。</p> <p>N/A</p>

原案	修正案
<p data-bbox="43 205 1004 368">A I サービスプロバイダ及びビジネス利用者は、消費者的利用者及び間接利用者を含むステークホルダに対しアカウントビリティを果たすよう努める。</p> <p data-bbox="43 428 134 468">論点</p> <ul data-bbox="43 478 866 639" style="list-style-type: none"><li data-bbox="43 478 714 525">ア) アカウントビリティを果たす努力</li><li data-bbox="43 592 866 639">イ) A I に関する利用方針の通知・公表</li></ul>	<p data-bbox="1050 205 2013 368">A I サービスプロバイダ及びビジネス利用者は、消費者的利用者及び間接利用者を含むステークホルダに対しアカウントビリティを果たすよう努める。</p> <p data-bbox="1050 428 1141 468">論点</p> <ul data-bbox="1050 478 1873 639" style="list-style-type: none"><li data-bbox="1050 478 1721 525">ア) アカウントビリティを果たす努力</li><li data-bbox="1050 592 1873 639">イ) A I に関する利用方針の通知・公表</li></ul>

構成変更無し

## ⑩ーア) アカウンタビリティを果たす努力

原案	修正案
<p>A I サービスプロバイダ及びビジネス利用者は、人々と社会から A I への信頼を獲得することができるよう、消費者的利用者、間接利用者、A I の利活用により影響を受ける第三者等に対し、利用する A I の性質及び目的等に照らして、相応のアカウントビリティを果たすよう努めることが期待されるのではないか。</p>	<p>A I サービスプロバイダ及びビジネス利用者は、人々と社会から A I への信頼を獲得することができるよう、消費者的利用者、間接利用者、A I の利活用により影響を受ける第三者等に対し、利用する A I の性質及び目的等に照らして、それぞれが有する知識や能力の多寡に応じ、<b>A I システムの特性について情報提供と説明を行う他、多様なステークホルダーとの対話を通じて様々な意見を聴取する等、</b>相応のアカウントビリティを果たすよう努めることが期待される。</p> <p>消費者的利用者は、A I の判断結果について疑義を感じた場合には、必要に応じて、同サービスを提供した開発者、AIサービスプロバイダ、及びビジネス利用者に問い合わせを行うことが期待される。</p>

## ⑩ーイ) A I に関する利用方針の通知・公表

原案	修正案
<p>A I サービスプロバイダ及びビジネス利用者は、消費者的利用者や間接利用者らがA I の利活用について、適切に認識することができるよう、A I に関する利用方針（A I を利用している旨、適正な利活用の範囲及び方法、利活用に伴うリスク、相談窓口など）を通知又は公表することが期待されるのではないか。</p> <p>利用する技術の性質及び用途に照らして、どのような場合に、どのような内容の利用方針を通知又は公表することを期待することが適当か。</p>	<p>A I サービスプロバイダ及びビジネス利用者は、以下の場合に、消費者的利用者や間接利用者らがA I の利活用について適切に認識することができるよう、以下の事項を含むA I に関する利用方針を作成し、問い合わせがあった場合には開示を行うこと、加えて、個人の権利・利益に重大な影響を及ぼす可能性のある場合には自発的に通知又は公表することが期待される。</p> <p>[ A I 利用方針を提示するケース ]  以下の場合に利用方針を提示することが適当か。  (案1) すべての場合  (案2) 個人の権利・利益に重大な影響を及ぼす可能性のある場合  (案3) 人間の判断の介在が予定されていない場合  (案4) 自動的に学習が行われ、自律的に変化する場合</p> <p>[ A I 利用方針に含まれる事項 (例) ]</p> <ul style="list-style-type: none"> <li>• A I を利用している旨</li> <li>• 適正な利活用の範囲及び方法</li> <li>• 利活用に伴うリスク</li> <li>• 相談窓口</li> <li>• 個人情報を取得する場合の目的など</li> </ul> <p>また、通知または公開は、利用開始前だけでなく、A I の動作に変更が生じたときや利用終了時も含め実施すること（特にA I の動作変更に伴い想定されるリスクに変更が生じる場合など）が期待される。</p> <p>消費者的利用者は、A I の判断結果について疑義を感じた場合には、必要に応じて、同サービスを提供した開発者、AIサービスプロバイダ、及びビジネス利用者にお問い合わせを行うことが期待される。</p>

## 参考

