# 「人間中心のＡＩ社会原則」（案）に関する意見募集に対して提出された意見

# 目　　次

u　**個人Ａ（20190115）**

Ø　該当箇所：４．１ ***ＡＩ社会原則***

***ご意見（全角 1000*** 文字以内でおまとめください）：４．１項に､「（８）ＡＩ多様性確保の原則」を追加すること***を提案します。***
***この提案理由を説明します。***
***海を航行するボートの中で多数の乗客が一斉に右舷側に移動すると、ボートが右に傾いて転覆する事も有り得ます。また、つり橋の上で、多数の人が同じタイミングで同期して一定周期で一斉にジャンプをし続けると、共振現象が発生して、つり橋が大きく揺れて橋が崩落する事も有り得ます。***
***同じ仕組みの多数のＡＩに、ネットワーク経由で同じデータが同時に与えられますと、全部のＡＩが同じタイミングで同じ反応をします。同じ反応をするＡＩに制御された活動が社会の多くの場所や分野で一斉に同じタイミングで発生しますと、「過負荷によるシステムの停止や破壊」や「バランスの崩壊」が発生します。これは、社会に大混乱や大損失を与えます。***
***したがって、ＡＩに多様性を確保して、ＡＩの一斉動作による破壊やバランス崩壊を防がねばならないということになります。ＡＩ多様性の確保のためには、同じアルゴリズムや同じ知識を持つＡＩは、分野別に存在できる個数および比率の上限値を設けるという規制が必要でしょう。***

Ø　該当箇所：４．１ ***ＡＩ社会原則***

***ご意見（全角 1000*** 文字以内でおまとめください）：４．１項に､「（９）ＡＩ正帰還ループ排除の原則」を追加す***る事を提案します。***
***この提案理由を説明します。***
***ＡＩ単体では、ＡＩ社会原則に合致していたとしても、複数個のＡＩがネットワーク経由で結合して、ポジティブフィードバックループ（正帰還ループ）を形成したならば、破壊的な現象が発生する事もあります。***
***ＡＩにおける情報処理と通信の速度は、人間の反応速度では全く追いつかないものですので、上記のような正帰還ループが、人間の１回の瞬きの間に１０００万回のサイクルを実行する事もあります。そうなると、その正帰還ループに組み込まれているＡＩに制御されている装置や資産や自然は、急激な変化をさせられることもあります。***
***ＡＩによる金融恐慌や、大規模停電や、環境汚染などが発生する可能性があります。リーマンショックの際には、ＡＩによる正帰還サイクルで金融の暴走が起きたとの報告もあります。ネットワークにＡＩが結合している世界では、ＡＩ正帰還サイクルになり得る結合の事前検知、ＡＩ正帰還サイクルの発生検知、ＡＩ正帰還サイクルの強制遮断措置などの安全対策も社会実装しておくことが必須であると思います。***
***これが、ＡＩ正帰還ループ排除の原則です。***

Ø　該当箇所：３ **Society 5.0 実現に必要な社会変革「AI-Ready な社会」**

***ご意見（全角 1000*** 文字以内でおまとめください）：（１）「人」のＢ）においては､「創造性」や「アイデア」の豊富な***人間が望ましいという趣旨を記載しています。***
***（３）「産業構造」においても、多様な人々が「アイデア」を実現できることや、「創造力」を産業を通じて発揮することが求められるとしています。***
***（４）「イノベーションシステム」においても、「新たな発想」が求められるとしています。***

このように、ＡＩの普及に伴なって、人間には創造性が求められると、「人間中心のＡＩ社会原則(案）」はしています。
しかし、教育基本法第２条で「創造性を培い」と、創造性の育成を目標に掲げていながら、実態としては創造性を育成できる教科もなければ、教師もいないままの数十年間でした。すなわち、創造性を評価する方法の確立、創造性を評価する体制の構築と運用、創造性の高い人に対する充分な社会的評価や処遇を与える制度の構築と運用を具体的に行なわねば、ＡＩの普及とともに、人間は便利さに流されるだけとなり、創造性を高める人間は増加しないと考えます。
そこで、次の提案をします。
内閣府で推進している知財創造教育の中の創造性教育を、教育基本法第２条での「創造性を培い」の実現のための施策とするとともに、「人間中心のＡＩ社会原則(案）」における人間の創造性の発揮のための施策にも位置付けることを、「人間中心のＡＩ社会原則(案）」の中に明記する事です。

u **個人 B（20190117）**

Ø 該当箇所：（**6**） ***公平性、説明責任及び透明性の原則***

***「AI-Ready な社会」においては、AI の利用によって、人々が、その人の持つ背景によって不当な差別を受けたり、人間の尊厳に照らして不当な扱いを受けたりすることがないように、公平性及び透明性のある意思決定とその結果に対する説明責任（アカウンタビリティ）が適切に確保されると共に、技術に対する信頼性（Trust）が担保される必要がある。***

***? AI の設計思想の下において、人々がその人種、性別、国籍、年齢、政治的信念、宗教等の多様なバックグラウンドを理由に不当な差別をされることなく、全ての人々が公平に扱われなければならない。***

***? AI を利用しているという事実、AI に利用されるデータの取得方法や使用方法、AI の動作結果の適切性を担保する仕組みなど、状況に応じた適切な説明が得られなければならない。***

***? 人々が AI の提案を理解して判断するために、AI の利用・採用・運用について、必要に応じて開かれた対話の場が適切に持たれなければならない。***

***? 上記の観点を担保し、AI を安心して社会で利活用するため、AI とそれを支えるデータの信頼性（Trust）を確保する仕組みが構築されなければならない。***

***ご意見（全角 1000*** 文字以内でおまとめください）：技術に対する信頼性（**Trust**）***の担保や AI に利用されるデータの取得方法や使用方法についての説明責任、および AI とそれを支えるデータの信頼性（Trust）を確保する仕組みの構築について言及していることは評価できる。しかし、AI はデータの質によってその結果が左右されることが多く、単にデータの信頼性だけではなく、「偏りの無い、精緻かつ正確なデータ」を利用することが求められる。これからの社会、「網羅的（偏りの無い）、精緻かつ正確なデータ」を整備することの重要性についても原則のなかに盛り込むべきと考える。***

u 個人Ｃ（20190203）

Ø 該当箇所：**4.1*（２）教育･リテラシーの原則***

***一番最後の段落の「行政や学校（教員）に負担を押し付けるのではなく、民間企業や市民も主体性をもって取り組んでいくことが望ましい」について***

***ご意見（全角* 1000** 文字以内でおまとめください）： ***現場の教員を例示しているのはよく理解できる。英語教育の導入で小学校の教員は業務が増えていることからして、ＩＴ教育の導入もおそらく現場の教員にしわ寄せがいくことが予想されるからである。***

***しかし、「行政」というのは具体的には誰のことを想定しているのかが不明。内閣府の経験や意見として、具体的にはどの行政機関が業務を押し付けられると予想しているのか。***

Ø 該当箇所：**4.1*（３）プライバシー確保の原則***

***ご意見（全角* 1000** 文字以内でおまとめください）： ***そもそも、今回の意見募集に際して住所を記載する必要性が乏しい。電磁的方法での意見提出であればメールアドレスがわかれば連絡可能だからである。必要性が無い（或は低い）目的で市民の個人情報を公権力が集めることに違和感を覚えないのは問題。内閣府の人間自身のプライバシーへの配慮の欠如を如実に示すものと考える。厚労省の不正統計問題も含めて、もっと「行政」は自分の身を正すべき。そうしなければ、「民間企業や主体も主体性をもって取り組んでいくことが望ましい」云々を公表しても、国民の心に響かない。***

***のみならず、プロファイリングによって画一的な機械的判断がなされないようにするためには、きめ細やかなプロファイリング***
をすることになり、各人の個人情報を大量に収集することが必要になる。日本は個人情報保護規制が緩く（例：ＧＤ
***ＰＲと比較せよ）、そのような規制の元でどのように実効性のあるプライバシー保護を強化するのかにつき、もっと詰めてい***
くべき（例：課徴金の導入）。

Ø 該当箇所：**4.1*（５）公正競争確保の原則***

***ご意見（全角* 1000** 文字以内でおまとめください）： ***画一的でないきめこまやかなプロファイリングを行うためには、高度なプログラムが必要だと思う。しかし、一部の大手ＩＴ企業のみがそのようなプログラムを有しており、今回の「ＡＩ社会 原則」の導入がそれ以外の企業がプロファイリングを行うことへの事実上の参入障壁になり、大手ＩＴ独占が進むのではないか。その懸念への対処策を今後真剣に詰めていくべき。一定のライセンス料を払えば他社もそのプログラムを利用できるようにすべきでは***？ ***（標準必須特許の仕組みをご参照）***

***また、複数の企業のＡＩが同様のメカニズムで「合理的な」な価格決定をする結果、カルテルと類似の効果が生じるのではないか。たとえば、金融市場では、ある事象発生をトリガーに通貨／株式を売却するという設定のＡＩを複数社が使用しており、その事象発生と同時に複数社が同時に売却をしたという事例を聞いたことがある（日経新聞。結果として、***

*その通貨／株式が大きく変動するという現象が起きた旨の解説が新聞には書かれていた記憶がある）。*

*しかし、この現象は、独禁法では規制が難しい。なぜならば、まず、上記の金融市場の例を見ても分かるように、価格決定についての、複数の競業企業間での通謀は無いからである。通謀がないにもかかわらず、カルテル類似の現象が起こる可能性は、ＡＩのアルゴリズムを各社が独自で作るのが（ＩＴ人材の供給不足で）難しく結局、競業する企業が同一のＡＩのアルゴリズムを大手ＩＴから買って（or ライセンス料を支払って使わせてもらって）同じような価格決定をすることになれば、より高まる。*

- **個人 D（20190205）**
  - 該当箇所：**P7**

***人々の格差や弱者を生み出さないために、幼児教育や初等中等教育において 幅広くリテラシー等の教育の機会が提供されるほか、社会人や高齢者の学び直しの機会の提供が求められる。***

***ご意見（全角 1000*** 文字以内でおまとめください）：社会人、教員の学び直しの機会が必要。
***第４次産業革命で、AI による無人化など技術的失業など懸念される。***
***・GAFAM 問題など、米国企業による富の集中***
***・AI など活用する企業の増加などにより持つもの、持たざる者の格差は広がっていく***
***・企業のライフサイクルの短さが個人の学習、再教育などに追いつかない***
***といった、自分たちではどうしようもない漠然とした不安がある。***
***収入のセーフティネットの必要性（ベーシックインカムの試験導入、検討）***
***再教育を安心して受けることのできる体制づくりが必要（指導者確保、育成、学習時間、収入面の不安払拭など）***
***ただ、再教育が、技術などの進歩に追いつくか、市場競争に勝てるかといった疑問もある。***

u 個人 E（20190207）

Ø 該当箇所：1 ページ目、第五パラグラフ

「この文書において、我々は、「AI」という用語について特定の技術を指すのではなく、広く「高度に複雑な情報システム一般」を指すものとして捉え」

ご意見（全角 1000 文字以内でおまとめください）：AI を「高度で複雑な情報システム一般」と定義するとした場合、「高度」「複雑」についての具体的な判断基準を定まっていないので、第４章で記述される各原則の適用対象が明確に定まりません。マルチステークホルダーの意見を十分に吸い上げたうえで具体的な定量的基準を明記すべきです。あるいは、現時点の技術動向を考え、「機械学習システム」と限定した方が良いとも考えます。

また、現状のままでは、4 ページ目の第一パラグラフにある「、IoT、ロボティックス、超高速広帯域通信網等と並んで AI がある。」という記述との整合性が取れていません。IoT やロボティクスは「高度で複雑な情報システム一般」の構成要素になるからです。

さらに、定義については、技術発展によって適宜変更が必要であると見込まれるため、3～5 年程度での見直し規定を導入すべきだと考えます。

Ø 該当箇所：3 ページ目、第一パラグラフ

「AI を人類の公共財として活用し、」

ご意見（全角 1000 文字以内でおまとめください）：「公共財」という言葉は経済学においては、非競合性と非排他性を有するために民間主体による供給が望めない財を意味します。AI は基本的には民間による研究開発をベースとするものですから、「公共財」にかえて「共通資産」などの別の用語を使う必要があります。

Ø 該当箇所：8 ページ目、（４）の第三パラグラフ

「社会は、常に AI の利用における持続可能性に留意すべきである。社会は、特に、単一あるいは少数の特定 AI に一義的に依存してはならない。」

ご意見（全角 1000 文字以内でおまとめください）：「AI の利用における持続可能性」の意味が不明です。後段の文脈からすると、「特定の AI ベンダーにロックインされないようにすること」と解釈されますが、その場合は「持続可能性」ではなく、「選択可能性」という用語が適切だと考えます。そのうえで、本項目については「（５）公正競争確保の原則」に含めるべきと思います。

Ø 該当箇所：9 ページ目、（７）第一パラグラフ

「徹底的な国際化・多様化と産学官民連携を推進するべきである。」

ご意見（全角 1000 文字以内でおまとめください）：「徹底的な国際化・多様化」が人材配置の目的として掲げられていますが、これは適切な人材配置の結果として達成される可能性のある状況を示すものであり、施策の目的として設定するのは不適切だと思います。適切な人材採用の結果として、特定の国・地域の出身者がマジョリティとなるケースを排除する必要はないと思います。一方、研究については、その成果の他国への展開と、さまざまなアプローチの追求いう意味で、「徹底的な国際化・多様化」を目的とすることについては賛成します。

u **salesforce. com Co.,Ltd.（20190207）**

7 February 2019

Council for Science, Technology and Innovation, Cabinet Office
The Central Common Government Office
No.8, 1-6-1 Nagatacho Chiyoda-ku,
Tokyo, 100-8914, Japan

Dear Cabinet Office

Salesforce is pleased to provide this submission to the Cabinet Office in response to its paper *"Social Principles on Human Centric AI."*

Salesforce welcomes discussion on this important topic and how to advance Japan's Society 5.0.

We note Artificial Intelligence (AI) is the replication of human intelligence in machines where computers learn to mimic human movement through robotics, visual understanding with computer vision and language capability with natural language processing.

While the benefits of AI are significant and include improved efficiency and productivity, solves complex problems, and allows focus for more higher level tasks, Salesforce acknowledges there are global ethical concerns that AI could replace jobs and other functions that humans currently perform, or could be used in ways that reinforce bias or discrimination.

**Salesforce**

Salesforce began in 1999 when it was founded with a vision for a new kind of company—one with a new technology model, a new business model and a new philanthropic model. The new technology model was based on multi-tenant cloud infrastructure and the business model was pay-as-you-go subscription for enterprise technology. The philanthropic model is the 1-1-1 Model, which dedicated 1% of Salesforce's equity, 1% of Salesforce's product and 1% of Salesforce employees' time back to communities around the world.

Salesforce is a provider of software as a service ("SaaS") and platform as a service ("PaaS") offerings and is the global leader in Customer Relationship Management (CRM) software. Customer trust is our number one value. Our success depends on the delivery of reliable services to our customers in

Australia, and around the globe.

Salesforce has been operating in Japan since 2000. Through our ecosystem of customers, partners, and developers, we expect to help create 153,387 new jobs and contribute over $106 billion in net new business revenues by 2022**.**

**Salesforce and Artificial Intelligence**

Built into the Salesforce Platform, Einstein is a layer of artificial intelligence that democratizes the power of AI for every Salesforce user. Salesforce Einstein empowers everyone to deliver smarter, more personalized and predictive customer experiences by automatically discovering relevant insights, predicting future behavior, proactively recommending best next actions and even automating tasks.

Salesforce is delivering more than 1 billion AI-driven predictions to its customers every day. Those predictions are powered by Einstein-branded capabilities, which include sales and service forecasting, and has the power to improve and transform the way societies live and work, and our interactions with one another.

Salesforce is of the view that AI systems should treat all people fairly; AI systems should empower everyone and engage people; AI systems should perform reliably and safely; AI systems should be understandable; and AI systems should be secure and respect privacy.

At Salesforce, we have determined that this ethical and humane use of technology, especially within this context of the Fourth Industrial Revolution, must be clearly addressed, not only by us, but by our entire industry.

Our industry has reached an inflection point that must be supported by a strong set of guiding values. Technology is not inherently good or bad. It's what we do with it that matters. And that's why we're making the ethical and humane use of technology, a strategic initiative at Salesforce.

Salesforce believes AI best serves society when it is:

- **Inclusive**: Respects the societal values of all those impacted, not just those of the creators. It must protect human rights, honor diversity, and promote equality and equity.
- **Responsible**: Respects laws and regulations. Strives to adhere to the highest security and safety protocols.
- **Accountable:** Seeks to address feedback and mitigate harm.
- **Empowering:** We will target our innovation efforts toward enabling human flourishing.

- **Transparent:** We strive to ensure our customers understand the recommendations and predictions Einstein makes.

**Recommendations**

Salesforce supports the seven principles outlined by the Cabinet Office concerning the *"Social Principles of Human-Centric AI."*

In addition Salesforce would like to provide additional recommendations for the Cabinet Office to consider:

- **Public-Private partnerships**. In doing so we wish to mention that AI is constantly evolving and improving, and there is a role in Public-Private partnerships. By leveraging these partnerships we can share learning between industries concerning the development of AI and further promote diversity and inclusion.
- **Principles-based vs regulation.** Flexible regulatory approach. Regulation could sometimes create unintended consequences. To this end a principles-based rather than regulatory approach is recommended.
- **Transparency.** One of the social principles covers transparency. Where appropriate those organizations using AI need to clearly articulate how it is being used, and algorithms are being applied responsibly.
- **Chief Officer.** The Japanese Government should consider appointing a "Chief ethical and humane use officer". (Salesforce appointed such an officer effective from this year).
- **Advisory Council.** The Japanese Government should consider appointing an independent Advisory Council with public and private sector representatives to advise on the development of the Social Principles moving forward as technologies develop.

Thank you for the opportunity to provide comment to this important initiative. Should you require further information please contact me at sgrigorian@salesforce.com

Yours sincerely

Sassoon Grigorian
**Director, Government Affairs & Public Policy, APAC**

u　**LINE 株式会社（20190208）**

Ø　該当箇所：全体

***ご意見（全角 1000 文字以内でおまとめ***ください）：今回の「人間中心の ***AI の社会原則」にかかる検討は、AI の適切な開発と社会実装に向けた前向きな一歩であり、このような検討を行うことについて支持します。***

***AI の適正利用にかかる分野において、日本政府が国際的なリーダーシップを取りながら今後も検討を進めていくことは、大変喜ばしいことと考えます。***

***また、日本政府が「人間中心の AI の社会原則」を国際的に発信することにより、各国における議論を促し、AI に関する技術や倫理の進歩に寄与することを望みます。***

***世界とのハーモナイゼーションが求められる中、他のグローバルリーダーの声にも耳を傾けつつ、更なる議論が継続される事を望みます。***

u　インテル株式会社（20190212）

（別紙様式）

# 意見書

平成31年2月12日

内閣府 政策統括官
（科学技術･イノベーション担当）
社会システム基盤担当 殿

郵便番号 100-0005
（ちよだく まるのうち 3－1－1 こくさいびる 5かい）
住所　千代田区丸の内 3-1-1 国際ビル 5 階
（いんてるかぶしきかいしゃ）
氏名　インテル株式会社
（なかがわ よしかつ）
代表者　中川 義克
電話番号 03-6268-4854
電子メールアドレス yoshikatsu.nakagawa@intel.com

「人間中心のAI社会原則」（案）について、別添のとおり、意見を提出します。

注 1 法人又は団体にあっては、その名称及び代表者の氏名を記載することとする。
注 2　用紙の大きさは、日本工業規格Ａ列４番とすること。別紙にはページ番号を記載すること。

（別添様式）

課題名：「人間中心のＡＩ社会原則」

| 該当箇所（ページ数等） | 御意見　（御参考までに日本語による説明を加えます） |
|---|---|
| **全般について** | We support that the development of Social Principles will help to promote an “AI-Ready society.” For your further consideration finalizing this document, we offer the recommendations in each section as follows.<br>[**社会原則を創ることが**“AI-Ready **な社会**”**を推進する上に大事であるという考えに賛同します。本案をより良い形で完成させる為に、御検討・御参考の対象になると思われる点を以下にコメントいたします。**] |
| 1 Introduction<br>[1 **はじめに**] | We understand that this chapter articulates that redesigning the current society and transforming into the “AI-Ready Society” is relevant to realize “Society 5.0”, future society of Japan. We would recommend including a summary description of “Society5.0” in this section not only for the benefit of Japanese citizens but also for global understanding.<br>[**この章では現在の社会をデザインし直して、**AI-Ready **な社会へと移行させることが**“Society5.0”**の実現に不可欠であると述べていると理解しました。その理解が正しいという前提で、日本国民からのみならず、グローバルな理解を得るためには、政府が構築を目指す**”Society5.0”**の概要がここに含められることを推奨します。**] |
| 2 Fundamental Philosophy”<br>[2 **基本理念**] | We recommend clarifying why this fundamental philosophy is important as relates to an AI-Ready Society. We presume this philosophy relates to building an “AI-Ready Society”. If true, this section may be better served with the title, “Fundamental philosophy for building an AI-Ready Society”.<br>[ “AI-Ready **な社会**“**にとって、この基本理念がなぜ大事であるかを明確にするべきかと思います。これらは**”AI-Ready **な社会“を構築する上での理念と理解しました。この解釈が正しければ、この章のタイトルを例えば**「AI-Ready **な社会を構築するための基本理念**」**とすることを推奨します。**] |
| 3 AI Ready Society: Essential social revolution to achieve Society 5.0<br>[3 Society 5.0 **実現に必要な社会変革**「AI-Ready **な社会**」] | We understand that this chapter introduces a vision of “AI-Ready Society” (what is “AI-Ready Society”) from the five areas highlighted: “Human”, “Social Systems”, “Industry Structures”, “Innovation Systems” and “Governance”. We recommend modifying the introductory part of this section to explain how |

| | |
|---|---|
| | these five areas, together, are the essential areas to consider in achieving Society 5.0.<br>[**この章では** AI-Ready **な社会のビジョン** (AI-Ready **な社会とは何か**) **を** 5 **つの観点**：**「人」「社会システム」「産業構造」「イノベーションシステム」「ガバナンス」で定義しています。したがって、この章の序文として、これら** 5 **つ全てが同等に** Society5.0 **を実現するうえで、本質である点を明確に述べるように編集し直す事を推奨します。**] |
| 3, (3) Industry Structures<br>[3,(3)**産業構造**] | We support that industry structures should be "flexible and open internationally." We recommend including a broad principle supporting that there should be no market access barriers against innovative AI applications and services; and also policies developed in Japan should support the free movement of data across borders, which will be critical to the continued development and optimization of AI systems.<br>[**産業構造は国際的にフレキシブルでオープンであるべきと考えます。したがって、このセクションには、**AI **の創造的な応用や** AI **サービスの市場投入に対して、いずれの障壁も設けないこと、さらに日本で策定される政策は** AI **システムの継続的な発展と最適化に極めて重要となる境界を越えた自由なデータ流通を支えるものであることが基本であることも記すべきと考えます。**] |
| 3, (5) Governance<br>[3 **ガバナンス**] | We support the goal to establish an agile system in which diverse stakeholders – government, industry, academia and the general public- collaborate to assess, decide and implement rules, systems, and standards to achieve appropriate governance.<br>We recommend including verbiage to specifically call out a goal for the harmonization of domestic & international standards and principles as they are developed in Japan<br>[**政府、企業、大学や一般の人々など多様なステークホルダーが適切なガバナンスを実施するためには、規則、システム、標準の評価、決定、実装に関して、これらステークホルダーが協働できる臨機応変な仕組みを確立することが目標であるという考えに賛同します。その上で、日本が創った各原則が国際的にハーモナイズされることを目指すという意思を明確に強調する表現を含めることを推奨します。**] |
| 4-4.1, (2) Education<br>[4-4.1, (2) **教育・リテラシーの原則**] | We support and agree with the need, as articulated, to develop an educational environment for all with respect to AI.<br>We recommend changing the following statement at the end of the section, "In order to develop such an educational |

| | |
|---|---|
| | environment, it is desirable that companies and citizens work on their own initiative, not to burden administrations and schools (teachers)"<br>To the following:<br>"In order to develop such an educational environment, government should work with both industry and academia to assess future needs in a world that increasingly uses AI, and develop curriculum that will help benefit society in general and the workforce in particular."<br>[AI に関して全ての人々に教育環境を提供するという考えに賛同します。その上で、本セクション最後の段落：<br>「このような教育環境の整備に向けて、行政や学校（教員）に負担を押し付けるのではなく、民間企業や市民も主体性をもって取り組んで いくことが望ましい。」という文は、<br>「このような教育環境の整備に向けて、行政は企業や学校とともに、AI の活用が増大する世界における将来のニーズを見極め、社会一般および特定の業務に携わる人々の利益に供するカリキュラムの開発に努めるべきである。」<br>という内容に変えることを推奨します。] |
| 4-4.1, (3) Privacy<br>[4-4.1, (3) プライバシー確保の原則] | We agree and support personal data protection by not infringing freedom, dignity and equality; by ensuring accuracy, legitimacy and individual manageability; by considering importance and sensitivity.<br>For applying these principles in practice, we would like to share our recommendations, as follows, for your reference:<br>New legislative and regulatory initiatives should be comprehensive, technology neutral, and should enable the free flow of data.<br>Organizations should embrace risk-based accountability approaches.<br>Automated decision-making should be fostered while augmenting with safeguards to protect individuals.<br>Governments should promote access to data by making open government database; actively supporting the creation of reliable data base; fostering incentives for data sharing between the public and private sector and among industry players; by contributing to the creation of international voluntary |

| | |
|---|---|
| | standards; and by promoting diversity in datasets to reduce bias.<br> Funding research in security is essential to protect privacy.<br> It takes data to protect data: In order to detect biases or cyber threats and to protect personal data, AI needs to process personal data.<br>[**自由、尊厳、平等が妨げられず、正確性、正当性、個人の関与が確保され、更にその重要性、要配慮性が考慮されるように個人データの保護を行うという原則に賛同します。**<br>**この原則を実装する上で必要と考えるポリシーを御参考までに紹介 いたします。**<br>**- 新たに発動される法律や規則は包括的かつ技術的に中立であり、データの自由な流通を可能にするものとなるべきである。**<br>**- 各種団体はリスクを考慮した説明責任を採用するべきである。**<br>**- 自動的な意思決定能力の向上は、個人を保護するセーフガードを拡充させながら実現されていくべきである。**<br>**- 政府は、オープンな政府データベースの構築、信頼できるデータベ ース構築への積極的なサポート、公的部門と私的部門間や産業界の各企業間でのデータ共有を促すインセンティブの育成、国際的なボランティア標準の開発、データセットの多様化推進によるバイアスの低減化などを通じて、データへのアクセスを推進していくべきである。**<br>**- プライバシー保護のために欠かせないセキュリティの研究開発へ投資するべきである。**<br>**・データを保護するにはデータを扱う必要があるという認識に立つ。つまり、バイアスやサイバー脅威を検知し、個人データを保護する為には、**AI **が個人データを処理する必要がある。**] |
| 4-4.1, (6) Fairness, Accountability, and Transparency<br>[4-4.1, (6) **公平性、説明責任及び透明性の原則**] | We agree and support fair and transparent decision making and accountability for AI systems and results.<br>For applying these principles in practice, we would like to share our recommendations, as follows, for your reference:<br>Standing for "Accountable Artificial Intelligence": Governments, industry and academia should apply the Information Accountability Foundation's principles to AI. Organizations implementing AI solutions should be able to demonstrate to regulators that they have the right processes, policies and resources in place to meet those principles. |

| | |
|---|---|
| | Transparent decisions: Governments should determine which AI implementations require algorithm explainability to mitigate discrimination and harm to individuals.<br>[**公平で透明性が保たれた意思決定がなされ、**AI **システムおよびその動作結果に関しては適切な説明がなされねばならないという原則 に賛同します。以下、この原則を実装する上での必要と思われるポリシーを御参考までに紹介いたします。**<br>**- 説明可能な** AI **を支持する： 政府、企業、大学は** Information Accountability Foundation(IAF)**の** AI **に対する原則を適用すべきである。**AI **ソリューションを実装･適用する組織は、政府等の規則管轄者に対して、そのソリューションが、これらの原則を満たす正しいプロセス、ポリシー、リソースを有していることを実証できるよう にするべきである。**<br>**- 透明性のある意思決定：政府は差別を軽減し、個々人を傷つけな い為に、どの** AI **の実装がアルゴリズムの説明を必要とするのかを判断すべきである。**] |
| 4-4.2 Development and Utilization Principles of AI<br>[4-4.2 AI **開発利用原則**] | We agree and support the need to build international consensus "AI Principles" given that the development & utilization principles of AI are currently being discussed across industry, academia, civil society and nations.<br>[**現在、産業界、大学、市民社会、各国で** AI **の開発および利用に関する原則が議論されている事を考慮し、国際的なコンセンサスを得た"**AI **原則"が形成されねばならないという考えに賛同します。**] |

u　電子情報技術産業協会技術戦略部会（20190212）

平成 31 年 2 月 12 日

内閣府　政策統括官
（科学技術・イノベーション担当）
社会システム基盤担当　御中

一般社団法人 電子情報技術産業協会
技術戦略部会

「人間中心の AI 社会原則」（案）について意見を提出します。

**課題名：「人間中心の AI 社会原則」**

| 該当箇所（ページ数等） | 御意見 |
|---|---|
| **全体を通して** | 本原則は非常に重要な原則であり、人間中心の AI 社会実現に向け、各ステークホルダーが留意すべき基本原則の全体像についてバランスよく的確に纏められており、敬意を表します。<br>一方、我が国からは、AI 関連の原則として、内閣府策定の本「AI 社会原則案」、総務省 AI ネットワーク社会推進会議策定の「AI 開発ガイドライン」、「AI 利活用原則案」、経済産業省の「AI・データの利用に関する契約ガイドライン」などが提案されております。弊協会でも、2018 年 5 月に「Society 5.0 実現のための人工知能（AI）の社会実装に向けて」とする提言(https://home.jeita.or.jp/press_file/20180705133445_ljo7s08kQ0.pdf)で、AI は社会課題の解決に貢献し、人々の生活をより豊かにするために使うべきとしました。さらに、同 10 月には、社会実装に向けたひとつの実行策として、CEATEC JAPAN 2018 でパネルディスカッション「AI が人々の豊かな生活を実現～SDGs・Society 5.0 実現のための AI の社会実装に向けて～」(https://www.ceatec.com/news/ja-webmagazine/ja-010)を開催し、AI の適切な理解を一般に広める活動も実施してきました。<br>今後、より広く市民、また多様な価値観を有する世界に受け入れられるよう、国内外の様々なガイドラインとの整合や相違点の確認を含め、AI 原則の標準化に向けた工夫、動きなどを加速していただくこと、さらには、生活・社会・産業の動向を受けた将来のアップデートもご検討いただくこと、を期待しております。 |
| **p.1 表題**<br>**<u>人間中心の AI 社会原則</u>** | 人間中心という文言について、人間対 AI という観点で、人の不要な不安を和らげる訴求力があり、p6 4.1(1)にその趣旨の記載もあります。しかし、人間対他の生命、人間対環境と言った人文科学的な視点では、誤解を招く懸念もありえます。多様な文化に配慮し、人を支えるなど、その趣旨を最初の頁に脚注で明記してもよいのではないか、と考えますが、如何でしょうか？ |

| | |
|---|---|
| p.1　<u>1.はじめに</u><br>「中でも我が国は、少子高齢化、人手不足、過疎化、財政支出増大等に直面し、社会課題のいわば先進国となっている。」 | 左記の我が国の論理は、本 AI 関連の原則を纏める上で必ずしも必要ではないと思います。本案の趣旨に多様な国々、特に発展途上国などにも賛同いただくためには、(1)特に記載はしない、もしくは、(2)我が国の論理との誤解を少なくし、国際的にも賛同いただけるように、文の順序を入れ替え、最初の「現代社会は、・・・直面している。」の直後、「AI はこれらの問題を解き」の前に、本文章を置き、課題に直面する現代社会の具体例を説明する形とする、などとされては如何でしょうか？ |
| p.1　<u>1.はじめに</u><br>「本文書の全体構成を図１に示す。」 | 本原則の導入部として極めて重要な図で、多様な読者に、本原則の高邁な理念、全体像をまずご理解いただけるよう、各章の簡潔な概要・趣旨も付記されていると良いと感じました。如何でしょうか？<br>理由は下記です。ご案内の通り、AI は事例データから知識を獲得するため、現状では「ルールが制御できない」、「知識の訂正が困難」といった課題があります。事例データが低品質である場合には、正当性や公平性を満たさない偏見や差別、プライバシーを脅かす個人情報の開示につながり、利活用者の信頼を損なわれ、社会に多大な便益を与える AI の開発、利活用が阻害される懸念があります。P9 に「上記の観点を担保し、AI を安心して社会で利活用するため、AI とそれを支えるデータの信頼性（Trust）を確保する仕組みが構築されなければならない」との記載もあります。本原則のこれらの狙いを最初に明記しておくことが有効ではないか、と考えた次第です。 |
| p.1　<u>1.はじめに</u><br>「<u>あらゆる面</u>で社会をリデザインし」 | 「あらゆる」という表現は少し強すぎるように考えます。「<u>に関して</u>リデザインし」という表現では如何でしょうか? |
| p.1 注釈 2<br>「本原則における「AI」は、「高度に複雑な情報システム一般」を意味する。」 | 本原則が、SI 事業全般に拡大適用されないよう、ある程度の範囲の限定が必要ではないか、と考えます。 |
| p.4　<u>3 Society 5.0 実現に必要な社会変革「AI-Readyな社会」</u><br>(1)「人」（A）<br>「主として統計的バイアス、社会の様態によって生じるバイアス及び AI 利用者の悪意によるバイアスの 3 種類がある」 | データバイアスとして挙げられている 3 種類について、用語だけでは何を指しているかわかりにくいため、補足説明があると良いのではないか、と考えます。また、脚注で、「国際的には、例えば IEEE P7003（アルゴリズムのバイアス検討の標準化）において、バイアスを unintended、unjustified、inappropriate に分類して対応策を検討する動きもある。」などの追記もご検討いただけないでしょうか。 |

| | |
|---|---|
| p.4　<u>3 Society 5.0 実現に必要な社会変革「AI-Ready な社会」</u><br>(2)社会システム | 様々な社会システムの発展と相互接続性を担保する柔軟なアーキテクチャと、社会システムに共通のデータ利活用基盤を整備する必要性が的確に述べられています。正当性や公平性、プライバシー、セキュリティのためには、AI がどのような判断根拠に基づいた判断を下したかが分かる透明性やアカウンタビリティが求められますが、AI に対する信頼が広く醸成され、AI による開発者・利用者の便益を増幅し、人間中心の AI 社会が早期に実現するには、一般のご理解を頂くことが重要と考えます。このため、データ利活用基盤をどう整備すればよいのかの例示、例えば、以下の例示などの脚注追記をご検討頂けないでしょうか？<br>(A)各適用分野における社会の要請に基づいた広範な議論を通し、利用者・開発者それぞれの観点で AI 開発・利活用原則の遵守レベルを規定し、過度にならない適正な利用・責任範囲を明確化していくこと、<br>(B)倫理規範の逸脱を検出・検証し、解消を図るためにデータ品質や学習の履歴を管理する技術の開発と、レベル分けされた AI システムの透明性や事後検証性を確認・認定する仕組みの整備を通じて、開発者から利用者の AI の取捨選択や、AI が相互接続され社会システムによる責任所在・範囲の判断に必要十分な情報を提供する仕組みを構築すること。 |
| p.5 <u>3 Society5.0 実現に必要な社会変「AI-Ready な社会」</u><br>(4)「イノベーションシステム」<br>「大学・研究機関・企業、さらに一般の人々に至るまで、分野や立場を超えて AI の研究開発、利活用及び評価に参加し、互いに刺激し合いながら、イノベーションが次々に生まれる環境ができていることが必要である」 | AI によるイノベーション創出をするためにどんな環境が必要なのか、例えば、「AI によるイノベーションを創出するために、大学・研究機関・企業、さらに一般の人々に至るまで、互いに刺激し合いながら、分野や立場を超えて AI の研究開発、利活用及び評価に参加する環境ができていることが必要である。」という書き方の方が、よりご趣旨が明確にならないでしょうか？ |
| p.5 <u>3 Society5.0 実現に必要な社会変革「AI-Ready な社会」</u><br>(5)「ガバナンス」<br>「また、多様なステークホルダーだけではなく、社会的に声の挙げにくい人たちの声も拾い上げ」 | 「社会的に声の挙げにくい人たち」も多様なステークホルダーに包含されていると思われます。そこで、例えば「活発に発言する人たちだけではなく、社会的に声の挙げにくい人たちの声も拾い上げ」といった表現に改めては如何でしょうか？ |

| | |
|---|---|
| p.5 <u>3 Society5.0 実現に必要な社会変革「AI-Ready な社会」</u><br>(5)「ガバナンス」：<br>「技術的手段を含む企業の自主的な取組によるなど、柔軟かつ実効的な方法がとられていることが求められる。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が求められる。」 | 日本としてプライバシー確保を謳っていることは良いと考えますが、ご案内のように、他国には「プライバシーに配慮しない手段で開発した」AI 技術もあるとされています。非常に難しい問題ですが、何らかの規制、倫理の導入も含め、この課題に対して我が国が Society 5.0 の実現を通じてどう取組んでいくべきと考えているのか、世界に向けたメッセージの追記についても、ご検討頂けないでしょうか？ |
| p.6 <u>4 人間中心の AI 社会原則</u><br>(4 行目)<br>「社会(特に、国などの立法・行政機関)が留意すべき『AI 社会原則』」 | 多様なステークホルダーに向けてのものであることを考慮しますと、「社会（特に、国などの立法・行政機関、但しそれらに限られるものではない）が留意すべき『AI 社会原則』」、という表現にされた方がより良いのではないかと思いますが、如何でしょうか？ |
| p.6 <u>4 人間中心の AI 社会原則</u><br>(1)人間中心の原則<br>「AI は、人々の能力を拡張し、多様な人々の多様な幸せの追求を可能とするために開発され、社会に展開され、活用されるべきである。AI が活用される社会において、人々が AI に過度に依存したり、AI を悪用して人の意思決定を操作したりすることのないよう、<br>我々は、リテラシー教育や適正な利用の促進などのための適切な仕組みを導入することが望ましい。」 | 過度な自動化や、行き過ぎを抑制するためには、AI 技術の成熟度や社会的影響を考慮し、適宜人が介在できる余地を設けることも重要ではないかと考えます。色々な議論があるかと思いますが、その旨の、脚注などでの記載をご検討頂けないでしょうか？ |
| p.6 <u>4 人間中心の AI 社会原則</u><br>(1)人間中心の原則<br>「AI の利用がもたらす結果については、問題の特性に応じて、AI の開発・提供・ | 非常に慎重な表現にして頂いておりますが、極力誤解のないように、過度な責任が特定のステークホルダーに課されることがないよう、文言をさらにご検討頂けないでしょうか？ |

| | |
|---|---|
| 利用に関わった種々のステークホルダーが適切に分担して責任を負うべきである。」 | |
| p.6 4 人間中心の AI 社会原則<br>(2)教育・リテラシーの原則<br>「AI の正確な理解と、社会的に正しい利用ができる知識と倫理を持っていなければならない」 | 経営者は AI の詳しいアルゴリズムを知る必要はないと理解していますが、高いレベルのことまでを要求していると捉えられると、AI 利用の妨げになることが懸念されます。このため、「AI の正確な理解」がどの程度のものを指しているのかの補足、例示があると良いのではないかと考えますが、如何でしょうか？ |
| p.7 4.人間中心の AI 社会原則<br>(2)教育・リテラシーの原則：<br>(2 つ目のブレット)<br>「AI の持つ公平性・公正性・プライバシー保護に関わる課題」「全ての人が文理の境界を超えて学ぶ必要がある。」 | データにも関わることなので、「AI・データの持つ公平性・公正性・プライバシー保護に関わる課題」としたほうが宜しくはないでしょうか？<br>また、「全ての人が文理の境界を超えて学ぶ必要がある。」の部分ですが、数理、データサイエンス、AI の基礎は“読み書きそろばん”のように基本的な教育内容になると考えますので、例え ば、「全ての人が文理の境界を超えて学ぶ環境・機会を提供する。何故ならば、Society5.0 においては、新しい”読み書きそろばん”として数理、データサイエンス、AI の基礎は文理関係なく学習する必要があると考えるからである。」というような例示内容について、ご検討頂けないでしょうか？ |
| p.7 4.人間中心の AI 社会原則<br>(3)プライバシー確保の原則<br>「パーソナルデータは、～きめ細やかに検討される必要がある。」 | この部分では、他の原則に比して具体的な検討事項が求められています。今後ガイドラインにおいて追加要求することがなければ、「保護法制に遵守して利活用する必要がある。」の記載に留めても良いのではないでしょうか？ |
| p.8 4.人間中心の AI 社会原則<br>(4)セキュリティ確保の原則(2 つ目のブレット)<br>「社会は特に、単一あるいは少数の特定 AI に一義的に依存してはならない」 | より深くご理解いただくため、なぜ依存してはならないのかの、簡素な理由の記載があると良いのではないかと考えます。例えば、「(5)公正競争確保の原則」1 つ目、2 つ目のブレットのような書き方は如何でしょうか？ |
| p.8 4.人間中心の AI 社会原則<br>(5)公正競争確保の原則 | 資源の集中もあってはならないように読めるので、例えば、「特定の国に AI に関する資源が集中した場合においても、その支配的な地位を利用した不当なデータの収集や主権の侵害が行わ |

| | |
|---|---|
| 「特定の国に AI に関する資源が<u>集中することにより</u>、その支配的な地位を利用した不当なデータの収集や主権の侵害が行われる社会であってはならない。」 | れる社会であってはならない。」、という趣旨の記載をご検討頂けないでしょうか？ |
| p.8 <u>4.人間中心の AI 社会原則</u><br>(5)公正競争確保の原則<br>「特定の企業に AI に関する資源が<u>集中することにより</u>、その支配的な地位を利用した不当なデータの収集や不公正な競争が行われる社会であってはならない。」 | 上記と同様の理由で、「特定の企業に AI に関する資源が<u>集中した場合においても</u>、その支配的な地位を利用した不当なデータの収集や不公正な競争が行われる社会であってはならない。」、という趣旨の記載をご検討頂けないでしょうか？<br>なお上記(P8)2 つをまとめ、「特定の<u>国もしくは企業</u>に AI に関する資源が<u>集中した場合においても</u>、その支配的な地位を利用した不当なデータの収集や主権の侵害が行われる社会であってはならない。」とする案は如何でしょうか？ |
| p.8 <u>4.人間中心の AI 社会原則</u><br>(6)公平性、説明責任及び透明性の原則<br>「「AI-Ready な社会」においては、AI の利用によって、人々が、その人の持つ背景によって不当な差別を受けたり、人間の尊厳に照らして不当な扱いを受けたりすることがないように、公平性及び透明性のある意思決定とその結果に対する説明責任（アカウンタビリティ）が適切に確保されると共に、技術に対する信頼性（Trust）が担保される必要がある。」 | 非常に重要な問題であると考えます。<br>(1)AI の結果に対するアカウンタビリティや信頼性を確保するためには、「AI 技術特有の不確実性を考慮」する必要があると考えられますので、その旨の明記をご検討頂けないでしょうか？<br>(2)説明責任については、説明責任の例示、例えば、AI を利用している事実や、データの取得方法や使用方法、動作結果の適正性などを説明することが期待されることなど、過剰な期待(誤解)が生じないように、補足などでの追記をご検討頂けないでしょうか？ |
| p.9 <u>4.人間中心の AI 社会</u>原<u>則</u><br>(6)公平性、説明責任及び透明性の原則<br>「上記の観点を担保し、AI を安心して社会で利活用するため、AI とそれを支えるデータの信頼性（Trust）を | 本原則は非常に重要で、内容に賛同いたします。特に 4 つ目のブレットについては AI システムに用いるデータは国境を越える自由な流通が不可欠であり、情報の利活用と保護のバランスの取れた両立を図りつつ、産業界やアカデミア等多様なステークホルダーとの対話を通じ、政府・国際機関が促進に向けてルール策定などに積極的に取り組んで頂くことを期待します。<br>このためには、AI に対する開発・利用者、特に一般市民、発展途上国、開発者などから、本原則について広く信頼を醸成することが必須と考えます。そこで、①AI 社会においては、多様な地 |

| | |
|---|---|
| **確保する仕組みが構築されなければならない」** | **域・文化・分野が接続されると考えられるため、相反する観点が生じうる可能性があり、多様性を許容することが重要であること、②AI が適性に学習したかどうか、あるいは、その AI によって公平・公正なサービスが行われたかを判断するために、(a)AI がどのような判断根拠に基づいた判断を下したかを検証し、問題がある場合には是正するための仕組みを提供・公開すること、さらには、(b)AI の判断の透明性について、直接的に公開するのではなく、認証等を行う機関・仕組みなどのオープン性とクローズ性を考慮すること、などが重要と考えます。このため、一般のご理解を頂く上で重要と考えられるこれらの項目について、例え ば、下記の観点での脚注例示などについてご検討頂けないでしょうか？<br>　(A)利用者・開発者それぞれの観点での AI 利活用原則の遵守レベルの規定、<br>　(B)学習させるデータの質による AI の判断結果の変化を検知・検証・修正できる仕組みの提供、<br>　(C)開発者の知的財産上の権利を保護するためのオープン性とクローズ性の考慮。** |
| **p.9 (7)「イノベーションの原則」** | **　イノベーションに寄与すべき原則は理解できますが、AI 以外の技術も当てはまるように読めるので、更に AI ならではの論点の強調について、ご検討頂けないでしょうか？** |
| **p.10 <u>5 おわりに</u><br>「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、今後柔軟に進化・発展させるものである」** | **　内容に賛同いたします。記載されている内容についてスピード感をもって行われることを期待致します。** |

**以上**

**【提出者（連絡担当者）の連絡先】**

**[情報削除]**

u　**個人Ｆ（20190214）**

Ø　該当箇所：「１．はじめに」または、「２．基本理念」の主文に加筆すべき事項として

***関連して次の部分***

***「４．１（１）人間中心の原則」***

***【本文１行目】「ＡＩは、人々の能力を拡張し」の部分***

***【各論１点目】「ＡＩは、人間の労働の一部を代替するのみならず」の部分***

***ご意見（全角 1000*** 文字以内でおまとめください）：ＡＩ社会原則を議論する前提として、『自動化・省力化・無人***化』に関するこれまでの歴史的流れを確認しておくべきでしょう。この流れは、人類が選んだ既存の方向性であり、ＡＩが新たに創り出した方向性ではありません。すでに始まっていることであり、ＡＩはこの既存の方向への推進力の加速度を劇的に増幅させるに過ぎません。とりわけ一般市民（無力な受益者）に認識してもらうべき最初の重要事項は、これから被る激変がごく自然な成り行きであり、その進行を止めることはもちろん速度を緩めることすら極めて困難であるという事実でしょう。***

***次に、ＡＩ社会を考える"手法"として肝要なのは、すでに始まっている"（ＡＩ化）過渡期"と末路たる"（ＡＩ化）成熟期"とに分けて議論することです。「人々の能力を拡張し」「人間の労働の一部を代替する」という"恩恵"は過渡期では妥当な表現かもしれませんが、成熟期では適切な表現ではありません。このように過渡期と成熟期を分けることによって、この議論における無用の混乱を予防できます。***

***この上で、過渡期では、ＡＩ権力の偏在化が加速することが容易に予想できますから、"無力な受益者たる一般市民へのﾒｯｾｰｼﾞ"と、"専門開発者（ＡＩ権力者候補）に対するﾒｯｾｰｼﾞ"を分けて、この原則を表現する必要があるでしょう。一方、成熟期に実現するであろう「人的貢献不要社会（＝不労社会）」では、"あまりにも無力な受動者（受益者）"と"ＡＩｲﾝﾌﾗ（生産手段）の独占者"とに分離して、「原則」の議論を整理すると伝わりやすいでしょう。***

Ø　該当箇所：「４．１（２）教育・ﾘﾃﾗｼｰの原則」の全般

***【各論１点目、１行目】「…格差や弱者を生み出さないために、幼児教育や初等中等教育において幅広くﾘﾃﾗｼｰ等の教育機会が提供される…」***

***【各論３点目、５行目】「…落伍者を出さないためのインタラクティブな教育環境…」***

***・・・など***

***ご意見（全角 1000*** 文字以内でおまとめください）：（ＡＩ社会への）過渡期では開発者・技術保有者らが「新資***本家」（ＡＩｲﾝﾌﾗ・情報・ﾉｳﾊｳの所有者・運用者）を形成するでしょうが、ＡＩ権力の偏在化のさらなる進行により、成熟期にはＡＩｲﾝﾌﾗは一部の集団による独占へと帰着してしまう危険性を否定できません。成熟期には新資本家の大多数も消え去るでしょう。***

***一方、"ﾏﾝ・ﾏｼﾝ・ｲﾝﾀｰﾌｪｲｽ"の飛躍的な発展によって操作性等のハードルが下がりつつある現状を見れば、成熟期にはもちろん過渡期にすら、ＡＩ機器にｱｸｾｽ・利用するための基礎知識といったﾘﾃﾗｼｰの必要性を強調する意味はありません。さらにいえば、現代において企業や専門機関が独占している大量の専門知識・ﾉｳﾊｳを、その大量さと独占ゆえに知ることもできず、その高度な専門性ゆえに理解することもできない"無力な受益者"は、すでに「信頼するのみ」の情報弱者です。***

***したがって、これはﾘﾃﾗｼｰや教育の問題ではなく、"信頼される側"の新資本家やＡＩｲﾝﾌﾗ独占集団が"責任を果たす"かどうかというＳＲ論（社会的責任論）の問題であることを認識すべきでしょう。現在多数存在している企業や専門機関などの"責任主体"は、ＡＩ過渡期を通じて急速に淘汰され、ＡＩ成熟期に残るであろう極少数の"責任主体"が、***

*「信頼するのみ」の多くの"無力な受益者"からの「（無数の個別の）信頼を裏切らないようにするための仕掛け・仕組み・システム」を、今から講じておく必要があるでしょう。*
*また、ＡＩ化の成熟期によってもたらされる「人的貢献不要社会（＝不労社会）」は、いわば究極の平等という恩恵をもたらすことからも、"格差"や"落伍者"の心配は不要であることは明らかでしょう。むしろ、ＡＩ過渡期におけるベーシックインカムやワークシェアリングといった対策につながる原則を提示しておく価値は極めて大きいといえます。いずれにしても、教育・リテラシーの問題ではなく、ＡＩインフラを独占する極少数の"責任主体"の側に課す義務の問題と理解して、その対策を講じておくべきでしょう。*

Ø　該当箇所：「３　***（５）ガバナンス」***
***【４行目】「政府、企業、大学、研究機関、一般の人々等、多様なステークホルダーが共同してルール、制度、標準化、行動規範等のガバナンスについて問題を設定し、影響を評価し、意思決定を行うと共に実装できる体制が整っていることが必要である」の部分***
***【１０行目】「…企業の自主的な取組…」の部分***

***ご意見（全角 1000*** 文字以内でおまとめください）：この点は極めて重要で、柔軟性に富みクイック・レスポンスができる***具体的な体制として、（期間を区切った）ネット上の意見交換広場の常設や、議論を整理してその成果物を遅滞なく発信できる専任のチェックパーソンの常設も有効でしょう。***
***さらに、「企業の自主的な取組」を効果的に推進する仕組みとして、多くの企業・組織に装備されているマネジメントシステムの有効活用を推奨します。***

Ø　該当箇所：「４．１（５）公正競争確保の原則」
***【各論１点目、２点目】「特定の国（企業）に AI に関する資源が集中することにより、その支配的な地位を利用した不当なデータの収集や主権の侵害が行われる社会であってはならない」***
***…および、***
***「４．２ ＡＩ開発利用原則」***
***【３行目】「…被規制的で非拘束的な枠組みとして国際的に共有されることが重要…」の部分***

***ご意見（全角 1000*** 文字以内でおまとめください）：ＡＩ成熟期のＡＩ権力の寡占化・独占化が想定される（そう***ならないシナリオやそうさせない方法が提示・立証されていない）以上、さらにそれが国内に存在せず、国外のいずれかの国あるいは組織・個人の掌中にある危険性も否定できない以上、この"４．２ ＡＩ開発利用原則"の文章ではあまりにも脆弱で、かつ楽観的に過ぎるでしょう。具体的には、国連の経済社会理事会などを通じて、拘束的・規制的なコミットメントを含む議論のリーダーシップを発揮していくための国内ガバナンスを原則に含めていくべきでしょう。***

- 日本 IBM（20190214）
  - 該当箇所：Ｐ１ SDGs への貢献

ご意見（全角 1000 文字以内でおまとめください）：当文書は、総理や大臣が国際会議の場で発言もしくは議論をリードすることを目的とした、ハイレベルの取りまとめと理解しており、その観点では、必要な点が網羅されていると思います。ただ、「SDGs への貢献」に関しては、本文でなく総理や大臣のスピーチ中でも結構ですが、日本がＡＩ技術で SDGs（17 原則のいずれか）に貢献している事例を紹介すると、より良いと思います。 特に新興国も参加する G20 では、日本が AI 技術を活用して国際社会をリードしている印象を与えられると思います。

  - 該当箇所：Ｐ４「人」。/ Ｐ６「教育・リテラシー」など数箇所。

ご意見（全角 1000 文字以内でおまとめください）：当原則を取りまとめた後、次の段階として、具体的な実行などに関して、更なる議論を進められると考えています。
次ステップとして、更なる深い議論が必要と考える点をいくつか挙げさせていただきます。

その１点目です。

- 教育が重要なことに異論はございません。 ただ、具体的な取り組みに関しては、能力認定制度を制定するのか、ＡＩの研修を義務化するのか、など更なる議論が必要と思います。

  - 該当箇所：Ｐ８「セキュリティ確保」

ご意見（全角 1000 文字以内でおまとめください）：次ステップとして、更なる深い議論が必要と考える点、２点目です。

- セキュリティの確保が重要なことに異論はございません。 ただ、技術的に対応が難しいものも含め、コストの観点も重要であり、企業にとって対応の目安となるガイドラインの制定など、更なる議論・取り組みが必要と思います。

  - 該当箇所：Ｐ８－９「公平性、説明責任及び透明性の原則」

ご意見（全角 1000 文字以内でおまとめください）：次ステップとして、更なる深い議論が必要と考える点、３点目です。

- 安心して利用できるＡＩを提供している事業者を、どのように選別するか。 安心なＡＩを提供していない事業者に対して、国・行政としてどう対応するのか、といった点も、今後具体的に詰めていく必要があると考えます。

u　**個人 G（20190214）**

Ø　該当箇所：「我々は、」

***ご意見（全角 1000*** 文字以内でおまとめください）：文中の我々とはどの範囲を指すか？

Ø　該当箇所：３ **Society 5.0 *実現に必要な社会変革「AI-Ready な社会」***

***（1）「人」***

***A) AI の長所・短所をよく理解している。特に AI の情報リソースとなるデータ、アルゴリズム、又はその双方にはバイアスが含まれること及びそれらを望ましくない目的のために利用する者がいることを認識する能力を人々が持つことが重要である。なお、データのバイアスには、主として統計的バイアス、社会の様態によって生じるバイアス及び AI 利用者の悪意によるバイアスの 3 種類があることを認識していることが望まれる。***

***B) AI の利用によって、多くの人々が創造性や生産性の高い労働に従事できる環境が実現できることが望ましい。そのためには、出自、文化、趣向等の観点で、多様な人々が各々の目指す多様な夢やアイデアを AI の支援によって実現する能力を獲得できることが期待される。このことを実現するための教育システム及びそれらの達成に資する社会制度が実現されなければならない。***

***C) データや AI の基礎教養から実装及び設計等の応用力を、幅広い分野の横断的、複合的及び融合的な枠組みで身につけた人材が十分に存在する。そのような人材は、社会のあらゆる活動の原動力となり、かつその人々の能力が AI を活用した生活環境の構成に寄与することが期待される。このような生活環境の整備によって、多くの人々がより豊かで充実した人生を送れるような社会制度が実現されている。***

***ご意見（全角 1000*** 文字以内でおまとめください）：社会原則上のＡＩ利用者側の人間に期待される能力及び役***割や社会制度を考える上で、ＡＩのバイアスの特性的な特徴の広義の意味で指標が必要だと考えます。***

１：ＡＩバイアス***(偏り)は、人間中心の社会原則でありながら、主体の人間側の提供するデータ(個々人の社会常識や倫理観)や、特定の偏った規範集団や、故意や無意識な結果的に悪意となってしまった想定外(人間中心ではない)の処理結果となる可能性があり、初期時点データやアルゴリズムには多くの一見ガベージに見えるものが多発する可能性が考えられます。***

２：ＡＩのバイアス（偏り）への数値化視覚化の指標が必要です。

３：上記のＢとＣ部分ですが、現社会制度では知的生産階級が企業・社会組織の中での地位が非常に低く、また、***主能力意外の補助能力を重要視する企業・社会制度と組織風土の抜本的な変革が必要です。***

Ø　該当箇所：４ ***人間中心のＡＩ社会原則***

***４．１ AI 社会原則***

***（7） イノベーションの原則***

***AI を効率的かつ安心して社会実装するため、AI に係る品質や信頼性の確認に係る手法、AI で活用されるデータの効***

***率的な収集・整備手法、AI の開発・テスト・運用の方法論等の AI 工学の確立が推進されなければならない。***

***ご意見（全角 1000*** 文字以内でおまとめください）：１：日本の文化風習に基づいた倫理・道徳の調和状態(***推奨基準)を満たしているＡＩかどうかの判別や認定基準が必要と考えます。***

２：状態(***健全性)をテストするＡＩ認定機関が必要と考えます。***
　ア：利用者に上記***認定(基準値内)された日本の汎用ＡＩかどうかの識別方法***
　イ：バイアスが大きく発生したＡＧＩに対する対策(***防御方法)***

u　**富士通株式会社（20190214）**

Ø　該当箇所：**1.** ***はじめに*** **(P.1)**

***この文書において、我々は、「AI」という用語について特定の技術を指すのではなく、広く「高度に複雑な情報システム一般」を指すものとして捉え…***

***ご意見（全角*** **1000** 文字以内でおまとめください）：本原則（案）***のAI の定義は「高度に複雑な情報システム一般」とされており、対象となる情報システムの範囲が非常に広く曖昧なものとなっています。総務省の「国際的な議論のための　AI 開発ガイドライン案」では AI について「AI ソフト及び AI システムを総称する概念」と定義しており、本原則においても同じ定義を適用すべきであると考えます。***

Ø　該当箇所：**4-1. AI *社会原則（6）公平性、説明責任及び透明性の原則*（P. 9）**

***「AI を利用しているという事実、AI に利用されるデータの取得方法や使用方法、AI の動作結果の適切性を担保する仕組みなど、状況に応じた適切な説明が得られなければならない」***

***「上記の観点を担保し、AI を安心して社会で利活用するため、AI とそれを支えるデータの信頼性（Trust）を確保する仕組みが構築されなければならない」***

***ご意見（全角*** **1000** 文字以内でおまとめください）：「～なければならない」としていますが、開発者に対する責務と捉え***られ、誤解を生みだす可能性があります。AI を活用したサービスは、用途によって求められる動作結果の適切性や信 頼性が異なります。全ての AI を活用したサービスにおいて過剰な動作結果の適切性や信頼性を求めることは AI の開発や普及を阻害する懸念があります。従って、「用途に応じた動作結果の適切性や信頼性を確保する仕組みが構築されるよう努めなければならない」という記載が望ましいと考えます。***

Ø　該当箇所：**5.** ***おわりに*** **(P.10)**

***「…我が国は、本原則を政府、関係企業、団体等で共有し、政策等に反映させるべきである」***

***ご意見（全角*** **1000** 文字以内でおまとめください）：本原則を政策等に反映させる際、規制やガイドライン等を検討す***る場合は改めて関係ステークホルダーを交えた議論を行っていただきたいと考えます。***

u 個人 H（20190214）

Ø 該当箇所：【表題】

［はじめに］

［基本理念］

［Society5.0 実現に必要な社会変革「AI-Ready な社会」］

ご意見（全角 1000 文字以内でおまとめください）：【表題】「人間中心の AI 社会原則(案)」を「格差社会を克服する人間中心の AI 社会原則（案）」とする。格差社会を克服することにより、持続可能な社会が構築される。

［はじめに］

「直面し、社会的課題のいわば先進国」という表現は皮肉的な表現、「直面しています。」としたほうが良い。

・第３段落について、「我が国は、AI の活用により、経済発展と社会的な課題を目指すが、それは格差社会を克服することが第一義的に考えられなければならない。すなわち富の再配分と地域間格差が AI 利用の再優先課題である。Society5.0 についてはバラ色の世界であるかのような表現が見受けられるが、格差社会を克服しすることにより、多様性を認めることができ、人間中心で心が通じ合う社会を目指すものである。」と変更すべき。地球規模以下は OK です。

・第４段落について

「イノベーション」や「AI-Ready」を削除

「AI はその社会への影響が大きいため、適切な社会実装が求められる。すなわち経済的合理性の前提として、倫理や社会正義についての検討が必要である。昨今、戦争のための無人戦闘機やドローン、労働者を弱みを見つけるためのビデオカメラの設置、消費者のインターネット履歴による執拗な商品アピールなどについては経験する話である。Socity5.0 や AI-Ready 社会の構築が最終目的なのか、その都度、倫理や社会正義について議論が必要である。」

［基本理念］

第一段落

「社会のあり方の質的変化を通じて〜真のイノベーションを通じて」を削除

第二段落

（２）について

「多様な幸せを追求するためには、格差社会の是正が前提でありその上で、多様な生き方や創造性が発揮される。格差社会の是正が、AI の適切な開発と展開となる。」に変更

（３）について

「科学・技術立国としての」は削除、それ以降は「その科学的・技術的蓄積を AI によって強化し、格差社会を是正することにより持続可能な社会をつくることが必要である」に変更

［Society5.0 実現に必要な社会変革「AI-Ready な社会」］

・題目の変更

「AI によって社会変革『格差の無い社会』を実現しよう」

・最初の部分

『AI の定義が、高度に複雑な情報システム一般』と定義している。

IoT,ロボテックス、超高速通信網等も AI では？「超高速通信網や人工知能などの AI がある」となる。

Ø 該当箇所：【Society5.0 実現に必要な社会変革「AI-Ready な社会」］

［人間中心の AI 社会原則］

ご意見（全角 1000 文字以内でおまとめください）：【Society5.0 実現に必要な社会変革「AI-Ready な社会」】
（１）について
問題点を人とか我々に責任を転嫁、問題点の取組は国が率先すべき
（２）社会システム
「これには、社会的に～あらたな価値の実現や、」を削除
以下「AI の進化によってもたらされる可能性のある負の側面（不平等や格差の拡大、社会的排除等）を第一義的に捉えてその対応が求められる。」に変更「そのためには～必要がある」が続く。
（３）産業構造について「労働・雇用環境」を削除
（４）イノベーションについて
「（４）イノベーション」の項目全部の削除
イノベーションという言葉を否定する理由は、イノベーション進めると資本主義自身が崩壊する。崩壊しないようにするためには、規制緩和、労働力の流動化が必死となり、格差社会の実現から遠のく政策になるからである。
（５）ガバナンスについて
「イノベーションシステム」の削除
ステークホルダーといいう表現で多くの人がフラットに評価・意思決定をするような表現である。
まず前提として格差社会のなくすることが目的であり、企業・研究機関においては、倫理や社会正義を前提にした経済合理性について、検討をする必要がある。儲かれば良いという事を排除し、社会正義の実現に邁進することが必要である。
【人間中心の AI 社会原則】
・最初の部分
「我々は、『AI-Ready な社会』」を「我々は『格差のない社会』」に変更
・４．１AI 原則について
「AI －Ready な社会」を「格差のない社会」に変更
（１）について
「（１）人間中心の原則」を「（１）格差社会を克服する、人間中心の原則」に変更
第二段落「適切な仕組を導入することが望ましい。」を「適切な仕組を政府・国会が中心となり導入する」に変更
第二パラグラフ
人間関係や人々の繋がり、心の問題など、実際の人と人の繋がりが重視される場面（保育、教育、介護、障害者・高齢者支援等）においては、AI は導入をさせないなど、明確な線引きが必要である。
「AI の利用にあたって、自らがどのように利用するか判断と決定を行うことが求められるが、難しいことが多い。その人の状態に合わせた AI の利用が求められ、貧困層や情報弱者、技術弱者については「本人のペースに合わせた、支援が必要であり、支援に対する個人負担は無償であることが望ましい。」と変更する。

Ø 該当箇所：【人間中心の AI 社会】
（２）教育・リテラシー
（５）構成競争確保の原則

ご意見（全角 1000 文字以内でおまとめください）：・（２）教育・リテラシーの原則について
第一段落「知識と倫理」を「倫理や社会正義、もちろん経済合理性の前提としてであり、」に変更
（２）教育とリテラシー
第二段落、第三パラグラフの最終文以降に次の文を付加

「教育も必要であるが、AI 導入で改善されるべき問題は格差社会の是正であり、持続可能な社会を維持していくためには所得の再配分機能が必要である。教育も行うが、それで終わりではなく、格差社会を終わらせるためにという視点が必須である。」
（５）「AI の利用によって、富や社会に対する影響力が一部のステークホルダーに不当過剰に偏る社会であってはならない。」を
「AI の利用によって、富や社会に対する影響力が一部のステークホルダーに不当過剰に陥る社会であってはならない。特に私企業は経済合理性や利益第一主義が企業存続のための原則と成っている。私企業については、倫理や社会正義からのチェックが必要である。」に変更
（６）「AI-Ready な社会」を「格差のない社会」に変更
第一段落
第二パラグラフの後に以下の文書を挿入
「『AI に利用されるデータの取得方法や使用方法、AI の動作結果の適切性を担保する仕組みなど、状況に応じた適切な説明』や『人々が AI の提案を理解して判断するために、AI の利用・採用・運用について、必要に応じて開かれた対話の場』については、企業内であれば、労使という関係で、話し合いの場が設けられなければならない。」
第三パラグラフはそれに続く
（７）「（７）イノベーションの原則」に書かれていることすべて削除
４．２AI 開発利用原則について
第二段落を以下のように変更
「AI の開発利用原則については～ことから、我々は地に足のついた議論を行い、各国の独自性を考慮にいれ、それが規制的なものや拘束的な枠組みでも許容する多様性が重要である。それでこそ多様性を認める社会が構築で来るのである。」
５．おわりに
第一段落
「AI＝Ready な社会」を「格差の無い社会」に変更
第二段落を以下のように変更
「また、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、国際的な議論を行い、それによって SDGs の実現を支える社会像を共有し、国際社会の協調的かつ創造的な新たな発展に寄与するものである。

以上コメントをつけなかった部分は、原案を踏襲する。

Ø 該当箇所：

ご意見（全角 1000 文字以内でおまとめください）：●人間中心の AI 社会原則（案）についての私の意見のまとめ

・AI の利用は格差社会を克服するためのものであり、その成果は、格差社会の克服のために使われるべきである。
・多様な生き方や創造性が発揮されるためには、格差社会の是正が前提であり、それにより持続可能な社会が構築できる。
・AI による新技術の開発や利用は、倫理や社会正義を前提にして、初めて経済合理性の検討がなされるべきである。
・特に私企業においては、経済合理性や利益第一主義が企業存続のための原則であり倫理や社会正義からのチェックが必要であり、労使という視点でチェックされるべきである。

***・人間関係や人々の繋がり、心の問題など、実際の人と人の繋がりが重視される場面（保育、教育、介護、障害者・高齢者支援等）での AI 導入に関して、明確な線引きが必要である。***
***・貧困層や情報弱者、技術弱者については本人のペースに合わせた、AI 支援が必要であり、支援に対する個人負担は無償であることが望ましい。***

***以上***

（別紙様式）

# 意見書

平成 31 年 2 月 14 日

**内閣府 政策統括官**
**（科学技術・イノベーション担当）**
**社会システム基盤担当 御中**

〒100-0014

東京都千代田区永田町 2 丁目 17 番 17 号 AIOS 永田町 312 号

一般財団法人情報法制研究所 AI 問題タスクフォース

代表：鳥海不二夫

電話番号：03-6205-8183

電子メールアドレス tori@jilis.org

**「人間中心のＡＩ社会原則」（案）について、別添のとおり、意見を提出します。**

（別添様式）

課題名：「人間中心のＡＩ社会原則」

| 該当箇所（ページ数等） | 御意見 |
| --- | --- |
| P。1<br>この文書において、我々は、「AI」という用語について特定の技術を指すのではなく、広く「高度に複雑な情報システム一般」を指すものとして捉え、社会に与える影響を議論した上で、AI 社会原則の一つの在り方を提示し、AI の研究開発や社会実装において考慮すべき問題を列挙する。来るべき Society 5。0 がより良いものとなるためには、関係ステークホルダーが対話しながら協力していくことが必要不可欠である。 | 「高度に複雑な情報システム」では曖昧すぎる。<br>技術者が AI といった場合は、自律的に動く AI と、機械学習を主に指すが、ここでの定義はどちらも指していないため、AI を対象とした社会原則よりも広い対象ということになる。<br>ここでは、道具としての AI を対象としていると考えると、強い AI や汎用 AI のような自律型システムは定義から明示的に外したほうが良い。<br>また、明らかに AI とは認められないシステムについても定義から外すべきである。<br>万が一、定義を明確にできない場合は、ケースバイケースで対応することを明記するべきである。<br>将来の技術についての対応を考慮して幅の広い書き方になっているのだとすれば、一定の期間で本原則を見直すことによって、その都度対象となる技術も同時に見直すことで対応すればよいのではないか。<br>また、本報告書が国際的な議論の場においては日本政府の提案の原型となる以上、総務省の AI ネットワーク等、政府の他の部局での議論との整合性をとり、定義を設定すべきである。 |

| | |
|---|---|
| 【提案】見直し規定の導入 | AI をめぐる急速な技術進歩を考えると今回提案されている原則が長期的に有効性を保つことは想定できない。3 年もしくは 5 年程度の期間で見直しを行う旨のシステムを用意しておくべきであると思われる。<br>また、見直しのタイミングが明確になるような文言が必要である。<br>なお、本コメントにおいて指摘している多くの点が数年単位の見直しを想定することで対応可能になると考えられる。 |
| P1 「はじめに」<br>「SDGs（Sustainable Development Goals）で掲げられている 17 目標を解決し」 | 左記のように記述しているのにも係わらず、「基本理念」より後には、資源を含む環境問題や気候変動など SDGs 全てに十分に言及できているわけではなく誤解を生む。<br>「SDGs（Sustainable Development Goals）で掲げられている 目標を解決し」<br>といった表現に留めるてはどうか。 |
| P。3 その科学的・技術的蓄積を AI によって強化し、そのような社会を作ることに貢献する責務がある。 | AI によって強化するとはどのような意味か？「高度に複雑な情報システムによって強化する」では意味が通らない。<br>ここでは、「AI 技術によって科学技術の発展を加速させ」などがよいと考えられる。 |
| P3 人間の尊厳が尊重される社会（Dignity）<br>「人間の様々な能力をさらに発揮することを可能とし、より大きな創造性を発揮したり、やりがいのある仕事に従事したりすることで、物質的にも精神的にも豊かな生活を送ることができるような、人間の尊厳が尊重される社会を構築する必要がある。」 | AI を使ったシステムの普及が、効率化を進め、結果として一部の職業や労働が無くなる恐れがあるという懸念はあるが、ここで触れられている「やりがいのある仕事に従事したりすること」と「人間の尊厳が尊重されること」の関係が十分に説明できておらず、やや飛躍している。<br>「やりがいのある仕事に従事したりすること」を削除する事が望ましい。 |

| | |
|---|---|
| P4 「3 Society 5。0 実現に必要な社会変革「AI-Ready な社会」 」 (1)<br>「人」B)<br>「多くの人々が創造性や生産性の高い労働に従事できる環境が実現できることが望ましい。」 | AI を使ったシステムの普及が、効率化を進め、結果として一部の職業や労働が無くなる恐れがあるという懸念はあるが、ここで触れられている「高い労働に従事できる」という記述は、やや理論の飛躍がある。<br>「多くの人々がより大きな創造性や高い生産性を発揮できる環境が実現できることが望ましい」と労働に限定しない事が望ましい。 |
| P。5<br>(4) 「イノベーションシステム」 | イノベーションとは、シュンペーター的な文脈では「既存技術の新たな結合による新たな価値の創生」を意味する。その場合は、イノベーションシステムではなく、「イノベーションを支援する環境」等とすべきである。<br>一方、イノベーションを一般的な用法としての「革新的な技術」という意味で用いるのであれば、誤解を防ぐ意味からも「革新的な技術を生み出すシステム」と修正すべきである。 |
| P5 ガバナンス<br>「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が求められる。」 | 仮の本分章を世界に向けての発信であれば、「国内におけるガバナンスに加えて」と国内のガバナンスに限定せず「世界各国におけるガバナンスに加えて」とすべきではないか。<br>また、本原則をより実効性のあるものとするためには、ガバナンスに対する監査基準を示すことが望ましい。 |
| P。6<br>(1) 人間中心の原則 | すでに人間社会が技術に依存している中、AI に特化して問題にするべきなのかどうか。<br>なぜ今 AI について議論しなければいけないのかを最初に明確にすべきである。 |

| | |
|---|---|
| P。6<br>AI の利用者側は、AI が従来のツールよりはるかに複雑な動きをするため、その概要を理解し、正しく利用できる素養を身につけていることが望まれる。一方、AI の開発者側は、技術の基礎を習得していることが当然必要であるが、それに加えて、社会で役立つ AI の開発の観点から、AI が社会においてどのように使われるかに関するビジネスモデル及び規範意識を含む社会科学や倫理等、人文科学に関する素養を習得していることが重要になる。 | 「AI の利用者」「AI の開発者」という言葉が出ているが、総務省の AI ネットワーク社会推進会議における定義との整合性を取る必要がある。<br>　現状では「AI の利用者」がどのような人々を想定しているのか、「AI の開発者」が誰を想定しているのか不明瞭であるため、AI ネットワークでの定義との関係を明確にすべきではないか。 |
| P7　教育・リテラシーの原則 | AI の利用者および提供に係わる人に焦点を当てているが、利用しない人に対する教育も必要である。AI 社会において、AI を利用していないが他者が利用する事で影響をうける第三者（受益者等）も、AI が実装される社会で生活するために必要な知識についても身につけるべきである。例えば、車社会で考えると、ドライバーだけでなく、歩行者にも交通ルールを教えることと同様である。AI を利用していないが他者が利用する事で影響をうける第三者に対する教育についても追記するのが相応しいのではないか。 |
| P。7<br>「社会は、これまでの教育環境における成功体験に拘ることなく、常に最適な形へと柔軟に変化し続ける意識を全体として共有する。」 | 本記述について支持する。<br>一方で、「常に最適な形へと柔軟に変化する」ためには固定化された原則をもってしては困難であることからも、本原則の数年単位での見直しが必要であると考える。 |
| P7<br>「このような教育環境の整備に向けて、行政や学校（教員）に負担を押し付けるのではなく、民間企業や市民も主体性をもって取り組んでいくことが望ましい。」 | このような配慮について記述がある事を支持する。<br>一方で、教育制度の確立が遅れた場合、直ちに AI 開発者は本原則にしたがう事が出来る AI 提供者に過度な責任がかかる事は、提供側の萎縮に繋がる恐れがある。<br>よって、教育環境の迅速な整備と普及が望まれる旨の追記があるべきではないか。 |

| | |
|---|---|
| P7　(3)　プライバシー確保の原則<br>「パーソナルデータを利用する AI は、当該データのプライバシーにかかわる部分については、正確性・正当性の確保及び本人が実質的な関与ができる仕組みを持つべきである。」 | プライバシーとはどういった事柄を指すのかが定義されておらず、読み手が理解できない恐れがある。ここで表現するプライバシーを定義するか、別の表現に変えるべきではないか。<br>また、現行の文案は、開発者など提供者にかかるものであると誤解を生じさせる。社会的枠組みに関する原則を記述したAI 社会原則の中で記述するのであれば、「パーソナルデータを利用する AI は、例え国や政府における利用であっても、当該データのプライバシーにかかわる部分については、正確性・正当性の確保及び本人が実質的な関与ができる仕組みを持つべきである。」のように、誤解のないように記述すべきではないか。 |
| P。8（5）　公正競争確保の原則 | 対中国、対 GAFA を想定したものであると考えられるが、<br>はじめに、に<br>「中でも我が国は、少子高齢化、人手不足、過疎化、財政支出増大等に直面し、社会課題のいわば先進国となっている。我が国は、AI の活用により、経済発展と共に社会課題を解決する Society5。01 の実現を通しつて、日本の社会と経済の活性化を実現し、国際的にも魅力ある社会を目指すと共に、地球規模での SDGs への貢献も果たしていく。」<br>とあるため、国内に向けたメッセージのように感じられる。<br>国際的なメッセージであることを明確にするためにも、<br>おわりに、にある<br>「国際的な議論の場において、・・・」<br>以下をはじめに、に持ってくることによって、本原則の立ち位置を明確にすべきである。 |
| P8<br>「その利活用と保護のバランスについては、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がある。」 | パーソナルデータの扱いに関してこのような記述があることを支持する。 |
| P8<br>「AI の設計思想の下において、人々がその人種、性別、国籍、年齢、政治的信念、宗教等の多様なバックグラウンドを理由に不当な差別をされることなく、全 | 文中に”設計”とあるように、現行の文案は、開発者など提供者にかかるものであると誤解を生じさせる。AI の利用範囲の広さを考えれば、利用者の利用方法すべてを開発者が想定することは困難であ |

| | |
|---|---|
| ての人々が公平に扱われなければならない。」 | る。そのような場合に生じた問題について、その責を開発者に負わせるような記述は避けるべきである。AI 社会原則としては、「AI の設計思想の下において、」の部分を「AI の社会実装において、」としてはどうか。 |
| P9<br>「AI を効率的かつ安心して社会実装するため、AI に係る品質や信頼性の確認に係る手法、AI で活用されるデータの効率的な収集・整備手法、AI の開発・テスト・運用の方法論等の AI 工学の確立が推進されなければならない。」 | AI 社会原則を考えたときに、工学的なアプローチだけでは不足していないか。AI 工学を含め、倫理的側面、経済的側面など幅広い学問の確立および発展が推進されるべきではないか。 |
| P.9<br>「我々は、開発者及び事業者において、基本理念及び上記の AI 社会原則を踏まえた AI 開発利用原則を定め、遵守するべきと考える。」 | すでに指摘しているように「人間中心の AI 社会」を論ずるにあたり、開発者，事業者のみならず利用者に関しても考慮する必要がある。<br>AI 開発利用原則における「AI 利用者」の立場を明確にするべきではないか。<br>なお、利用者と事業者は明確に異なるものである。 |
| P9　AI 開発利用原則<br>・開発者及び事業者において、基本理念及び上記の AI 社会原則を踏まえた AI 開発利用原則を定め、遵守するべき。<br>・AI 開発利用原則については、現在、多くの国、団体、企業等において議論されていることから、我々は早急に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みとして国際的に共有されることが重要 | 本項は以下のように解釈できるが、それでよいのか。<br>・開発者および事業者が原則を踏まえながら、独自の開発利用原則を定めればよい。<br>・開発利用原則の制定には国際的でオープンな議論があればよい。<br>・開発利用原則は定めても守らなくてもよい。<br>以上のことから企業ホームページにおいて、以下のような対応が考えられる。<br>「当社における AI 開発利用原則は URL＊＊に示す通りです。AI 開発利用原則の制定に関わる議論は＋＋国際会議を通じて議論した結果です。当社において開発利用原則は規制や拘束にあたりません。」 |
| P.10<br>「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、今後柔軟に進化・発展させるものである」 | 具体的にどのようなタイミングで進化・発展させるのかを明記されたい。<br>また，進化発展のみならず、改定、全面的な見直しも含めて考えるべきではないか。 |

| | |
|---|---|
| 全体について | 米国の PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE は、”In general, the approach to regulation of AI-enabled products to protect public safety should be informed by assessment of the aspects of risk that the addition of AI may reduce, alongside the aspects of risk that it may increase. If a risk falls within the bounds of an existing regulatory regime, moreover, the policy discussion should start by considering whether the existing regulations already adequately address the risk, or whether they need to be adapted to the addition of AI. Also, where regulatory responses to the addition of AI threaten to increase the cost of compliance or slow the development or adoption of beneficial innovations, policymakers should consider how those responses could be adjusted to lower costs and barriers to innovation without adversely impacting safety or market fairness.”と記述し、AI の追加によって減少する可能性のあるリスクと、増加する可能性のあるリスクの評価を求め、リスクが既存の規制制度の範囲内にある場合は、既存の規制がすでにリスクに適切に対処しているかどうか、または AI の追加に適応する必要があるかどうかを検討することから政策議論を始めるべきと指摘し、AI を包括的に規制することに対して謙抑的な立場を取っています。<br><br>EU の DRAFT ETHICS GUIDELINES FOR TRUSTWORTHY AI は 、 ”The High-Level Expert Group on AI (“AI HLEG”) believes in an approach to AI ethics that uses the fundamental rights commitment of the EU Treaties and Charter of Fundamental Rights as the stepping stone to identify abstract ethical principles, and to specify how concrete ethical values can be operationalised in the context of AI. The EU is based on a constitutional |

| | |
|---|---|
| | commitment to protect the fundamental and indivisible rights of human beings1, ensure respect for rule of law, foster democratic freedom and promote the common good. Other legal instruments further specify this commitment, like the European Social Charter or specific legislative acts like the General Data Protection Regulation (GDPR). “と記述し、欧州連合基本権憲章を中心とした EU 法に基づく倫理観を AI に実装することを求めています。<br><br>対して本案では、「AI はこれらの問題の解を導き、SDGs（Sustainable Development Goals）で掲げられている 17 目標を解決し、持続可能な世界の構築するための鍵となる技術」と位置付けながら、「人間中心の AI 社会原則」制約を課すことによって、なぜ、どのようにして、AI がSDGs の 17 目標を達成するための手段として効果を増すのかの説明がなされていません。また、SDGs17 目標と「人間中心の AI 社会原則」7 原則の対応関係も明確化されていません。<br><br>米国が自由な研究開発と市場適用の推進とアプリケーション個別のリスク評価と対応する既存規制の改善による AI の社会適応を志向し、EU がEU 法の価値基準への準拠を求めているところ、国連の掲げる SDGs17 目標の達成を促進する AI 活用を掲げる方向性には賛同できますが、SDGs17 目標達成にいかに「人間中心の AI 社会原則」が資することができるかを掘り下げて論じていない現行案は目的（ゴール）設定と手段の選択が明瞭な米欧と比較していたずらに国内での研究開発及び市場化を遅らせるだけの「非拘束的」原則となる虞があると考えます。<br><br>せめて SDGs17 目標達成が日本の国益といかに結びついているのか追加で解説が必要と考えます。 |

[意見書・要旨]

BSA は、本原則が、責任ある AI の開発、導入、利用に取り組む上で、重要かつ前向きな一歩であると考えています。特に、AI が潜在能力を十分に発揮する上で、マルチステークホルダーが重要な役割を果たすという本原則の認識は、極めて重要な洞察です。「Society 5.0」のビジョンを推進するには、究極的には、社会全体を巻き込むアプローチが求められます。そのために、BSA は「AI-Ready な社会」の推進に向けた社会理念の策定を強く支持するとともに、本ビジョンの進展に役立つよう提言をします。

以下は意見の要約であり、詳しくは提出意見をご参照ください。

1. AI 原則は、特定の利用事例に関連するその分野特有のニュアンス及び AI の開発、展開、利用に関与する様々なステークホルダーを考慮した柔軟なものであるべき です。
2. 健全な AI 政策の策定には、マルチステークホルダーの関与のためのプロセスと仕組みが不可欠です。
3. AI 原則は、強固で競争力ある AI エコシステムを創出するために、市場アクセスと技術中立性を推進すべきです。
4. AI 原則は、国内及び越境データの自由な移転を明示的に保証すべきです。
5. AI 原則は、プライバシーに関してリスクベースのアプローチを採用し、過度に規範的になることを避け、現実の損害の防止を重視すべきです。
6. AI 原則は、セキュリティを推進し、AI ツールにおける信頼の基本的構成要素として安全性を組み込むべきです。
7. AI 原則は、AI システムにおいて、公平性、説明責任、透明性を達成する全ての場合に通用するアプローチは存在しない、ということを認識すべきです。
8. AI 原則は、機密でない政府データをより利用可能とし、データ分析に対する障壁を取り除くことにより、イノベーションを推進すべきです。

February 14, 2019

__Respectfully to__: **Council for Science, Technology and Innovation, Cabinet Office**

## BSA COMMENTS ON DRAFT SOCIAL PRINCIPLES ON HUMAN-CENTRIC AI

We offer below a series of recommendations to help advance this vision. These recommendations are summarized in the following paragraph, with more details set-out in the **Annex**:

1. **AI principles should be flexible** to account for context-specific nuances that are implicated by specific use cases and for the multiple stakeholders involved in the development, deployment, and use of AI;
2. **Multi-stakeholder engagement processes and structures are critical** to the development of sound AI policy;
3. **AI principles should promote market access and technology-neutrality** to create a robust and competitive AI ecosystem;
4. **AI principles should explicitly ensure the free movement of data within and across borders**;
5. **AI principles should take a "risk-informed" approach to privacy that avoids over-prescriptiveness and maintains focus on preventing actual harms**;
6. **AI principles should promote security and incorporate safety as a fundamental component of trust in AI tools**;
7. **AI principles should acknowledge that there is no one-size-fits all approach for achieving fairness, accountability, and transparency in AI systems**; and
8. AI Principles should promote innovation by increasing the **availability of non-sensitive government data and eliminating barriers to data analytics**.

## Annex: Detailed Recommendations

| Paragraph (Pg. No) | Recommendations/Comments |
|---|---|
| General Comments | BSA \| The Software Alliance (**BSA**) [1] thanks the Cabinet Office for the opportunity to submit comments on the Draft Social Principles on Human-Centric AI (**Draft Principles**).<br><br>BSA is the leading advocate for the global software industry before governments and in the international marketplace. Our members are at the forefront of software-enabled innovation that is fuelling global economic growth, including cloud computing and AI products and services. As leaders in AI development, BSA members have unique insight into both the tremendous potential that AI holds to address a variety of social challenges and the governmental policies that can best support the responsible use of AI and ensure continued innovation.<br><br>BSA welcomes the Draft Principles. They provide an insightful and important articulation of how principles and policies surrounding AI development and utilization should place humans at the center of these considerations. The Draft Principles also provide an appropriate and balanced discussion on how AI innovation can enhance human dignity and support sustainable development goals.<br><br>The responsible use of AI has the potential to spur tremendous economic growth across every sector, improve human decision-making, and enable cutting-edge breakthroughs on some of the world's most pressing challenges. Conversely, AI services that are not developed or used responsibly, like other ground-breaking innovations, create a risk of unintended consequences or malicious uses. Governments are therefore rightly focused on developing thoughtful policies to address both the opportunities and risks associated with AI.<br><br>BSA and its members are attuned to these challenges and are committed to developing AI in a manner that will maximize the benefits and minimize the risks associated with the deployment of the technology. BSA has published a range of materials related to AI, including AI Policy Principles, an AI Primer, and examples of AI being applied in different sectors.[2]<br><br>BSA has also worked closely with governments around the world on AI policy development and has provided strategic advice from industry's perspective on how government policy approaches should encourage responsible use of AI in order to foster trust |

[1] BSA's members include: Adobe, Akamai, Amazon Web Services, Apple, Autodesk, AVEVA, Bentley Systems, Box, Cadence, Cisco, CNC/Mastercam, DataStax, DocuSign, IBM, Informatica, Intel, MathWorks, Microsoft, Okta, Oracle, PTC, Salesforce, Siemens PLM Software, Slack, Splunk, Symantec, Synopsys, Trend Micro, Trimble Solutions Corporation, Twilio, and Workday.

[2] https://ai.bsa.org/

| | |
|---|---|
| | among businesses and consumers and ensure continued innovation.<br><br>In our view, the Draft Principles represent an important and positive step towards addressing responsible development, deployment and use of AI. The Draft Principles' recognition that multiple stakeholders have important roles to play in unlocking the full potential of AI is a critically important insight. Ultimately, advancing the vision of "Society 5.0" will require an all-of-society approach. To that end, we are strongly supportive of the goal of developing Social Principles to help promote an "AI-Ready society." We offer below a series of recommendations to help advance this vision. These recommendations are summarized in the following paragraph, with more details set-out in the **Annex**:<br><br>1. **AI principles should be flexible** to account for context-specific nuances that are implicated by specific use cases and for the multiple stakeholders involved in the development, deployment, and use of AI;<br>2. **Multi-stakeholder engagement processes and structures are critical** to the development of sound AI policy;<br>3. **AI principles should promote market access and technology-neutrality** to create a robust and competitive AI ecosystem;<br>4. **AI principles should explicitly ensure the free movement of data within and across borders**;<br>5. **AI principles should take a "risk-informed" approach to privacy that avoids over-prescriptiveness and maintains focus on preventing actual harms**;<br>6. **AI principles should promote security and incorporate safety as a fundamental component of trust in AI tools**;<br>7. **AI principles should acknowledge that there is no one-size-fits all approach for achieving fairness, accountability, and transparency in AI systems**; and<br>8. AI Principles should promote innovation by increasing the **availability of non-sensitive government data and eliminating barriers to data analytics**. |
| Section 2, Pg. 3: on **Dignity, Diversity and Inclusion, and Sustainability** | BSA strongly agrees that AI systems and applications should be developed, deployed, and used in ways that **promote *Dignity, Diversity & Inclusion, and Sustainability***. BSA supports broad access to the benefits of AI, particularly because many applications of AI will contribute greatly to improving underserved members of society.<br><br>***Recommendations:*** |

| | |
|---|---|
| | BSA recommends including:<br><br>· Acknowledgement that AI principles should be flexible enough to account for context-specific nuances, including considering the differences between uses that are consumer-facing and those that are developed as enterprise solutions or that use data solely from machine-to-machine communication. |
| Section 3, Paragraph 1, pg. 4, on the need to consider "**interaction with technological progress**";<br><br>and<br><br>Section 3, Paragraph 6, pg.4, on "**Social Systems**" | BSA supports the Draft Principles' recognition that AI principles should be applied to the entire "social system" including multiple sectors (e.g. medical, financial, and energy) and should be flexible, so that the respective sectors can evolve and cope with the evolution of AI technology and its unique effects on these sectors.<br><br>Principles and policies directed at AI should take into account the diverse set of underlying technologies and use cases that comprise the AI ecosystem. Exciting and beneficial new applications of AI are constantly being developed, making it difficult to predict the full range of potential uses cases. It is therefore important that any principles or considerations identified at a single point in development time not remain static but evolve as the technology and its application develops. Hence a principles-based approach to governing AI is preferred to one that involves binding regulation.<br><br>***Recommendation:***<br><br>· The Draft Principles should include overt recognition that AI principles should be flexible enough to account for the multiple stakeholders involved in the development, deployment, and use of AI, and the context-specific nuances that are implicated by the diverse range of AI applications (e.g., including considering the differences between uses that are consumer-facing and those that are developed as enterprise solutions or that use data solely from machine-to-machine communication). |
| Section 3, Paragraph 8, pg. 5, on the need to create "**industry structures**" that are conducive to innovation and internationally open | BSA supports the recognition that industry structures should be "flexible and open internationally."<br><br>***Recommendations:***<br>· As an overt recognition of the importance of the digital economy and trade policies that drive job creation, competitiveness, and innovation, the Draft Principles should include a broad principle that there should be **no market access barriers and no discrimination against innovative AI applications and services**.<br><br>· In addition, policies should also support **the free movement of data across borders**, given the importance of international data transfers to the development of AI. In addition, free movement of data also allows for businesses at every level, including small and medium business, to have access to cutting-edge AI services, many of which rely on the seamless international transfer of data. |

| | |
|---|---|
| Section 3, Paragraph 12, pg. 5 on the establishment of an agile system of “diverse stakeholders” to be engaged on AI governance issues; and paragraph 13 on “international collaboration systems”. | BSA welcomes the document’s focus on multi-stakeholder involvement. As AI systems are used in many different ways, it is important to ensure that best practices and recommendations are developed in collaboration with all involved stakeholders, and grounded in the necessary technical considerations. In addition, such best practices and recommendations should incorporate an international view.<br><br>***Recommendation:***<br><br>The document can include further consideration on the importance for governments to put in place specific process or structures that ensure multi-stakeholder engagement as AI policies or regulations are developed. For example, encouraging the introduction of an advisory council, or other public-private collaborations, to advise on the development and review of relevant principles as technologies develop.<br><br>In addition, this document should also include a statement on the importance of aligning domestic efforts to internationally-recognized standards and principles. |
| Section 4.1(3), pg. 7 and 8 on Privacy | BSA supports the implementation of privacy practices that protect individual rights and build trust in AI systems and applications. Privacy best practices include those that increase the transparency of personal data collection and use; enable and respect informed choices by providing governance over that collection and use; provide consumers with control over their personal data; provide robust security; and promote the use of data for legitimate business purposes.<br><br>Frameworks for privacy best practices should be risk-based, principle-driven, and eschew overly prescriptive requirements. Incorporating flexibility into such frameworks will ensure individuals are able to exercise appropriate control over their personal information and, at the same time, not stymie the ability of companies to provide innovative products and services.<br><br>***Recommendations***<br><br>· Sub-Paragraph 3 rightly recognizes that personal data should be “properly protected according to its importance and sensitivity.” Other portions of this section would benefit from a better alignment with this important principle. For instance, we recommend a clarification in Paragraph 1 that not all forms of AI pose heightened risks regarding the use of personal data. Rather than suggesting that “more careful treatment of personal data” is needed in the context of AI writ large, it would be helpful to clarify that heightened scrutiny is warranted in circumstances where AI could consequentially “affect the rights and benefits of individuals.” Rather than focusing on the nature of the data itself, or the particular technology or system managing the data (e.g. AI), the Draft Principle should instead focus on the risk of harm to the individual.<br><br>***In cases where the use of AI could result in*** |

| | |
|---|---|
| | ***greater risk of harm to individuals, mechanisms or frameworks should be put in place to address such scenarios."***<br><br>· As currently drafted, Sub-paragraph 2 suggests that all AI systems should include a mechanism for ensuring accuracy and enabling users to be "substantially involved in the management of his or her data." While we support the Sub-paragraph 2 recommendation in principle, we suggest that it too should be revised to allow for a risk-informed approach to determining whether (and what types of) redress mechanisms are needed to address potential risks of consequential harm that a specific AI system poses.<br><br>**<u>We recommend the following modifications to Sub-paragraph 2</u>**:<br><br>*"AI that uses personal data should* ***~~have a fair mechanism~~ incorporate mechanisms ~~that~~ ensures to address risks of consequential harms to the public. Such mechanisms may include tools to ensure*** *accuracy and legitimacy and enable the person* ***~~himself/herself~~*** *to* ***~~be substantially involved in the management of his or her privacy~~ exercise control over their personal*** *data* ***where appropriate***" |
| Section 4.1(4), paragraph 1, pg.8 on Security | BSA is a strong advocate of cybersecurity and resilience. As AI and other digital technologies increasingly enable a globally connected economy, we recognize that society must be vigilant in addressing increased security risks. BSA advocates for policies that strengthen enhanced security measures outlined in our International Cybersecurity Policy Framework.[3] While we broadly support the Draft Principles' discussion on Security, we suggest the following improvements:<br><br>· Paragraph 1, on the *impossibility of AI systems to respond to "rare events or deliberate attacks."* In an effort to future-proof these principles, it is helpful to avoid staking out absolute positions about the potential for AI to address specific problems. It may be more accurate to instead state the following:<br><br>"**~~within the scope of today's technologies, it~~ is ~~impossible for~~ there may be circumstances where** AI is unable to respond appropriately to…"<br><br>· BSA is fully committed to the highest standards in AI development and deployment. Trust can only be earned through safety in practice. We recommend also including the concept of "safety as a fundamental component of trust in AI tools." |
| Section 4.1(5), pg. 8 on fair competition | BSA supports policies that foster fair competition among trading partners and between private firms. AI remains a burgeoning field with quickly evolving market dynamics. Domestic competition |

[3] https://bsacybersecurity.bsa.org/wp-content/uploads/2018/04/BSA_cybersecurity-policy.pdf

| | |
|---|---|
| | policies should therefore remain technology-neutral, avoiding the creation of AI-specific rules. International competition for AI services should be encouraged by eliminating barriers to digital trade and ensuring that foreign markets are kept open. This includes ensuring that cross-border data transfers are permitted, and that data localization mandates and other protectionist measures that favour domestic technologies and producers are prohibited.<br><br>We recommend the following in particular:<br><br>· Promote competition by ensuring the free movement of data within and across borders. Given the importance of cross-border data to the modern economy, governments must use privacy or security policies only as necessary, and never as disguised market access barriers. |
| Section 4.1(6), pg. 8 on fairness, accountability and transparency | BSA agrees that the principles of fairness, accountability, and transparency (**FAT**) are critical to the development of trustworthy AI. BSA members are constantly developing and updating policies and technological solutions to reduce the impact of bias in AI processes — as well as any other processes — to ensure trust and confidence in their products. BSA members are fully committed to ensuring that AI systems respect fundamental rights and norms.<br><br>The ideal mechanisms for building FAT into any particular AI system will vary depending on a variety of factors, so guidance on these issues must remain sufficiently flexible to accommodate different use cases and means of deployment. For instance, the level of transparency that is necessary to support public trust in an AI system that provides restaurant recommendations based upon user-inputted criteria is unlikely to require an intricate level of transparency or explanation. In contrast, when an AI system is deployed in a context that affects consumers' eligibility in consequential areas, such as access to credit or housing, the public will rightly have far greater expectations about the effectiveness of measures the system's developer has implemented to ensure the systems decisions are consistent with high standards of fairness, accountability, and transparency.<br><br>It is also important to recognize that there may be contexts in which efforts to safeguard one ethical principle might come at the expense of others. For instance, there may be circumstances in which design choices aimed at preserving the fairness and accountability of an AI system could involve trade-offs with the degree to which the underlying model can be made transparent or explainable to the public. For example, in designing a fraud detection system, the need to ensure that it is operating in an accurate and unbiased manner may necessitate limited disclosures to the public about how it operates. Indeed, there are instances where too much transparency can have the unintended consequence of making an algorithm vulnerable to manipulation. Moreover, research has shown that disclosing the algorithms, source codes, or associated data sets is ineffective in helping to provide explanations, in part because they cannot be meaningfully understood in isolation. |

| | |
|---|---|
| | ***Recommendations:***<br><br>Sub-paragraph 3, on "*appropriate explanations should be provided…*", is overly prescriptive and sits at odds with the rest of the document that is principle-based. We recommend the sub-paragraph instead state that "~~**appropriate explanations should be provided such as**~~ **assessments for what is an appropriate level of explanation should be carried out, and such explanations may include** the fact that AI is being used…". |
| Section 4.1(7), pg. 9 on Innovation | Data is critical to the development of AI. As a result, a sound approach to data policy is intrinsically linked to the ability for societies to innovate and apply AI. BSA has written about the need for sound data policy to spur AI innovation, providing key recommendations.[4] In particular, government data sets are a resource that can improve the training of AI models, create more inclusive societies, and serve as a catalyst for economic growth. It is likewise important to eliminate unnecessary barriers that may prevent researchers from harnessing data to which they have lawful access for the purpose of training AI systems.<br><br>***Recommendation:***<br><br>In addition to addressing issues related to efficient collection and maintenance of data utilized by AI, the Draft Principles should encourage governments to **make non-sensitive data freely available to the public in machine-readable formats**.<br><br>The Draft Principles should also acknowledge the importance of ensuring that AI researchers are able to use content and data to which they have lawful access for training AI algorithms and performing data analytics and digital analysis. |
| Conclusion | We would like to thank the Cabinet Office again for the opportunity to comment on the Draft Principles. We appreciate your kind consideration of our above comments. For any questions or if any point of clarification is required on any part of this submission, please feel free to contact us. |

[4] See https://ai.bsa.org/wp-content/uploads/2018/05/BSA_2018_AI_DataPolicy.pdf

（別紙様式）

# 意見書

平成３１年２月１４日

内閣府 政策統括官
（科学技術・イノベーション担当）
社会システム基盤担当　御中

郵便番号　150-0043
住所　東京都渋谷区道玄坂 1-12-1
渋谷マークシティ　W22 階
氏名　BSA ｜ ザ・ソフトウェア・アライアンス [1]
（日本代表事務所　代表者　角田　良平）
電話番号 03(4360)5474
電子メールアドレス　tomokon@bsa.org

「人間中心のＡＩ社会原則」（案）について、別添のとおり、意見を提出します。

[1] BSA の活動には、Adobe, Akamai, Amazon Web Services, Apple, Autodesk, AVEVA, Bentley Systems, Box, Cadence, Cisco, CNC/Mastercam, DataStax, DocuSign, IBM, Informatica, Intel, MathWorks, Microsoft, Okta, Oracle, PTC, Salesforce, Siemens PLM Software, Slack, Splunk, Symantec, Synopsys, Trend Micro, Trimble Solutions Corporation, Twilio, Workday が加盟企業として参加しています。詳しくはウェブサイト（http://bsa.or.jp）をご覧ください。

（別添様式）

課題名：「人間中心のＡＩ社会原則」

| 該当箇所（ページ数等） | 御意見 |
|---|---|
| 全体について | BSA ｜ ザ・ソフトウェア・アライアンス（BSA）は、人間中心の AI 社会原則（案）（以下「本原則」）に関するコメントを提出する機会に感謝します。<br><br>BSA は、各国政府の前で、また国際的な市場において、グローバルなソフトウェア業界のための主要な提唱者です。BSA の会員企業は、クラウドコンピューティングや AI の製品・サービスなど、世界的な経済成長を牽引する、ソフトウェアによって実現されるイノベーションの最前線にいます。BSA の会員企業は、AI 開発におけるリーダーとして、AI が様々な社会的課題に対処する大きな可能性及び AI の責任ある利用と継続的なイノベーションの確保を最も良く支援できる政府の政策について独自の洞察力を持っています。<br><br>本原則は、AI の開発と活用をめぐる原則及び政策において、どのように人間をその考察の中心に置くべきかにつき洞察力ある重要な意見を提示しており、BSA は本原則を歓迎します。また、本原則は、AI イノベーションが、いかに人間の尊厳を高め、かつ、持続可能な開発目標を支援するかについて、適切かつバランスのとれた議論を行っています。<br><br>AI の責任ある利用は、あらゆる分野にわたって多大な経済成長をもたらし、人間の意思決定を改善し、世界で最も差し迫った課題のいくつかにおいて最先端のブレークスルーを起こす可能性があります。他方、他の画期的なイノベーションと同様、AI サービスが責任を持って開発または利用されない場合、意図しない結果又は悪意ある利用といったリスクをもたらします。従って、 |

| | |
|---|---|
| | 各国政府は、AI に関連する好機とリスクの双方に対処するための思慮深い政策の策定に注力しており、これは的確です。<br><br>　BSA 及び会員企業は、これらの課題を認識し、技術導入に伴う利益を最大限にしながらリスクを最小限に抑える方法で AI を開発することにコミットしています。BSA は、AI 政策原則、AI 入門書及び様々な分野で利用されている AI の事例を含む、AI に関連する一連の資料を公表しています [2]。<br><br>　また、BSA は、世界中の政府と緊密に連携して AI 政策の策定に取り組んでおり、企業と消費者の信頼を醸成し、継続的なイノベーションを確保するために、政府の政策アプローチが責任ある AI の利用をどのように奨励すべきかについて、産業界の視点から戦略的な提言をしてきました。<br><br>　BSA は、本原則が、責任ある AI の開発、導入、利用に取り組む上で、重要かつ前向きな一歩であると考えています。特に、AI が潜在能力を十分に発揮する上で、マルチステークホルダーが重要な役割を果たすという本原則の認識は、極めて重要な洞察です。「Society 5.0」のビジョンを推進するには、究極的には、社会全体を巻き込むアプローチが求められます。そのために、BSA は「AI-Ready な社会」の推進に向けた社会理念の策定を強く支持するとともに、本ビジョンの進展に役立つよう提言をします。まず、以下に提言の要約を列挙し、その後、本原則の具体的箇所について詳細を記載します。<br><br>1.　AI 原則は、特定の利用事例に関連するその分野特有のニュアンス及び AI の開発、展開、利用に関与する様々なステークホルダーを考慮した柔軟なものであるべきです。 |

[2] https://ai.bsa.org/

| | |
|---|---|
| | 2. 健全な AI 政策の策定には、マルチステークホルダーの関与のためのプロセスと仕組みが不可欠です。<br>3. AI 原則は、強固で競争力ある AI エコシステムを創出するために、市場アクセスと技術中立性を推進すべきです。<br>4. AI 原則は、国内及び越境データの自由な移転を明示的に保証すべきです。<br>5. AI 原則は、プライバシーに関してリスクベースのアプローチを採用し、過度に規範的になることを避け、現実の損害の防止を重視すべきです。<br>6. AI 原則は、セキュリティを推進し、AI ツールにおける信頼の基本的構成要素として安全性を組み込むべきです。<br>7. AI 原則は、AI システムにおいて、公平性、説明責任、透明性を達成する全ての場合に通用するアプローチは存在しない、ということを認識すべきです。<br>8. AI 原則は、機密でない政府データをより利用可能とし、データ分析に対する障壁を取り除くことにより、イノベーションを推進すべきです。 |
| <u>２．基本理念</u><br><br>「人間の尊厳が尊重される社会」「多様な背景を持つ人々が多様な幸せを追求できる社会」「持続性ある社会」について（3 頁、第 2 段落以降） | BSA は、AI システム及びアプリケーションが、人間の尊厳が尊重される社会、多様な背景を持つ人々が多様な幸せを追求できる社会及び持続性ある社会を推進する形で開発、展開、使用されるべきことについて賛同します。AI の適用は、多くの場合、社会において十分なサービスを受けられていない人々の状況の改善に大きく貢献することから、BSA は、AI の利益が幅広く享受されることが重要であるとの方向性を支持します。<br><br>**提言**<br>・ 本原則に、消費者向けに使用する場合と企業向けソリューションを開発する際 |

| | |
|---|---|
| | に使用する場合との違いや machine to machine 通信によるデータ使用の場合の違いを考慮すること等、その分野特有のニュアンスを考慮した柔軟なものであるべきとの認識を含めることを提言します。 |
| <u>３．Society5.0 実現に必要な社会変革「AI-Ready な社会」</u><br><br>「技術の進展との相互作用に留意しながら考える必要」性（4 頁、第 1 段落）及び「社会システム」（4 頁、(2)「社会システム」）について | 本原則は、医療、金融、エネルギー等様々な業界を含む社会システム全体に適用されつつ、それぞれの業界が、AI 技術の進化及び業界における独自の効果とともに進化して対応できるよう柔軟であるべきとの認識を示しており、BSA はこれ支持します。<br><br>AI に関する原則と政策は、AI エコシステムを構成する多様な基礎技術と利用事例を考慮に入れたものであるべきです。魅力的で有用な新しい AI アプリケーションが絶えず開発されており、潜在的な利用事例の全てを予測することは困難です。従って、策定時に決められた原則及び考慮点が将来も変更されずに残存するのではなく、技術及びその適用例の発展に応じて進化することが重要です。従って、AI については、拘束力ある規制を策定するアプローチよりも、原則に基づき規律するアプローチを採用すべきです。<br><br>**提言**<br>· 本原則に、AI の開発、展開、及び使用に関与するマルチステークホルダー及び多様な分野に及ぶ AI 適用例により示される分野特有のニュアンス（例えば、消費者向けに使用する場合と企業向けソリューションを開発する際に使用する場合の違いや machine to machine 通信のデータ使用する場合の違い等）を考慮に入れて十分に柔軟なものであるべき、との明確な認識を含めることを提言します。 |
| イノベーションに貢献し国際的に開かれた「産業構造」構築の必要性<br>（5 頁、(3)「産業構造」） | BSA は、産業構造が「柔軟で国際的に開かれたものになっていることが必要である」という認識を支持します。 |

| | |
|---|---|
| | **提言**<br>・ 雇用創出、競争力、イノベーションを推進するデジタル経済及び貿易政策の重要性についての明確な認識の下、**市場参入障壁がなく、革新的なAIアプリケーションやサービスが差別的取扱いを受けない**という一般的原則を本原則に含めるべきです。<br><br>・ また、国際的なデータの移転がAIの発展に重要であることに鑑み、政策は**自由な越境データ移転を支持したものである**べきです。さらに、自由なデータ移転は、中小企業を含むあらゆる規模の企業が、最先端のAIサービス（その多くは、途切れない国際的なデータ移転を前提とし ています。）にアクセスすることを可能にします。 |
| AI ガバナンスに関わる「多様なステークホルダー」による体制及び「国際協力体制」の構築について（5 頁、(5)「ガバナンス」第 2 段落） | BSA は、本原則がマルチステークホルダーとの協働を重要視していることを歓迎します。AI システムは多様な方法で使用されるため、ベストプラクティス及び提言は、関係する全てのステークホルダーとの協力及び必要な技術的考慮要素に基づいて策定されることが重要です。また、ベストプラクティス及び提言には、国際的な視点を盛り込むべきです。<br><br>**提言**<br>・ 本原則に、AI政策や規則の策定過程におけるマルチステークホルダーの関与を確保するための具体的なプロセスや仕組みを政府が整備することの重要性を含めることを提言します。これには、例えば、技術の進展に伴い、関連する原則の策定及び見直しについて助言を行う諮問委員会や官民連携の仕組みを取り入れることの奨励が挙げられます。<br><br>・ また、本原則は、国内における取組みを国際的に認められた標準及び原則と整 |

| | |
|---|---|
| | 合させることの重要性に関して言及すべきです。 |
| **4　人間中心の AI 社会原則**<br>**「プライバシー」について**<br>**（7 頁、（3）プライバシー確保の原則）** | BSA は、個人の権利を保護し、AI システム及びアプリケーションへの信頼性を高めるプライバシー慣行の実践を支持します。プライバシーに関するベストプラクティスには、パーソナルデータ収集及び利用の透明性を高めること、収集及び利用に関するガバナンスを提供することで情報に基づいた選択を可能にし、また、かかる選択を尊重すること、消費者にパーソナルデータのコントロールを提供すること、強固なセキュリティを提供すること並びに正当な事業目的のためのデータ利用を促進することが含まれます。<br><br>プライバシーに関するベストプラクティスの枠組みは、リスクベース、かつ、原則を推進するものであるべきで、過度に規範的な要件を避けるべきです。柔軟性を枠組みに組み込むことで、企業が革新的製品やサービスを提供する能力を妨げずに、個人が個人情報を適切にコントロールすることを可能にします。<br><br>**提言**<br>・「(3)プライバシー確保の原則」箇条書き第3項目において、パーソナルデータは「その重要性・要配慮性に応じて適切な保護がなされなければならない」ことを正しく認識しています。もっとも、他の部分の記載について、この重要な原則との整合性をより高めるべきです。<br><br>例えば、「(3)プライバシー確保の原則」最初の段落において、全てのAIが、パーソナルデータ利用に関するリスクを高めるわけではないことを明確に示すよう提案します。そのうえで、AIにつ |

いて漠然と「単なる個人情報を扱う以上の慎重さ」が必要であることを示唆するよりも、AIが「個人の権利や利益に影響を及ぼす」可能性がある状況においてはより厳格な精査が正当化されることを 明確にする方が有益であると考えます。また、本原則は、データ自体の性質、あるいは（AIを含む）データを管理する特定の技術やシステムに焦点を当てるの ではなく、個人に害を及ぼすリスクに焦点を当てるべきです。そこで、以下を本原則に追記することを提案します。

「AI の使用が個人に害を及ぼすリスクを高める可能性がある場合には、そのような状況に対処するための仕組みや枠組みを整備すべきである。」

- 「(3)プライバシー確保の原則」箇条書き第2項目では、全てのAIシステムは、正確性の確保とともに、本人が「実質的に」自身のデータの管理に「関与する」ことを可能にする仕組みを含むべきであることを示唆しています。BSAは、箇条書き第2項目を概ね支持しますが、特定のAIシステムが、その処理結果として与え得る損害の潜在的なリスクに対処 するための救済の仕組み（及びその種類）が必要かどうかを判断するという、リスク情報に基づくアプローチを許容するように本原則を修正することを提案します。**従って、箇条書き第2項目については以下の修正を提言します。**

「パーソナルデータを利用する AI は、~~当該データのプライバシーにかかわる部分については~~ **その処理結果として公衆に与え得る損害に対処する仕組みを実装すべきである。当該仕組みには、**正確性・正当性 ~~の~~ **を確保する手段及び適切な場合にはパーソナルデータに対するコントロールを本人が** ~~実質的な関~~

| | |
|---|---|
| | ~~与ができる仕組みを持つべきである~~ **行うことができる手段を含む場合がある。** |
| セキュリティについて<br>（8 頁、（4）セキュリティ確保の原則） | BSA はサイバーセキュリティと強靭性を強く提唱しています。BSA は、AI などのデジタル技術がグローバルにつながる経済を可能にするにつれ、社会がセキュリティリスクの増大に注意を払わなければならないことを認識しています。BSA は、インターナショナル・サイバーセキュリティ・ポリシーフレームワーク[3]に概説される拡張されたセキュリティ措置を強化する政策を提唱しています。BSA は、本原則におけるセキュリティに関する記載について総論的に支持しますが、以下の改善を提案します。<br><br>・ 第 1 段落に記載される、AI システムが「希少事象や意図的な攻撃」に対応することができないことについて、本原則の将来に渡る有効性を確保する上で、AI の特定の問題に対する対応可能性について絶対的な立場をとることを避ける方が賢明と考えます。当該箇所については、以下のように記載する方がより正確であると考えます。<br><br>「~~少なくとも現在想定できる技術の範囲では、~~希少事象や意図的な攻撃に対して AI が ~~常に~~ 適切に対応することは~~不可能で~~ **できない場合もあり**…」<br><br>・ BSA は、AI の開発と展開における最高水準に全面的にコミットしています。信頼は、安全性の実践を通じてのみ獲得することができます。また、「AI ツールにおける信頼の基本的構成要素としての安全性」の概念を含めることを推奨します。 |

[3] https://bsacybersecurity.bsa.org/wp-content/uploads/2018/04/BSA_cybersecurity-policy.pdf

| | |
|---|---|
| 公正競争について<br>（8 頁、（5）公正競争確保の原則） | BSA は、貿易相手国間及び民間企業間の公正競争を促進する政策を支援します。AI は、急速に進化する市場ダイナミクスを有する成長分野であり続けます。従って、国内の競争政策は、AI 特有の規則の策定を避け、技術中立性を維持すべきです。AI サービスの国際競争は、デジタル貿易に対する障壁を取り除き、外国市場の開放を確保することにより促進されるべきです。これには、越境データ移転が許されていること並びにデータローカライゼーションの義務付け及び国内の技術や製造者を優遇するその他の保護主義的措置の禁止を確実に行うことが含まれます。<br><br>**提言**<br>BSA は特に以下を加えるよう提言します。<br>· 国内及び越境データの自由な移転を確保することによる競争の促進。現代の経済における越境データの重要性に鑑み、政府はプライバシー又はセキュリティ政策を真に必要な場合のみに用い、それらを市場参入障壁として用いてはならない。 |
| 公平性、説明責任及び透明性について<br>（8 頁、（6）公平性、説明責任及び透明性の原則） | BSA は、公平性、説明責任及び透明性（FAT）の原則が信頼される AI の開発に不可欠であることに同意します。BSA メンバー企業は、AI プロセス及び他のプロセスにおけるバイアスの影響を低減し、製品への信頼と信用を確保するために、方針及び技術的解決策の継続的な策定及び見直しを行っています。BSA 会員企業は、AI システムが基本的な権利及び規範を尊重するよう確保することについて全面的にコミットしています。<br><br>それぞれの AI システムに FAT を組み込むための最適な仕組みは、様々な要因により異なるため、この問題についてのガイダンスは、様々な利用事例や実装方法に適応できるよう十分に柔軟なものでなければなりません。例えば、ユーザーの評価に基づきレストランを薦める AI システムにおいて一般市民からの信頼を得るために必要な透明 |

| | |
|---|---|
| | 性のレベルは、複雑な透明性または説明が要求されるレベルのものではないでしょう。これに対して、AI システムが、借入れや入居審査など、処理結果がその分野における消費者の適格性に影響を及ぼす状況で使用される場合、一般市民は、システムによる判断が、確実に公平性、説明責任及び透明性についての高い水準と合致するようシステム開発者が採用した措置の有効性について、はるかに大きな期待を有するものであり、これは当然のことです。<br><br>　また、ある倫理原則を守ることが他の倫理原則を犠牲にする状況があり得ることを認識することも重要です。AI システムの公平性及び説明責任を保つことを目的にした設計上の選択が、基礎となるモデルを公衆に透明性をもって説明可能にすることについて代償を払うといったトレードオフを伴うこともあり得ます。例えば、不正検出システムを設計する際、それが正確かつバイアスのない方法で動作していることを確保するためには、その動作方法の公衆への開示を制限しなければならないかもしれません。実際、過度な透明性が、アルゴリズム操作の脆弱性を高めるという意図しない結果を招く場合があります。また、アルゴリズム、ソースコード又は関連データセットを開示しても、断片的で、それだけでは意味を理解できないため、説明の提供として役に立たないことが研究結果により示されています。<br><br>**提言**<br>・　箇条書き第 2 項目の「状況に応じた適切な説明が得られなければならない。」との記載は過度に規範的であり、本原則の他の部分が原則ベースで記載されていることと異なっています。従って、この部分を削除し、**「いかなるレベルの説明が適切であるかの評価を実施すべきであり**、当該説明には AI が使用されている事実を含める、ということも考えられ |

| | |
|---|---|
| | る。」と変更すべきと考えます。 |
| イノベーションについて<br>（9 頁、（7）イノベーションの原則） | データは AI の発展に重要です。従って、データ政策への健全なアプローチは、AI を革新し適用する社会の能力と本質的に結びついています。BSA は、AI イノベーションを促進する健全なデータ政策の必要性について取りまとめ、重要な事項の提言をしています[4]。特に、政府が有するデータセットは、AI モデルの学習を改善し、より包摂的な社会を構築し、経済成長の触媒として機能を果たすために有用な資源です。またそれと同様に、AI システムの学習を目的として、合法的に入手できるデータを研究者が利用することを妨げ得る不必要な障壁を取り除くことが重要です。<br><br>**提言**<br>・ 本原則は、AI によって利用されるデータの効率的な収集及び維持に関連する問題並びに**非機密データを機械可読形式で公衆に自由に利用可能にすることの**政府に対する奨励を盛り込むべきです。<br><br>・ 本原則は、AI 研究者が、AI アルゴリズムの学習、データ分析及びデジタル分析の実施を目的として、確実にコンテンツ及びデータを合法的に入手し使用できることの重要性を明確に記載すべきです。 |
| 結語 | BSA は本原則に関する意見を提出する機会を与えられたことについて感謝申し上げます。本意見についてご検討いただけますと幸いです。本意見について、ご質問、ご不明な点がございましたらご連絡下さい。 |

4 https://ai.bsa.org/wp-content/uploads/2018/05/BSA_2018_AI_DataPolicy.pdf

u　**全国銀行協会（20190214）**

［文書番号削除］
平成 31 年２月 14 日

内閣府　政策統括官(科学技術・イノベーション担当)
　社会システム基盤担当　御中

一般社団法人全国銀行協会

「『人間中心の AI 社会原則』（案)」に対する意見について

今般、標記原則（案）（平成 30 年 12 月 27 日公表）に対する意見を別紙のとおり取りまとめましたので、何卒ご高配を賜りますようお願い申しあげます。

以　上

（別紙）

「人間中心のＡＩ社会原則」（案）に対する意見

| No. | 該当箇所 | 意見等 | 理由等 |
|---|---|---|---|
| 1 | 1 はじめに | 脚注2について、「高度に複雑な情報システム一般」は定義として広すぎるため、より具体的な定義としてはいかがか。 | 左記のとおり。 |
| 2 | 4.1 AI 社会原則(3)プライバシー確保の原則 | 2つ目の矢羽根に記載された、「正確性･正当性の確保及び本人が実質的な関与ができる仕組み」とは具体的にはどのようなものか。個人情報保護法に則ればよいのか。 | 趣旨確認のため。 |
| 3 | 4.1 AI 社会原則(3)プライバシー確保の原則 | 「単なる個人情報を扱う以上の慎重さ」とは具体的にはどのようなものか。 | 左記のとおり。 |
| 4 | 4.1 AI 社会原則(6)公平性、説明責任及び透明性の原則 | 「AIの動作結果の適切性」についての「結果の適切性」とは何か。また、「適切な説明」とは具体的にどのような説明を想定しているか。 | 趣旨確認のため。 |
| 5 | 4.1 AI 社会原則(6)公平性、説明責任及び透明性の原則 | AIの学習を経た結果として、人間のある特定のバックグラウンドが有意に影響を与えるように見えるアルゴリズムとなった場合はどのように説明すればよいか。 | 趣旨確認のため。 |

以　上

**Opinion Letters**

Feb 14 2019

Dear Council for Science, Technology and Innovation, Cabinet Office

［署名削除］

I will be sending you the opinion letters on the draft of Social Principles of Human-centric AI as below.

Subject : [Social Principles of Human-centric AI]

| Identify the paragraph(Page no.) | Your opinion |
|---|---|
| Overall vision | First, we would like to commend the Japanese government for developing this important principles document. In particular, Japan's vision of human-centered AI, where it is a tool to enable humans' to live more convenient and flourishing lives, is commendable. We appreciate the structure of the document, which includes guidance both for developers and utilizers of AI technologies as well as broader principles for how society can become AI-ready. |
| Section (4) Innovation systems (p.5)- "Therefore, all kinds of data including real space data, should be instantaneously and securely available at the AI analyzable level. It is required that everyone can provide data at ease and can obtain the benefits from the data that they provided." | While we support the objective of increasing access to data both for development and analysis of AI systems, to avoid misunderstanding, we would recommend clarifying that this data provision should be in accordance with privacy principles and trade secrets. |

| | |
|---|---|
| Section (5) Governance (p.5)- “The diverse stakeholders should be cautious not to exclude the social minorities when they build the systems to solve technological/social problems. The governance can be achieved not only by law but also by technological architectures. Industries should play the independent roles to implement those architectures flexibly and effectively. In addition to the domestic governance, we should implement the international collaboration systems to cope with the problems that cross national borders. | We applaud both the commitment to diversity as well as recognition that new modes of multi-stakeholder governance involving both legal, technological and other solutions, with international cooperation, will be needed to address the nuanced, diverse and rapidly evolving nature of AI technologies. |
| Section 4.1(2) Education (p. 7)- “Literacy education provides the following contents: 1) Data used by AI are usually contaminated by bias, 2) AI is easy to generate unwanted bias in its use, and 3) The issues of impartiality, fairness, and privacy protection which are inherent to actual use of AI.” | While the issues mentioned here are important (bias, fairness and privacy), these may not represent an exhaustive list of AI literacy education. For example, education on the scope and limitations of AI technologies beyond bias (e.g. where do we expect AI to perform more or less accurately for particular tasks) and methods of ensuring security and preventing overreliance on automated decision inputs may also be useful. Publicly available tools for exploring the performance of models could be useful to explore. |
| Section 4.1(3) Privacy (p.7)- “AI that uses personal data should have a mechanism that ensures accuracy and legitimacy…” | While we agree with the spirit of this principle, it is important to note that in many AI systems, accuracy cannot be 100% ensured. Our expectations about accuracy might vary depending on the potential impact on a data subject’s well-being: for example, we may tolerate a lower degree of accuracy for a book recommendation system based on data about someone’s favorite novels (where the harm of inaccuracy is low) compared to a system determining a data subject’s likelihood of defaulting on a loan. |

| | |
|---|---|
| Section 4.1(6) (p.8)- "Appropriate explanations should be provided such as the fact that AI is being used, the method of obtaining and using the data used in AI… | As Japan's paper rightly notes, AI is not a precisely defined term but rather refers to a diverse range of practical engineering techniques. Thus, disclosing that AI is being used may not be a useful explanation in every context. For example, while users may want to know that they are talking to a bot (generated using AI), they may have less interest in knowing that the automatic adjustments on their cameraphones were made using AI rather than some other technology. |
| and the mechanism to ensure the appropriateness of the operation results of AI according to the situation AI is used." | While we agree that these mechanisms are important, they may not be easily digestible to all users of AI systems so disclosing them may not provide a helpful user experience in all cases. |
| Section 4.1(7) Innovation (p.9)- To ensure the sound development of AI technology, it is necessary to establish an accessible platform in which data from all fields can be mutually utilized across borders with no monopolies, while ensuring privacy and security. In addition, research and development environments should be created in which computer resources and highspeed networks are shared and utilized, to promote international collaboration and accelerate AI research. | While we agree that promoting the accessibility and free flow of data within and across borders is an admirable goal, it would be useful to have more clarity on what is envisaged by the data platform referenced here. We also recommend adding a caveat "while ensuring privacy, security *and protection of trade secrets and proprietary information*". |