

別添 1
(報告書 (案) 本体)

報告書2018 (案)

—AIの利活用の促進及びAIネットワーク化の健全な進展に向けて—

平成 30 年●月●日

AI ネットワーク社会推進会議

目次

要旨	I
はじめに	1
序章	3
1. 「AIネットワーク化」と「智連社会」	
2. 本推進会議における取組	
第1章 AIネットワーク化をめぐる最近の動向	11
1. 国内の動向	
2. 海外の動向	
3. 国際的な議論の動向	
第2章 AIネットワーク化が社会・経済にもたらす影響	24
1. AIネットワーク化に関するシナリオ分析（分野別評価）	
2. AIネットワーク化の進展に伴い形成されるエコシステムの展望	
3. AIネットワーク化の進展に関する考察	
4. 教育・人材育成	
5. AIネットワーク化の進展に関する指標	
第3章 AIの利活用において留意することが期待される事項	45
1. 背景・経緯	
2. 関係する主体の整理	
3. AI利活用原則案	
4. AI利活用原則案の各原則の内容に関する論点	
5. 今後の検討	
第4章 今後の課題	64
結びに代えて	79

- <別紙 1> AI ネットワーク化が社会・経済にもたらす影響（分野別評価）
- <別紙 2> AI ネットワーク化の進展に伴い形成されるエコシステムの展望
- <別紙 3> AI ネットワーク化の進展において想定される課題
（ネットワーク化の観点から）
- <別紙 4> 教育・人材育成に関する主な意見
- <別紙 5> 報告書 2018（案）に関する意見募集に寄せられた主な意見に対する考
え方
- <別紙 6> AI ネットワーク社会推進会議及び分科会構成員一覧
- <別紙 7> 開催経緯

要 旨

AIに関する技術開発と利活用が急速に進展する中、今後、AIシステムがインターネット等を通じて他のAIシステム等と接続し連携する「AIネットワーク化」が進展していくことにより、社会的な課題の解決など人間や社会・経済に多大な便益がもたらされることが期待される。他方、AIの判断のブラックボックス化や制御喪失などのリスクが懸念されるとともに、人々のAIに対する不安などが、AIの開発及び利活用の促進やAIネットワーク化の健全な進展の阻害要因となるのではないかと懸念されている。

これらのことを踏まえ、AIの便益を増進させ、リスクを抑制するとともに、人々のAIに対する不安を取り除き、信頼を醸成するため、AIの開発者、利用者やデータ提供者が留意することが期待される事項等を整理することが求められる。

総務省では、今後のAIネットワーク化の進展を見据え、平成28年10月から「AIネットワーク社会推進会議」を開催し、本推進会議は、平成29年7月に、AIの開発者が留意することが期待される事項を整理した「国際的な議論のためAI開発ガイドライン案」を含む『報告書2017』を取りまとめた。また、『報告書2017』の公表後、AIの利用者やデータ提供者が留意することが期待される事項に関する検討を行うとともに、その検討に資するようAIネットワーク化が進展した社会の将来像を描きAIの利活用における課題等を抽出するために、AIネットワーク化の進展に伴い形成されるエコシステムの展望に関する検討等を行い、平成30年7月に、本報告書を取りまとめた。

序章では、本推進会議における検討の基調をなすAIネットワーク化の展望を示した上で、目指すべき社会像として掲げる人間中心の「智連社会」(Wisdom Network Society : WINS [ウインズ])及びその基本理念を紹介している。

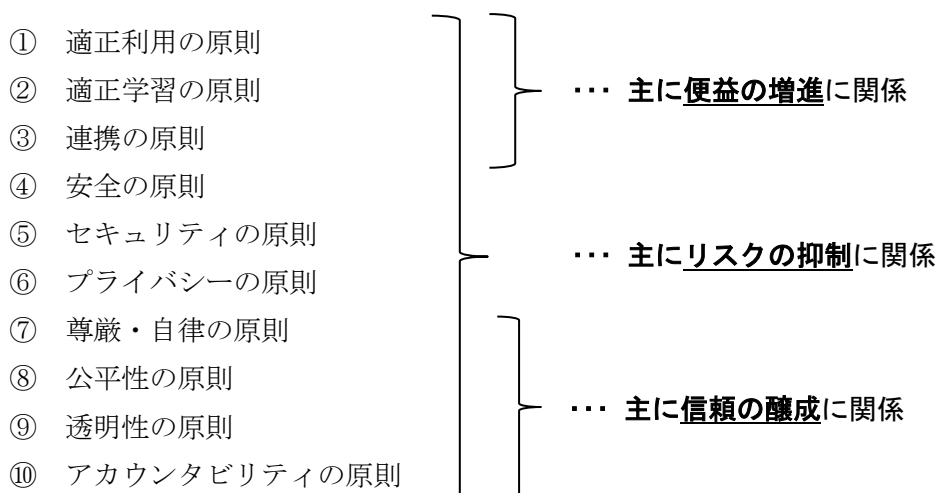
第1章では、『報告書2017』の公表以降の最近のAIネットワーク化に関する国内外の動向及び国際的な議論の動向を概観している。

第2章では、社会・経済の様々な分野におけるAIシステムの利活用の場面を想定したシナリオを作成し、AIネットワーク化が社会・経済の各分野にもたらすインパクト及びリスクの評価(分野別評価)を行い、AIネットワーク化の進展に伴い形成されるエコシステムの展望を示すとともに、ネットワーク化の観点からの考察を行い、AIの開発及び利活用の促進並びにAIネットワーク化の健全な進展等に関する課題を整理している。また、教育・人材育成やAIネットワーク化の進展に関する指標についての検討状況を示している。

分野別評価においては、「行政」、「移動」、「居住」の3分野を対象として評価を行った。エコシステムの展望においては、「都市部における利用」、「地方部における利用」、「家庭内における利用」、「企業における利用」の4つの類型ごとに想定した利用シーンにおいて展望したAIの利活用等をもとに、「完全自動運転」、「医療診断」、「人事評価・採用」、「金融(保

険)、「行政」の5つの事例に関するケーススタディを行い、AIの利活用の便益及び課題を整理した。ネットワーク化の観点からの考察においては、ネットワーク化に着目した課題(AIネットワーク化の健全な進展を阻害し得る要因)を整理した。

第3章では、これらの課題等を踏まえ、AIの利活用において留意することが期待される事項について、検討の背景・経緯を概観した上で、AIの利活用において関係する主体を整理するとともに、「AI利活用原則案」を取りまとめ、各原則の内容に関する論点を整理している。「AI利活用原則案」の項目は、次のとおりである。



第4章では、第2章及び第3章における検討の結果等を踏まえ、今後の課題を次のように整理している。

- AIの開発及び利活用の促進並びにAIネットワーク化の健全な進展に関する事項
 - AI開発ガイドライン(仮称)の策定、AIの利活用に関する指針の策定 等
- AIネットワーク上を流通する情報・データに関する事項
 - セキュリティ対策、プライバシー及びパーソナルデータの保護 等
- AIネットワーク化が社会・経済にもたらす影響に関する事項
 - AIネットワーク化が社会・経済にもたらす影響に関するシナリオ分析 等
- AIネットワーク化が進展する社会における人間をめぐる課題に関する事項
 - 人間とAIとの関係の在り方に関する検討、ステークホルダ間の関係の在り方に関する検討 等

本報告書が、AIネットワーク化に関する社会的・経済的・倫理的・法的課題をめぐる国内外の検討に寄与し、AIの開発及び利活用の促進並びにAIネットワーク化の健全な進展が図られ、人間中心の「智連社会」の実現に向けた道標となることを願いたい。

はじめに

最近では、人工知能（Artificial Intelligence：A I）について、毎日のように新聞などで報道がなされ、テレビや雑誌でもA Iに関する特集が数多く企画されるなど、A Iに関する人々の関心は急速に高まっている。また、A Iを活用した消費者向けの製品やサービスの普及が見られ、人々の生活にとってA Iが身近なものとして浸透しつつある。A Iは、今後、個人、地域社会、各国、国際社会の抱える様々な課題¹の解決のため、その利活用が期待されており、特にA Iシステムがインターネット等を通じて他のA Iシステム等と接続し連携する「A Iネットワーク化」が進展することにより、人間及びその社会や経済に多大な便益が広範にもたらされることが期待される。我が国においても、人口減少・少子化を迎える中、新しい社会・経済を構築していく上で、A Iの利活用は非常に有効であるものと考えられている。

他方、A Iの判断のブラックボックス化や制御喪失などのリスクとともに、人々のA Iに対する不安などが、A Iの開発及び利活用の促進やA Iネットワーク化の健全な進展の阻害要因となるのではないかと懸念されている。

A Iの便益を人々や社会が享受するためには、A Iの開発をさらに推し進めるとともに、利活用の裾野をより一層広げていくことが必要である。そのためには、まずもって人々のA Iに対する不安を取り除き、信頼を醸成する取組を進めることが必要である。A Iの開発者や利用者が、それぞれA Iの開発や利活用において、A Iの便益を増進させリスクを抑制する観点から期待される取組を行うとともに、開発者や利用者に加えて、政府や業界団体等を含めた関係するステークホルダがA Iの開発及び利活用の促進やA Iネットワーク化の健全な進展に向けた環境整備のための取組を行うことにより、A Iに対する社会的な受容性の向上がもたらされるものと考えられる。

こうした中、平成28年4月に日本で開催されたG7（先進7か国）香川・高松情報通信大臣会合において、ホスト国である日本は、A I開発原則のたたき台を紹介し、各国関係閣僚による議論が行われた。その結果、G7において「A I開発原則」及びその内容の解説からなる「A I開発ガイドライン」の策定に向け、引き続きG7各国が中心となり、OECD（経済協力開発機構）等国際機関の協力も得て議論していくことで合意した。その後、欧米各国においても、A Iに関する社会的・経済的・倫理的・法的課題をめぐる検討が本格化し、国際的な議論が加速している。

以上のような背景と問題意識に鑑み、総務省情報通信政策研究所では、平成28年10月から「A Iネットワーク社会推進会議」を開催し、「A Iネットワーク社会推進フォーラム」（国際シンポジウム）²における議論等を踏まえ、平成29年7月に、A Iの開発者が留意す

¹ 国際社会の抱える課題については、国連の「持続可能な開発目標」（SDGs）（http://www.un.org/ga/search/view_doc.asp?symbol=A/70/L.1）などを参照。

² 総務省が平成29年3月に東京大学において開催した国際シンポジウム。日米欧の産学民官のステークホルダが参加し、A Iネットワーク化のガバナンスの在り方について活発な議論が行

ることが期待される事項を整理した「国際的な議論のためA I 開発ガイドライン案」を含む『報告書 2017』³を取りまとめた。

報告書 2017 の公表後、本推進会議は、A I の利用者やデータ提供者が留意することが期待される事項に関する検討を行うとともに、その検討に資するようA I ネットワーク化が進展した社会の将来像を描きA I の利活用における課題等を抽出するために、A I ネットワーク化の進展に伴い形成されるエコシステムの展望に関する検討等を行ってきた。

本推進会議は、報告書 2017 の公表後、平成 29 年 10 月から平成 30 年 5 月までに開催した 8 回の会合（本推進会議の会合（親会）3 回のほか、親会の下に置かれた環境整備分科会及び影響評価分科会⁴の合同分科会の会合 5 回）における検討や平成 30 年 1 月から同年 4 月までに開催した事業者等からのヒアリング等⁵を踏まえて、本報告書（案）を作成した。その後、平成 30 年 6 月 12 日から同月 26 日までにかけて、本報告書（案）に関する意見募集を行った上で、環境整備分科会・影響評価分科会 合同分科会（第 6 回）及び本推進会議（第 10 回）の会合での議論を経て、本報告書を最終的に取りまとめた。

本報告書は、本推進会議における検討と関連する国内外の動向及び国際的な議論の動向を概観し、社会・経済の様々な分野におけるA I システムの利活用の場面を想定してA I ネットワーク化が社会・経済の各分野にもたらす影響の評価を行い、A I ネットワーク化の進展に伴い形成されるエコシステムの展望を示すとともに、ネットワーク化の観点からの考察を行い、それらの検討において抽出された課題等を踏まえて、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、各原則の内容に関する論点を整理した上で、今後の課題を整理している。

本報告書が、A I ネットワーク化に関する社会的・経済的・倫理的・法的課題をめぐる国内外の検討に寄与し、A I の開発及び利活用の促進やA I ネットワーク化の健全な進展が図られ、人間中心の「智連社会」⁶の実現に向けた道標となることを願いたい。

われた。

³ A I ネットワーク社会推進会議「報告書 2017 - A I ネットワーク化に関する国際的な議論の推進に向けて-」（平成 29 年 7 月 28 日。以下「報告書 2017」という。）

⁴ 本推進会議の発足当初は、親会の下に開発原則分科会及び影響評価分科会が置かれていたが、平成 29 年 10 月に分科会を改組し、現在、環境整備分科会及び影響評価分科会が置かれている。

⁵ 本推進会議において、平成 30 年 1 月から同年 4 月にかけて、特にA I の利活用の取組が進んでいる又は期待されると考えられる「自動運転」、「医療」、「人事評価・採用」、「金融」及び「製造過程・バックオフィス業務の自動化」の 5 分野について、主としてA I の利活用の取組を進めている事業者等（行政機関、業界団体を含む 17 者）からヒアリングを行った。なお、ヒアリング内容等を公表すると、各者のA I に関する取組内容や事業戦略等が明らかになり、事業活動に支障を来すおそれがあるため、事業者名を含めて、内容は公表していない。

⁶ 智連社会（Wisdom Network Society）とは、A I ネットワーク化の進展の結果として、人間が主体的に技術を使いこなすことによりA I ネットワークと共生し、データ・情報・知識を自由かつ安全に創造・流通・連結して「智のネットワーク」（Wisdom Network）を形成することにより、あらゆる分野におけるヒト・モノ・コト相互間の空間を越えた協調が進展し、もって創造的かつ活力ある発展が可能となる社会である。詳細は、序章 1. 参照。

序章

本推進会議の前身たる「A I ネットワーク化検討会議」（以下「検討会議」という。）においては、目指すべき社会像として人間中心の社会像「智連社会」（Wisdom Network Society：WINS [ウインズ]）を提唱して、その基本理念を呈示するとともに、A I ネットワーク化の進展が「智連社会」の形成にどのような影響を与えるのかということやその際の人間とA I ネットワークとの関係に関する評価及びA I ネットワーク化が社会・経済にもたらす影響（インパクト⁷及びリスク）の基礎的な評価を行った上で、今後の課題を整理した（平成28年4月に『中間報告書』⁸、同年6月に『報告書2016』⁹を公表。）。

これらの成果を踏まえ、本推進会議においては、A I の開発において留意することが期待される事項を整理した「国際的な議論のためのA I 開発ガイドライン案」（以下「A I 開発ガイドライン案」という。）を取りまとめるとともに、A I システムの具体的な利活用の場面を想定してA I ネットワーク化が社会・経済にもたらす影響の評価等を行った上で、今後の課題を整理した（平成29年7月に『報告書2017』¹⁰を公表。）。

本章においては、本推進会議における検討の基調をなすA I ネットワーク化の展望、目指すべき社会像としての「智連社会」及びその基本理念並びにA I ネットワーク化の進展が「智連社会」に及ぼす影響を提示するとともに、本推進会議の取組の概要を紹介する。

1. 「A I ネットワーク化」と「智連社会」¹¹

○ A I ネットワーク化

A I システムが今後インターネットその他の情報通信ネットワークと接続されていくことにより、その影響（インパクト及びリスク）が大きくなるとともに、空間を越えて即座に波及し得るようになるものと考えられる。そこで、検討会議及び本推進会議においては、A I システムの普及及びその利活用の進展に伴う問題を検討するに当たり、これを単に個々のA I システムをめぐる問題として捉えるのではなく、「A I ネットワーク化」¹²（A I システムがインターネットその他の情報通信ネットワークと接続され、

⁷ 主に良い影響や便益のことを意味している。本報告書において以下同じ。

⁸ A I ネットワーク化検討会議「中間報告書 A I ネットワーク化が拓く智連社会（WINS [ウインズ]）－第四次産業革命を超えた社会に向けて－」（平成28年4月15日。以下「中間報告書」という。）

⁹ A I ネットワーク化検討会議「報告書2016 A I ネットワーク化の影響とリスク－智連社会（WINS [ウインズ]）の実現に向けた課題－」（平成28年6月20日。以下「報告書2016」という。）

¹⁰ 前掲注（3）参照。

¹¹ 本節は、基本的に報告書2017の記述を踏襲しているが、その後の議論を踏まえて、記述を改めているところがある。

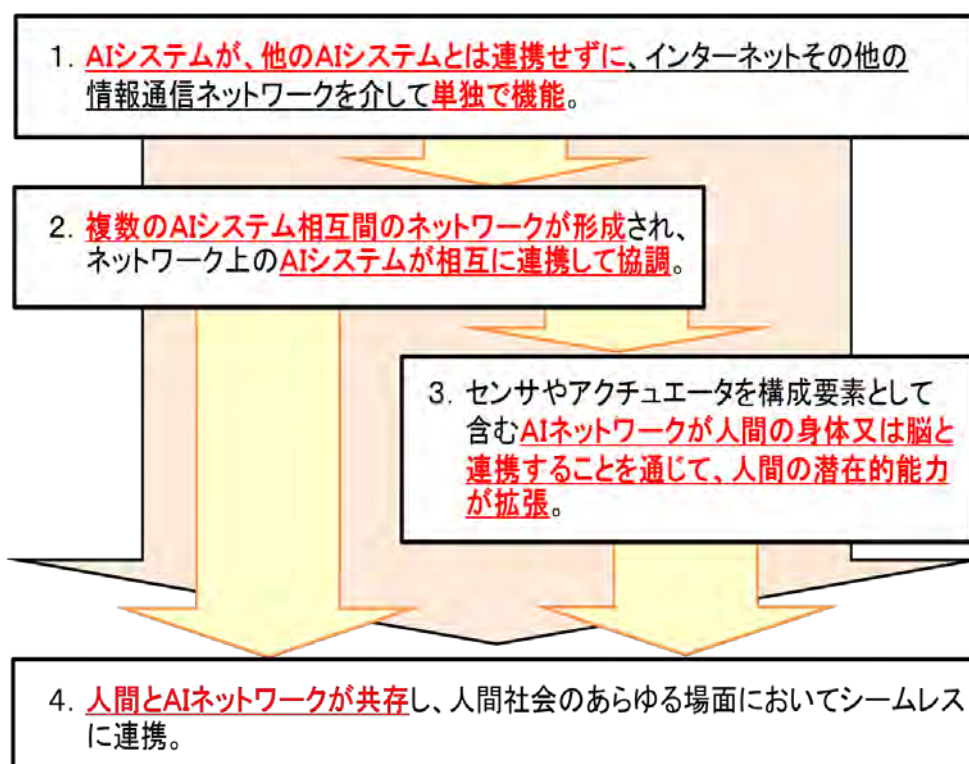
¹² この「A I ネットワーク化」は、

- ・ A I ネットワークの構築
- ・ A I ネットワークの高度化（情報通信ネットワークを介したA I システム相互間の連携

A Iシステム相互間又はA Iシステムと他の種類のシステムとの間のネットワーク（以下において「A Iネットワーク」という場合がある¹³。）が形成されるようになることをいう。以下同じ。）をめぐる問題として捉えて検討を進めてきた。

検討会議においては、A Iネットワーク化が社会・経済にもたらす影響を評価し、そのガバナンスの在り方等関連する課題を検討する際の参照軸とするため、A Iネットワーク化の進展段階を次図のように理念的に整理している。

「AIネットワーク化」の進展段階



第一段階は、単独のA Iシステムがインターネットその他の情報通信ネットワークに接続され、他のA Iシステムとは連携することなく機能するという段階である。それぞれのA Iシステムは、クラウド上に構築されたり、センサーネットワークと接続されるなどインターネットその他の情報通信ネットワークに接続されているが、他のA Iシステムとは連動していない。この段階は、既に社会の様々な場面において実現が進みつつある。

等)

からなるものとして理解できる（中間報告書5頁参照。）。

¹³ 検討会議においては、この「A Iネットワーク」のことを「A Iネットワークシステム」と呼んでいた（なお、検討会議においては、「A Iネットワークシステム」を「A Iを構成要素とする情報通信ネットワークシステム」と定義していた（中間報告書5頁参照。）。

第二段階は、インターネットその他の情報通信ネットワークを介して複数のA Iシステムのネットワークが形成され、当該ネットワーク上のA Iシステムが相互に連携して協調するという段階である¹⁴。

第三段階は、センサやアクチュエータを構成要素として含むA Iネットワークが人間の脳や身体などと連携して、人間の身体機能が拡張される段階である¹⁵。なお、A Iネットワークと人間の脳や身体などとの連携については、A Iが実装されたチップを脳や身体に埋め込むような形態だけではなく、今後の技術の発展により、A Iネットワークと人間との間で円滑なインタラクションが図られる様々な形態が想定される。この段階において、人間の能力が拡張され、さらにそれらも含めた高度に知的な能力を有するシステムの連携が発生する Internet of Ability (I o A)が具体的な姿となって見えてくるものと考えられる。

これらの進展が広く進み、A Iネットワークが社会全体に浸透していくことにより、第四段階として、人間とA Iネットワークとが共存し、人間社会のあらゆる場面において、情報通信ネットワークを介してA Iシステムと他のA Iシステム、人間とA Iシステム、さらには人間と人間とがシームレスに連携するという段階に至るものと整理されている。

○ 目指すべき社会像「智連社会」

検討会議においては、目指すべき社会像として、「智連社会」(Wisdom Network Society: W I N S [ウインズ])を掲げた。これは、人間が主体的に技術を使いこなし、他者との連携を図ることで、社会全体の幸福へとつながる社会像と捉えることができる。

A Iネットワーク化の進展に伴い、A IとA I、A Iと人間、さらには人間と人間がシームレスに連携・協調できるようになることにより、ネットワーク化される「知能」、すなわち、ネットワーク化されるA Iの知能又はこれを活用する人間の知能により「データ」・「情報」・「知識」を学習・解析して新たな「データ」・「情報」・「知識」を創造・流通・連結することが可能となる社会の到来が予測される。同時に、極めて高度なA Iネットワークに対する依存度の増大が懸念されるが、これに対して、「データ・情報・知識に基づき、知能を活用することにより、人間や社会の在り方を構想し、その実現に向けた課題を解決するための人間の能力」としての「智慧」(智) (Wisdom)の必要性が

¹⁴ A Iシステム相互間の連携・協調がもたらす便益を指摘するとともに、目的が競合し又は対立するA Iシステム相互間の交渉・協調の在り方を今後の課題として指摘するものとして、長尾顧問(京都大学名誉教授)「A Iネットワーク化の本質と将来」(A Iネットワーク社会推進フォーラム(特別講演)配付資料)2-3頁(平成29年)参照。

¹⁵ 第三段階への進展は、第一段階から第二段階への進展とは異なる性格の進展によるものであり、必ずしも第二段階の後でなければならないものではない。また、第三段階を経ずに第四段階に進展することもある。

高まる。ここで、多くの人々がA Iネットワーク社会における智慧を獲得することができれば、人間が主体的に技術を社会全体で高度に使いこなし、問題を解決していくための人間の「智慧」の連携・協調が進んでいくものと見込まれる。すなわち、人間は、A Iネットワークを利活用することにより、高度な問題解決能力としての各々の「智慧」（智）を連結し、「智のネットワーク」（Wisdom Network）を形成していくことが期待されるのである¹⁶。

「智連社会」という社会像は、このような問題意識に基づいて構想したものであり、A Iネットワーク化が進展した場合において、その力を健全に活用する姿として、人間が主体的に技術を使いこなすことによりA Iネットワークと共生し、データ・情報・知識を自由かつ安全に創造・流通・連結して「智のネットワーク」を形成することにより、あらゆる分野におけるヒト・モノ・コト相互間の空間を越えた協調が進展し、もって創造的かつ活力ある発展が可能となるという人間中心の社会像である¹⁷。

智連社会は、A Iネットワーク化が進展して「人間とA Iネットワークとが共存」する段階において目指すべき社会像として掲げたものである。この「人間とA Iネットワークとが共存」する段階とは、A Iネットワークが人間を包むように存在し、人間とA Iネットワークとがシームレスに連携し、人間にとってA Iネットワークが不可欠な存在となるという客観的な状況をいうものである。これに対し、「智連社会」という社会像は、単に「人間とA Iネットワークとが共存」するという客観的な状況を描写するにとどまるものではなく、人間が主体的に技術を使いこなすことによりA Iネットワークと共生するという人間の営みにより形成される社会像である。目指すべき社会像として「智連社会」を掲げるということは、A Iネットワークが社会の中心となるのではなく、あくまでも人間が社会の中心となり、人間がA Iネットワークを主体的に使いこなす社会を目指すべき¹⁸との考えを含意している。

¹⁶ ここで、A Iネットワークは、A Iシステム及びこれと接続される情報通信ネットワークからなる物理的存在であるのに対し、「智のネットワーク」は、人間がA Iネットワークという物理的存在の機能を活用することにより得られる能力（この能力は、人間が、データ・情報・知識に基づき、知能を活用することにより、人間や社会の在り方を構想し、その実現に向けた課題を解決するための能力であり、「智慧」（智）にほかならない。）が相互に連携し合い、相互に協調し合う様子を捉え、観念的に描写する概念である。

¹⁷ 人間中心の社会像を表すものとして、「Society 5.0」がある。「Society 5.0」とは、サイバー空間とフィジカル空間（現実社会）を高度に融合させることにより、地域、年齢、性別、言語等による格差なく、多様なニーズ、潜在的なニーズにきめ細かく対応したモノやサービスを提供することで経済的発展と社会的課題の解決を両立し、人々が快適で活力に満ちた質の高い生活を送ることのできる、人間中心の社会である（科学技術イノベーション総合戦略2016はじめに（3）〈<http://www8.cao.go.jp/cstp/sogosenryaku/2016.html>〉参照。）。

「Society 5.0」を通じて実現することが期待されている「超スマート社会」と、A Iネットワーク化を通じて目指すべき社会像として掲げる智連社会（Wisdom Network Society：WINS）との関係については、中間報告書13-14頁注（4）参照。

¹⁸ このことに関連して、A Iは社会における人間の様々な活動のエンハンスを目指すべきものであることを指摘するものとして、長尾・前掲注（14）1頁参照。

智連社会

智連社会(Wisdom Network Society【WINS】)は、人間が主体的に技術を使いこなすことによってAIネットワークと共生し、**データ・情報・知識を自由かつ安全に創造・流通・連結して「智のネットワーク」(Wisdom Network)を形成することにより、あらゆる分野におけるヒト・モノ・コト相互間の空間を越えた協調が進展し、もって創造的かつ活力ある発展が可能となる人間中心の社会像。**

人間が主体的に技術を使いこなすことによって
AIネットワークと共生し

人機共生

データ・情報・知識を自由かつ安全に創造・流通・
連結して「智のネットワーク」を形成することにより

総智連環

あらゆる分野におけるヒト・モノ・コト相互間の空間
を越えた協調が進展し

協調遍在

もって創造的かつ活力ある発展が可能となる社会

「智連社会」においては、人間がAIネットワークを主体的に使いこなし、人類や地球の将来を見据えた人間や社会の在り方を構想して、その実現に向けた課題を解決するための人間の能力としての「智慧」を連結して「智のネットワーク」を形成することにより可能となる智慧の連携と協調により実現される。このような智慧の連携と協調という「智連」の本質に鑑みると、智連社会の実現に向けたAIネットワーク化の健全な進展に関する課題や智連社会における人間が社会の在り方に関する課題の解決には、個人の努力だけではなく、社会におけるコミュニティの役割が重要になるものと考えられる¹⁹。

○ 基本理念

検討会議においては、「智連社会」を形成するに当たってのつとるべき基本理念として、次の8項目が掲げられている。

・ すべての人々による恵沢の享受

すべての人々が、AIネットワークを容易にかつ主体的に利用する機会を有し、個々の能力を創造的かつ最大限に発揮し、又は拡張することが可能となり、もってAIネットワークの恵沢をあまねく享受できること。

¹⁹ このことに関し、濱田顧問（東京大学名誉教授）「AIネットワーク化と智連社会」（AIネットワーク社会推進フォーラム（特別講演）配付資料）9頁（平成29年）参照。

- ・ **人間の尊厳と個人の自律**
個人が人間としての尊厳をもった自律的な主体としてA I ネットワークを安心して安全に利活用することにより、豊かさと幸せを感じられること。
- ・ **イノベーティブな研究開発と公正な競争**
イノベーティブな研究開発と公正な競争を通じて、多様で高度なA I ネットワークが提供されること。
- ・ **制御可能性と透明性**
A I ネットワークに関する制御可能性と透明性が技術的・制度的に確保されること。
- ・ **ステークホルダの参画**
A I ネットワークの在り方に関する意思決定に当たり、多様なステークホルダが民主的に参画できること。
- ・ **物理空間とサイバー空間の調和**
A I ネットワークを利活用して物理空間とサイバー空間を連結し、両者の調和を図ることにより、ヒト・モノ・コト相互間の空間を越えた協調の実現を可能とすること。
- ・ **空間を越えた協調による活力ある地域社会の実現**
A I ネットワークを利活用してヒト・モノ・コト相互間の空間を越えた協調が地域内・地域間で進展することにより、活力ある地域社会が実現すること。
- ・ **分散協調による地球規模の課題の解決**
人類が、A I ネットワークを基盤として構築する智のネットワークにより、地球規模の課題を国際的な分散協調により解決できること。

これらの基本理念は、A I ネットワーク化の健全な進展を通じて実現することを目指すべき社会像たる「智連社会」が備えるべき性質を様々な角度から掲げるものであり、今後のA I ネットワーク化のガバナンスの在り方を検討するに当たって立脚すべきものとして評価できよう。

2. 本推進会議における取組

検討会議においては、A I ネットワーク化の展望並びに目指すべき社会像及びその基本理念の整理と併せて、今後の課題を整理した。今後の課題の中には、A I の開発研究に当たって国際的に参照されるべき「A I 開発原則」及びその内容の解説からなる「A I 開発ガイドライン」の策定に向けた議論の推進、A I ネットワーク化が社会・経済にもたらす影響の評価等が掲げられた。その過程において、平成 28 年 4 月 29 日及び 30 日に G 7 香川・高松情報通信大臣会合が開催され、高市総務大臣（当時）から、「A I 開発原則」に関し中間報告書に掲げる 8 項目からなるたたき台を配付し、OECD 等において「A I

開発ガイドライン」の策定に向けた国際的な議論を進めるよう提案した結果、各国から賛同が得られたところである。

本推進会議は、検討会議における検討やG7香川・高松情報通信大臣会合の成果等を踏まえ、社会全体におけるAIネットワーク化をめぐる社会的・経済的・倫理的・法的課題を総合的に検討することを目的として、検討会議を発展的に改組したものである。

上記のとおり、本推進会議においては、これまで、AIの開発において留意することが期待される事項を整理したAI開発ガイドライン案を取りまとめるとともに、AIシステムの具体的な利活用の場面を想定してAIネットワーク化が社会・経済にもたらす影響の評価等を行った上で、今後の課題を整理した（報告書2017）。

これらの成果を踏まえ、本推進会議においては、引き続きAIネットワーク化が社会・経済にもたらす影響の評価を進めるとともに、AIネットワーク化の進展に伴い形成されるエコシステム²⁰の展望やAIの利活用において留意することが期待される事項に関する検討等を行った。

本推進会議における主たる取組については、次のとおりである²¹。

○ AIの開発において留意することが期待される事項

➤ 国際的な議論のためのAI開発ガイドライン案の作成²²【報告書2017】

→ G7、OECD等における国際的な議論のフォローアップ及び国際的な議論への貢献²³【継続】

○ AIの利活用において留意することが期待される事項

➤ AI利活用原則案の作成及び各原則の内容に関する論点整理²⁴【報告書2018²⁵】

→ 各原則の内容に関し整理された論点の検討²⁶【報告書2018を踏まえ検討】

²⁰ 本来は生態系を表す用語である。デジタル、経済、ビジネスといった文脈で用いられる場合には、動植物の食物連鎖や物質循環といった生物群の循環系という元の意味から転化して、経済的な依存関係や協調関係、あるいは、企業間の連携関係などを表すことがある。本報告書においても、このような意味で用いており、その構成要素はAIの開発者、AIサービスプロバイダ、一般の個人の利用者まで含む。これらのステークホルダ間を循環する情報がAIによって生成され、更新されていくという生態系を本報告書では意図している。

²¹ 本推進会議においては、これらの事項に限らず、必要に応じて、AIネットワーク化の推進に向けた社会的・経済的・倫理的・法的課題に関連する事項を検討することとしている。

²² 報告書2017第2章及び別紙1参照。

²³ 第4章（1）参照。

²⁴ 第3章参照。

²⁵ ここでは、本報告書のことを「報告書2018」と記載している。本章において以下同じ。

²⁶ 第4章（2）参照。

- **AIネットワーク化が社会・経済にもたらす影響の評価（シナリオ分析）**
 - 先行的評価（10 ユースケース）²⁷【報告書 2017】
 - 分野別評価（3 分野（まちづくり、健康、モノ））²⁸【報告書 2017】
 - 分野別評価（3 分野（行政、移動、居住））²⁹【報告書 2018】
 - 6 分野（危機管理、教育・学び、仕事、財、趣味・娯楽、カネ）³⁰【継続】
- **AIネットワーク化の進展に伴って形成されるエコシステムの展望等**
 - 利用シーンにおけるAIの利活用、ケーススタディ等³¹【報告書 2018】
 - AIネットワーク化の進展に伴って形成されるエコシステムの展望に関する継続的な検討³²【報告書 2018 を踏まえ検討】
- **AIネットワーク化の進展に伴う影響の評価指標及び豊かさや幸せに関する評価指標の設定**
 - AIネットワーク化に関する指標の検討³³【報告書 2018】
 - AIネットワーク化に関する指標の検討³⁴【報告書 2018 を踏まえ検討】
- **AIネットワーク化に対応した教育・人材育成**
 - 雇用・働き方への影響に関する検討【報告書 2017】
 - AIネットワーク社会に向けた教育・人材育成の在り方等の検討³⁵【報告書 2018】
 - AIネットワーク社会に向けた教育・人材育成の在り方等の検討³⁶【報告書 2018 を踏まえ検討】

²⁷ 報告書 2017 第3章2. 及び別紙3 参照。

²⁸ 報告書 2017 第3章3. 及び別紙4 参照。

²⁹ 第2章1. 参照及び別紙1 参照。

³⁰ 第4章（11）参照。

³¹ 第2章2. 及び3. 並びに別紙2 及び別紙3 参照。

³² 第4章（11）参照。

³³ 第2章5. 参照。

³⁴ 第4章（12）参照。

³⁵ 第2章4. 参照。

³⁶ 第4章（16）参照。

第1章 AIネットワーク化をめぐる最近の動向

AIに関する動きはめまぐるしく、AIに関連する報道を目にしない日がないほどである。本章においては、主として報告書2017公表後のAIネットワーク化をめぐる動向を概観する³⁷。

なお、報告書2017公表以前は、海外においては、学术界や産業界などの取組が中心であったが、報告書2017公表以降は、OECDやG7における国際的な議論が活発に行われていることなどを背景に、学术界・産業界などの取組に加えて、政府レベルでのAIに関する社会的・経済的・倫理的・法的な課題についての取組が広がりつつある。

1. 国内の動向

(1) 人間中心のAI社会原則検討会議

政府は、AIをよりよい形で社会実装し共有するための基本原則となる「人間中心のAI社会原則」を策定し、同原則をG7、OECD等の国際的な議論に供するため、AIに関する倫理や中長期的な研究開発・活用などについて、産学民官による幅広い視野からの調査・検討を行うことを目的として、人工知能技術戦略会議の下に、「人間中心のAI社会原則検討会議」を設置し³⁸、平成30年5月8日に第1回会合が開催された。今後、継続的に検討を進め、平成30年度内に「人間中心のAI社会原則」を策定することとしている。

(2) 高度情報通信ネットワーク社会推進戦略本部（IT総合戦略本部）

高度情報通信ネットワーク社会推進戦略本部（IT総合戦略本部）データ活用基盤・課題解決分科会道路交通ワーキングチームにおいて、高度自動運転システムの実現に向けた政府全体の制度整備の方針（大綱）を策定するため、「自動運転に係る制度整備大綱サブワーキングチーム」が設置され、平成29年10月5日に第1回会合が開催された。以降、平成30年3月23日までに5回の会合が開催され、自動運転に係る制度整備大綱（案）を策定した上で、当該案の内容に即して、高度情報通信ネットワーク社会推進戦略本部・官民データ活用推進戦略会議が同年4月17日に同大綱を決定した³⁹。

大綱においては、①安全性の一体的な確保（関係省庁連携による客観的な指標に基づく走行環境条件の設定）、②自動運転車の安全確保の考え方（安全性要件のガイドライン化（2018年夏目処）、保安基準など）、③交通ルールの在り方（2020年実用化等を見

³⁷ 報告書2017公表以前のAIネットワーク化をめぐる動向については、報告書2017第1章参照。

³⁸ 内閣府、総務省、文部科学省、厚生労働省、経済産業省、国土交通省が合同で開催する会議として設置された。

³⁹ 次に掲げるURLのウェブサイトに掲載。

<http://www.kantei.go.jp/jp/singi/it2/kettei/pdf/20180413/auto_drive.pdf>

据えた交通ルールの検討)、④責任関係(民事責任(下記(3)の検討内容を含む)、刑事責任、自動運転車の走行中のデータ保存に係る検討)、⑤運送事業に関する法制度との関係等に対する制度整備等の方向性が取りまとめられている。

(3) 自動運転における損害賠償責任に関する研究会

国土交通省は、自動運転における自動車損害賠償保障法(昭和30年法律第97号)の損害賠償責任の在り方について検討することを目的として、「自動運転における損害賠償責任に関する研究会」を設置し、第1回会合が平成28年11月2日に開催された。以降、平成30年3月20日までに6回の会合が開催され、同日付けで報告書を取りまとめ公表した⁴⁰。

報告書においては、(レベル4までの自動運転車が混在する)高度自動運転システム⁴¹の導入初期である2020-25年を想定し、次に掲げる5つの論点が整理されている⁴²。

- ① 自動運転システム利用中の事故における自動車損害賠償保障法の「運行供用者責任」をどのように考えるか。
- ② ハッキングにより引き起こされた事故の損害(自動車の所有者が運行供用者責任を負わない場合)について、どのように考えるか。
- ③ 自動運転システム利用中の自損事故について、自動車損害賠償保障法の保護の対象(「他人」)をどのように考えるか。
- ④ 「自動車の運行に関し注意を怠らなかったこと」について、どのように考えるか。
- ⑤ 地図情報やインフラ情報等の外部データの誤謬、通信遮断等により事故が発生した場合、自動車の「構造上の欠陥又は機能の障害」があるといえるか。

(4) 保健医療分野におけるAI活用推進懇談会

厚生労働省は、AIの特性を踏まえ、その活用が患者・国民にもたらす効果を明らかにするとともに、保健医療等においてAIの導入が見込まれる領域を見据えながら、開発推進のために必要な対応及びAIを用いたサービス等の質・安全性確保のために必要な対応等を検討することを目的として、「保健医療分野におけるAI活用推進懇談会」を設置し、平成29年1月12日に第1回会合が開催された。以降、同年3月29日までに4回の会合が開催され、同年6月27日に報告書を取りまとめ公表した⁴³。

⁴⁰ 次に掲げるURLのウェブサイト在所掲。

<<http://www.mlit.go.jp/common/001226452.pdf>>

⁴¹ SAE(Society of Automotive Engineers)レベル3以上の自動運転システムをいう(官民ITS構想・ロードマップ2017)。

⁴² 現行の自動車損害賠償保障法においては、免責3要件(①自己及び運転者が自動車の運行に関し注意を怠らなかったこと、②被害者又は運転者以外の第三者に故意又は過失があったこと、③自動車に構造上の欠陥又は機能の障害がなかったこと)を立証しない限り、自動車所有者などの運行供用者が損害賠償責任を負うこととなっている。

⁴³ 次に掲げるURLのウェブサイト在所掲。

報告書においては、技術の強みを持つこと及び解決する必要のある課題を持つことの両面から、A I 開発を進めるべき重点領域として、ゲノム医療、画像診断支援、診断・治療支援、医薬品開発、介護・認知症対応、手術支援の6点を選定した上で、個別のA I 技術の開発促進に加え、A I の有効性・安全性確保のために適切なルール整備が必要であるとの見解が示されている。また、A I 活用促進のために、十分な質・量を兼ね備えたデータベースを整備することを起点とし、開発の推進と活用が進むことによる保健医療の向上が循環的に進む「保健医療循環A I 開発エコシステム」が重要であるとされている。一方、保健医療分野におけるA I の限界とあるべき姿として、A I による判断が患者・国民のためにならない可能性があることに言及し、そのために、A I はあくまでコンピュータによる人間の知的活動の支援のためのものであり、A I による予測を含めた各種状況を踏まえて人間が最終判断を下す必要性は引き続き残ること、将来の保健医療関係者は、人工知能を保健医療の質の向上を増幅させるものとしてうまく使いこなしていく知識・技術が必要であることが指摘されている。

(5) A I ・データ契約ガイドライン検討会

経済産業省は、契約によるデータの利活用やA I の利用・開発を促進するため、それらの契約を締結する際の法的論点などを整理することを目的として、平成29年12月に「A I ・データ契約ガイドライン検討会」を設置した。以降、同検討会の下に設置した作業部会において具体的なユースケース⁴⁴に基づき検討を重ね、平成30年3月20日の第3回会合において素案を提示した後、意見募集（パブリックコメント）を経て、同年6月●日に「A I 編」及び「データ編」の2編からなる「A I ・データの利用に関する契約ガイドライン」を取りまとめ公表した⁴⁵。

A I 編は、特に機械学習技術を念頭に、そうした技術を利用したソフトウェアの開発契約又は利用契約を作成するに当たっての考慮要素、トラブルを予防する方法等に関する基本的な考え方等について整理したもので、技術特性（学習済みモデルの内容・性能等が契約締結時に不明瞭であること、その内容・性能等が学習データセットに依存していること、その生成に際して特にノウハウの重要性が高いこと、各種生成物に対する再利用の需要が存在すること等）を踏まえた契約の在り方等が説明されている。特に、開発契約に関しては、A I 技術の特性を考慮した「探索的段階型」の開発方式（①アセスメント段階、②P o C⁴⁶段階、③開発段階、④追加学習段階を設け、ステップ・バイ・

<<http://www.mhlw.go.jp/file/05-Shingikai-10601000-Daijinkanboukouseikagakuka-Kouseikagakuka/0000169230.pdf>>

⁴⁴ ユースケースとして、実際の事案において契約形態及び契約時に課題となること（相談事項）を事業者から作業部会に示してもらうことにより、その充実を図っており、当該ガイドラインを検討する上での題材として取り扱っている。

⁴⁵ 次に掲げるURLのウェブサイトに掲げる。

<●●>

⁴⁶ Proof Of Concept の略。概念実証やコンセプト実証などといわれ、新しい概念や理論、原理

ステップで検証と確認を経ながら開発を進める方式)が提案され、段階ごとに論点やモデル契約が提示されている。このほか、外国企業との間で契約を締結する場合の考慮要素等が示されている。

データ編は、契約類型別にデータの取扱いに関する法的論点や契約での取決め方等について整理したもので、契約の類型を「データ提供型(一方当事者から他方当事者へのデータ提供)」、「データ創出型(複数当事者が関与して創出されるデータの取扱い)」、「データ共用型(プラットフォームを利用したデータの共用)」の3つに分類した上で、契約条項例や条項作成時の考慮要素(クロス・ボーダー取引、個人情報の取扱いを含む)がそれぞれ説明されている。

(6) 産業競争力懇談会

一般社団法人産業競争力懇談会(COCN: Council on Competitiveness - Nippon)は、平成29年度の推進テーマの1つとして、「人工知能間の交渉・協調・連携」に関する検討を産学官連携の下で行い⁴⁷、最終報告書を平成30年2月21日に公表した⁴⁸。

報告書においては、対象とするユースケースを「製造バリューチェーン分野」(加工、組立から物流に至るまでの一連のサプライチェーンの受発注契約の支援)及び「交通・人流分野」(交通管制連携、モビリティ・サービス(交通リソースの融通))の2つ⁴⁹に絞り、両分野(のユースケース)における課題分析を行うとともに、想定されるステークホルダへのヒアリングを計16件実施している。その結果として、ネットワーク外部性⁵⁰の突破が多く、ステークホルダの主たる課題として挙げられており、

- ・ 最小実装によるスモールスタートを初段とした各ユースケースでの交渉・協調・連携の社会実装を官民が連携して段階的に推進していくことが重要であること
 - ・ その基本的な段取りを①プレイヤー間でのシミュレーション、②限定環境下(特区など)での一部プレイヤー間での実動作、③限定環境下でのプレイヤー拡大、④一般環境へ拡大の4つとし、各前段の開発内容がそのまま使えること
 - ・ 各ユースケースでこれらのステップを横断的に進められることが好ましいこと
- などの提言がなされている。

などが実現可能であることを示すための簡易な試行のことをいう。

⁴⁷ 本プロジェクトは、「人工知能間の自動交渉技術による挙動調整・利害調整を行う仕組みの社会実装」を目標として平成28年度より検討が進められてきたもので、コンセプトの提唱、政府が掲げる将来像における適用シーンの抽出、ユースケース群の選定、社会実装における分野共通の課題分析が行われてきていた。

⁴⁸ 次に掲げるURLのウェブサイトに掲載。
<<http://www.cocn.jp/theme98-L.pdf>>

⁴⁹ それぞれ、本推進会議の分野別評価における「モノ」及び「まちづくり」とも関係が深い分野である。「モノ」及び「まちづくり」の評価の結果については、報告書2017別紙4参照。

⁵⁰ 電話やFAXのように、相互運用可能な形態で一定の普及率を超えるまではユーザーに利益がない性質のこと。

(7) 日本経済団体連合会

一般社団法人日本経済団体連合会は、産業界の現状として日本企業のAI及びその関連分野における国際的な産業競争力を向上させるべく、未来産業技術委員会の下に「AI活用原則タスクフォース」を設置し⁵¹、第1回会合を平成30年1月17日に開催した。タスクフォースにおいては、①社会的インパクト、②公共財(Social Goods)としてのAI及び③産業競争力の向上を論点として議論を行っており、平成30年内の取りまとめに向け、AIの展開戦略とともに、AI-readyな組織・社会の構築の必要性等を打ち出す方向で検討が進められている。

2. 海外の動向

(1) 米国電気電子学会(IEEE)

米国電気電子学会(IEEE)の自律的及び知的システムの倫理(Ethics of Autonomous and Intelligent Systems)に関するグローバル・イニシアティブは、2017年12月12日に報告書「倫理的に調整された設計 第2版」を公表した⁵²。第2版は、AIの倫理的及び社会的な実装の在り方についての公衆の議論の発展、標準(IEEE P7000 シリーズ等)及び関連する認証プログラムの策定の喚起、各国の政策及びグローバルな政策の形成の促進を目指して、第1版⁵³の公表(2016年12月)後に寄せられたフィードバックなど⁵⁴を踏まえて取りまとめられたものである。

第2版においては、第1版に引き続き、AIの倫理的な設計、開発及び実装において参照されるべき一般原則として、人権、幸福、アカウントビリティ、透明性、悪用への警戒の5つ項目が掲げられている。また、自律型知的システム(AIS:Autonomous and Intelligent System)に人間の規範や価値を実装するための方法論に基づきAISに関する経済的・人道的・法的課題について整理するとともに、汎用人工知能や自律型兵器に特有の問題についても論じられている。なお、第2版では、第1版に盛り込まれている8つの章⁵⁵に加え、新たに、⑨感情(affective)コンピューティング、⑩政策、

⁵¹ タスクフォースの主査は、本推進会議の北野構成員(株式会社ソニーコンピュータサイエンス研究所代表取締役社長兼所長)が務めている。

⁵² The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, Ethically Aligned Design: A Vision for Prioritizing Human Wellbeing with Autonomous and Intelligent Systems, Version Two - Request For Public Discussion (2017).

⁵³ The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems, Ethically Aligned Design: A Vision for Prioritizing Human Wellbeing with Artificial Intelligence and Autonomous Systems, Version One - Request For Public Discussion (2016).

⁵⁴ 我が国からも、本推進会議の江間環境整備分科会・影響評価分科会構成員が中心となり国内でワークショップを開催し、意見を取りまとめた上でIEEEに提示している。

⁵⁵ ①倫理的に調整された設計をするための一般原則、②AISへの価値の埋め込み、③倫理的な研究及び設計を導く方法、④汎用人工知能(AGI:Artificial General Intelligence)

⑪ A I Sにおける伝統的倫理感、⑫ I C Tにおける複合現実、⑬幸福 (Well-being) をテーマとする5つの章が追加されている。

さらに、第2版の公表に先立って、I E E Eは、2017年11月17日に第1版において示された論点等を踏まえ、A Iの倫理に関し、次に掲げる3件の標準プロジェクトを新たに承認した⁵⁶。

- ・ I E E E P7008TM-ロボット、知能及び自律的システムのための倫理的に駆動されたナッジの基準：ロボットやA Iが利用者に倫理的なナッジ（選択肢の提示等）を与えられるよう設計するための基準を策定。
- ・ I E E E P7009TM-自律的及び準自律的システムのフェールセーフの設計基準：A I等の障害・事故が生じた際の安全性を確保するための設計の基準を策定。
- ・ I E E E P7010TM-倫理的なA I及び自律的システムのための幸福の計測基準：A I等に影響を受ける人間の幸福を計測するための基準を策定。

また、2018年5月23日に次に掲げる2件の標準プロジェクトを新たに承認した⁵⁷。

- ・ I E E E P7011TM-ニュース情報源の信頼性特定・評価のプロセスに関する基準：A I等に影響を及ぼすニュース記事の精度（真実度）を評価するための基準を策定。
- ・ I E E E P7012TM-機械読取可能な個人のプライバシー条項に関する基準：個人のプライバシー条項を読み了解し同意する手段を、第三者が操作する機械が個人に提供するための基準を策定。

（2）米国情報技術工業協議会（I T I）

米国情報技術工業協議会（I T I）⁵⁸は、2017年10月24日にA Iの可能性を最大限に引き出すことを目的とした「A I政策原則」を公表した⁵⁹。

原則においては、（責任ある開発・利活用を促進するための）産業界の責任に関する項目として、責任ある設計及び実装、安全と制御可能性、頑健で代表的なデータの利用（対バイアス）、解釈可能性、自律性に応じたA Iシステムに関する責任が掲げられている。また、（A Iエコシステム構築推進に関する）政府への期待に関する項目として、A Iの研究開発への投資、柔軟な規制手法の検討、インターネットのイノベーション及

及び人工超知能（A S I：Artificial Superintelligence）の安全性及び恩恵、⑤パーソナルデータと個々のアクセス制御、⑥自律的兵器システム、⑦経済的・人道的問題、⑧法律の8章である。

⁵⁶ https://standards.ieee.org/news/2017/ieee_global_initiative.html

このほかI E E E P7000シリーズとして、A Iの倫理的設計、透明性、バイアス、プライバシー、個人データ等に関する標準プロジェクトも既に始動している。

⁵⁷ http://standards.ieee.org/news/2018/p7011_p7012.html

⁵⁸ 米国の情報技術関連の業界団体であり、IBM、Microsoft、Google、Amazon、Facebook、Appleなど米国の主要I C T企業のほか、キヤノン、富士通、パナソニック、トヨタ、東芝など日系企業も参加している<<https://www.itic.org/about/member-companies>>。

⁵⁹ 次に掲げるU R Lのウェブサイトに所掲。

<<https://www.itic.org/public-policy/ITIAIPolicyPrinciplesFINAL.pdf>>

びセキュリティの促進、サイバーセキュリティとプライバシー、グローバルな基準及びベストプラクティスが掲げられている。さらに、公私協働に関する項目として、アクセスの民主化と平等な機会の創出、科学・技術・工学・数学（STEM）教育の促進、労働の置き換えに対する考慮、公私協働が掲げられている。

（３）UNI グローバルユニオン

サービス業や情報産業等の労働者を代表する国際的な労働組合であるUNI グローバルユニオンは、2017年12月11日に労働者の利益を擁護し職場における権力の健全な均衡を維持する観点から、AIの倫理に関する「倫理的AIのための上位10原則」を公表した⁶⁰。

10原則は、①透明性、②「倫理的な記録装置」の実装、③人々と地球への奉仕、④人間による指揮アプローチの採用（AIシステムが人事に用いられた際の労働者の説明を求める権利を含む）、⑤性別に左右されず、バイアスのないAI、⑥AIシステムの便益の分配、⑦（雇用環境の変化への）適切な移行の確保並びに基本的自由及び権利の支援、⑧グローバルガバナンスのメカニズムの構築、⑨ロボットへの帰責の禁止、⑩AI軍拡競争の禁止、から構成されている。

（４）EU第29条データ保護作業部会

EU第29条データ保護作業部会は、EU一般データ保護規則（General Data Protection Regulation：GDPR）の解釈の明確化を目的として、2017年10月3日に同規則における「自動処理による個人に関する意思決定（decision）及びプロファイリングに関する規定」に関するガイドラインを公表し、改訂した上で、2018年2月6日に採択した⁶¹。

ガイドラインにおいては、AI等の技術発展によりプロファイリング及び自動処理による意思決定が容易になっており、プロファイリング及び自動処理による決定が、個人や経済・社会に便益をもたらすと同時に、個人の権利及び自由に重大なリスクを及ぼし得るとの認識が示されている。

その上で、①プロファイリング及び自動処理による意思決定一般に関する指針（透明性、公正性、目的の限定、データ最小化、正確性等）及び②（GDPRにより原則禁止とされる）法的効果又はそれに類する重大な影響を個人に及ぼす自動処理のみによる意思決定に関する指針（「決定に服さない権利」、例外規定等）が示されている。関連して、プロファイリングの文脈で、データ主体の「情報を与えられる権利」、「アクセス権」、

⁶⁰ 次に掲げるURLのウェブサイトに掲載。

<http://www.thefutureworldofwork.org/media/35420/uni_ethical_ai.pdf>

⁶¹ Guidelines on Automated individual decision-making and Profiling for the purposes of Regulation 2016/679, wp251rev.01

<http://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053>

「異議申立権」等が具体化されている。また、子供を対象とした行動ターゲティング広告等のプロファイリング及び自動処理による決定について、適切なセーフガード措置を取るべきである等の厳格な指針が示されている。

(5) 欧州委員会

欧州委員会の独立諮問機関である欧州科学・新技術倫理グループ (the European Group on Ethics in Science and New Technologies : EGE) は、2018年3月9日にAI・ロボティクス・自律システムの設計、利活用及びガバナンスなどに関する国際的な理解を確保していくよう欧州委員会に対し要求する旨の声明文を公表した⁶²。

声明文においては、議論の第一歩として、EU条約・基本権憲章に含まれる価値に基づく9つの倫理原則(①人間の尊厳、②自律性、③責任、④正義・公平・連帯、⑤民主主義、⑥法の支配とアカウントビリティ、⑦セキュリティ・安全性・心身の整合性、⑧データ保護とプライバシー、⑨サステナビリティ)が提案されている。

また、欧州委員会は、同日に次に掲げる事項をタスクとするハイレベル専門家グループの募集を開始した。

- ・ 産官学に加え、消費者団体や労働組合、市民団体などEU内外の複数のステークホルダを含めたAIに関するアライアンス「European AI Alliance」の設立の助言を行うこと
- ・ 上記EGEの声明を踏まえたAIガイドライン案を策定し、アライアンスに照会しつつ、2018年末までに公開すること

さらに、欧州委員会は、同年4月25日に、ハイレベル専門家グループを同年5月までに、「European AI Alliance」を同年7月までに、それぞれ設立する予定であることを公表するとともに⁶³、欠陥製品における消費者及び生産者のための法的な明瞭性確保に向け、技術開発の観点から、製造物責任(PL)指令の解釈に関するガイダンスを2019年中頃までに発行する予定であることを公表した⁶⁴。

おって、欧州委員会は、同年5月17日に完全自動運転に関する発表を行い、2030年代に欧州で完全自動運転社会を実現していくための工程の1つとして、AI倫理ガイドラインを策定する旨を掲げている⁶⁵。

⁶² 次に掲げるURLのウェブサイトにも掲。

<http://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf>

⁶³ 次に掲げるURLのウェブサイトにも掲。

<http://europa.eu/rapid/press-release_MEMO-18-3363_en.htm>

⁶⁴ 次に掲げるURLのウェブサイトにも掲。

<http://europa.eu/rapid/press-release_IP-18-3362_en.htm>

⁶⁵ Connected & Automated Mobility - For a competitive Europe

<<https://ec.europa.eu/transport/sites/transport/files/3rd-mobility-pack/3rd-mobility-pack-factsheets-automatedconnected.pdf>>

(6) 仏国

仏与党（LREM）所属のセドリック・ビラニ下院議員は、2018年3月28日にAI普及のための報告書を発表した。報告書においては、AIの普及が最も進んでいる国々として、米国、中国、英国、カナダ、イスラエルを挙げ、欧州／フランスが緊急に挽回すべきである旨が強調されており、そのための具体策として、AI関係の研究者の初任給の倍増、公共部門によるプロジェクト入札の手續の簡素化と迅速化により、独創的な研究を進めやすくする環境作りなどが提案されている。また、AI開発に際して注力すべき分野として、医療、輸送、環境、防衛の4分野が挙げられている。さらに、AIがもたらす倫理上の問題についても議論を深めるよう求めており、AI倫理に関する独立行政機関の設立が提唱されている⁶⁶。

マクロン大統領は、この報告書を受けて、同月29日にAIの普及に向けた施策を発表した。新技術分野でAIが新たな発展のカギを握ることを見据えて、諸外国との競争に打ち勝つべく、テコ入れを図るため、

- ・ 大統領任期終了の2022年までに15億ユーロの予算をAI振興のために充当
- ・ フランス国立情報学自動制御研究所（INRIA）が中心となり、4～5箇所のAI研究所のネットワークを構築し、民間と協力しつつ研究推進と人材育成を推進
- ・ 研究者によるスタートアップ設立の手續を簡素化
- ・ データについて、官民両方でアクセスを簡易化、特に大手企業の一部データについては、公益性を理由に利用を認めさせるという制度を欧州レベルで導入
- ・ AI開発の重要分野として、「保健」と「モビリティ」の2分野を選定
- ・ 倫理面では、国際的な規模での専門家会議の設置を提言、国内では、公的機関のアルゴリズム開示に取り組む

などの施策が示されている。

(7) 英国

英国政府は、2017年11月22日にAI等における安全で倫理的かつ画期的なイノベーションを可能にすることを目的とした「データ倫理イノベーションセンター（Centre for Data Ethics and Innovation）」を新設する予定であることを公表した⁶⁷。本センターを通じて、実用的なユースケースの検討において世界をリードしていくとしている。

また、英国上院のAI特別委員会は、2018年4月16日にAIに関する報告書「英国における人工知能（AI）：英国はAIを活用し、そして活用できる準備ができている

⁶⁶ この独立行政機関については、政府から諮問を受けるだけでなく、一般市民の請願も受け付ける開かれた機関とすべきであるとされている。

⁶⁷ <https://www.gov.uk/government/news/autumn-budget-2017-25-things-you-need-to-know>

か (A I in the UK: ready, willing and able?)」を公表した⁶⁸。

報告書においては、今後の英国経済成長の原動力（課題解決や生産性向上等）としてのA Iの必要性を強調した上で、その実現のためにはA Iの倫理的側面を中核においたアプローチが重要であるとされている。さらに、国内外において採用されるべき分野横断的なA Iの倫理行動規範 (Code of Practice) となる「A Iコード」の策定を含め、A Iが英国社会及び経済にもたらす可能性を引き出すことを支援する一方で、A Iがもたらす可能性がある脅威やリスクから社会を守るために必要な74項目にわたる勧告が盛り込まれている。「A Iコード」の策定に当たり、次に掲げる5つの原則が提示されている。

- ① 人工知能は、人類共通の利益のために開発されるべきである。
- ② 人工知能は、透明性及び公平性の原則に基づいて運用されるべきである。
- ③ 人工知能は、個人・家族・コミュニティのデータに関する権利やプライバシーを侵害する形で利用されるべきではない。
- ④ すべての市民は、人工知能と共に精神的・感情的・経済的な繁栄を享受できるような教育を受ける権利を持つべきである。
- ⑤ 人間を傷つけ、破壊し、欺く自律的な力は人工知能に埋め込まれるべきではない。

併せて、同報告書において、上記「データ倫理イノベーションセンター」の役割として、データ信託 (data trusts) によってデータが管理されている人々の意見の反映確保、A Iシステムの透明性に関する要件に係るガイダンスの作成、分野横断的なA Iコードの導入などを行う予定であることが示されている。

加えて、下院科学技術委員会は、同年5月23日にアルゴリズムにおける意思決定の在り方に関する報告書 (Algorithms in Decision-making) を公表した⁶⁹。

報告書においては、アルゴリズムのアカウントビリティを担保するために特に透明性が重要であることが指摘されるとともに、政府に対して、公共サービスで使用するアルゴリズムの監督を行うとともに、民間とのパートナーシップの開発・展開を調整する閣僚を特定すべき旨が要求されている。これに加えて、上記「データ倫理イノベーションセンター」の役割として、アルゴリズムによるバイアスの評価、透明性確保のためのツールの検証などを行うべき旨が指摘されている。

(8) 米国

ホワイトハウスは、2018年5月10日、米国がA Iについて主導的立場をとるための必要な政策を議論するため、100名を超える政府機関、産業界、学术界の代表を集めた

⁶⁸ 次に掲げるURLのウェブサイトにも掲げ。

<<https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf>>

⁶⁹ 次に掲げるURLのウェブサイトにも掲げ。

<<https://publications.parliament.uk/pa/cm201719/cmselect/cmsctech/351/351.pdf>>

AIサミットを開催した⁷⁰。

サミットにおいては、国家レベルのAIのR&Dエコシステムの支援、AIの便益を享受するための人材育成、AIイノベーションに対する阻害要素の除去、特定領域におけるインパクトの高いAI応用の可能性を論点として議論が行われた。また、研究開発の省庁間での優先順位を調整し、連邦政府のAI投資計画を推進するAI特別委員会を設立することが公表された。

(9) 中国

中国政府は、2017年7月20日に2030年までに同国を理論・技術・応用の全分野で世界の最高水準まで引き上げ、世界の主要なAIのイノベーションの拠点となることを目標に掲げた「次世代人工知能発展計画」⁷¹を公表した。

計画においては、第一段階である2020年までにAIの技術を「世界先進水準まで引き上げる」ことが目標として掲げられている。その具体策として、政府工業・情報化部は、2017年12月14日にAIの発展に必要となるハード・ソフト両面の基盤強化、人材育成などを含む「次世代人工知能産業の発展促進に関する三年行動計画」(2018-20年)を発表した。

また、AIに関するサービスの開発が急速に進展する中、標準化の推進も重視されており、その一環として、工業・情報化部配下で情報通信産業の標準化を担当する中国電子技術標準化研究所は、2018年1月18日に産学各団体の協力の下、国内外のAIに関する標準化の動向を取りまとめた白書を作成し公表⁷²するとともに、AI標準化組織及び専門諮問組織の設立を発表した。白書においては、技術面の進歩及びその標準化と並行して、安全性、倫理面(人間の尊厳、人間・自然の利益の最大化)、プライバシーの政策、法、及び標準の策定が重要である旨が指摘されている。

3. 国際的な議論の動向

(1) OECD(経済協力開発機構)

OECDは、2017年(平成29年)10月26日から27日にかけて、フランス・パリにおいて、総務省との共催により、AIの普及が社会にもたらす機会と課題、政策の役割と国際協調について議論を行うべく、AIに関する政策をテーマとする国際カンファレンス「AI: Intelligent Machines, Smart Policies」を開催した⁷³。カンファレン

⁷⁰ <https://www.whitehouse.gov/wp-content/uploads/2018/05/Summary-Report-of-White-House-AI-Summit.pdf>

⁷¹ 「新一代人工智能发展规划」

[〈http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm〉](http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm)

⁷² 次に掲げるURLのウェブサイトにも掲げられている。

[〈http://www.cesi.ac.cn/images/editor/20180124/20180124135528742.pdf〉](http://www.cesi.ac.cn/images/editor/20180124/20180124135528742.pdf)

⁷³ <http://www.oecd.org/going-digital/ai-intelligent-machines-smart-policies/>

スには、世界各国の産学民官の有識者が一堂に会し、活発な議論が行われた。本推進会議から、須藤議長及び平野幹事・環境整備分科会長が参加し、A I 開発ガイドライン案など本推進会議の検討状況及び成果を紹介した。加盟各国の参加者やOECD事務局らから、A I 開発ガイドライン案などを踏まえつつ、国際社会においてA I に関する議論と検討を進めることについて賛同する姿勢が示された。

また、2017年（平成29年）11月21日から22日にかけて、デジタル経済政策委員会（Committee on Digital Economy Policy：CDEP）が開催され、OECDのA I に関する今後の取組について議論が行われた⁷⁴。本推進会議から、森川幹事（同委員会副議長（当時））及び実積構成員（同委員会次期副議長（当時））が参加した。委員会における議論の結果、OECDにおける今後の取組として、①2018年（平成30年）5月開催の次回会合に向け、事務局が分析レポートを作成すること、②加盟国の意向を踏まえつつ、2019年（平成31年）以後、理事会勧告作成に向けた作業に着手すること⁷⁵について、加盟国の同意が得られた。

さらに、2018年（平成30年）5月16日から18日にかけてCDEPが開催され、事務局より「社会におけるA I」と題する分析レポートが報告された。分析レポートにおいては、本推進会議の検討状況及びA I 開発ガイドライン案のOECD、G7等への展開について紹介されている。分析レポートは、今次会合における議論を踏まえた修正がなされた後、同年11月の会合で最終化される予定である。

（2）G7

2017年（平成29年）9月25日から26日にかけて、イタリア・トリノで開催されたG7情報通信・産業大臣会合の閣僚宣言において、A I の進歩が経済及び社会に莫大な便益をもたらすことを認識するとともに、デジタル経済におけるイノベーション及び成長を主導する人間中心のA I というビジョンを共有し、「附属書2：我々の社会のための人間中心のA I に関するG7マルチステークホルダー交流」⁷⁶に記載するマルチステークホルダーの交流を通じて同ビジョンを一層発展していくことが合意された。なお、附属書2において、A I 開発ガイドライン案が言及されている。

また、2018年（平成30年）3月27日から28日にかけて、カナダ・モントリオール

⁷⁴ 委員会においては、OECD事務局から、①平成29年10月に開催されたA I に関する国際カンファレンス（総務省共催）の結果、②OECDにおける今後の取組について説明が行われたところ、事務局の説明に対し日本を含め12箇国が支持を表明した。我が国からは、A I 開発ガイドライン案や影響評価に係る分析など、本推進会議の取組を紹介した。OECDの今後の取組については、①マルチステークホルダーによる対話の必要性、②理事会勧告がA I の開発を萎縮させる規制的、拘束的なものであってはならないことを指摘した上で、引き続きOECDにおける議論に貢献していく意向を表明した。

⁷⁵ 今後の作業の方向性について加盟国の同意が得られたものであり、理事会勧告の作成については、2018年（平成30年）11月以降、CDEPで改めて議決／承認が必要となる。

⁷⁶ 閣僚宣言サマリー附属書2：<http://www.g8.utoronto.ca/ict/2017-G7-ICT-Annex2-AI.pdf>

で開催されたG7イノベーション大臣会合において、「未来の仕事に備える (Preparing for Jobs of the Future)」をテーマに、IoT、ビッグデータ、AI等の新たなイノベーションが社会・経済や労働市場に及ぼす影響について議論が行われ、その成果が、議長サマリーの形で取りまとめられた。なお、「附属書B：AIに関するイノベーション会合声明」⁷⁷において、AI開発ガイドライン案が言及されている。

今後は、G7による将来の政策協議に情報を入力するため、OECDの支援を受け、AIに関するマルチステークホルダーによる対話・連携を促進していく予定である。また、時限設置されるイノベーションWGの支援により、同年秋にカナダが主催するAIに関するマルチステークホルダー会合を開催する方向で調整が進められる予定である。

⁷⁷ 議長サマリー附属書B：<http://www.g8.utoronto.ca/employment/2018-labour-annex-b-en.html>

第2章 AIネットワーク化が社会・経済にもたらす影響

1. AIネットワーク化に関するシナリオ分析（分野別評価）

（1）背景・経緯

検討会議において、AIネットワーク化の進展に伴いロボットに関し想定され得るリスクのシナリオを試行的に例示しつつ分析を行い、シナリオに即したリスク対処の在り方を検討した。

ここで示されたシナリオは、あくまでも試行的に例示されたものに過ぎず、今後、社会の各分野のステークホルダの参画を得つつ、AIネットワーク化の進展等に応じて、リスク・シナリオを拡充するとともに、その内容やリスク評価等を不断に見直すなど、リスク・シナリオについて継続的に検討を行い、リスク・シナリオの共有を図り、シナリオに基づくリスク対処を進めていくことが求められるとされている⁷⁸。

また、報告書2016において、今後の課題として、リスク・シナリオの作成・共有に向けて、

- ・ AIネットワークの利活用の場면을想定した各種のリスクに関するシナリオの作成
- ・ シナリオに基づくリスク対処（リスク評価、リスク管理、リスク・コミュニケーション）の推進

等が提言されている⁷⁹。

検討会議を発展的に改組した本推進会議においては、親会の下に環境整備分科会及び影響評価分科会が置かれ⁸⁰、このうち主に影響評価分科会において、AIネットワーク化が社会・経済にもたらす影響（インパクト及びリスク）の評価に関する検討を行ってきた。

本推進会議においては、報告書2016の提言を踏まえて、AIネットワーク化に関するリスクのみならず、インパクト（主に良い影響、便益）についても併せて評価することが重要であるとの認識の下、AIネットワーク化が社会・経済にもたらすインパクト及びリスクに関し、AIシステムの具体的な利活用の場面（ユースケース）を想定したインパクト及びリスクの評価（シナリオ分析、将来展望）として、

- ・ 国際的な議論のためのAI開発ガイドライン案の検討に向けた先行的評価
- ・ AIシステムを利活用する分野ごとの分野別評価⁸¹

⁷⁸ 詳細については、報告書2016第4章参照。

⁷⁹ 報告書2016第5章11. 参照。

⁸⁰ 前掲注（4）参照。

⁸¹ 分野別評価は、その結果から得られる示唆をAIの利活用において留意することが期待される事項等の検討に役立てることを主たる目的として実施するものである。報告書2017においては、先んじて実施した3分野（まちづくり、健康、モノ）の評価の結果を掲げている。

を実施した⁸²。

なお、インパクト及びリスクの評価の対象とするA Iシステムの具体的な利活用の場面（ユースケース）について、分野別評価においては、A Iシステムを利活用する分野の構成を利用者の視点から（ユーザ・セントリック・アプローチ）区分した上で本推進会議が整理したユースケースを用いた⁸³。

これら成果を踏まえ、報告書 2017 において、今後の課題として、

- ・ A I ネットワーク化が社会・経済にもたらす影響に関するシナリオ分析の継続的な実施及び国際的な共有が提言されている⁸⁴。

（２）分野別評価

① 評価の枠組み

分野別評価における分野の区分については、報告書 2017 において、次のように整理している⁸⁵。

A I ネットワーク化の進展に伴いA Iシステムの利活用をめぐるエコシステムが複雑化していくものと考えられることに鑑み、将来社会の在り方に関するフォアキャスト的なアプローチとバックキャスト的なアプローチの双方を採り入れ、次の事項を念頭に置き、次表のように整理した。

- ・ A I ネットワーク化が主としてA Iシステムの「利用者」（特に最終利用者）にどのような便益やリスクをもたらす得るのかを重視し、利用者を「個人」、「公共」及び「産業」の3つに区分した上で、「個人」については生活の各場面、「公共」については公益の確保に関連する場面、「産業」について商流が形成される場面に着目して分野の区分を設定する。
- ・ 分野別評価においては、A I ネットワーク化の進展により、領域が融合していくことや領域横断的なA I ネットワークの利活用が進むことが見込まれることを踏まえて分野の区分を設定する。
- ・ A I ネットワーク化が社会的なニーズへの対応や課題の解決に貢献するものであることから、社会的なニーズ又は課題を踏まえて分野の区分を設定する。

⁸² 先行的評価の結果については、報告書 2017 第 3 章 2. 及び別紙 3 参照。分野別評価（まちづくり、健康、モノ）の結果については、報告書 2017 第 3 章 3. 及び別紙 4 参照。

⁸³ 先行的評価においては、国内外における関連する議論を進めている機関等が提示している既存のユースケースを基礎として本推進会議が整理したユースケースを用いた。

⁸⁴ 報告書 2017 第 4 章（10）参照。

⁸⁵ 「パブリック・ガバナンス」については、分かりやすさの観点から、本報告書において、名称を「行政」と変更している。

また、分野別評価においては、利用者の視点から（ユーザ・セントリック・アプローチ）評価を行うこととしていることに鑑み、主として製品・サービスの供給面の領域を横断し又は領域が融合していくものと見込まれるユースケースを用いることとし、「領域横断前の段階」と「領域横断後の段階」の2段階に分けて評価を行うこととしている。

分野別評価（分野の区分）

大分類	中分類	小分類	領域横断・融合（例）	主な社会的ニーズ/課題
公共	まちづくり		■地域振興やまち全体としての情報保有、コミュニティ形成 公共インフラ+防災+スマートシティ+居住+移動など	持続可能な都市の構築 移民・難民問題への対応
	行政		■所得の再分配等の社会保障、政策立案、財・サービスの多様化に 応じた税の徴収 行政+金融+財など	廣域対策、財政の健全化
	危機管理		■災害やパンデミック等への対応における連携 公共インフラ+防災+行政+医療+物流+金融など	レジリエントなインフラ整備
個人	健康	例えば、 ◇世代軸 ◇時間軸	■健康情報や生活情報をめぐるAIネットワーク化 医療+生活支援+農林水産+移動+居住+保険など 在宅ケア、育児や介護を支援するAIネットワーク化 遠隔医療+遠隔介護+福祉情報支援など	健康的な生活の確保・福祉の推進 健康長寿社会の実現
	教育・学び		■能力、スキル、関心、意欲等をめぐるAIネットワーク化 学校教育+研修+生涯学習+コミュニティなど	生涯学習の機会の促進 知の資産の持続的創出
	仕事		■能力、適性、仕事観、労働市場等をめぐるAIネットワーク化 労働+財+ワークライフバランス(趣味・娯楽、居住、移動等)など	働きがいのある人間らしい仕事の実現 女性のエンパワメント
	財		■ライフステージに応じた様々な財・サービスの提供 カネ+モノ+金融+行政など	年金制度等の改善
	移動		■移動のためのAIネットワーク化 運輸・物流+医療+介護+防災+観光・旅行など	少子高齢化時代における交通インフラの構築
	居住		■住環境をめぐるAIネットワーク化 スマートハウス・スマートシティ+建設+防災など	既存住宅の活用やリフォームの活性化
	趣味・娯楽		■心を豊かにするAIネットワーク化 豊かさ創造+サービス業(観光・旅行等)+教育・学びなど	文化芸術資源の活用
産業	モノ		■インダストリー4.0やインダストリアル・インターネット等による産業構造 の変化(第6次産業、スマートファクトリー、モノのサービス化等) 製造業など全産業	生産性の向上 持続可能な農業・エネルギー供給
	カネ		■フィンテックを中心とする取引の高度化 金融+保険+小売+サービス業+生活支援など	開かれた強靱な金融システムの構築

なお、評価の対象とするインパクト及びリスクの考え方については、報告書 2017 において、次のように整理しており、本報告書においても、これを踏襲している。

インパクト及びリスクの評価に当たっては、インパクトについては、AIシステムを利活用することにより、従来人間には不可能又は困難であったことが可能となることや、効率化や精度の向上等が図られることをインパクトとして評価することとした。他方、リスクについては、AIシステムの利活用に伴い新たに生ずるリスク(AIネットワーク化固有のリスク)だけではなく、AIシステムを利活用しない場合と比べて増加するリスクやAIシステムを利活用しない場合と比べて軽減するものの、

利活用する場合における事故等による被害の規模や波及がある程度大きいリスクについても評価の対象とすることとした⁸⁶。

なお、セキュリティに関するリスク（A Iシステムがハッキングされること等により機能不全に陥るリスク等）、不透明化のリスク（多重かつ複雑に連携しているA Iシステムの一部の動作が不透明化することに伴い、ネットワークを介してつながるA Iシステムの動作全体が不透明化するリスク等）、制御喪失のリスク（A Iシステムが攻撃を受けたり、不具合が生じたりすることにより、制御が困難又は不可能となり、当該A Iシステムとネットワークを介してつながるA Iシステム全体が機能不全に陥るリスク等）については、各ユースケースに共通するリスクと考えられることから、個々のユースケースの評価において個別には記載しないこととしている⁸⁷。

② 分野別評価

分野別評価について本報告書においては、上記12分野のうち、次の3分野に関するユースケースを対象として評価を行った⁸⁸。

- ・ 公共：行政に関するユースケース
- ・ 個人：移動に関するユースケース
- ・ 個人：居住に関するユースケース

これらのユースケースについて、次に掲げる居住に関するユースケースの例のように、A Iシステムの利活用に関するインパクト及びリスクの評価を行った。各ユースケースの評価の結果の詳細は、別紙1のとおりである⁸⁹。

⁸⁶ インパクト及びリスクのいずれについても、あくまでも典型的に想定され得るものの一部を例示として記載しているだけであり、網羅的なものではないことに留意することが必要である。

⁸⁷ これらのリスクへの対処が重要であることは言うまでもない。むしろ、共通するリスクであるからこそ、これらのリスクへの対処が特に重要となる。

⁸⁸ 報告書2017において、公共：まちづくり、個人：健康、産業：モノの3分野について評価を実施した。他の分野については、本報告書の公表後、本推進会議において評価を行う予定である。

⁸⁹ 各ユースケースに記載されているA Iシステムの利活用については、現在の法制度等を必ずしも前提とはせず、将来的な利活用の可能性を展望して記載している。このため、現在の法制度等の下では実現が困難な利活用も含まれており、そのような利活用に関し制度的な課題の抽出等に今後役立つ考えである。また、現在の実用化されている技術や研究が進められている技術の水準では実現が困難であると見込まれるものであっても、将来的な利活用の可能性を展望して記載している。さらに、実用化に当たっては、経済的なコスト等を勘案することとなる点に留意が必要である。

居住に関するユースケース（概要）

AIシステム相互間のネットワークが形成されることにより、行動に合わせた家事の自動化、嗜好等に合ったレシピの提案・自動調理や住宅マッチングなど、それぞれのAIシステムが単独で機能する場合に比べて、**快適な居住環境を作ることができる**ようになる。

さらに、AIシステム相互間のネットワークが、医療、スマートハウス、見守り、行政、金融などの領域を横断して形成されることにより、自宅における健康状態の推定、病気の予防・早期発見、病気や怪我の際の遠隔診断、高齢者等の見守りサービスが提供されるとともに、大規模な災害が発生した際の仮設住宅・復興住宅の割当て、保険金の支払いの迅速化なども図られ、領域が横断する前に比べて、**災害時を含めて安心して安全に生活できる環境が整備される**ようになる。

【領域横断前の段階】（領域内でのAIシステム相互間のネットワークが形成され、領域内における利用者の便益が飛躍的に増大）

＜シナリオ上想定される利活用（主なものの例示）＞

- ①位置情報や行動履歴等から帰宅時間を予測し、帰宅時間に合わせて、室内温度の調整、入浴の準備などを行うとともに、掃除ロボットや洗濯ロボットなどが家事を自動で行う。
- ②行動履歴等から把握した嗜好や健康状態を考慮したレシピを提案し、冷蔵庫等の在庫状況をもとに食材を自動発注し、食材が届いた後に料理ロボットが調理する。
- ③趣味・嗜好や家族構成、経済状況等に応じて、最適な住居地、居住形態（戸建かマンションか、購入か賃貸か等）、間取り等を提案し、マッチングを行う。

【領域横断後の段階】（AIシステム相互間のネットワークが領域横断的に形成され、領域間を連携する高度なサービスが実現）

＜シナリオ上想定される利活用（主なものの例示）＞

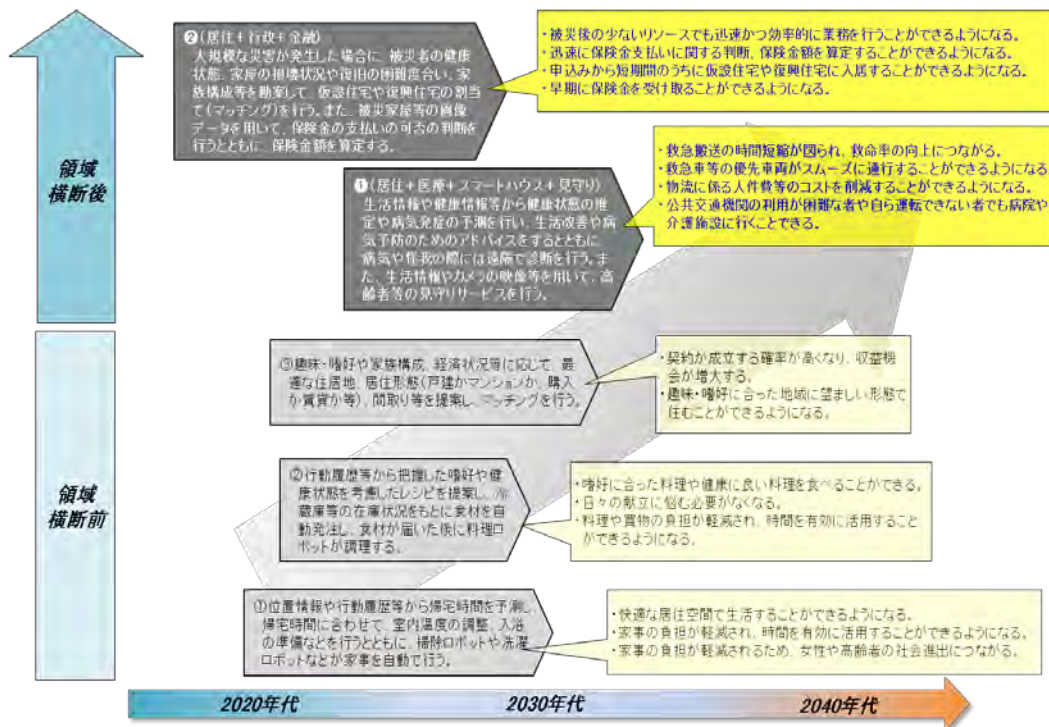
①（居住＋医療＋スマートハウス＋見守り）

生活情報や健康情報等から健康状態の推定や病気発症の予測を行い、生活改善や病気予防のためのアドバイスをするとともに、病気や怪我の際には遠隔で診断を行う。また、生活情報やカメラの映像等を用いて、高齢者等の見守りサービスを行う。

②（居住＋行政＋金融）

大規模な災害が発生した場合に、被災者の健康状態、家屋の損壊状況や復旧の困難度合い、家族構成等を勘案して、仮設住宅や復興住宅の割当てを行う。また、被災家屋等の画像データを用いて、保険金の支払いの可否の判断を行うとともに、保険金額を算定する。

居住に関するユースケース（将来ビジョン）



居住に関するユースケース（評価サマリ）

（注）想定される利活用のうち、主なものについて記載

【領域横断前の段階（サマリ）】

＜インパクト評価＞		＜リスク評価＞			
	シナリオ上想定される利活用	シナリオ上想定されるインパクト(例)	シナリオ上想定されるリスク(例)	リスクが顕在化した場合に想定される被害(例)	リスク評価(例)
①	位置情報や行動履歴等から帰宅時間を予測し、帰宅時間に合わせて、室内温度の調整、入浴の準備などを行うとともに、掃除ロボットや洗濯ロボットなどが家事を自動で行う。	・快適な居住空間で生活することができるようになる。 ・家事の負担が軽減され、時間を有効に活用することができるようになる。 ・家事の負担が軽減されるため、女性や高齢者の社会進出につながる。	・ハッキング等により、ロボットが制御不能に陥るおそれがある。 ・ハッキング等により、パーソナルデータ（位置情報や行動履歴等）が流出し、悪用されるおそれがある。 ・AIシステム間の連携が十分でなく、室温の調整ができなかったり、ロボットが正常に動かないおそれがある。 ・AIシステムに家事を依存しすぎのおそれがある。	・ロボット同士が衝突したり、人間に危害を加えたり、家具を破壊するおそれがある。 ・プライバシーが侵害されたり、犯罪に悪用されるおそれがある。 ・災害発生時等においてAIシステムが利用できなくなった場合、生活に支障を来すおそれがある。	AIシステム間の意思疎通能力の評価 AIシステムへの依存に関する影響の評価 情報流出時の影響の評価 プライバシー侵害の評価
②	行動履歴等から把握した嗜好や健康状態を考慮したレシピを提案し、冷蔵庫等の在庫状況をもとに食材を自動発注し、食材が届いた後に料理ロボットが調理する。	・嗜好に合った料理や健康に良い料理を食べることができる。 ・日々の献立に悩む必要がなくなる。 ・料理や買物の負担が軽減され、時間を有効に活用することができるようになる。	・誤ったデータや偏ったデータに基づく学習や学習不足の結果などにより、適切な提案・マッチングができないおそれがある。 ・ハッキング等により、センシティブな情報（年収や預貯金額等）が流出し、悪用されるおそれがある。 ・センシティブな情報（年収や預貯金額等）が本人同意なく、第三者に提供、共有されるおそれがある。	・注文の取消しや再注文など余計な手間がかかったり、食事までに余計な時間がかかるおそれがある。 ・プライバシーが侵害されたり、犯罪に悪用されるおそれがある。	情報の発生確率等の評価
③	趣味・嗜好や家族構成、経済状況等に応じて、最適な居住地、居住形態（戸建かマンションか、購入か賃貸か等）、間取り等を提案し、マッチングを行う。	・契約が成立する確率が高くなり、収益機会が増大する。 ・趣味・嗜好に合った地域に望ましい形態で住むことができるようになる。	・誤ったデータや偏ったデータに基づく学習や学習不足の結果などにより、適切な提案・マッチングができないおそれがある。 ・ハッキング等により、センシティブな情報（年収や預貯金額等）が流出し、悪用されるおそれがある。 ・センシティブな情報（年収や預貯金額等）が本人同意なく、第三者に提供、共有されるおそれがある。	・適切な提案・マッチングができないことにより、収益機会が増大が見込まれなくなるおそれがある。 ・プライバシーが侵害されたり、犯罪に悪用されるおそれがある。	犯罪の発生確率等の評価
リスク管理(例)		リスクコミュニケーション(例)			
・機微情報の取得又は活用に関する本人同意の確保、名寄せの制限の徹底 ・複数のAIシステム間の連携に関する仕組みの確立（インターフェースの標準化やプロトコルの変換等）		・プライバシーポリシーなどの公表や周知啓発、社会的受容性の確認 ・リスク認知の差異や情報の非対称性に留意したリスク情報に関するコミュニケーションや対話			
雇用・働き方への影響(例)		その他留意すべき点(例)			
・学習データやパラメータ調整、情報セキュリティ等に関連する雇用が創出される。 ・家電メーカーやロボットメーカーの売上が増加し、雇用が創出される。 ・不動産事業者等の売上が増加し、雇用が創出される。ただし、全体として人口減少の傾向にあることに留意が必要である。		・複数のエージェント（業者）がいる場合、調整方法等を決めておく必要がある。 ・AIシステムが生成したレシピの知的財産制度上の取扱いが問題となるおそれがある（特に2次利用する場合）。 ・生活者の嗜好に応じてレシピを提案すると、いつも特定のレシピが採用されることが多くなり、これを学習することにより、ますます同様のレシピが提案されることとなる可能性がある。			

【領域横断後の段階（サマリ）】

＜インパクト評価＞		＜リスク評価＞			
	シナリオ上想定される利活用	シナリオ上想定されるインパクト(例)	シナリオ上想定されるリスク(例)	リスクが顕在化した場合に想定される被害(例)	リスク評価(例)
①	（居住＋医療＋スマートハウス＋見守り） 生活情報や健康情報等から健康状態の推定や病気発症の予測を行い、生活改善や病気予防のためのアドバイスをするとともに、病気の怪我の際には遠隔で診断を行う。また、生活情報やカメラの映像等を用いて、高齢者等の見守りサービスを行う。	・健康状態を把握することができ、健康管理、病気の予防・早期発見に役立つ。 ・自宅に居ながら、かかりつけ医や専門医の診断・診療を受けることができるようになる。 ・安心して生活することができる。 ・体調が急変した場合でも、迅速に対応してもらえる。	・ハッキング等により、個人情報（健康状態やカメラの映像等）が流出し、悪用されるおそれがある。 ・AIシステム間の連携が十分でなく、適切なアドバイスを行うことや遠隔診断を行うことができないおそれがある。 ・個人情報（健康状態やカメラの映像等）が本人同意なく、第三者に提供、共有されるおそれがある。	・病気の発症を見逃したり、重篤化を招くおそれがある。 ・プライバシーが侵害されたり、犯罪に悪用されるおそれがある。	AIシステムによる誤った判断が及ぼす影響範囲の評価 AIシステムへの信頼に対する影響の評価
②	（居住＋行政＋金融） 大規模な災害が発生した場合に、被災者の健康状態、家屋の損壊状況や復旧の困難度合い、家族構成等を勘案して、仮設住宅や復興住宅の割当て（マッチング）を行う。また、被災家屋等の画像データを用いて、保険金の支払いの可否の判断を行うとともに、保険金額を算定する。	・被災後の少ないリソースでも迅速かつ効率的に業務を行うことができるようになる。 ・迅速に保険金支払いに関する判断、保険金額を算定することができるようになる。 ・申込みから短期間のうちに仮設住宅や復興住宅に入居することができるようになる。 ・早期に保険金を受け取ることができるようになる。	・誤ったデータや偏ったデータに基づく学習や学習不足の結果などにより、適切な割当てができないおそれがある。 ・画像認識の精度等が十分でない場合、適切な保険金支払いに関する判断、保険金額の算定ができないおそれがある。 ・通信の遅延や情報通信ネットワークの遮断により、AIシステムが正常に機能しないおそれがある。	・優先的に入居すべき者が入居できなかったり、入居の必要性が高くない者が入居するなどのミスマッチが生ずるおそれがある。 ・本来支払う必要のない保険金を支払うおそれがある。 ・早期に仮設住宅や復興住宅に入居することができないおそれがある。 ・受け取ることができるはずの保険金を受け取ることができなかったり、受け取るまで期間が長くなるおそれがある。	AIシステム間の意思疎通能力の評価 通信環境がAIシステムの動作に与える影響の評価
リスク管理(例)		リスクコミュニケーション(例)			
・専門家による学習データのチューニング等による精度向上、継続的なモニタリングの実施 ・AIシステムによる予測の精度の乖離を許容する範囲の設定		・AIシステムの利用に関する関係者間における双方向的な対話の場の形成 ・利用者に対するインシデントに関する情報共有と対応策の説明・共有			
雇用・働き方への影響(例)		その他留意すべき点(例)			
・学習データやパラメータ調整、情報セキュリティ等に関連する雇用が創出される。 ・高齢者の健康維持などに役立つため、高齢者の社会進出が促進される。 ・災害時において、住宅の割当てに関連する業務の職員を他の人手が足りていない業務に振り分けることが可能となる。		・特に高齢者に有益な利活用と考えられるため、高齢者のAIシステムに関するリテラシーの向上が重要である。 ・災害が少ない地域では利活用されないおそれがある。			

③ シナリオ分析から得られた示唆

分野別評価におけるシナリオ分析から、次のような示唆を得ることができた⁹⁰。

- 様々な領域横断的・融合的なA Iシステムの利活用が可能となり、A Iシステム相互間の連携におけるインパクトに加えて、社会の様々な分野に広範にわたって大きなインパクトがもたらされる。

＜領域横断・融合の例＞

- 行政+スマートシティ+居住、行政+仕事+豊かな人生
- 移動+医療・介護+物流+小売、
移動+教育+仕事+医療・介護+物流+小売など
- 居住+医療+スマートハウス+見守り、居住+行政+金融
- 各ユースケースにおいて、次に掲げる例のように共通して想定されるリスクが見受けられることから、それらのリスクへの対処が重要である。

＜共通して想定されるリスクの例＞

- ハッキングや不正操作等により、A Iシステムが機能不全に陥り、事故が発生するおそれがある。また、事故が発生した場合、A Iシステムがどのような判断をしたのか検証できないと、責任の所在を明らかにすることができないおそれがある。
- A Iシステム間の連携が十分でなく、最適なマッチングや調整ができなかったり、適切なサービスを受けることができなかったりするおそれがある。
- 個人情報やパーソナルデータが、ハッキング等により流出したり、本人同意なく第三者に提供、共有されるおそれがある。
- 各ユースケースにおいて、次に掲げる例のようにA Iネットワーク化の健全な進展に向けた今後の検討に当たって留意すべき点が抽出された。

＜今後の検討に当たって留意すべき点の例＞

- 自動運転により取得される移動履歴やカメラ映像等について、プライバシーに配慮しつつ、データの積極的な利活用が期待される。
- プロファイリング結果により、就職や転職、結婚などにおいて、不当に不利な立場に陥ることのないようにセーフティネットの検討が必要である。
- 特に高齢者に有益なA Iシステムの利活用と考えられるため、高齢者のA Iシステムに関するリテラシーを向上させる方策が求められる。

⁹⁰ ここでは、特に領域横断的・融合的な観点から得られた示唆について記載している。

2. AIネットワーク化の進展に伴い形成されるエコシステムの展望

(1) 背景・経緯

① AIの利活用に関する指針の策定に向けた検討

本推進会議において、AIの研究開発に当たって国際的に共有されるべき「AI開発原則（仮称）」及びその解説からなる国際的な指針たる「AI開発ガイドライン（仮称）」の策定に向けて、G7やOECD等における国際的な議論の用に供するため、AI開発ガイドライン案を整理した⁹¹。

また、報告書2017において、AIシステムが、特段の措置が講じられているものを除いて、学習等により、利活用の過程を通じて出力やプログラムが将来に向けて継続的に変化することが想定されるものであることやネットワークに接続されたAIシステムの便益及びリスクが、空間を越えて、特に国境を越えて、即座に広く波及することがあり得ること等を踏まえて、AIネットワーク化の健全な進展を通じたAIシステムの便益の増進及びリスクの抑制について、「AI開発ガイドライン（仮称）」を国際的に共有することに加えて、利用者がそれぞれの属性等に応じて留意することが期待される事項について国際的に共有を図ることが必要である旨が提言されている。

さらに、このAIの利活用に関するガバナンスの在り方に関する検討に当たっては、その前提として、AIネットワーク化が社会・経済にもたらすインパクト及びリスクに関する問題の所在を把握することに加えて、AIネットワーク化の進展を通じて形成される関係者間のエコシステムに関する展望を得ることが必要である旨が提言されている⁹²。

② 本推進会議における検討

これらの提言を踏まえて、本推進会議において、AIの利活用において留意することが期待される事項等について検討を行うとともに、その前提となるAIネットワーク化の進展に伴い形成されるエコシステムの展望（以下「エコシステムの展望」という。）に関する検討を行ってきた。

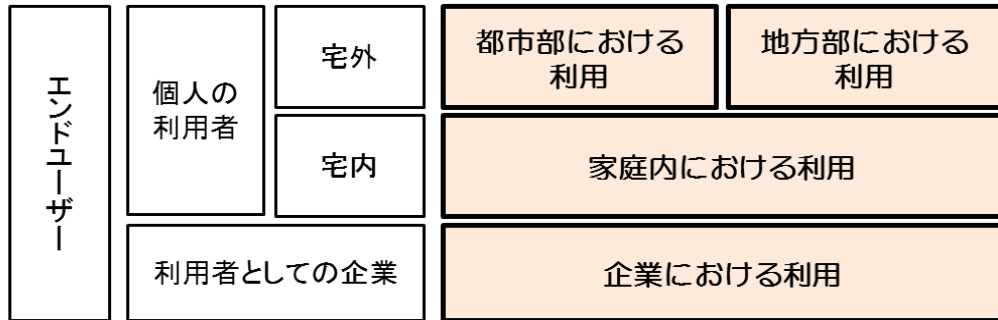
本推進会議におけるエコシステムの展望に関する検討に当たっては、

- ・ 主に利用者の視点からAIの利活用に着目し、次に示すようにAIの利用場面を4つに分類した上で、それぞれの分類ごとにAIの利用シーンを展望し、

⁹¹ 報告書2017別紙1参照。

⁹² 報告書2017第4章（2）参照。

＜AIの利用シーンの分類＞



- その利用シーンをもとに、次に掲げる5つの事例に関するケーススタディを行い、AIの利活用による便益及び課題を整理した。⁹³

- ケース①：完全自動運転
- ケース②：医療診断
- ケース③：人事評価・採用
- ケース④：金融（保険）
- ケース⑤：行政

（2）AIネットワーク化の進展に伴い形成されるエコシステムの展望

① AIの利用シーン

AIの利用場面を4つに分類（都市部における利用、地方部における利用、家庭内における利用、企業における利用）した上で、次に掲げるように、それぞれの分類において4つの主な利用シーンを想定してAIの利活用を展望した⁹⁴。

【利用場面】

【利用シーン】

- 都市部における利用：移動、介護、観光・旅行、教育・人材育成
- 地方部における利用：移動、医療、仕事、コミュニティ
- 家庭内における利用：医療、家事、安全・快適な居住環境
- 企業における利用：ヒト、モノ、カネ、情報

なお、時間軸を意識して、それぞれの利活用の実現時期を、「既に実用化されているもの・近い将来実現しそうなもの」、「中期的に実現しそうなもの」、「遠い将来に実現しそうなもの」の3段階に分けている。

⁹³ 本報告書の公表後、本推進会議において、これらの成果を踏まえて、AIネットワーク化の進展に伴い形成されるデータの流通に関するエコシステムやAIネットワークを通じて提供されるサービス等に関するエコシステムの展望についての検討を行う予定である。

⁹⁴ それぞれの主な利用シーンについては、分類された利用場面における利活用に限られるものではなく、また、重複するものもあるが、想定される特徴的なAIの利活用を記載するため便宜的に分けている。

これらの利用の場面について、次に掲げる都市部における利活用の例のように、A Iの利活用を展望した。それぞれの利用の場面におけるA Iの利活用の展望の詳細は、別紙2-1のとおりである⁹⁵。

都市部におけるAIの利用シーン（例）



② ケーススタディ

上記①のA Iの利用シーン等をもとに、5つの事例（完全自動運転、医療診断、人事評価・採用、金融（保険）、行政）に関するケーススタディを行い、A Iの利活用による便益及び課題を整理した⁹⁶。

例えば、完全自動運転については、次に掲げるような便益が期待される。

⁹⁵ 各利用シーンに記載されているA Iの利活用については、現在の法制度等を必ずしも前提とはせず、将来的な利活用の可能性を展望して記載している。このため、現在の法制度等の下では実現が困難な利活用も含まれており、そのような利活用に関し制度的な課題の抽出等に今後役立つ考えである。また、現在の実用化されている技術や研究が進められている技術の水準では実現が困難であると見込まれるものであっても、将来的な利活用の可能性を展望して記載している。さらに、実用化に当たっては、経済的なコスト等を勘案することとなる点に留意が必要である。

⁹⁶ 便益、課題のいずれについても、あくまでも典型的に想定され得るものの一部を例示として記載しているだけであり、網羅的なものではないことに留意することが必要である。

- ・ 人間は運転する必要がなく、自動車での移動において、移動時間を有効に活用することができるようになる。
- ・ 高齢者や障害者の方にとって、手軽に移動することができる手段が確保されることとなり、病院や買物などに容易に出かけることができるようになる。
- ・ 深夜や早朝などにおける長距離トラックや長距離バスの運転をする必要がなく、働き方やワーク・ライフ・バランスを見直すことなどができるようになる。
- ・ 特に地方部などにおける路線バスの運転手不足などの問題を改善することができ、路線の廃止・縮小を回避することができるようになる。

このような便益が期待される一方で、次に掲げるような課題が想定される。

- ・ 自動運転車間で交渉・調整が成立しない場合、車線変更や合流等が適切に行うことができないおそれがある。[連携]
- ・ AIシステムがハッキング等された場合、そのAIシステムを実装している自動運転車が正常に機能しなくなるだけでなく、ネットワークを介して、次々と他の自動運転車にも影響が及び、事故や交通の混乱が生ずるおそれがある。[連携、セキュリティ、安全]
- ・ 長期間AIソフトのアップデート等をしなかった場合や学習したことがない(想定できない)場面に遭遇した場合、自動運転車が周囲の状況を正しく認識することができず、障害物等を誤って認識してしまい、事故を起こすおそれがある。[セキュリティ、安全]
- ・ 自動運転車が事故を起こしてしまった場合、AIがどのように判断をしたのか検証できないと、責任の所在を明らかにできず、適切な責任分配がなされないおそれがある。[安全、ブラックボックス化]

完全自動運転の例のように、医療診断、人事評価・採用、金融（保険）及び行政に関し、それぞれのケーススタディにおいて便益及び課題を整理した。各ケーススタディの詳細は、別紙2-2のとおりである⁹⁷。

それぞれのケーススタディにおいて示された課題について、次章におけるAIの利活用において留意することが期待される事項の検討に資するよう、次表のように類型を整理した上で、それぞれの課題が、主にどの類型に属するものであるかを示した⁹⁸。

⁹⁷ 各ケーススタディで想定されているAIの利活用については、現在の法制度等を必ずしも前提とはせず、将来的な利活用の可能性を展望して記載している。このため、現在の法制度等の下では実現が困難な利活用も含まれており、そのような利活用に関し制度的な課題の抽出等に今後役立つ考えである。

⁹⁸ 表の右側に示す凡例を各課題に対応するように付して、どの類型に属する課題であるか紐付

また、このほか、A I の利活用を促進するため、あるいは、A I ネットワーク化の健全な進展を図るためには、教育・人材育成の在り方、事故等による被害者の救済の在り方、A I の利活用を促進するための法制度等の在り方などについても検討が必要であるものと考えられる。

【主として生命・身体の安全、権利・利益等を守るための課題】

<ul style="list-style-type: none"> 生命・身体・財産の安全に関する課題 (事故の防止など) 	: 安全
<ul style="list-style-type: none"> A I による判断の正当性や公平性に関する課題 (差別、生命倫理との関係など) 	: 正当性・公平性
<ul style="list-style-type: none"> プライバシーに関する課題 (プライバシーの尊重、プロファイリングなど) 	: プライバシー

【主として人間とA I との関係等に関する課題】

<ul style="list-style-type: none"> 人間とA I との役割分担等に関する課題 (人間の判断の介在、関係者間の協力など) 	: 役割分担
<ul style="list-style-type: none"> A I に対する受容性に関する課題 (利用者に対する説明責任など) 	: 受容性

【主として技術的な観点からの解決が求められる課題】⁹⁹

<ul style="list-style-type: none"> A I の判断のブラックボックス化に関する課題 (事故が発生した場合の原因究明など) 	: ブラックボックス化
<ul style="list-style-type: none"> セキュリティに関する課題 (ハッキング対策など) 	: セキュリティ
<ul style="list-style-type: none"> A I 間の連携に関する課題 (A I 間の交渉・調整など) 	: 連携

【主としてデータに関する課題】

<ul style="list-style-type: none"> A I が学習するデータに関する課題 (データの正確性など) 	: データ
--	-------

(注) 右側は凡例

くようにして示している。

⁹⁹ それぞれの課題の解決に当たっては、基本的に社会的な観点、技術的な観点の両面からのアプローチが必要であるものと考えられるが、その中でも特に技術的な観点からのアプローチが求められるものを分類している。

3. AIネットワーク化の進展に関する考察

上記1. 分野別評価や上記2. エコシステムの展望に関する検討において、様々なAIの利活用を展望した。これらのAIの利活用を踏まえて、ネットワーク化の観点からの考察（主としてAIネットワーク化の進展において想定される課題（AIがネットワーク化することにより生ずる又は増幅する課題）の検討）を行った。

(1) AIネットワーク化による便益の増進

AIネットワーク化の進展により、様々な分野において多大な便益がもたらされることについては、分野別評価やエコシステムの展望（特に②ケーススタディ）で示したとおりである¹⁰⁰。

また、報告書2017においても、先行的評価におけるシナリオ分析から得られた示唆として、同じ領域内においても、様々なAIシステム相互間の連携が可能となることにより、AIシステム相互間の連携前におけるインパクトに加えて、特に連携に係るAIシステムを利活用する業務全体を通じたリアルタイムでの最適化など更に大きなインパクトがもたらされることが述べられている¹⁰¹。

(2) AIネットワーク化の進展において想定される課題（ネットワーク化の観点から）

AIネットワーク化の特徴として、次に掲げるようなことが挙げられる。

- ・ 様々な事業者間における多種多様なAIの接続（複雑な連携）
- ・ 様々な事業者の多種多様なAIを取りまとめる基盤的なAIの構築（情報や権限の集中）
- ・ 様々な事業者間における多種多様なAIによる情報の連携（情報の共有・拡散）

このような特徴を活かして、新しいサービスの開発・提供、最適化や効率化、コスト削減が可能となるなど多大な便益がもたらされる一方で、ネットワークという観点に着目した場合、次のような課題（AIネットワーク化の健全な進展を阻害し得る要因）¹⁰²

¹⁰⁰ 分野別評価の詳細については、別紙1参照。ケーススタディの詳細については、別紙2-2参照。

¹⁰¹ 先行的評価の結果については、報告書2017第3章2.及び別紙3参照。

¹⁰² 主にAIネットワーク化の進展段階の第二段階（複数のAIシステム相互間のネットワークが形成され、ネットワーク上のAIシステムが相互に連携して協調する段階）を想定している（AIネットワーク化の進展段階については、序章1.参照。）。また、ここで掲げる課題については、AIシステムが学習等により自らの出力やプログラムを変化させたり、自律性を有することなどが想定されることを踏まえ、従来のICTシステムのネットワーク化と比べて、AIネットワーク化によりリスクが新たに生じたり増幅する可能性があるため、ネットワーク化の観点から想定される課題として整理している。

なお、本章2.（2）② ケーススタディにおいて整理する課題についても、ネットワーク化により生ずる（増幅する）課題も含まれている場合があり、ここで整理する課題と重複するものもある。

に留意すべきであるものと考えられる。

- ・ 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
- ・ AIシステム間の連携・調整が成立しないなどのおそれ
- ・ AIの判断・意思決定を検証できないおそれ（ネットワーク化により、システム全体としてブラックボックス化するおそれ）
- ・ 少数のAIの影響力が強くなりすぎるなどのおそれ（少数のAI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）
- ・ 領域横断での情報の共有と特定の基盤的なAIへの情報の集中によるプライバシー侵害のおそれ
- ・ AIが想定外の動作を行うなどのおそれ

例えば、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については、

- AIシステムが相互に連携して、協調・調整することになるため、例えば、ある事業者において、ハッキングやシステム障害等により一部のAIシステムが正常に機能しなくなった場合や災害等により通信が遮断された場合、あるいは、ある一部のAIシステムの判断・予測等に誤りが生じた場合や虚偽又は不適切なデータが流通した場合などに、他のAIシステムに次々に影響が波及し、ネットワーク全体が正常に機能しなくなったり、期待された効果（適切な判断・予測、マッチング等）が得られなくなったりするおそれがある。

といった課題が想定され、このような課題が生ずる事例として、次の掲げるようなものが想定される。

- ・ セキュリティ対策等が不十分な事業者が提供しているAIシステムを活用した自動運転車がハッキングされた場合、連携している別のAIシステムが誤作動又は機能不全に陥り、交通事故や交通障害等が発生するおそれがある。[連携、セキュリティ、安全]
- ・ サプライチェーン全体として最適化を調整している場合、ある事業者のAIシステムが生産量等を誤って予測した場合や虚偽のデータを流通させた場合、全体としての最適化が実現できない（在庫不足／過剰在庫、原材料不足等）おそれがある。[連携、データ]
- ・ 被災地の緊急救命において、救急車や医療機関等の間でリソース配分の最適化の調整を行う場合、ある医療機関のAIシステムが通信の遮断により機能不全に陥ると、適切な情報共有、リソース配分が実現されず、適切な救命活動ができなくなるおそれがある。[連携、安全]

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の例のように、上記に掲げられているAIネットワーク化の進展において想定される課題（AIネットワー

ク化の健全な進展を阻害し得る要因) について、その事例等を整理した。それぞれの課題に関する整理の詳細は、別紙 2-3 のとおりである。

なお、ここで掲げられる課題についても、上記 2. (2) ② ケーススタディで掲げられた課題と同様に、次章における A I の利活用において留意することが期待される事項の検討に資するよう、それぞれの課題が、主にどの類型に属するものであるかを示した¹⁰³。

4. 教育・人材育成

(1) 背景・経緯

報告書 2016 において、A I ネットワーク化の進展を通じて異業種間の融合による新たな商品・サービス、ビジネスモデルが創出されるようになるものと展望されるが、そのような変化は、そこで働く人々の働き口、働き方、働く内容を左右し得ることとなり、ひいては、人間の生き方にも影響を及ぼし得るものと考えられると指摘されている¹⁰⁴。

また、報告書 2017 において、「失われる雇用から新しく創出される雇用への円滑な移行、新しく創出される雇用への適応のための教育や人材育成が重要となるものと考えられる。この点については、社会に出る(就職する)までの教育、社会に出た(就職した)後の再教育・人材育成のいずれもが重要であり、学校教育、企業内の人材育成(OJT(On the Job Training))とともに従来の発想や方法からの転換が求められる。また、今後、自営的就労¹⁰⁵の可能性が高まることに鑑みると、個々人が自ら技能習得しなければならなくなるため、自助・自学のための環境整備を進めることも重要である」と指摘されている¹⁰⁶。

(2) 本推進会議における検討

これらの指摘を踏まえて、本推進会議において、A I ネットワーク社会に向けた教育・人材育成の在り方に関する検討を行うとともに、A I の利活用により教育・人材育成にどのような影響をもたらすのかという観点から検討を行った。

教育・人材育成に関する検討に当たっては、一般的に、人間が行った方がよいことと A I に任せた方がよいことなどを仕分けした上で、人間が身に付けるべき能力やスキルは何か、そのような能力やスキルを身に付けるためにどのような教育や人材育成を行うべきか等について検討するというアプローチをとることが考えられる。

しかしながら、本推進会議においては、人間と A I との役割分担を考える際に、仮説を作り、その仮説に基づいて議論を行うよりも、複合的かつ複線的に議論を行うこ

¹⁰³ 課題の類型については、本章 2. (2) ② ケーススタディと同様である(34-35 頁参照)。

¹⁰⁴ 報告書 2016 第 2 章 2. 参照。

¹⁰⁵ 場所的・時間的拘束が少なく、自分で仕事のオーガナイズができる働き方のこと。詳細は、報告書 2016 第 2 章 2. 参照。

¹⁰⁶ 報告書 2017 第 3 章 4. 参照。

とが重要であるとの認識のもと、本報告書においては、構成員の意見を列挙する形で検討状況を示すこととする。

なお、構成員の意見については、次に掲げる6つのテーマに分類して示すこととする¹⁰⁷。

- | | |
|-------------------|------------|
| ① 人間とA Iとの関係・役割分担 | ④ 職業スキルの習得 |
| ② 社会の変化及び人間の技能・能力 | ⑤ 生涯教育 |
| ③ 学校教育 | ⑥ A Iの利活用 |

ここでは、それぞれのテーマにおける主な意見の概要を示すこととする。なお、主な意見の詳細は、別紙3のとおりである。

① 人間とA Iとの関係・役割分担

人間とA Iの役割分担については、人間中心ということ意識する必要があるといった意見や価値判断やコミュニケーション、創造性をポイントとして挙げる意見などがあつた。

② 社会の変化及び人間の能力・技能

人間社会の変化を考える上では、人間自身がA Iとのインタラクションを通して変わっていくことを1つの軸として考える必要があるといった意見や人間に求められる能力として、教養、コミュニケーション能力、数学、哲学を挙げる意見などがあつた。

③ 学校教育

現状の学教教育に対して強い危機感を示す意見や「自分で学ぶ」という意識改革が重要であるといった意見などがあつた。

④ 職業スキルの習得

職業スキルを身につける上でも、「自ら学ぶ」、「学び続ける」ことが重要であるといった意見や今後は企業内での人材育成が難しくなるといった意見などがあつた。

⑤ 生涯教育

リカレント教育が重要であるといった意見やリカレント教育を含め全体をデザインする必要があり、早く議論を始めるべきであるといった意見などがあつた。

⑥ A Iの利活用

A Iを用いて個人に応じた教育パターンを作ることができるのではないかとといった意見や学ぶ時間や自己実現のための時間を確保するためにA Iを利活用できるとよいといった意見などがあつた。

¹⁰⁷ 複数のテーマに跨がる意見もあるが、最も関係が深いと考えられるテーマに分類している。

教育・人材育成については、本推進会議において継続的に検討を行っていく予定であり、本報告書における取りまとめは、今後の検討に当たって、キックオフ的なものと位置付けられる。本報告書の公表後、これらの意見を踏まえて、本推進会議において、さらに検討を深めることとする。

5. AIネットワーク化の進展に関する指標

(1) 背景・経緯

報告書 2016 において、AI ネットワーク化の進展が産業構造や雇用にもたらす影響を含めて、社会の変化がどの程度進展しているのかを把握することが重要であり、また、AI システムを研究開発する者、サービスを提供する者、利用者等の様々な関係ステークホルダがいる中でも、とりわけ政府が、適切な政策を企画立案し実施する上で、これらの状況を把握することは極めて重要であるという問題意識のもと、AI ネットワーク化の進展が社会にもたらす影響を評価するための指標の設定に向けた検討の方向性を検討した¹⁰⁸。

また、報告書 2017 において、評価指標の設定について、国際比較にも用いることができるような定量的に測定し、又は客観的に把握することができる指標の設定に向けて継続的に検討することが必要である旨が提言されている¹⁰⁹。

(2) 本推進会議における検討

① 検討の目的及びスコープ

これらのことを踏まえ、本推進会議において、特に政府が、政策を企画・立案するとともに、関係するステークホルダが、AI の開発及び利活用の促進等に向けた環境整備に取り組むに当たり、社会の変化の状況等を把握するために参考となる指標を策定すること、また、国際的に共有され、適切に国際比較¹¹⁰を行うことができる指標の策定に向けた国際的な議論に貢献することを目的として検討を行った。

検討に当たっては、次に掲げる指標を対象とし、また、国際比較に用いることができる指標の策定を視野に入れ、当面は、G7 各国を対象に比較することができる指標の策定に向けて検討を行った¹¹¹。

¹⁰⁸ 報告書 2016 第 3 章参照。

¹⁰⁹ 報告書 2017 第 4 章 (11) 参照。

¹¹⁰ 国全体としてはインフラや各指標が低い場合であっても、一部の層・地域が先進的な取組を行っているような国については、メディアに取り上げられるなど注目されやすいといったギャップが生じ得ることに留意が必要である。

¹¹¹ G7 各国を対象とする指標の策定を検討した後、G20 各国、OECD 加盟国などに対象の範囲を拡大することを検討する。

- ・ AI ネットワーク化の進展の度合いに関する指標
- ・ AI ネットワーク化が社会・経済にもたらす影響（インパクト及びリスク）¹¹²の度合いに関する指標

② 検討の方向性

この2つの指標の検討に当たっては、次表に示す方向性に基づいて検討を行うこととした。ただし、AI ネットワーク化が社会・経済にもたらす影響（インパクト及びリスク）の度合いについては、検討を進める中で、何が要因となってインパクトやリスクとなったのか、その因果関係を把握すること及びその影響の度合いを算出（定量化）することが極めて困難であるとの意見等に基づき、当面は、AI ネットワーク化の進展の度合いに関する指標について検討を行うこととした。

検討の対象		検討の方向性			
AIネットワーク化の進展の度合いに関する指標		<ul style="list-style-type: none"> ○ AIネットワーク化の進展段階を踏まえ、進展段階に応じた指標を検討 <ul style="list-style-type: none"> ・ AIネットワーク化の前提となる環境整備に関する指標【事前】 ・ AIネットワーク化の初期段階に関する指標【第1段階】 ・ AIネットワーク化の進展に関する指標【第2段階以降】 ○ 国際比較の対象として有用な項目を検討 ○ 必要に応じて、アンケート調査等により新しいデータを収集 			
AIネットワーク化が社会・経済にもたらす影響（インパクト及びリスク）の度合いに関する指標	インパクトに関する指標	<ul style="list-style-type: none"> ○ 国内外の関連する検討・調査を参考にしつつ、国際比較に用いることが可能か（適しているか）などの観点から踏まえ、有用な項目の検討及び関連する統計の調査・整理 			
	リスクに関する指標				
	<table border="1"> <tr> <td>産業</td> </tr> <tr> <td>個人</td> </tr> <tr> <td>公共</td> </tr> </table>	産業	個人	公共	
産業					
個人					
公共					
		<ul style="list-style-type: none"> ○ ミスリードにならないよう分野共通的な指標の在り方（定性的なものを含む。）について検討 ○ 分野別評価やAIネットワーク化の進展に伴い形成されるエコシステムの展望に関する検討において抽出された課題等を参考にしつつ検討 			

AI ネットワーク化の進展の度合いに関する指標については、AI ネットワーク化の進展段階¹¹³を踏まえ、それぞれの進展段階において指標としての適切な項目が異なるものと考えられることから、次に示すように、進展段階に応じた指標について検討を行うこととした。

- ・ 事前の段階 : AI ネットワーク化の前提となる環境整備に関するもの
 - AI ネットワーク化事前指標

¹¹² インパクトについては、個々の利用者が享受する利益としての「便益」のみならず、便益がもたらされる結果として社会全体にもたらされる価値としての「恵沢」を含むものとする。なお、例えば、社会における失業の増加、平均所得の低下等は、負の影響（ネガティブインパクト）とする。また、リスクについては、リスクが顕在化した場合に想定される被害を含むものとする。

¹¹³ 序章1. 参照。

- ・ 第一段階 : A I ネットワーク化の初期段階に関するもの
 - A I ネットワーク化初期指標
- ・ 第二段階以降 : A I ネットワーク化の進展に関するもの
 - A I ネットワーク化本指標¹¹⁴

③ A I ネットワーク化の進展の度合いに関する指標

A I ネットワーク化事前指標、A I ネットワーク化初期指標及びA I ネットワーク化本指標について、現時点においては、「第一段階」が進みつつある状況であること、「第二段階」以降の取組事例がないためA I ネットワーク化本指標に関する有用な統計データがないこと等を踏まえ、当面はA I ネットワーク化事前指標及びA I ネットワーク化初期指標について検討することとした。

A I ネットワーク化事前指標及びA I ネットワーク化初期指標については、次表に示すように既存統計の有無等を調査し、その状況に基づき今後の取組方針を定めた。

AI-NW化の進展段階	指標の類型	指標の概要	既存統計の有無	検討方針(案)
第1段階	AIネットワーク化事前指標	AIネットワーク化の前提となる環境整備に関する指標	<ul style="list-style-type: none"> ○ ITU、OECD等がICTに関する国際統計を作成 ○ 英オックスフォード大学が、公的部門に関する「政府AI準備指標」を作成 (注) AIに直接関連する指標は少ない。 	<ul style="list-style-type: none"> ○ 指標として有用な項目の洗い出し・整理 ○ アンケート調査における有用な項目の検討
	AIネットワーク化初期指標	AIネットワーク化の初期段階に関する指標	<ul style="list-style-type: none"> ○ AIに関する国際比較が可能な統計は極めて少ない ○ 米スタンフォード大学が、AIに関する新たな統計「AI Index2017」を作成 (注) 既存の統計では指標の設定は困難。統計を作成する等の方法はある。 	<ul style="list-style-type: none"> ○ 指標として有用な項目の洗い出し・整理 ○ アンケート調査における有用な項目の検討
第2段階以降	AIネットワーク化本指標	AIネットワーク化の進展に関する指標	<ul style="list-style-type: none"> ○ 既存の統計はない (注) 現時点では、新規の統計の作成は困難。 	<ul style="list-style-type: none"> ○ エコシステムの展望に関する検討等を踏まえ、指標として必要な項目を検討

A I ネットワーク化事前指標及びA I ネットワーク化初期指標について、既存統計の有無等に関する状況は、それぞれ次頁及び次々頁に示すとおりであり、今後の取組方針として、いずれも指標として有用な項目の洗い出し・整理、アンケート調査に必要な項目の検討を行うこととしている¹¹⁵。

¹¹⁴ 今後の検討によっては、進展段階に応じて細分化することもあり得るものと考えられる。

¹¹⁵ 今後の検討に当たって、GPU (Graphics Processing Unit) の出荷台数、量子コンピュータの開発状況、SDN (Software Defined Network) の導入状況などが指標として有用な項目になるのではないかといった意見があった。

AIネットワーク化事前指標については、国際電気通信連合（ITU）やOECD等が国際統計の調査を実施していることを参考に、次表のように有用な項目を整理した上で、主として「対象国」を基準として、項目案を抽出した。

今後は、これらの項目の統計データを収集、正規化するとともに、アンケート調査における有用な項目を検討し、また、必要に応じて、項目ごとの重要性等を考慮しウェイト付けして統合¹¹⁶する作業を進めることが重要である。

項目		選定基準			統計等
上位概念	下位概念	指標	対象国	重要性	
利活用	利用者の基礎能力	ICTスキル	○ (150ヶ国以上)	△	ITU「ICT開発指数(ICTスキル)」(注1)
		企業におけるAI活用意識	× (日、米など 一部の先進国)	○	アイアト「AIに関する意識調査」
		AIに対する社会受容性	× (日、米、独におけるロボットの受容性に関する調査あり)	○	野村総合研究所「ロボット・AI技術の導入をめぐる生活者の受容性と課題」
		政府におけるAI関連の取組	×	○	既存統計なし(「go」ドメインでの「AI」関連のキーワードヒット数を集計)
	アプリケーションレイヤー	クラウドサービス利用率(企業)	× (日、米、加など先進国の一部で欠損値あり)	△	OECD「企業におけるクラウドコンピューティング利用率」
		オンラインでの商品・サービス購入率(個人)	× (日、米、加など先進国の一部で欠損値あり)	△	OECD「個人」のオンラインでの購買率」
インフラの状況	ネットワークレイヤー	インターネット通信速度	○ (150ヶ国以上)	△	ITU「インターネット利用者の通信速度(bits/s)」
		インターネット加入数	○ (150ヶ国以上)	△	ITU「インターネット加入数」
		ブロードバンド契約者数	○ (150ヶ国以上)	△	ITU「100人当たりのブロードバンドの契約数」
		無線ブロードバンド契約者数	○ (150ヶ国以上)	△	ITU「100人当たりの無線ブロードバンド契約数」
	基盤レイヤー	スーパーコンピューター処理能力	○ (ランキングしているか否かで統計作成可能)	○	TOP 500 List
		GPU処理性能・価格	×	○	公開されている統計なし(NVIDIA等の主要メーカーの商品の性能・価格を調査)
	端末レイヤー	携帯電話契約者数	○ (150ヶ国以上)	○	ITU「100人当たりの携帯契約者数」
		PC所有者数(注2)	○ (150ヶ国以上)	×	ITU「PC所有者数」
		IoT競争力	○ (主要先進国に関して統計が整備されている)	○	総務省「IoT国際競争力指標」
		インベション環境	Global Innovation Index	○ (127ヶ国)	△
研究・開発等	研究	人工知能関連の論文数	○ (OECD加盟国以上)	○	Web of Science等の論文検索データベース(独自にデータを取得・集計する必要あり)
	開発者人材	数学リテラシー	○ (OECD加盟国)	○	OECD「国際学習到達度調査(PISA)数学リテラシー」(注3)
	データ流通	データ流通量	×	○	(国ごとに比較可能な統計なし)
	オープンデータ	オープンデータ	○ (OECD加盟国)	○	OECD「OURdata Index」
	プライバシー	プライバシー	○ (主要先進国に関して統計が整備されている)	○	EMC Privacy Index(注4)

(注1)「大人の情報リテラシー」、「中等教育就学率」、「大学及び専門教育就学率」をそれぞれ1/3で重み付けしたもの。
 (注2)今後モバイル端末への移行が想定され、PC所有者数の減少が予想されることに留意が必要。
 (注3)15歳が対象で3年ごとに更新。
 (注4)民間の調査であり毎年更新されるものではない。

※評価基準
 「対象国」
 ○：OECD加盟国以上
 △：OECD加盟国より少ないがG7各国を含む
 ×：G7各国を一部含まない
 「重要性」
 ○：影響が大きい
 △：あまり影響を与えない
 ×：影響が小さい

¹¹⁶ ウェイト付けに当たっては、その方法、根拠について説得性を有することが重要である。

AIネットワーク化初期指標については、国際比較可能な統計がほとんど存在しないことから、次表のように項目を整理したものの、ほとんど有用な項目を抽出することはできなかった。

今後は、アンケート調査における有用な項目を検討することが重要である。

項目		選定基準		統計等	
上位概念	下位概念	指標	対象国		重要性
利活用 の 状況	利用者の基礎能力	ITリテラシーの程度	× (各国ごとに集計)	△	各国ごとにアンケート調査等を実施しているが、国際的な指標はなし
	利用者の意識	AIに対する意識	× (各国ごとに集計)	○	各国の言語で公開情報をもとに自然言語分析が行われている
	アプリケーション レイヤー	サービス提供者におけるAIの導入率	× (日、米、独における企業のAI導入率に関する調査結果あり)	○	MM総研「人工知能技術のビジネス活用概況」(2017年)
		各産業におけるAIの導入率	× (日本における企業のAI導入率に関する調査結果あり)	○	MM総研「人工知能技術のビジネス活用概況」(2017年)
		消費者のAIの利用率 (1週間当たりの利用回数)	-	○	AIの利用率に関する統計はなし(主要ベンダーのホームスピーカーの地域別出荷台数など個別のアプリケーションに関する調査はある)
		AIスタートアップ数	×	△(注1)	スタートアップ関連のデータベースCrunchbaseのAPから自動取得
基盤 レイヤー(注3)	スーパーコンピューター処理能力	○(ランキング入しているか否かで統計作成可能)	○	TOP 500 List	
	GPU処理性能・価格	×	○	公開されている統計なし(NVIDIA等の主要メーカーの商品の性能・価格を調査)	
インフラ の 状況	ネットワーク レイヤー	エッジコンピューティングの普及度合い	×	○	無料の統計は公開されていない(米国の調査会社の有料レポートあり)
		5Gの普及率	×	○	各国で比較可能な統計なし
	端末レイヤー	スマートフォンの普及率	○(40ヵ国以上)	○	米国の調査会社Anaysys Mason「The Connected Consumer Survey」
	ロボットの普及率	○	○	国際ロボット連盟が世界の産業用ロボット稼働台数、統計を公表	
研究・ 開発等	特許	特許出願数	△(既存の調査はあり)	○	
	開発者人材(研究者)	大学におけるAI関連講座の受講者数	×(各国に集計)	○(注2)	
	開発者人材(エンジニア)	AIに関わる資格の取得状況	×	○	資格制度が民間団体によって創設され始めているが、取得者数が公表されていない
	投資	AIスタートアップへの投資額	○	△(注1)	スタートアップ関連のデータベースCrunchbaseのAPから自動取得
	データ流通	データ流通量	×(国ごとに比較可能な統計なし)	○	
		オープンデータ	○(OECD加盟国)	○	OECD's OURdata Index
	プライバシー	○(主要先進国に関して統計が整備されている)	○	EMC Privacy Index(注3)	

(注1) 米国のデータベースであり、非英語圏の国の数値が低くなるおそれがある。

(注2) 大学の規模(入学者数等)に左右され易いと考えられるため、割合(例えば、人口比率)とする考え方もある。

(注3) 民間の調査であり毎年更新されるものではない。

(注4) 基盤レイヤー及びデータ流通については、環境整備の段階と共通するものと考えられるため、事前指標と同一の項目となっている。

※評価基準

「対象国」

○: OECD加盟国以上

△: OECD加盟国より少ないがG7各国を含む

×: G7各国を一部含まない

「重要性」

○: 影響が大きい

△: あまり影響を与えない

×: 影響が小さい

第3章 AIの利活用において留意することが期待される事項

1. 背景・経緯

(1) 背景・経緯

平成28年4月29日及び30日に開催されたG7香川・高松情報通信大臣会合において、高市早苗総務大臣（当時）から、G7各国が中心となり、OECD等国際機関の協力も得て、AIネットワーク化が社会・経済に与える影響やAIの開発原則の策定などAIネットワーク化をめぐる社会的・経済的・倫理的課題に関し、産学民官の関係ステークホルダの参画を得て、国際的な議論を進めることの提案がなされ、各国から賛同が得られた。

同年6月20日に、検討会議の報告書2016が公表され、今後の課題のうち第一の課題として「研究開発の原則・指針の策定」が提言されている。

同年10月31日に、検討会議を発展的に改組した本推進会議が発足し、第1回会合が開催された。本推進会議には、親会の下に開発原則分科会及び影響評価分科会が置かれ¹¹⁷、このうち主に開発原則分科会において、AI開発ガイドライン案の作成に向けた検討を行った。同年12月28日から平成29年1月31日にかけて行ったAI開発ガイドライン案の論点に関する意見募集に寄せられた意見、同年3月13日及び14日に開催されたAIネットワーク社会推進フォーラム（国際シンポジウム）¹¹⁸における議論及び同年6月14日から同年7月7日にかけて行った報告書2017（案）に関する意見募集に寄せられた意見等を踏まえ、同年7月28日に、AI開発ガイドライン案を含む報告書2017が公表された。

報告書2017において、AIシステムが、特段の措置が講じられているものを除いて、学習等により、利活用の過程を通じて出力やプログラムが将来に向けて継続的に変化することが想定されるものであることやネットワークに接続されたAIシステムの便益及びリスクが、空間を越えて、特に国境を越えて、即座に広く波及することがあり得ること等を踏まえて、AIネットワーク化の健全な進展を通じたAIシステムの便益の増進及びリスクの抑制について、「AI開発ガイドライン（仮称）」を国際的に共有することに加えて、利用者がそれぞれの属性等に応じて留意することが期待される事項について国際的に共有を図ることが必要である旨が提言されている。

また、報告書2017において、AIの利活用に関するガバナンスの在り方に関する検討に当たっては、その前提として、AIネットワーク化が社会・経済にもたらすインパクト及びリスクに関する問題の所在を把握することに加えて、AIネットワーク化の進展を通じて形成される関係者間のエコシステムに関する展望を得ることが必要であ

¹¹⁷ 前掲注（4）参照。

¹¹⁸ 報告書2017第1章1.（1）及び参考資料参照。

る旨が提言されている¹¹⁹。

(2) 本推進会議における検討

A I の開発及び利活用の促進やA I ネットワーク化の健全な進展を通じて人間中心の「智連社会」¹²⁰を実現するためには、A I の便益の増進やリスクの抑制、A I に対する理解・信頼の醸成等の取組が必要である。

本推進会議における取組として、A I の開発において留意することが期待される事項としてA I 開発ガイドライン案を取りまとめた。他方、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。

このような問題意識を踏まえ、本推進会議において、A I の利活用において留意することが期待される事項に関する検討を行うとともに、その検討に資するようA I ネットワーク化が進展した社会の将来像を描きA I の利活用における課題等を抽出するために、A I ネットワーク化が社会・経済にもたら影響（分野別評価）やA I ネットワーク化の進展に伴い形成されるエコシステムの展望に関する検討、A I ネットワーク化の進展に関する考察等を行ってきた¹²¹。

なお、これらの事項は密接に関連するものと考えられることから、主として親会の下に置かれた環境整備分科会及び影響評価分科会¹²²の合同分科会において検討を進めてきた。

2. 関係する主体の整理

(1) 基本的な考え方

A I の利活用においては、利用者やデータ提供者など様々な主体の関与が想定されることから、次に掲げる考え方に基づいて関係する主体を整理することとする。

- ・ A I の利活用においては、自らの生活や趣味等にA I を用いる者（例：一般の個人ユーザーなど）、A I を用いてサービスを提供する者（例：医師や金融機関など）と、
いうように利用者には様々な立場の者がいることが想定される¹²³。

¹¹⁹ 報告書 2017 第 4 章（2）参照。

¹²⁰ 序章 1. 及び前掲注（6）参照。

¹²¹ 分野別評価については、前章 1. 参照。A I ネットワーク化の進展に伴い形成されるエコシステムの展望については、同章 2. 参照。

¹²² 前掲注（4）参照。

¹²³ A I 開発ガイドライン案においては、「利用者」とは、“A I システムを利用する者（最終利用者（エンドユーザ）のほか、他者が開発したA I ネットワークサービスを第三者に提供するプロバイダを含む。）をいう。”と定義されているのみであり、利用者の性質や属性等について

また、A Iの利活用に当たっては、A Iシステム又はA Iサービス¹²⁴（以下単に「A I」ということがある。）を直接的に利用しない場合でも、A Iを用いたサービス（例：医療サービスや金融サービスなど）を利用（A Iを間接的に利用）する場合が想定される。

- ・ A Iの利活用においては、開発者がA Iシステムのアップデートや追加的な学習等を行うことが多いという実態を踏まえて、「A Iサービス」を“A Iシステムの機能を提供するサービス”とし、「A I付随サービス」を“A Iシステムのアップデート又は追加的な学習等に係るサービス”とし、これにより開発者も利活用において一定の関与があり得ることを示すこととする¹²⁵。

- ・ これらを踏まえると、A Iの利活用において留意することが期待される事項の検討に当たっては、利用者等について一定の分類を試み、分類された利用者の類型に沿って検討を行うことが適当であるものと考えられる。

なお、A Iの提供及び利活用において、どのように関与しているのかを外形的に観察した上で、利用者等の分類を試みることとする。

- ・ 利用者等の分類について、
 - 特に消費者的利用者及び間接利用者（後述（（2）参照））について、A Iを利活用していることを認識している者と認識していない者とを区別すべきではないか。
 - 消費者的利用者について、A Iシステムを保有・管理している場合やA Iの入出力を調整することができる場合と、このような事情が認められない場合とを区別すべきではないか。
 - 子ども、高齢者、障害者など社会的弱者については、特別に留意することが期待される事項が考えられるのではないか。

といった意見があることを踏まえると、今後「A I利活用原則案」に掲げられた各原則の内容に関し整理された論点¹²⁶を検討する上では、利用者の主観（利用しているシステム又はサービスがA Iを用いたものであることを認識しているか否か）や性質、属性等も考慮する必要があるものと考えられる。

- ・ A Iの提供及び利活用の形態については、様々なケースが想定されることから、必要に応じて、今後の検討において関係する主体の整理を見直すこともあり得る。

言及していない。

¹²⁴ A I開発ガイドライン案においては、「A Iシステム」とは、“A Iソフトを構成要素として含むシステムをいう。例えば、A Iソフトを実装したロボットやクラウドシステムはこれに含まれる。”とされている。また、「A Iサービス」については定義されていない。

¹²⁵ この場合、開発者がA Iシステムのアップデートや追加的な学習等を行うと、開発者はA I付随サービスを提供する「A Iサービスプロバイダ」（後述（（2）参照））と位置付けられることとなる。

¹²⁶ A I利活用原則案及び各原則の内容に関し整理された論点については、後述（本章3．及び4．参照）。

(2) 関係する主体の整理

このような基本的な考え方にに基づきA I の利活用において関与が想定される者を次のように整理した。

- ・ **利用者**
A I システム、A I サービス又はA I 付随サービスを利用する者
- ・ **A I サービスプロバイダ**
A I 利用者のうち業としてA I サービス又はA I 付随サービスを他者に提供する者サービスプロバイダ
- ・ **最終利用者**
利用者のうち業としてA I サービス又はA I 付随サービスを他者に提供することなくA I システム又はA I サービスを利用する者
- ・ **A I ネットワークサービスプロバイダ**
A I サービスプロバイダのうち業として情報通信ネットワークを通じてA I サービス又はA I 付随サービスを他者に提供する者¹²⁷
- ・ **オフラインA I サービスプロバイダ**
A I サービスプロバイダのうち業として情報通信ネットワークを通じることなくA I サービス又はA I 付随サービスを他者に提供する者
- ・ **ビジネス利用者**¹²⁸
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
- ・ **消費者的利用者**¹²⁹
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）
- ・ **間接利用者**¹³⁰
最終利用者がA I システム又はA I サービスを利用して提供するサービス等を受ける者

¹²⁷ 現状においては、主にA I ネットワーク化の第一段階を念頭に、インターネット等を通じてA I サービス又はA I 付随サービスが提供されることが想定される。今後、技術の発展等に伴って、第二段階以降を念頭に、A I システム相互間のネットワークを通じたA I サービス又はA I 付随サービスの提供が増えていくものと考えられる（A I ネットワーク化の進展段階については、序章1．参照。）。

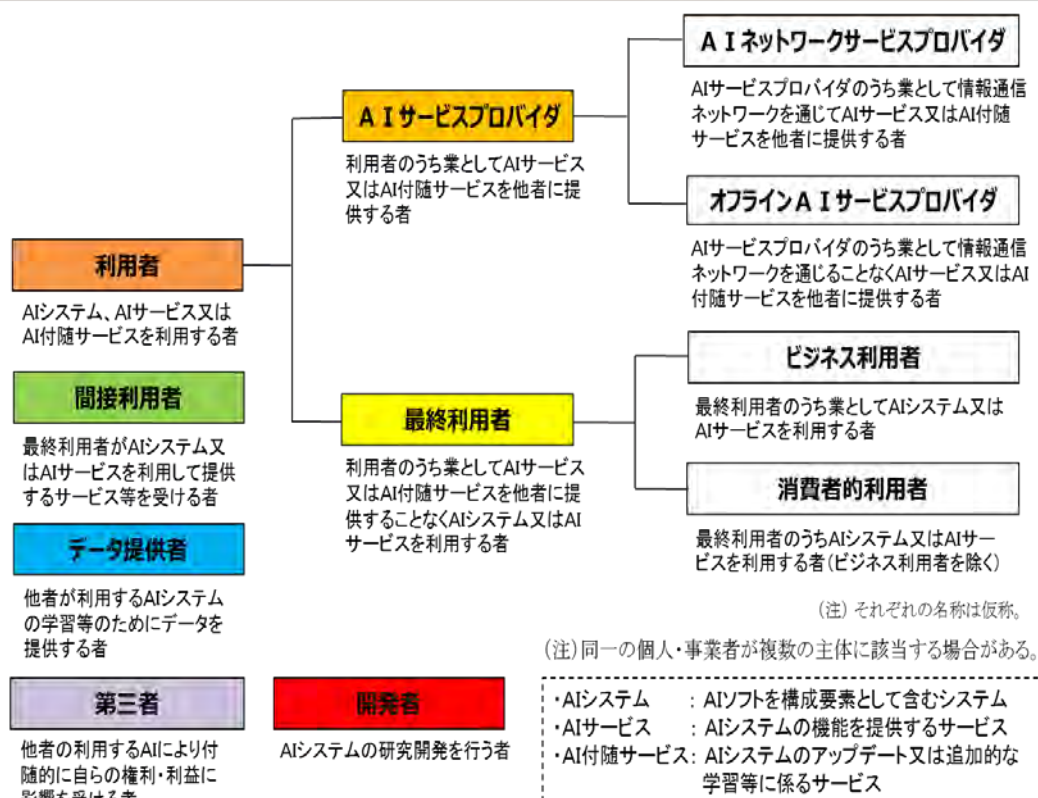
¹²⁸ 法人（企業や団体など）だけではなく、個人でも業としてA I システム又はA I サービスを利用する者である場合には、ビジネス利用者に該当する。

¹²⁹ 個人だけではなく、法人（企業や団体など）でも業としてA I システム又はA I サービスを利用しない者である場合には、消費者的利用者に該当する。

¹³⁰ あくまでA I を用いたサービスを利用（A I を間接的に利用）する者であり、「利用者」には含まれない。ただし、その名称から「利用者」に含まれると誤解されるおそれがあるため、名称を再検討すべきではないかとの指摘がある。

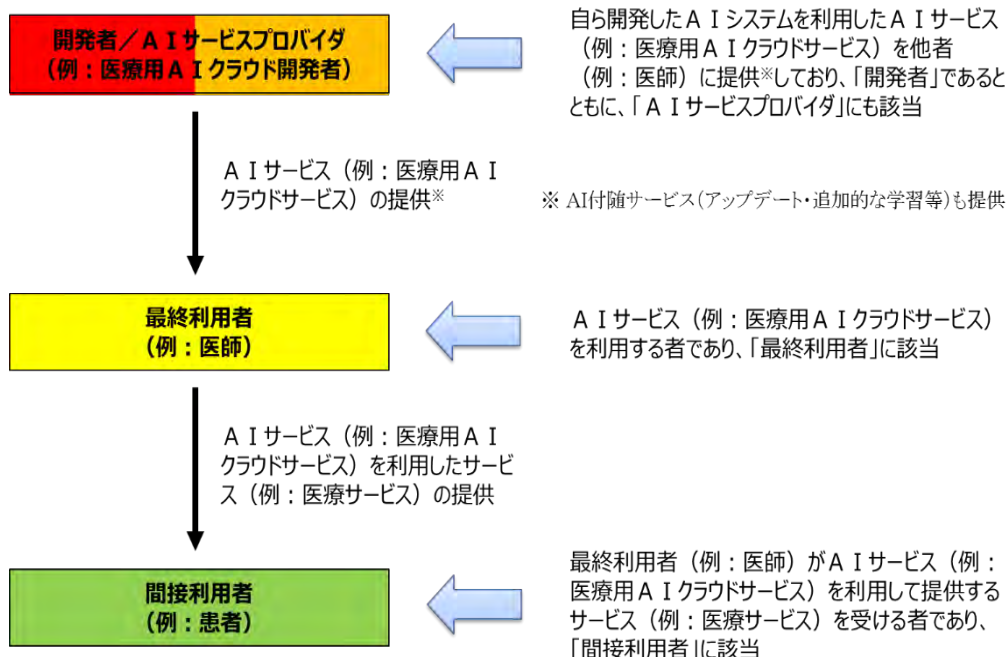
- ・ **データ提供者**
他者が利用するA Iシステムの学習等のためにデータを提供する者
- ・ **第三者**
他者の利用するA Iにより付随的に自らの権利・利益に影響を受ける者
- ・ **開発者**
A Iシステムの研究開発を行う者

関係する主体の整理

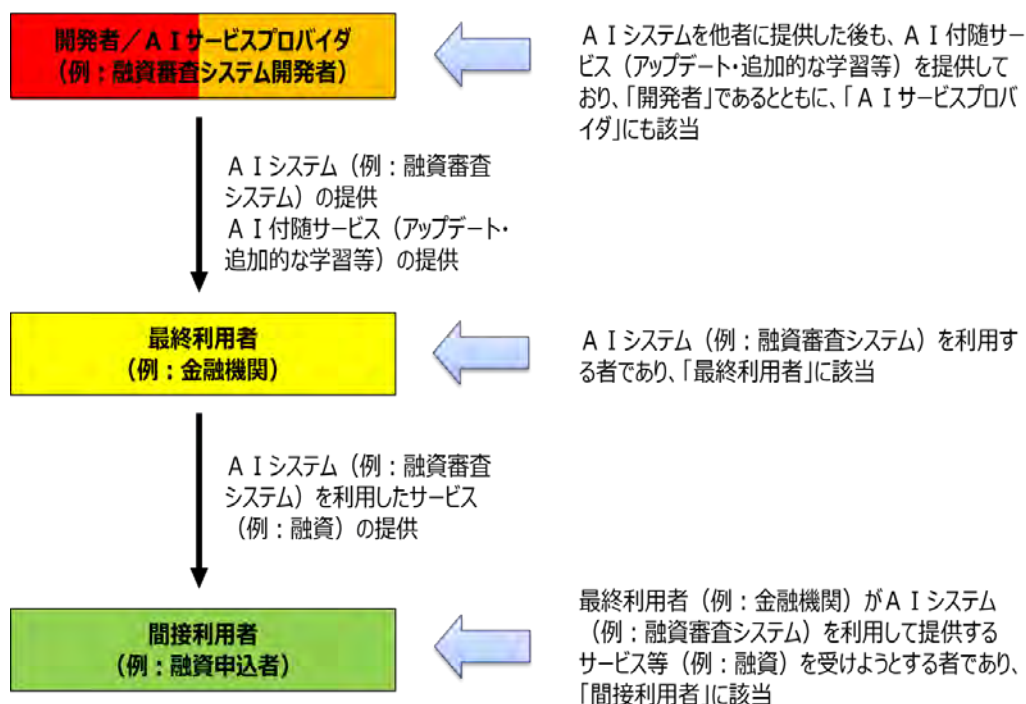


このような整理を踏まえて、A Iの利活用の事例に即した具体例は次のとおりである。

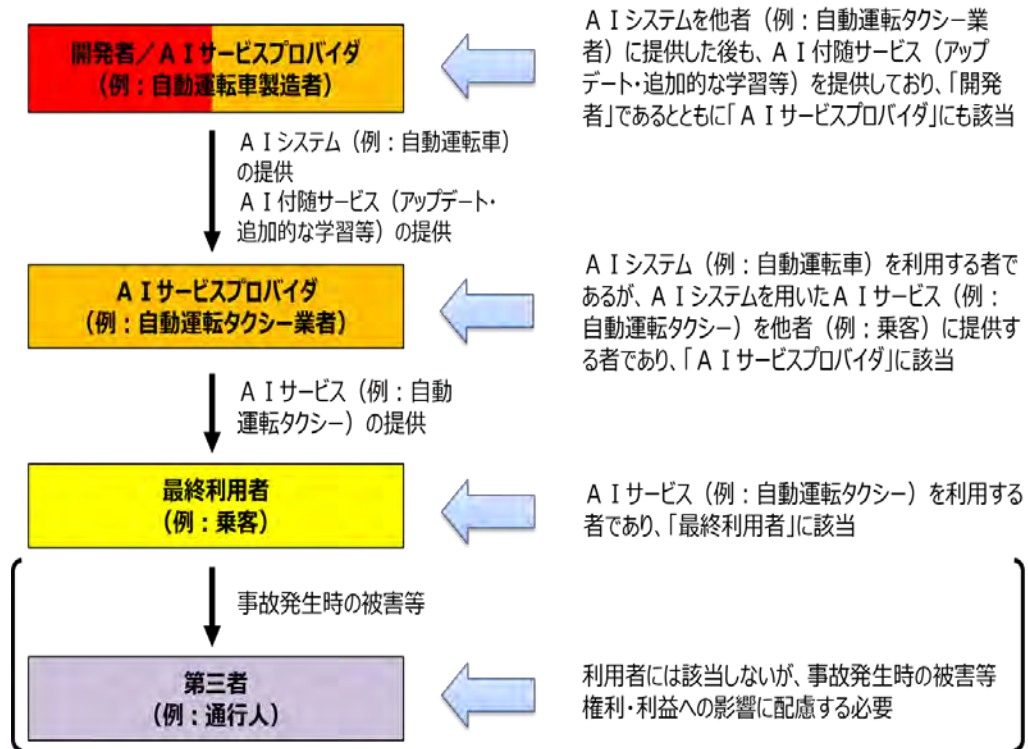
<例①：医療用クラウド>



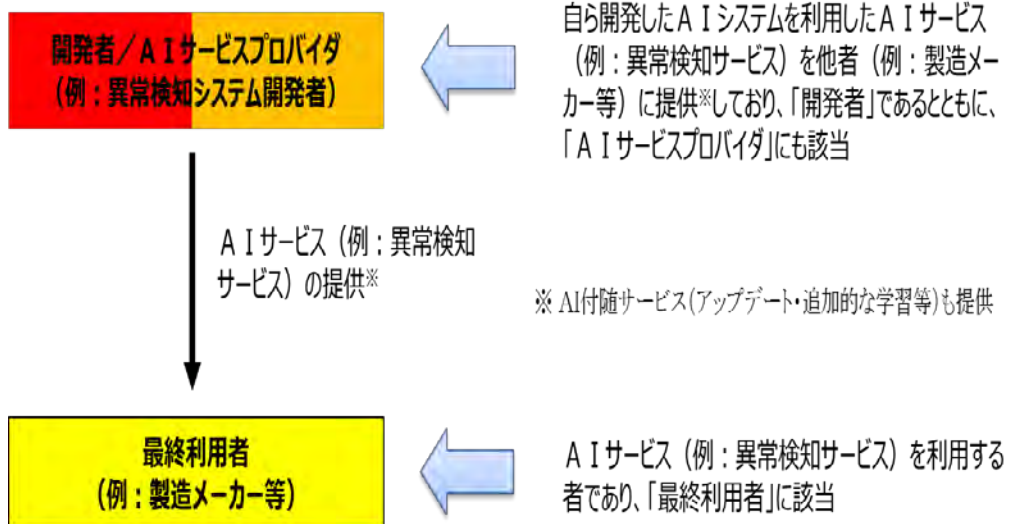
<例②：融資審査>



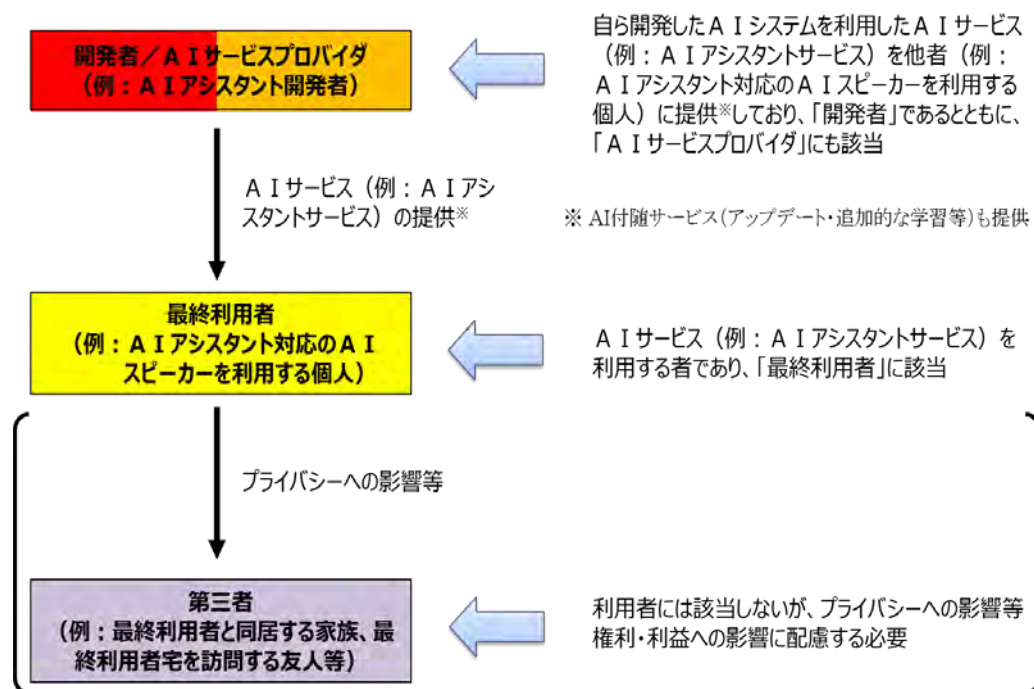
<例③：自動運転タクシー>



<例④：生産ラインの異常検知>



<例⑤：A I アシスタント（A I スピーカー）>



3. A I 利活用原則案

前章において分野別評価やエコシステムの展望、ネットワーク化の観点からの考察で課題を抽出し、それぞれの課題に応じた類型を整理した。これらのことを踏まえ、各類型に属する課題への対応を意識して、A I の利活用において留意することが期待される事項の検討を行った。

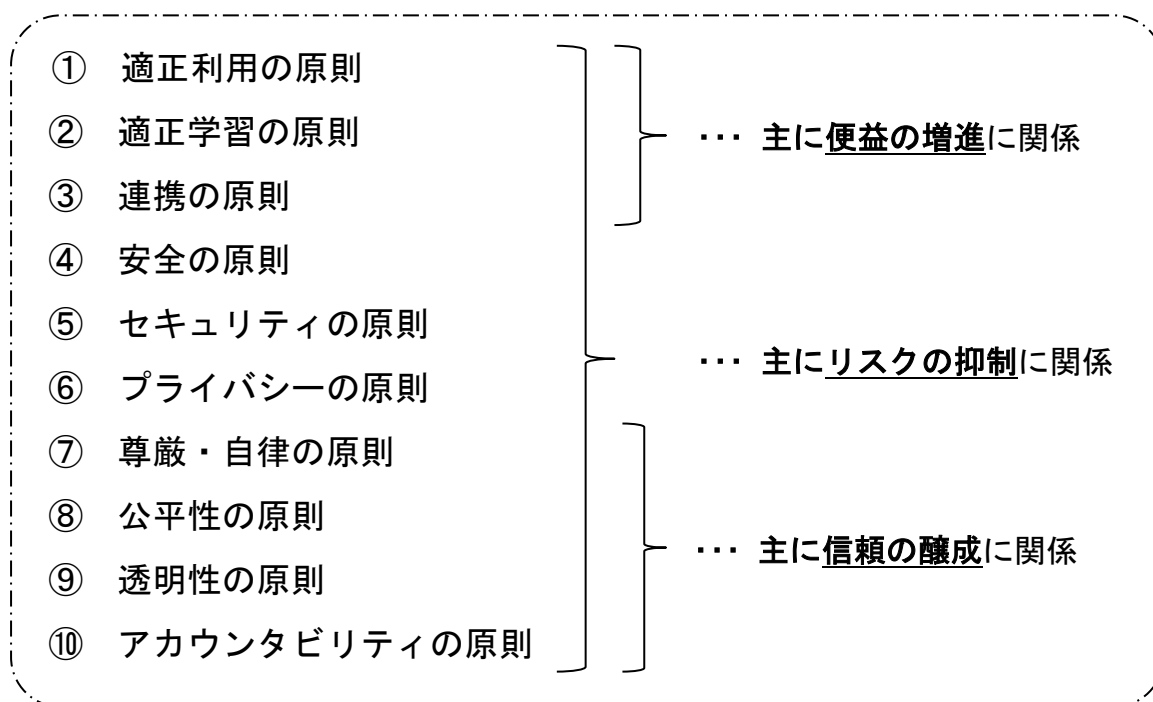
A I の利活用の促進やA I ネットワーク化の健全な進展に向けて、A I の便益の増進とリスクの抑制を図り、A I に対する信頼を醸成するため、利用者及びデータ提供者がA I の利活用において留意することが期待される事項を「A I 利活用原則案」として取りまとめた¹³¹。

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としている。

¹³¹ それぞれの原則が、主として、どのような課題に対応するためのものであるかを示すと次のとおりである（課題の類型については、前章2.（2）②と同様である（34-35頁参照）。）。

①適正利用の原則：安全、役割分担、②適正学習：データ、正当性・公平性、③連携の原則：連携、④安全の原則：安全、⑤セキュリティの原則：セキュリティ、⑥プライバシーの原則：プライバシー、⑦尊厳・自律の原則：正当性・公平性、⑧公平性の原則：正当性・公平性、⑨透明性の原則：ブラックボックス化、⑩アカウントビリティの原則：受容性

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。



① 適正利用の原則

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。

② 適正学習の原則

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。

③ 連携の原則

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。

④ 安全の原則

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配慮する。

⑤ セキュリティの原則

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。

⑥ プライバシーの原則

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。

⑦ 尊厳・自律の原則

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。

⑧ 公平性の原則

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。

⑨ 透明性の原則

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。

⑩ アカウンタビリティの原則

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タビリティを果たすよう努める。

A I 利活用原則案では、「2. 関係する主体の整理」で分類した主体ごとに、留意することが期待される事項を原則として取りまとめているが、例えば、A I サービスプロバイダ、最終利用者のいずれもが留意することが期待される事項については、「利用者は・・・留意する。」というように、大括り化して取りまとめている。しかしながら、A I サービスプロバイダと最終利用者とは、留意することが期待される事項の具体的な内容は異なるものと考えられる。

例えば、「適正利用の原則」に関して、A I サービスプロバイダは、適正な範囲及び方法でA I を利活用するに当たっては、A I サービスを他者に提供する観点から必要となる事項について開発者から情報提供・説明を受けることが必要となる。他方、最終利用者は、自己の業務や生活における利便性の向上等のためにA I を利用するに当たって必要となる事項についてA I サービスプロバイダから情報提供・説明を受けることが必要となる。このように、それぞれの場面で求められる情報提供・説明の内容が異なることに加え、A I サービスプロバイダは、開発者に説明を求める立場であるとともに、最終利用者からは説明を求められる立場にある。

これらのことを踏まえると、「適正利用の原則」に限らず、他の原則に関しても、A I

サービスプロバイダと最終利用者とは、留意することが期待される事項の具体的な内容が異なることが想定される。

本報告書においては、A I 利活用原則案を取りまとめるとともに、各原則の内容に関し継続して検討すべき論点を整理した¹³²。A I の利活用において、各原則の内容に関し、利用者等が留意することが期待されるものとして適切な方向性を示すためには、より具体性を持たせることが求められる。様々な立場の利用者が関与することや様々なA I システムやA I サービスの提供が見込まれること等を踏まえ、A I 利活用原則案の各原則の内容について詳細な検討を行うこととする。

4. A I 利活用原則案の各原則の内容に関する論点

A I の利活用の内容や形態は多岐にわたることや利用者の分類ごとに期待される事項が異なることなどに鑑み、A I 利活用原則案の各原則の内容については、さらに詳細な検討を要する。このため、本報告書においては、各原則の内容に関し継続して検討すべき論点を整理した。本報告書の公表後、整理された論点を踏まえ、最終的なアウトプットの取りまとめに向けて検討を行うこととする¹³³。

① 適正利用の原則

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。

【主な論点】

ア 適正な範囲・方法での利用

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。

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。

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¹³⁴が期待されるのではないか。

¹³² A I 利活用原則案の各原則の内容に関し整理された論点については、後述（本章4. 参照）。

¹³³ 最終的なアウトプットの取りまとめに向けた検討においては、利用者の分類（A I サービスプロバイダ、最終利用者等）によって有している能力や知識の多寡、必要な措置を実施する容易性等が異なることを踏まえて、利用者の分類を意識しつつ、留意することが期待される事項の具体的な内容を検討することとする。

¹³⁴ 例えば、開発者が提供するアップデート機能やA I 向けのセキュリティソフト等を用いるなどして確認することが考えられる。

イ AIの便益とリスクの適正なバランス

AIサービスプロバイダ及びビジネス利用者は、AI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AIの積極的な利活用を検討するなど、AI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と。

ウ AIソフトのアップデート及びAIの点検・修理等

利用者は、利活用の過程を通じて、AIの機能を向上させ、リスクを抑制するため、AIソフトのアップデート及びAI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と。

エ 人間の判断の介在

AIによりなされた判断について、必要かつ可能な場合（例：AI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と。その場合、人間の判断の介在の可否の基準をどのように考えるか。

また、アクチュエータ等を通じて稼働するAI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当か。

[基準として考えられる観点（例）]

- ・ AI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
- ・ AIの判断の信頼性の程度（人間の判断の信頼性との優劣）
- ・ 人間の判断に必要な時間的猶予
- ・ 利用者に期待される能力

オ 利用者間の役割分担

利用者は、それぞれがAI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と。

カ 関係者間の協力

利用者及びデータ提供者は、AI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と。

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。

② 適正学習の原則

利用者及びデータ提供者は、AIシステムの学習等に用いるデータの質に留意する。

【主な論点】

ア AIの学習等に用いるデータの質への留意

利用者及びデータ提供者は、利用するAIの特性及び用途を踏まえ、AIの学習等に用いるデータの質（正確性や完全性など¹³⁵）に留意することが期待されるのではないか。また、AI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か。

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。

イ 不正確又は不適切なデータの学習等によるAIのセキュリティ脆弱性への留意

利用者及びデータ提供者は、AI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ずるリスクに留意することが期待されるのではないか。

③ 連携の原則

AI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、AIシステム又はAIサービス相互間の連携に留意する。また、利用者は、AI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。

【主な論点】

ア 相互接続性と相互運用性への留意

AIネットワークサービスプロバイダは、AIネットワーク化の健全な進展を通じて、AIの便益を増進するため、AI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。

イ データ形式やプロトコル等の標準化への対応

AIサービスプロバイダ及びビジネス利用者は、AI相互間及びAI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。

ウ AIネットワーク化により惹起・増幅される課題への留意

AI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AIがインターネット等を通じて他のAI等と接続・連携することにより制御不能となる等、AI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。

¹³⁵ データが表そうとしている実体が正しく示されていること、データがすべて揃っていて欠損がないこと、データに不整合や矛盾がないことなどのことをいう。なお、利用するAIの性質や目的、データの種類等によって、求められる水準が異なり得ることに留意が必要である。

[基準として考えられる観点 (例)]

- ・ 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
- ・ AIシステム間の連携・調整が成立しないなどのおそれ
- ・ AIの判断・意思決定を検証できないおそれ（ネットワーク化により、システム全体としてブラックボックス化するおそれ）
- ・ 少数のAIの影響力が強くなりすぎるなどのおそれ（少数のAI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）
- ・ 領域横断での情報の共有と特定の基盤的なAIへの情報の集中によるプライバシー侵害のおそれ
- ・ AIが想定外の動作を行うなどのおそれ

④ 安全の原則

利用者は、AIシステム又はAI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配慮する。

【主な論点】

ア 人の生命・身体・財産への配慮

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AI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AIの点検・修理及びAIソフトのアップデートを行うことなどにより、AI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と。

また、利用者は、AI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と。

⑤ セキュリティの原則

利用者及びデータ提供者は、AIシステム又はAIサービスのセキュリティに留意する。

【主な論点】

ア セキュリティ対策の実施

利用者は、AI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。

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。

イ セキュリティ対策のためのサービス提供等

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と。

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

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と。[再掲：①適正学習の原則 主な論点 イ)]

⑥ プライバシーの原則

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。

【主な論点】

ア 他者のプライバシーの尊重

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。

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。

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

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と。

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

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と。

エ 自己等のプライバシー侵害への留意

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と。

オ パーソナルデータの流出の防止

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ることが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と。

⑦ 尊厳・自律の原則

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。

【主な論点】

ア 人間の尊厳と個人の自律の尊重

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。

イ A Iによる意思決定・感情の操作等への留意

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切か否か。

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。

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

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。

⑧ 公平性の原則

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。

【主な論点】

ア A Iの学習等に用いられるデータの代表性への留意

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¹³⁶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。

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。

イ アルゴリズムによる不当な差別への留意

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。

¹³⁶ サンプルされて利活用に使わ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。

ウ 人間の判断の介在

AI サービスプロバイダ及びビジネス利用者は、AI 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、AI 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か。

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。

⑨ 透明性の原則

AI サービスプロバイダ及びビジネス利用者は、AI システム又はAI 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¹³⁷。

【主な論点】

ア AI の入出力の記録・保存

AI サービスプロバイダ及びビジネス利用者は、AI 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¹³⁸。

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、AI 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。

イ 説明可能性の確保

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。

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。

¹³⁷ 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。この点に関し、競争政策上の問題として、「③ 連携の原則」にも関連するものと考えられる（「③ 連携の原則」については、57-58 頁参照。）。

¹³⁸ 個人情報やパーソナルデータについて、どのようなデータが使われているのか、データが正確に輸入されているのかなど確認を求められることが想定されるが、データの記録、保存、開示などに関しては、個人情報保護制度の問題として担保されるものである。

これに関連して、AI の利活用の促進のためには、個人情報やパーソナルデータに関するデータポータビリティを容易に行うことができることが望ましいものと考えられるが、この点については、個人情報保護制度の問題として検討されるべき課題として挙げられる。

(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が必要である。)

⑩ アカウンタビリティの原則

AI 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タビリティを果たすよう努める。

【主な論点】

ア アカウンタビリティを果たす努力

AI 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AI 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、AI の利活用により影響を受ける第三者等に対し、利用するAI の性質及び目的等に照らして、相応のアカウンタ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と。

イ AI に関する利用方針の通知・公表

AI 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AI の利活用について、適切に認識することのできるよう、AI に関する利用方針（AI 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と。

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切かと。

5. 今後の検討

今後、AI 利活用原則案の各原則の内容に関し整理された論点を踏まえ、最終的なアウトプットの取りまとめに向けて検討を行うこととする。

今後の検討においては、最終的なアウトプットを取りまとめるに当たって、AI 開発ガイドライン案のように、目的や基本理念を整理することが必要であるものと考えられる¹³⁹。この点に関し、AI 利活用原則案の各原則の内容に関し整理された論点において、ある原則の内容に関し論点として掲げられている事項が、他の原則の内容に関し論点として掲げられている事項の前提として整理することが可能であるのではないかとといった意見がある。各原則相互の関係も踏まえつつ、最終的なアウトプットの目的や基本理念等を検討することが適切であるものと考えられる。

また、先に述べたとおり、AI 利活用原則案は、主体を大括り化して取りまとめていることから、利用者の分類ごと、さらには、利用者の主観（利用しているシステム又はサービスがAI を用いたものであることを認識しているか否か）や性質、属性等も考慮した上

¹³⁹ AI 開発ガイドライン案においては、目的及び5つの項目からなる基本理念が掲げられている（報告書2017別紙1 1. 及び2. 参照。）。

で、各原則の内容に関し整理された各論点についてより詳細な検討を行うことが求められる¹⁴⁰。その際、特に消費者的利用者が留意することが期待される事項については、懸念される課題に対し、どのような対策を講ずればよいか、より具体的に示していくことが、A Iに対する信頼を醸成し、A Iの利活用の裾野を広げていく上で重要である。

おって、最終的なアウトプットの在り方についても、A I開発ガイドライン案のように、「ガイドライン」という形や名称で取りまとめることとするか、また、分野ごとの問題の性質等、国ごとの法制度等に相違があることに鑑み、分野共通のものとして取りまとめるのではなく、各分野のステークホルダ等が分野ごとのガイドライン等を策定する際に参照する基本的な指針（メタガイドライン）として取りまとめることとするか等について検討を行うことが必要となる。

これらの検討とともに、最終的なアウトプットを人々に周知し、普及させるため、利用者等の類型ごとに整理したり、特に消費者的利用者向けに分かりやすいメッセージを発信することが重要であるものと考えられ、「ハンドブック」や「マニュアル」などリテラシー教材（利用者の手引き）を作成し、それらに基づいてワークショップ等を行ったりする等の手段・方法等についても検討することが望ましい。

¹⁴⁰ 今後、各原則の内容に関し整理された各論点について詳細な検討を行うに当たっては、従来のICTシステムやサービス、ネットワークと比べて、AIシステムやAIサービス、AIネットワークであるが故に、どのようなことに留意することが期待されるのか、どのような対策を講ずればよいか、といった観点が重要である。

第4章 今後の課題

報告書 2017 に掲げられている「今後の課題」¹⁴¹並びに本推進会議におけるAIネットワーク化が社会・経済にもたらす影響に関する検討及びAIの利活用において留意することが期待される事項に関する検討等を踏まえ、今後の課題を次のとおり整理する。

これらの課題について検討を行うに当たっては、AIの研究開発や利活用、社会実装を進める中でリスクが顕在化してしまった事例¹⁴²を踏まえて、それらから得られる教訓を活かすことが重要である。

○ AIの開発及び利活用の促進並びにAIネットワーク化の健全な進展に関する事項

(1) AI開発ガイドライン（仮称）の策定

第1章において述べたように、本推進会議が作成したAI開発ガイドライン案が、G7やOECD等において共有され、国際的な議論が進められている¹⁴³。今後さらに国際的な議論が進められ、ステークホルダ間の合意形成の下でAI開発ガイドライン（仮称）が策定されることが期待される。

本推進会議においては、AI開発ガイドライン（仮称）の策定に向けた今後の国際的な議論をフォローアップするとともに、その動向に応じて、AI開発ガイドライン案を見直すなどして国際的な議論に継続的に貢献していくべきものとする¹⁴⁴。

(2) AIの利活用に関する指針の策定

前章において述べたように、本推進会議において、国際的な議論の用に供するため、「AI利活用原則案」を取りまとめた。今後、G7やOECD等において、AI開発ガイドライン（仮称）と併せてAIの利活用に関する指針の策定に向けた国際的な議論が進められることが期待される。

本推進会議においては、AIの利活用に関する国内外の議論の動向を踏まえて、AI利活用原則案を継続的に見直し、必要に応じて改定するとともに、各原則の内容に関し整理された論点を踏まえて、最終的なアウトプットの取りまとめに向けた検討を進めることが必要である。

¹⁴¹ 本章末尾参照。詳細については、報告書 2017 第4章参照。

¹⁴² 海外においては、自動運転車が死亡事故を起こしたり、AIスピーカーが勝手に（本人同意なく）家庭内の会話を知人に送信したりする事例が発生している。

¹⁴³ AI開発ガイドライン案については、欧州諸国の有識者からも、「一般論として、アプローチについて同意する。」「明確であり、かつ議論の継続のためのよい枠組みを整えているものとする。」といった好意的な評価を得ている（三部環境整備分科会構成員「英国・ドイツ・フランスにおけるAIの利活用等に関する動向調査」（提出資料）参照。）。

¹⁴⁴ 第1章1. で述べたように、国内において、国際的な議論に貢献するために、人工技術戦略会議の下に「人間中心のAI社会原則検討会議」が置かれた。同会議の動向にも留意することが必要である。

最終的なアウトプットの取りまとめに向けた検討に当たっては、次に掲げる事項に留意することが必要である¹⁴⁵。

- 目的や基本理念の整理が必要と考えられるところ、その検討に当たっては、A I 利活用原則案の各原則相互の関係を踏まえること。
- 利用者の主観（利用しているシステム又はサービスがA I を用いたものであることを認識しているか否か）や性質、属性等も考慮する必要があること。
- 利用者の分類（A I サービスプロバイダ、最終利用者等）によって有している能力や知識の多寡、必要な措置を実施する容易性等が異なることを踏まえて、利用者の分類を意識しつつ、留意することが期待される事項等の具体的な内容を検討する必要があること。
- 次に掲げる論点を踏まえ、最終的なアウトプットの在り方について検討する必要があること。
 - ・ 「ガイドライン」という形や名称で取りまとめることとするか。
 - ・ 各分野のステークホルダ等が分野ごとのガイドライン等を策定する際に参照する基本的な指針（メタガイドライン）として取りまとめることとするか。

（3）関係するステークホルダが取り組む環境整備に関する課題

A I の便益を人々や社会が享受するためには、A I の開発をさらに推し進めるとともに、利活用の裾野をより一層広げていくことが必要である。そのためには、人々のA I に対する不安を取り除き、信頼を醸成する取組を進めることが必要である。A I の開発者や利用者が、それぞれA I の開発や利活用において、A I の便益を増進させリスクを抑制する観点から期待される取組を行うとともに、開発者や利用者に加えて、政府や業界団体等を含めた関係するステークホルダがA I の開発及び利活用の促進やA I ネットワーク化の健全な進展に向けた環境整備のための取組を行うことにより、A I に対する社会的な受容性の向上がもたらされる。

これまで述べたように、A I の開発者が留意することが期待される事項については、A I 開発ガイドライン案を作成し、A I 開発ガイドライン（仮称）を策定するために国際的な議論を進めているところである。また、A I の利用者やデータ提供者が留意することが期待される事項については、A I 利活用原則案を取りまとめたところであり、今後、各原則の内容に関し整理した論点を踏まえて検討を行い、最終的なアウトプットを取りまとめることとしている。

これらの取組に加えて、開発者、利用者及びデータ提供者を含む関係するステークホルダ（政府、業界団体等）がA I の開発及び利活用の促進やA I ネットワーク化の健全な進展に向けた環境整備のために取り組むことが期待される事項を整理することが必

¹⁴⁵ 前章参照。

要であり、例えば次に掲げる課題が挙げられる¹⁴⁶。

- すべてのステークホルダが取り組むことが期待される課題
 - ・ AIの開発及び利活用の促進並びにAIネットワーク化の健全な進展に関するステークホルダ間の協力、ベストプラクティスの共有の在り方の検討
- 主として業界団体等が中心となって取り組むことが期待される課題
 - ・ AIの学習等に用いられるデータの形式やAI間の自動交渉のプロトコル等の標準化の在り方の検討
 - ・ AIの利活用、AIネットワーク化に必要な知識・技能等に関する啓発・教育
 - ・ AIの開発及び利活用並びにAIネットワーク化を推進するための人材育成
 - ・ AIの利活用、AIネットワーク化に係る市場競争環境整備、利用者の選択肢の多様性の確保の在り方の検討
 - ・ AIの事故等に関する被害者の救済（保険等）及び被害発生の防止の在り方の検討
- 主として政府が中心となって取り組むことが期待される課題
 - ・ 開発者、利用者及びデータ提供者を含む関係するステークホルダ間の対話や情報共有が促進されるための方策（対話の場の設置等）の在り方の検討
 - ・ AIの事故等が発生した場合の原因究明や再発防止等に関して関係者の協力が得られる仕組み（インセンティブの付与など）の検討
 - ・ AIの便益及びリスクを検証するための実験環境（特区、仮想実験空間等）の整備の在り方の検討
 - ・ AIの開発及び利活用並びにAIネットワーク化を推進するための経済政策や産業政策、国内産業の活性化のための対外政策の検討
 - ・ AIの開発及び利活用並びにAIネットワーク化を推進するための法制度等の在り方の検討
 - ア AIの事故等に関する法的責任の所在の在り方
 - イ 各分野におけるAIの利活用と業法等との関係の整理
 - ウ AIの学習等に用いられるデータ、AIにより作成されたモデル、AI創作物等に関する知的財産制度上の保護と利活用の在り方
 - エ AIの利活用に関する個人情報の保護と利活用の在り方
 - オ AIネットワークに関する通信の秘密の保護（AI間の通信の秘密の保護の範囲及び程度など）の在り方

¹⁴⁶ これらの課題について、（４）以降で掲げられる課題と重複するものがあり得る。

(4) A I システム又はA I サービス相互間の円滑な連携の確保

A I 開発原則¹⁴⁷の「連携の原則」において、開発者は、A I システムの相互接続性と相互運用性¹⁴⁸に留意することが掲げられるとともに、同原則の解説¹⁴⁹において、開発者は、関連情報の共有、国際的な標準や規格への準拠、データ形式の標準化、A P I を含むインターフェースやプロトコルのオープン化への対応、標準必須特許など相互接続性と相互運用性の確保に資する知的財産権のライセンス契約及びその条件についてのオープンかつ公平な取扱いに留意することが望ましい旨が掲げられている。

また、A I 利活用原則案¹⁵⁰の「連携の原則」において、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ことが掲げられている。

これらのことを踏まえると、A I 相互間の円滑な連携の確保に関連する課題としては、例えば次に掲げる課題が挙げられる。

- 開発者やA I サービスプロバイダが他の開発者やA I サービスプロバイダ等と共有することが期待される関連情報の範囲、共有の方法等の検討
- インターフェース、プロトコル、データ形式等の標準化又はオープン化の在り方の検討
- 標準必須特許など相互接続性と相互運用性の確保に資する知的財産権のライセンス契約及びその条件の在り方等に関する検討
- 連携及びこれに伴う責任の分配に関する当事者間の契約の在り方の検討
- 当事者間の紛争解決の在り方の検討

なお、これらの課題に取り組むに当たっては、ネットワーク上における様々なA I システムの混在（性質や性能が異なる多様なA I システムがネットワーク上に混在したり、新旧のA I システムがネットワーク上に混在したりすること等）の可能性に留意することが必要である。

また、本推進会議において、A I 利活用原則案の「連携の原則」に関し整理された論点¹⁵¹を検討するに当たっては、上記の事項に留意することが必要である。

¹⁴⁷ A I 開発原則については、報告書 2017 別紙 1 4. 参照。以下本章におけるA I 開発原則に関する記述において同じ。

¹⁴⁸ ここで相互接続性と相互運用性としては、自らの開発するA I システムが情報通信ネットワークと接続され、他のA I システム等と適切に協調して運用することが可能であることを一体的に捉えて念頭に置いている。

¹⁴⁹ A I 開発原則の解説については、報告書 2017 別紙 1 5. 参照。以下本章におけるA I 開発原則の解説に関する記述において同じ。

¹⁵⁰ A I 利活用原則案については、前章 3. 参照。以下本章におけるA I 利活用原則案に関する記述において同じ。

¹⁵¹ A I 利活用原則案の各原則の内容に関し整理された論点については、前章 4. 参照。以下本章におけるA I 利活用原則案の各原則の内容に関し整理された論点に関する記述において同じ。

(5) 競争的なエコシステムの確保

AI ネットワーク化の進展を通じて智連社会¹⁵²を実現するためには、

- ・ AI システム相互間又はAI システムと他の種類のシステムとの間のネットワークの形成
- ・ AI システムやAI サービスの提供
- ・ ネットワーク化されたAI により創造・流通・連結されるデータ・情報・知識の利活用

のそれぞれのエコシステムにおいて公正な競争が確保されることが必要である。

これらのエコシステムは、今後のAI や情報通信ネットワークに関する技術の高度化、市場におけるAI システムやAI サービスに対するニーズの変化、AI システムの利活用の進展等に応じて変遷していくものであり、予断を抱くべきものではないことから、競争的なエコシステムの確保に関する規制の創設の検討は将来に向けて謙抑的であるべきである。その一方で、競争的なエコシステムの確保に関し必要に応じて時宜に応じた適切な対応ができるよう備えておく見地からは、報告書 2016 において指摘されているように¹⁵³、関連する市場の動向を継続的に注視することが必要であるものと考えられる。

(6) 利用者の利益の保護

AI ネットワーク化が進展して、様々な物、サービスや社会のシステムがネットワーク化されたAI の利用可能性を前提として提供されるようになった社会における人間の「包摂」(inclusiveness) の見地に加えて、ネットワークに接続されたAI に関するいわゆるネットワーク効果の見地からは、あらゆる個人又は団体が、利用者として、関連するAI を手頃な条件で¹⁵⁴安心・安全に利用することができるよう、その利益が保護されることが求められる。

このため、AI ネットワーク化のガバナンスにおいては、AI システムの開発のガバナンス及び利活用のガバナンスのいずれについても、その全体を通じて直接的又は間接的に利用者の利益の保護に資するものであることが期待される。AI 開発ガイドライン案においては、このような問題意識の下、その目的として利用者の利益の保護を掲げているとともに、同ガイドライン案に掲げるAI 開発原則の個々の項目は、いずれも利用者の利益の保護に資するものとしている¹⁵⁵。

¹⁵² 序章 1. 及び前掲注 (6) 参照。

¹⁵³ 報告書 2016・50 頁参照。

¹⁵⁴ 手頃な条件での利用可能性は、公正な競争を通じて実現されることが基本である。

¹⁵⁵ AI 開発原則のうち「連携の原則」は、連携に係るAI システムと他のAI システム等との円滑な連携を通じて、連携に係るAI システムの便益を増進するとともにリスクを抑制することにより利用者の利益の保護に資するものである。「透明性の原則」、「制御可能性の原則」、「安全の原則」、「セキュリティの原則」、「プライバシーの原則」及び「倫理の原則」は、それぞれ関連するリスクを抑制することにより利用者の利益の保護に資するものである。「利用者

また、A I 利活用原則案においても、個々の原則は、いずれも利用者の利益の保護に資するものとしている¹⁵⁶。

これらのことを踏まえると、特に利用者の利益の保護を直接的に志向する課題としては、例えば次に掲げる課題が挙げられる。

- 開発者やA I サービスプロバイダ等から利用者に対する
 - ・ A I システム又はA I サービスの技術的特性に関する情報¹⁵⁷
 - ・ 継続的なアップデートへの対応の必要性に関する情報¹⁵⁸
 - ・ その他リスクの顕在化又は波及の抑止に資する情報の自発的な提供¹⁵⁹の在り方の検討
- A I サービスプロバイダ等がその提供の業務を通じて知り得た利用者又はその利活用に関するデータ等の取扱いの在り方の検討
- A I システムのリスクが顕在化した場合（事故の発生時等）における責任の分配や利用者等を保護する仕組み（保険等）等の在り方の検討

なお、本推進会議において、特にA I 利活用原則案の「アカウントビリティの原則」、「適正学習の原則」及び「適正利用の原則」等に関し整理された論点を検討するに当たっては、上記の事項に留意することが必要である。

（7）技術開発に関する課題

A I 開発原則及びA I 利活用原則案の個々の項目の実効性を高めるとともに、本章に掲げる課題の解決を図るためには、制度的な仕組みとともに、技術的な仕組みが有効に機能することが期待される。

例えば、次に掲げる技術の研究開発が進むことが期待される。

- A I システムの入出力の検証可能性を確保するための技術及びA I システムの判断結果の説明可能性を確保するための技術

支援の原則」及び「アカウントビリティの原則」は、それぞれ利用者の受容性を向上させることにより、利用者の利益の保護に資するものである。

¹⁵⁶ A I 利活用原則のうち「適正利用の原則」、「適正学習の原則」及び「連携の原則」は、それぞれA I の便益を増進するとともにリスクを抑制することにより利用者の利益の保護に資するものである。「安全の原則」、「セキュリティの原則」及び「プライバシーの原則」は、それぞれ関連するリスクを抑制することにより利用者の利益の保護に資するものである。「尊厳・自律の原則」、「公平性の原則」、「透明性の原則」及び「アカウントビリティの原則」は、それぞれ利用者の受容性を向上させるとともにリスクを抑制することにより、利用者の利益の保護に資するものである。

¹⁵⁷ A I 開発原則に掲げる事項に関する情報を含む。

¹⁵⁸ 報告書 2016・54 頁参照。

¹⁵⁹ これらの情報提供のうち開発者によるものについては、A I 開発原則のうち「アカウントビリティの原則」に関連するものとして、また、利用者によるものについては、A I 利活用原則案のうち「アカウントビリティの原則」に関連するものとして検討することが期待されよう。

- 他のAIに対する監督（監視、警告など）や制御（停止、ネットワークからの切断、修理など）を行うことのできる技術
 - AIシステムがアクチュエータ等を通じて稼動する際の本質安全（アクチュエータの運動エネルギーなど本質的な危険要因の低減）や機能安全（自動ブレーキなど付加的な制御装置の作動によるリスクの抑制）を確保するための技術
 - 利用者に選択の機会を適時適切に提供する機能（例えば、デフォルトの設定、理解しやすい選択肢の提示、フィードバックの提供、緊急時の警告、エラーへの対処など）に関する技術
 - AIネットワークのセキュリティを高める（セキュリティ攻撃のパターンの学習、学習データのノイズの排除等）ための技術やセキュリティ攻撃を受けた際の早期復旧のための技術
 - AIシステムの学習等に用いられるデータに含まれる偏見等に起因して個人が不当に差別されるリスクを排除するための技術
- なお、本推進会議において、特にAI利活用原則案の「透明性の原則」、「安全の原則」、「セキュリティの原則」及び「公平性の原則」等に関し整理された論点を検討するに当たっては、上記の事項に留意することが必要である。

○ AIネットワーク上を流通する情報・データに関する事項

（8）セキュリティ対策

AI開発原則の「セキュリティの原則」において、開発者は、AIシステムのセキュリティに留意することが掲げられている。

また、AI利活用原則案の「セキュリティの原則」において、利用者及びデータ提供者は、AIシステム又はAIサービスのセキュリティに留意することが掲げられている。

今後、AIの開発及び利活用の両面において、セキュリティの確保の在り方について検討していくことが必要である。

- AIシステムへのセキュリティの実装の在り方の検討
- AIシステムの学習等による利活用の過程を通じた変化に起因するセキュリティ上の問題への対処の在り方の検討
- 偽装・なりすまし等によりAIシステムが犯罪等に悪用されるリスクへの対処の在り方の検討
- AIの研究者とセキュリティの研究者との間の連携の促進に向けた検討

なお、本推進会議において、AI利活用原則案の「セキュリティの原則」に関し整理された論点を検討するに当たっては、上記の事項に留意することが必要である。

(9) プライバシー及びパーソナルデータの保護

A I 開発原則の「プライバシーの原則」において、開発者は、A I システムにより利用者及び第三者のプライバシーが侵害されないよう配慮することが掲げられている。

また、A I 利活用原則案の「プライバシーの原則」において、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ことが掲げられている。

今後、A I の開発及び利活用の両面において、個人情報・プライバシーの保護の在り方について、A I が取り扱うデータ・情報の性質や、データ・情報の流通と個人情報・プライバシーの保護とのバランス等に留意しつつ検討することが必要である。

- 個人情報の取得や活用に当たっての本人同意等の在り方、データ・情報の加工（匿名化、暗号化等）に関する検討
- A I ネットワーク上を流通するデータ・情報を利活用する価値と個人情報保護・プライバシーとのバランスに配慮した制度の在り方の検討¹⁶⁰
- A I システムの学習等による利活用の過程を通じた変化に起因する意図しないプライバシー侵害のリスクへの対処の在り方の検討
- プロファイリングが利用者にもたらす便益及びプライバシー侵害等のリスクを踏まえたプロファイリングに関するルールの在り方の検討

なお、本推進会議において、A I 利活用原則案の「プライバシーの原則」に関し整理された論点を検討するに当たっては、上記の事項に留意することが必要である。

(10) コンテンツに関する制度的課題

報告書 2017 におけるシナリオ分析に基づく留意事項として、「A I システムが取り扱うデータ・情報の性質」の中で「例えば、A I 生成物や学習用データ等に関し、知財制度上の取扱いが問題となる可能性がある。」ことが挙げている¹⁶¹。また、新たな情報財の保護・利活用の在り方に関し、A I 生成物の知的財産制度上の在り方について引き続き検討すべきとの指摘がある¹⁶²。

今後、A I ネットワーク上におけるコンテンツの流通と知的財産の保護とのバランスに配慮した制度の在り方等について検討することが必要である。

- 学習用データの作成の促進に関する環境の整備、学習済みモデルの適切な保護と利活用促進、A I 生成物の知的財産制度上の在り方の検討¹⁶³
- A I 学習用データの作成の促進に向けた、国及び地方公共団体等が保有するオー

¹⁶⁰ 前掲注 (138)

¹⁶¹ 報告書 2017 第 3 章 5. オ参照。

¹⁶² 報告書 2017 第 1 章 1. (3) ⑥参照。

¹⁶³ 知的財産戦略本部 検証・評価・企画委員会 新たな情報財検討委員会「報告書—データ・人工知能 (A I) の利活用促進による産業競争力強化の基盤となる知財システムの構築に向けて—」(平成 29 年 3 月) 参照。

ブンデータの推進

➤ AIにより自動集積されるデータベースの保護と利活用の在り方の検討

○ AIネットワーク化が社会・経済にもたらす影響に関する事項

(11) AIネットワーク化が社会・経済にもたらす影響に関するシナリオ分析

報告書 2017 において、今後の課題として、AI ネットワーク化が社会・経済にもたらす影響に関するシナリオ分析の継続的な実施及び国際的な共有が掲げられていたところ¹⁶⁴、本推進会議において、シナリオ分析（分野別評価）及びエコシステムの展望に関する検討を行った¹⁶⁵。

今後も、引き続き、シナリオ分析を行うとともに、その成果を国際的に共有することが必要である。また、エコシステムの展望について、これまでの検討を踏まえ、AI ネットワーク上を流通するデータやAI ネットワークを通じて提供されるサービス等に関するエコシステムを展望することが期待される。これらのエコシステムの展望に当たっては、AI の利活用の在り方は単一ではなく、複数のものが想定され得ることを念頭に置いて検討を行うことが望ましい。

➤ AI ネットワーク化が社会・経済にもたらす影響に関するシナリオ分析の継続的な実施及び国際的な共有

➤ AI ネットワーク上を流通するデータに関するエコシステムやAI ネットワークを通じて提供されるサービスに関するエコシステム等についての継続的な検討及び国際的な共有

(12) AI ネットワーク化の進展に伴う影響の評価指標及び豊かさや幸せに関する評価指標の設定

報告書 2017 において、今後の課題として、国際比較にも用いることができるような定量的に測定し、又は客観的に把握することができる指標の設定に向けた検討が掲げられていたところ¹⁶⁶、本推進会議において、AI ネットワーク化に関する指標について、特に次に掲げる2つの指標に関する検討を行った¹⁶⁷。

・ AI ネットワーク化の前提となる環境整備に関するもの： AI ネットワーク化
事前指標

・ AI ネットワーク化の初期段階に関するもの： AI ネットワーク化初期指標
なお、本報告書において、今後の取組方針として、指標として有用な項目の洗い出し・整理、アンケート調査における有用な項目の検討が示されている。

¹⁶⁴ 報告書 2017 第4章 (10) 参照。

¹⁶⁵ 分野別評価については、第2章1. 参照。エコシステムの展望については、同章2. 参照。

¹⁶⁶ 報告書 2017 第4章 (11) 参照。

¹⁶⁷ 第2章5. 参照。

今後も、引き続き、指標の設定に関する検討を行うことが期待される。

- 国際比較にも用いることができるような定量的に測定し、又は客観的に把握することができる指標の設定に向けた検討

(13) AIの利活用に関する社会的受容性の醸成

AI開発原則の「アカウントビリティの原則」において、開発者は、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ことが掲げられている。これは、AIシステムの利活用に当たっては、利用者を含む社会全体の理解が必要で、社会に受容されることが重要であるという認識に基づくものである。

また、AI利活用原則案の「アカウントビリティの原則」において、AI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ことが掲げられている。

今後、AIの開発及び利活用の促進並びにAIネットワーク化の健全な進展に向けて、社会的な受容性を注視するとともに、受容性を高める方策の在り方等について検討することが必要である。

- 社会におけるAIに関する受容度の継続的注視、受容性を高める方策の在り方の検討

なお、本推進会議において、AI利活用原則案の「アカウントビリティの原則」に関し整理された論点を検討するに当たっては、上記の事項に留意することが必要である。

○ AIネットワーク化が進展する社会における人間をめぐる課題に関する事項

(14) 人間とAIとの関係の在り方に関する検討

AI開発原則の「倫理の原則」において、開発者は、AIシステムの開発において人間の尊厳と個人の自律を尊重することが掲げられている。

また、AI利活用原則案の「適正利用の原則」において、利用者は、人間とAIとの適切な役割分担のもと、利用者間における適切な役割分担を踏まえ、適正な範囲及び方法でAIシステム又はAIサービスを利用するよう努めること、「尊厳・自律の原則」において、利用者は、AIシステム又はAI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こと、「公平性の原則」において、AI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、AIシステム又はAI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ことが掲げられている。

今後、人間中心とする智連社会に向けた人間とAIとの役割分担、AIシステムに委ねる判断の範囲の妥当性、AIネットワーク化の進展による人間の能力の拡張の在り方等について検討することが必要である。

- 専門職（医師、弁護士、会計士等）とAIシステムとの役割分担の在り方や重要な判断（診断、法令の解釈・適用、採用、人事評価、融資の審査等）に当たりAIシス

テムの機能に委ねてもよい事項の範囲に関する検討

- AI・ロボットによる人間の知的・身体的能力の拡張（エンハンスメント）に関する倫理的問題の検討
- AIシステムやAIサービスに対する過度な依存を回避する方策の検討
- 「AI依存」など人間の心理や子供の発育への影響等に関する継続的注視

なお、本推進会議において、特にAI利活用原則案の「適正利用の原則」、「尊厳・自律の原則」及び「公平性の原則」等に関し整理された論点を検討するに当たっては、上記の事項に留意することが必要である。

(15) ステークホルダ間の関係の在り方に関する検討

報告書 2017 におけるシナリオ分析に基づく留意事項として、「事故の発生等AIシステムのリスクが顕在化した場合の責任の分配の在り方」が挙げられている¹⁶⁸。

また、AI利活用原則案の「適正利用の原則」において、利用者は、人間とAIとの適切な役割分担のもと、利用者間における適切な役割分担を踏まえ、適正な範囲及び方法でAIシステム又はAIサービスを利用するよう努めること、「連携の原則」において、利用者は、関係者と協力して事故の未然防止及び事故の発生時等の対応に取り組むよう努めること、「アカウントビリティの原則」において、AI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ことが掲げられている。

今後、AIネットワーク化の進展を見据え、上記(14)人間とAIとの関係の在り方に関する検討とともに、AIのリスクの抑制等を図ることにより利用者及び第三者の利益を保護するために、ステークホルダの役割分担、権利義務の帰属の在り方、事故発生時等の責任の分配の在り方について検討することが必要である。

- AIのリスクが顕在化しないようにするためのステークホルダ間の役割分担の在り方の検討
- AIを利活用した取引における権利義務の帰属の在り方の検討
- AIのリスクが顕在化した場合のステークホルダ間の責任の分配の在り方の検討¹⁶⁹
- 特に最終利用者の利益を保護するためのアカウントビリティの在り方の検討
- AIの利活用をめぐる紛争の解決の在り方の検討

なお、本推進会議において、特にAI利活用原則案の「適正利用の原則」、「連携の原則」及び「アカウントビリティの原則」等に関し整理された論点を検討するに当たって

¹⁶⁸ 報告書 2017 第3章5. ウ参照。

¹⁶⁹ AIシステムのリスクが顕在化した場合の責任の分配の在り方の検討の一環として、AIの利活用等に伴う事故等による損害の賠償等に充てるための資財が乏しい個人、中小・零細事業者等によるAIの開発及び利活用を支援するためのファンドの設立や保険の在り方等についても併せて検討することが必要である。

は、上記の事項に留意することが必要である。

(16) A I ネットワーク化に対応した教育・人材育成及び就労環境の整備

報告書 2017 において、今後の課題として、A I ネットワーク化の進展に応じた就労や円滑な配置転換等に資する教育・人材育成の在り方等に関する課題が掲げられているところ¹⁷⁰、本推進会議において、教育・人材育成に関する検討を行い、次に掲げる6つのテーマごとに構成員の主たる意見を示す形で取りまとめた¹⁷¹。

- ① 人間とA I との関係・役割分担
- ② 社会の変化及び人間の技能・能力
- ③ 学校教育
- ④ 職業スキルの習得
- ⑤ 生涯教育
- ⑥ A I の利活用

なお、本報告書における取りまとめは、今後の検討に当たって、キックオフ的なものと位置付けられ、本報告書の公表後、これらの意見を踏まえて、本推進会議において、さらに検討を深めることとしている。

今後も、引き続き、A I ネットワーク化の進展に応じた就労や円滑な配置転換等に資する教育・人材育成の在り方等について検討すること及びその結果に基づき所要の対策を講ずることが必要である。

- 技術者、法的・倫理的・社会的問題に対処し得る人材、産業連携等に対処し得る人材の育成
- A I ネットワーク化の進展に応じた雇用・働き方の変化に合わせた教育の実施・人材の育成及び労働法制の在り方の検討

(17) A I に関するリテラシーの涵養及びA I ネットワーク・ディバイドの形成の防止

報告書 2017 において、高齢者のリテラシーの向上やディバイド形成の防止の在り方に関する課題が掲げられている¹⁷²。また、分野別評価において、A I ネットワーク化の健全な進展に向けた今後の検討に当たって留意すべき点として、高齢者のA I に関するリテラシーを向上させる方策が求められる旨が挙げられている¹⁷³。

今後、すべての人々がA I の恵沢をあまねく享受できるよう特に高齢者のリテラシーの向上やディバイド形成の防止の在り方について検討すること等が必要である。

¹⁷⁰ 報告書 2017 第4章 (15) 参照。

¹⁷¹ 第2章4. 参照。

¹⁷² 報告書 2017 第4章 (16) 参照。

¹⁷³ 第2章1. (2) ③参照。

- 特に高齢者など情報弱者のリテラシーの向上を図るための方策の検討
- AI を利活用する者と利活用しない（できない）者が混在すること、地域間でAI の利活用の状況が異なっていること等を踏まえ、社会全体がAI の利活用による便益を享受できるような環境の整備

(18) セーフティネットの整備

報告書 2017 において、大規模な失業が発生しないような方策の在り方等に関する課題が掲げられている¹⁷⁴。また、分野別評価において、AI ネットワーク化の健全な進展に向けた今後の検討に当たって留意すべき点として、プロファイリング結果により、就職や転職、結婚などにおいて、不当に不利な立場に陥ることのないようにセーフティネットの検討が必要である旨が挙げられている¹⁷⁵。

今後、労働市場の動向を注視するとともに、大規模な失業が発生しないような方策の在り方、AI によるプロファイリングの結果により、不当に不利な立場に陥らないような方策の在り方等について検討すること、セーフティネットを整備すること等が必要である。

- 労働市場の動向の継続的注視、AI ネットワーク化の進展に伴う所得の再分配等格差防止の在り方の検討
- AI によるプロファイリングの結果により、不当に不利な立場に陥らないような方策の在り方の検討

(19) その他の課題

報告書 2017 において、「その他の課題」として掲げられている事項¹⁷⁶についても、引き続き、注視し、又は検討することが必要である。

- 情報通信インフラの高度化の加速
- 経済発展・イノベーションの促進に向けた課題
- 地球規模課題の解決を通じた人類の幸福への貢献
- AI ネットワーク化のガバナンスの在り方

¹⁷⁴ 報告書 2017 第 4 章 (17) 参照。

¹⁷⁵ 第 2 章 1. (2) ③参照。

¹⁷⁶ これらの課題は、報告書 2016 において、「今後の課題」として掲げられたもののうち、報告書 2017 においては、「今後の課題」として具体的に取り上げられなかった課題である。

(参考) 報告書 2017 に掲げられている「今後の課題」¹⁷⁷

○ AI ネットワーク化の健全な進展に関する事項

- AI 開発ガイドライン (仮称) の策定
 - ・ AI 開発ガイドライン (仮称) の策定に向けた国際的な議論のフォローアップ
- AI 利活用ガイドライン (仮称) の策定
 - ・ 「AI 利活用ガイドライン (仮称)」の策定に関する検討
- AI システム相互間の円滑な連携の確保
 - ・ 関係ステークホルダ間で共有することが期待される関連情報の範囲、共有の方法等の検討
- 競争的なエコシステムの確保
 - ・ 関連する市場の動向の継続的注視
- 利用者の利益の保護
 - ・ 利用者に対する開発者等からの自発的な情報提供の在り方の検討、利用者を保護する仕組み (保険等) の在り方の検討
- 技術開発に関する課題
 - ・ 他のAI に対する監督 (監視、警告など) や制御 (停止、ネットワークからの切断、修理など) を行うことのできる技術等の研究開発の推進

○ AI ネットワーク上を流通する情報・データに関する事項

- セキュリティ対策
 - ・ セキュリティの実装の在り方の検討等
- プライバシー及びパーソナルデータの保護
 - ・ 本人同意等の在り方、データ・情報の加工 (匿名化、暗号化等) に関する検討等
- コンテンツに関する制度的課題
 - ・ 学習用データの作成の促進に関する環境整備、学習済みモデルの適切な保護と利活用促進等

○ AI ネットワーク化が社会・経済にもたらす影響に関する事項

- AI ネットワーク化が社会・経済にもたらす影響に関するシナリオ分析
 - ・ シナリオ分析の継続的な実施・国際的な共有

¹⁷⁷ ここに掲げられている課題のほか、情報通信インフラの高度化の加速、経済発展・イノベーションの促進に向けた課題、地球規模課題の解決を通じた人類の幸福への貢献、AI ネットワーク化のガバナンスの在り方、が今後の課題として掲げられている。

- AIネットワーク化の進展に伴う影響の評価指標及び豊かさや幸せに関する評価指標の設定
 - ・ 指標の設定に向けた検討
- AIシステムの利活用に関する社会的受容性の醸成
 - ・ 社会におけるAIシステムの利活用に関する受容度の継続的注視等
- **AIネットワーク化が進展する社会における人間をめぐる課題に関する事項**
- 人間とAIシステムとの関係の在り方に関する検討
 - ・ 専門職（医師、弁護士、会計士等）とAIシステムとの役割分担の在り方等の検討
- ステークホルダ間の関係の在り方に関する検討
 - ・ AIシステムのリスクが顕在化した場合の責任の分配の在り方等の検討
- AIネットワーク化に対応した教育・人材育成及び就労環境の整備
 - ・ AIネットワーク化の進展に応じた教育の実施・人材の育成等
- AIシステムに関するリテラシーの涵養及びAIネットワーク・ディバイドの形成の防止
 - ・ 特に高齢者など情報弱者のリテラシーの向上を図るための方策の検討等
- セーフティネットの整備
 - ・ 労働市場の動向の継続的注視、AIネットワーク化の進展に伴う所得の再分配等格差防止の在り方の検討等

むすびに代えて

本報告書は、本推進会議における検討と関連する国内外の議論を踏まえ、社会・経済の様々な分野におけるA Iシステムの利活用の場面を想定してA Iネットワーク化が社会・経済の各分野にもたらす影響の評価を行い、A Iネットワーク化の進展に伴い形成されるエコシステムの展望を示すとともに、ネットワーク化の観点からの考察を行い、それらの検討において抽出された課題等を踏まえて、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、各原則の内容に関する論点を整理した上で、今後の課題を掲げている。

第1章で見たとおり、G7香川・高松情報通信大臣会合（平成28年4月）を契機として、海外における議論やG7、OECDにおける国際的な議論が活発に行われているが、主として開発者の立場によるものが多いように見受けられる。本報告書において、国際的な議論のために「A I利活用原則案」が取りまとめたことを踏まえ、利活用に関する国際的な議論も行われることが期待される。

A Iネットワーク化がもたらす社会への影響は、A Iの高度化が加速度的に進展していくことなどに伴い、一度示された予測が続々と塗り替えられる形で、急速に進展しつつある。本報告書の内容は、現時点において検討した結果を述べているものであり、A Iネットワーク化の進展等に応じて不断の見直しを行うことが必要である。

また、第4章で整理した今後の課題は、多岐にわたっており、継続的かつ多面的な検討が必要である。これらの課題の検討に当たっては、産学民官の幅広い分野から関係ステークホルダの参画を得て国内外において検討を進めていくことが必要である。今後、緊急性、重要性等を勘案してプライオリティをつけ、継続的に検討が進められていくことを期待したい。