

報告書2017 (案)

—AI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●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. 本推進会議の目的	
第1章 AIネットワーク化をめぐる最近の動向	10
1. 国内の動向	
2. 海外の動向	
第2章 AI開発ガイドライン	23
1. 背景・経緯	
2. 整理された論点に関する意見募集の結果等を踏まえたAI開発ガイドライン案の検討の方向性	
3. AI開発ガイドライン案の概要	
4. 今後の展開	
第3章 AIネットワーク化が社会・経済にもたらす影響	31
1. 背景・経緯	
2. 先行的評価	
3. 分野別評価	
4. 雇用・働き方への影響	
5. 留意すべき事項	
第4章 今後の課題	48
結びに代えて	61

- <別紙 1> 国際的な議論のための A I 開発ガイドライン案
- <別紙 2-1> 整理された論点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 2-2> 報告書 2017 (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 3> A I ネットワーク化が社会・経済にもたらす影響 ~先行的評価~
- <別紙 4> A I 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 5> A I ネットワーク社会推進会議及び分科会構成員一覧
- <別紙 6> 開催経緯
- <参考資料> 「A I ネットワーク社会推進フォーラム」関連資料

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・経済に多大な便益をもたらされることが期待される一方、不透明化や制御喪失等のリスクも懸念されている。

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、国際的な議論のためのAI開発ガイドライン案の検討その他AIネットワーク化のガバナンスの在り方の検討を行うとともに、AIネットワーク化が社会・経済にもたらすインパクト（主に良い影響や便益）及びリスクの評価を行ってきた。

本報告書は、本推進会議における検討及び関連する国内外の動向を踏まえ、国際的な議論のためのAI開発ガイドライン案の基本的な考え方と内容を示すとともに、様々な分野における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の評価を行った上で、今後の課題を整理している。

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、AI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、総務省の主催により平成29年3月に開催された国際シンポジウム「AIネットワーク社会推進フォーラム」をはじめ最近のAIネットワーク化に関する国内外の動向を概観している。

第2章では、本推進会議が作成した国際的な議論のためのAI開発ガイドライン案（別紙1）について、検討の背景・経緯を概観した上で、検討の方向性を説明するとともに、その概要を紹介している。同ガイドライン案に掲げるAI開発原則は、次のとおりである。

- ① 連携の原則-----開発者は、AIシステムの相互接続性と相互運用性に留意する。
- ② 透明性の原則-----開発者は、AIシステム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。
- ③ 制御可能性の原則-----開発者は、AIシステムの制御可能性に留意する。
- ④ 安全の原則-----開発者は、AIシステムがアクチュエータ等を通じて利用者及び第三者の生命・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-----開発者は、AIシステムのセキュリティに留意する。
- ⑥ プライバシーの原則-----開発者は、AI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。

- ⑦ 倫理の原則-----開発者は、A I 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。
- ⑧ 利用者支援の原則-----開発者は、A I システムが利用者を支援し、利用者に選択の機会を適切に提供することが可能となるよう配慮する。
- ⑨ アカウンタビリティの原則-----開発者は、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。

第3章では、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（ユースケース）を想定したシナリオを作成し、A I 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を分析した上で、その結果を踏まえ、今後の課題を検討するに当たって留意すべき事項を次のように整理している。

- ・ 様々なA I システムが混在することに伴う問題
- ・ A I システムを利活用する者と利活用しない（できない）者との混在に伴う問題
- ・ 事故の発生等A I 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
- ・ データ・情報の流通と個人情報保護・プライバシーとのバランス
- ・ 人間とA I（ロボット）との役割分担 等

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。

- A I ネットワーク化の健全な進展に関する事項
 - A I 開発ガイドライン（仮称）の策定、A I 利活用ガイドライン（仮称）の策定、A I システム相互間の円滑な連携の確保 等
- A I 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
- A I ネットワーク化が社会・経済にもたらす影響の評価に関する事項
 - 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、A I 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 等
- A I 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A I システムとの関係の在り方に関する検討、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 等

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 開発ガイドラインの策定をはじめとするA I ネットワーク化のガバナンスの枠組みの形成に向けた国際的な議論を推進することにより、A I ネットワーク化を通じた人間中心の「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A Iに関する報道を目にしない日がないほど、A Iに関する技術開発と利活用が目まぐるしく進展している。また今後、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」の進展により、個人、地域社会、各国、国際社会の抱える様々な課題¹の解決が促されるなど、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ている。一方、A IやA Iネットワーク化については、不透明化や制御喪失などリスクも懸念されている。

こうした中、平成28年4月に日本で開催されたG7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。

少子高齢化などの課題を抱える日本は、A Iを積極的に開発し活用することにより様々な課題を解決するとともに、その知見を活かしつつA I開発ガイドライン案などA Iネットワーク化のガバナンスの枠組みについて国際的に議論を提起することにより、国際社会に大きく貢献することができるものと思われる。

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、国際的な議論のためのA I開発ガイドライン案の検討²その他A Iネットワーク化のガバナンスの在り方の検討を行うとともに、具体的な利活用の場面を想定したA Iネットワーク化の影響の評価を行ってきた。また、これまでの検討を踏まえ、総務省は平成29年3月に東京大学において国際シンポジウム「A Iネットワーク社会推進フォーラム」を開催した。フォーラムには、日米欧等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われた。フォーラムでは、A Iネットワーク化に関する社会的・経済的・倫理的・法的課題について多様な意見が示されたが、国内外のトップレベルの有識者から、A Iネットワーク化を通じた人間中心の「智連社会」の実現という理念が広く支持されるとともに、グローバルでオープンな議論を継続することを通じて合意形成を図っていくことに賛意が示された。

本推進会議は、平成28年10月から平成29年6月までに開催した15回の会合（本推進会議の会合5回のほか、本推進会議の下に置かれた開発原則分科会及び影響評価分科会の

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 国際的な議論のためのA I開発ガイドライン案の検討に当たっては、平成28年12月28日から平成29年1月31日までにかけて、本推進会議において整理したA I開発ガイドライン案の論点に関する意見募集を行った。

会合5回ずつ)における検討や国際シンポジウム「A I ネットワーク社会推進フォーラム」における議論等を踏まえて、本報告書(案)を作成した。その後、平成29年6月14日から同年7月7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、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議論を踏まえ、A I 開発ガイドライン案の基本的な考え方と内容を示すとともに、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った上で、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 開発ガイドラインの策定をはじめとするA I ネットワーク化のガバナンスの枠組みの形成に向けた国際的な議論を推進することにより、A I ネットワーク化を通じた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。

序章

1. 「A I ネットワーク化」と「智連社会」

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、A I ネットワーク化の進展を通じて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 : W I N S [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 ネットワーク化が社会・経済にもたらす影響（インパクト³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』⁴、同年 6 月に『報告書 2016』⁵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並びに目指すべき社会像としての「智連社会」及びその基本理念を紹介する。

○ 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」⁶（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、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⁷。）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。

³ 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。

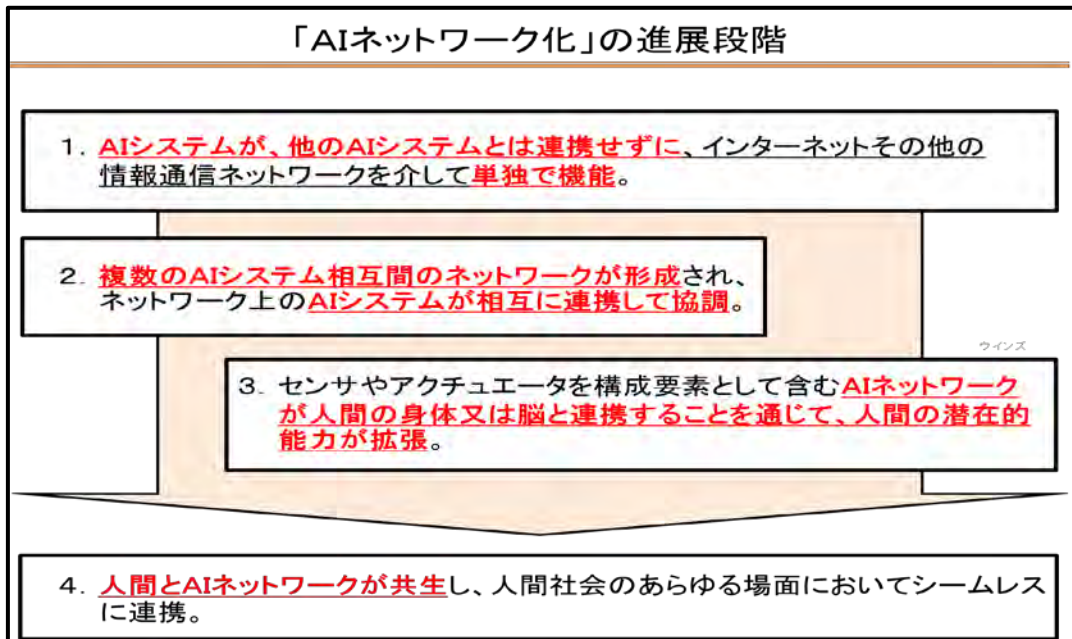
⁴ A I ネットワーク化検討会議「中間報告書 A I ネットワーク化が拓く智連社会（W I N S [ウインズ]）―第四次産業革命を超えた社会に向けて―」（平成 28 年 4 月 15 日。以下「中間報告書」という。）

⁵ 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（W I N S [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。）

⁶ 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等）

からなるものとして理解できる（中間報告書5頁参照。）。

⁷ 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。）。



第一段階は、単独のAI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AI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。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。

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AI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AIシステムが相互に連携して協調するという段階である⁸。

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AI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⁹。

これらの進展が広く進み、AI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AIネットワークとが共生し、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AIシステムと他のAIシステム、人間とAI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。

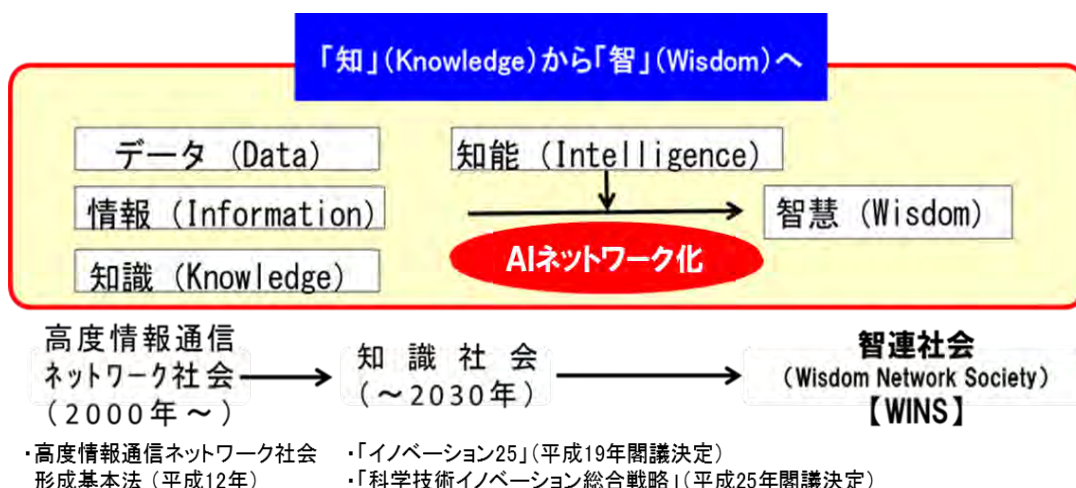
○ 目指すべき社会像「智連社会」

検討会議においては、AIネットワーク化が進展し、人間とAIネットワークとが共

⁸ AI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AI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）「AIネットワーク化の本質と将来」（AI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。

⁹ 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。

生ずる段階において目指すべき社会像として、「智連社会」(Wisdom Network Society : WINS [ウインズ]) を掲げた。この「智連社会」という社会像は、「高度情報通信ネットワーク社会」及び「知識社会」という「データ」・「情報」・「知識」(知) の流通及びその影響に着目した社会像の実現を経て、その次に実現を目指すべき社会像として構想したものであり、その内容は「データ」・「情報」・「知識」(知) を活用する人間の能力としての「智慧」(智) ¹⁰に着目したものである ¹¹。



「データ」・「情報」・「知識」・「知能」・「智慧」の関係	
データ (Data)	断片的な事実、数値、文字
情報 (Information)	データの組み合わせに意味を付与したもの
知識 (Knowledge)	データ・情報の体系的集積
知能 (Intelligence)	データ・情報・知識を学習し、解析することにより、新たなデータ・情報・知識を創造する機能
智慧 (Wisdom)	<u>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</u>

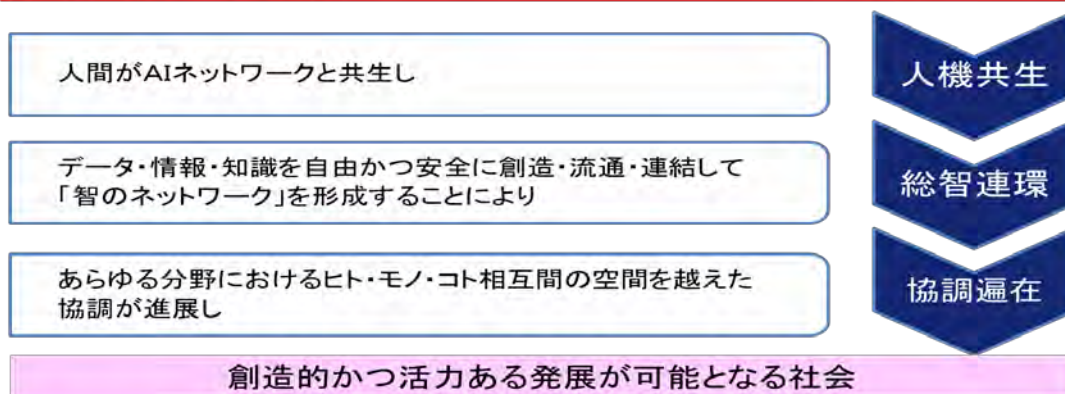
¹⁰ 人間の能力としての「智慧」に関し、濱田顧問 (東京大学名誉教授) 「A I ネットワーク化と智連社会」(A I ネットワーク社会推進フォーラム (特別講演) 配付資料) 2-5 頁 (平成29年) 参照。

¹¹ 中間報告書 12 頁参照。なお、データ (Data)、情報 (Information)、知識 (Knowledge)、知能 (Intelligence) 及び智慧 (Wisdom) の関係の詳細については、中間報告書 13 頁参照。

AIネットワーク化の進展に伴い、AIとAI、AI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AI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。それに伴い、「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)が高まるとともに、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、AIネットワークを利活用することにより、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¹²。

「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、AIネットワーク化の健全な進展の結果として、人間がAIネットワークと主体的に共生し、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。

智連社会(Wisdom Network Society【WINS】)は、人間が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。



智連社会は、AIネットワーク化が進展して「人間とAIネットワークとが共生」す

¹² ここで、AIネットワークは、AI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AI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。

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生」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生」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¹³との考えを含意している。

「智連社会」においては、人間がA I ネットワークを主体的に使いこなし、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A I 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¹⁴。

○ 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、A I 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A I ネットワークの恵沢をあまねく享受できること。

- ・ **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**

¹³ 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（8）1頁参照。

¹⁴ 濱田・前掲注（10）9頁（平成29年）参照。

AI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**

AI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**

AI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**

AI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、AI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、AI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AI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. 本推進会議の目的

検討会議においては、AI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、AIの開発研究に当たって国際的に参照されるべき「AI開発原則」及びその内容の解説からなる「AI開発ガイドライン」の策定に向けた議論の推進、AI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G7香川・高松情報通信大臣会合が開催され、高市総務大臣から、「AI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「AI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、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り、OECD等におけるAI開発ガイドラインの策定に向けた国際的な議論のための案の作成を進

めるとともに、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等を進めている。

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、主として本推進会議が発足（平成28年10月31日）した後のAIネットワーク化をめぐる動向を概観する¹⁵。

1. 国内の動向

(1) 「AIネットワーク社会推進フォーラム」(国際シンポジウム)

AI開発ガイドラインの策定などAIネットワーク化をめぐる社会的・経済的・倫理的・法的課題に関し、日米欧等の産学民官のステークホルダが意見交換し、今後のG7、OECD等における国際的な議論を推進することを目的として、平成29年3月13日及び14日に東京大学伊藤謝恩ホールにおいて、総務省の主催により「AIネットワーク社会推進フォーラム」(以下「フォーラム」という。)を開催した¹⁶。

フォーラムにおいては、本推進会議の顧問を含む構成員のほか、米、伊、独、仏、EUの政府関係者やOECDの代表者、AIの分野の国際的な議論をリードする主要団体¹⁷の代表者を含む国内外のトップレベルの有識者が参加し、意見交換を行った。

フォーラムにおける議論を通じて、主に次の事項に関し概ね意見の一致を見ることができたものと考えられる。

- ・ グローバルな議論を通じた合意形成
 - 多様なステークホルダによるオープンでグローバルな議論を通じた合意形成が重要である。
 - 継続的に議論を行っていくことが重要である。
- ・ 人間中心の社会の構築
 - AIネットワーク化の便益が広く享受される人間中心の社会を構築すべきである。
 - 教育、人材育成が重要である。
- ・ AIネットワーク化のガバナンス
 - 非規制的かつ非拘束的なアプローチ(いわゆるソフトロー)が適切である。
 - 開発者は、説明責任を果たすことが重要である。
 - 連携、透明性、制御可能性、安全性、プライバシー、セキュリティ、公正などの観点が必要である。

¹⁵ 本推進会議の発足以前のAIネットワーク化をめぐる動向については、報告書2016第1章参照。

¹⁶ 詳細は、参考資料参照。

¹⁷ Partnership on AI、Future of Life Institute(FLI)等。これらの動向については、本章2.参照。

なお、フォーラム後、同年4月5日に開催された日米インターネット・エコノミー対話民間会合において、米国側の参加者から、フォーラムにおいて議論された人間中心の社会を構築すべき等の考え方やAI開発ガイドライン策定の必要性について強い賛同が示されるとともに、標準化の重要性についての指摘があった。

(2) 「世界最先端IT国家創造宣言・官民データ活用推進基本計画」

平成29年5月30日付けで閣議決定された「世界最先端IT国家創造宣言・官民データ活用推進基本計画」¹⁸において、AIネットワーク化について、次のとおり記載された。

- ・ AIネットワーク化が社会・経済にもたらすインパクトやリスクの評価の国際的な共有、関連する社会的・経済的・倫理的・法的課題の解決に資するガバナンスの在り方に関するG7、OECD等の場における国際的な議論を通じた検討の推進。

(3) 関連する動向

① 人工知能と人間社会に関する懇談会

第5期科学技術基本計画¹⁹（平成28年1月22日閣議決定）に掲げられた「Society5.0」²⁰の実現の鍵である人工知能の研究開発及び利活用を健全に進展させるべく、人工知能と人間社会の関わりについて検討を行うため、内閣府特命担当大臣（科学技術政策）の下に「人工知能と人間社会に関する懇談会」が設置され²¹、平成

¹⁸ 「世界最先端IT国家創造宣言・官民データ活用推進基本計画」は、従来の「世界最先端IT国家創造宣言」に、官民データ活用推進基本法（平成28年法律第103号）第8条に規定する官民データ活用の推進に関する基本的な計画（官民データ活用推進基本計画）を加えた一体のものとして、高度情報通信ネットワーク社会推進戦略本部及び官民データ活用推進戦略会議が決定した上で、その内容に即して閣議決定がなされたものである。平成29年においては、高度情報通信ネットワーク社会推進戦略本部及び官民データ活用推進戦略会議が同年5月30日付けで決定した上で、同日付けで、当該決定の内容に即して同宣言・計画を決定する旨の閣議決定がなされた。

¹⁹ 科学技術基本法（平成7年法律第130号）に基づき、政府が長期的視野に立って体系的かつ一貫した科学技術政策を実行するために策定された基本計画。平成28年1月22日付けで、平成28年度～32年度の第5期基本計画が閣議決定された。

²⁰ 「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に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、AI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。

²¹ この懇談会には、本推進会議の構成員のうち、江間開発原則分科会・影響評価分科会構成員、大内影響評価分科会構成員、新保構成員、鈴木幹事及び松尾構成員が構成員に就いている。また、この懇談会の事務局の事務を掌る内閣府政策統括官（科学技術・イノベーション担

28年5月30日に第1回会合が開催された。以降、平成29年1月20日までに6回の会合が開催され、同年3月24日に報告書を取りまとめ、公表した²²。

報告書においては、移動、製造、個人向けサービス、対話・交流の4つの代表的分野について様々な事例（ケース）を挙げて検討する手法がとられ、人工知能技術は人間の知的能力と行為を補助し、一部を代替し拡張することを可能とすることから、持続可能社会の強力な推進力になることが期待できるとした上で、倫理的論点、法的論点、経済的論点、教育的論点、社会的論点、研究開発的論点について検討し考慮する必要があるとしている。

② 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）²³においては、平成28年度の推進テーマの一つとして「人工知能間の交渉・協調・連携による社会の超スマート化」に関する検討が産学官連携の下で行われることとなり²⁴、平成28年5月23日にキックオフ会合が開催された。以降、平成29年1月19日までに4回の会合が開催され、同年2月15日に報告書を取りまとめ、公表した²⁵。

報告書においては、製造バリューチェーン自動接続、スマートシティ：電力・水、スマートシティ：交通・人流、自動運転車・移動体の4つのユースケースが特に早急な取組が重要なユースケースとして選定されており、産業競争力強化のための提言及び施策として、

- ・ 相互接続性・運用性が本質になるアイテム群を戦略的大型NP（National Project）で一気に一体開発
- ・ 複数ユースケースを同時実装する特区等の指定により、必要な実証実験、詳細制度設計と改善、社会受容性の醸成、ネットワーク外部性に起因する普及の壁の突破
- ・ 国際協調の場での議論を牽引し、標準や制度を日本のニーズに合った形に国際的に収斂

を柱とした提言がなされている。

当）付参事官が本推進会議のオブザーバに就いている。

²² 次に掲げるURLのウェブサイト在所掲。

<<http://www8.cao.go.jp/cstp/tyousakai/ai/summary/index.html>>

²³ 国の持続的発展の基盤となる産業競争力を高めるため「科学技術政策」、「イノベーション政策」を提言として取りまとめ、政府に対して、その推進と支援の要請を行い、実現を図ることを目的として、平成18年に発足した民間団体（平成27年に一般社団法人となる。）。

²⁴ COCNによる本テーマの検討のプロジェクトには、本推進会議の三部開発原則分科会構成員及び落合影響評価分科会構成員がメンバーに、本推進会議の事務を統括する総務省情報通信政策研究所調査研究部長がオブザーバ/アドバイザーに就いている。また、COCNの事務局長が本推進会議のオブザーバに就いている。

²⁵ 次に掲げるURLのウェブサイト在所掲。

<<http://www.cocn.jp/thema94-L.pdf>>

また、報告書においては、本推進会議の議論が随所に参照されているほか、本推進会議とCOCNとの連携について指摘されている。

③ 人工知能技術戦略会議

平成28年4月12日に開催された「未来投資に向けた官民対話（第5回）」において、安倍内閣総理大臣から、人工知能の研究開発目標と産業化のロードマップを同年度中に策定するよう指示があり、この指示を受け、「人工知能技術戦略会議」が創設された²⁶。同会議が司令塔となって、総務省、文部科学省及び経済産業省が中心となり、AIを利用する側の産業の関係省庁とも連携し、人工知能技術の研究開発と社会実装に向けた検討を進めることとなり、同月18日に第1回会合が開催された。以降、平成29年3月17日までに5回の会合が開催され、同月31日に「人工知能技術戦略」及び「人工知能の研究開発目標と産業化のロードマップ」を取りまとめ、公表した²⁷。

「人工知能技術戦略」においては、人工知能とその他関連技術の融合による産業化のロードマップが示され、当面取り上げるべき重点分野²⁸である「生産性」、「健康、医療・介護」、「空間の移動」の各分野について、目指すべき社会像とその実現に向けたフェーズごとの産業化のイメージが取りまとめられている。また、ロードマップの実現に当たっては、産官学の叡智を結集し取り組むことが必要であるが、3センター²⁹をはじめとする国の機関は、基盤的な技術の開発、高度な人材育成、公的データの整備、ベンチャー支援など産学官のプラットフォームの役割を担うべきであるとして、それぞれ具体的な取組例が示されている。

④ 人工知能学会

人工知能学会の倫理委員会³⁰は、平成28年6月6日に開催された全国大会において、人工知能の研究開発者が遵守すべき倫理綱領の案を提示し、公開討論に付した。その後、意見募集で提出された意見等を踏まえて改訂作業を進め、同年12月に倫理

²⁶人工知能技術戦略会議の議長には、本推進会議の安西顧問が就いている。

²⁷ それぞれ次に掲げるURLのウェブサイトに掲載。

人工知能技術戦略：<<http://www.nedo.go.jp/content/100862413.pdf>>

ロードマップ：<<http://www.nedo.go.jp/content/100862412.pdf>>

²⁸ 産業化ロードマップとして当面、取り上げるべき重点分野を、①社会課題として喫緊の解決の必要性、②経済波及効果への貢献、③人工知能技術による貢献の期待、の観点から検討した結果、「生産性」、「健康、医療・介護」、「空間の移動」の3分野に加え、横断的な分野として「情報セキュリティ」の4つの分野が特定された。

²⁹ 情報通信研究機構 脳情報通信融合研究センター（C i N e t）、ユニバーサルコミュニケーション研究所（UCRI）、理化学研究所 革新知能統合研究センター（A I P）及び産業技術総合研究所 人工知能研究センター（A I R C）

³⁰ 人工知能学会倫理委員会には、本推進会議の松尾構成員が委員長に、江間開発原則分科会・影響評価分科会構成員、武田開発原則分科会構成員、西田幹事及び堀幹事・開発原則分科会技術顧問が委員に、栗原開発原則分科会構成員及び山川構成員がオブザーバに就いている。

要綱案の改訂版を策定した。さらに、倫理の専門家や編集委員会の意見を踏まえて改訂版に修正を施し、平成 29 年 2 月 28 日に『人工知能学会倫理指針』が理事会で承認され、公表した³¹。

本倫理指針は、人工知能学会員の倫理的な価値判断の基礎となるものとして定められた指針であり、同学会員はこれを指針として行動するよう心がけるものとされ、1. 人類への貢献、2. 法規制の遵守、3. 他者のプライバシーの尊重、4. 公正性、5. 安全性、6. 誠実な振る舞い、7. 社会に対する責任、8. 社会との対話と自己研鑽、9. 人工知能への倫理遵守の要請、から構成されている。特に第 9 条については、本指針に従って作られた人工物に対しても本倫理指針が適用されるという再帰性を含んでおり、特徴的なものであるとされている³²。

⑤ 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） データ流通環境整備検討会 AI、IoT 時代におけるデータ活用ワーキンググループ

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）データ流通環境整備検討会において、多種多様かつ大量のデータを安全・安心に流通・利活用できる環境整備に必要な措置の検討に資するため、AI、IoT 時代におけるデータ活用ワーキンググループが設置され³³、平成 28 年 9 月 30 日に第 1 回会合が開催された。以降、平成 29 年 2 月 24 日までに 9 回の会合が開催され、同年 3 月に中間とりまとめを公表した³⁴。

中間取りまとめにおいては、分野横断的なデータ流通を阻害する課題を解決し、パーソナルデータを含めた多種多様かつ大量のデータの円滑な流通を実現するためには、個人の関与の下でデータの流通・活用を進める仕組みである PDS（Personal Data Store）、情報銀行、データ取引市場が有効であるとして、観光分野、金融・フィンテック分野、医療・介護・ヘルスケア分野、人材分野、農業分野、防災減災分野（災害発生時）の 6 分野のユースケースを整理している。

⑥ 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

³¹ 次に掲げる URL のウェブサイト に所掲。

<<http://ai-elsi.org/archives/471>>

³² 人工知能学会倫理委員会委員長・松尾豊「人工知能学会 倫理指針について」（平成 29 年 2 月 28 日）参照。次に掲げる URL のウェブサイト に所掲。

<<http://ai-elsi.org/archives/471>>

³³ このワーキンググループには、本推進会議の大橋構成員、宍戸構成員・開発原則分科会長代理及び新保構成員が構成員に就いている。また、このワーキンググループの事務局の事務を掌る内閣官房情報通信技術（IT）総合戦略室参事官が本推進会議のオブザーバに就いている。

³⁴ 次に掲げる URL のウェブサイト に所掲。

<http://www.kantei.go.jp/jp/singi/it2/senmon_bunka/data_ryutsuseibi/dai2/siryou2.pdf>

知的財産戦略本部 検証・評価・企画委員会において、A I の利活用が様々な産業に広がるのが想定される中、A I 生成物とその生成過程である「学習用データ」や「学習済みモデル」等の新たな情報財の保護・利活用の在り方について検討を行うため、新たな情報財検討委員会が設置され³⁵、平成 28 年 10 月 31 日に第 1 回会合が開催された。以降、平成 29 年 3 月 13 日までに 7 回の会合が開催され、同年 3 月に報告書を取りまとめ、公表した³⁶。

報告書においては、A I の作成・利活用促進のための知的財産制度の在り方について、具体的に検討を進めるべき事項等として、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が挙げられるとともに、A I 生成物の知財制度上の在り方について、具体的な事例を継続的に把握しつつ、引き続き検討すべきなどとされている。

2. 海外の動向

(1) 米国ホワイトハウス

ホワイトハウスは、2016 年（平成 28 年）5 月 3 日付けで、A I がもたらす利益とリスクに関する検討を開始することを発表し、同年 5 月から 7 月にかけて、大学や N P O 等とともに、全米各地において計 4 回のワークショップを共催した³⁷。一連のワークショップでは、産業界、学界、市民社会、政府機関等のステークホルダが参加し、A I と法・統治の関わり、社会的利益のための A I、A I の安全性と制御、A I 技術と社会・経済の関わり等について議論が行われた³⁸。

この一連のワークショップにおける議論及びその後の意見募集の結果を踏まえ、同年 10 月 12 日、国家科学技術会議（N S T C）及び科学技術政策局（O S T P）が中心となって取りまとめた報告書「人工知能の未来に備えて」を公表した。

報告書においては、規制制度、研究開発、経済・雇用、公正性・安全性、安全保障など幅広い項目が検討され、各項目に対応して、連邦政府機関等に対する 23 の提言がまとめられている。報告書の結論部では、政府が果たすべき役割として、

- ・ 重要な問題について対話の場を設け、公の討論のためのアジェンダ設定を助けるこ

³⁵ この委員会には、本推進会議の喜連川幹事及び福井開発原則分科会・影響評価分科会構成員が委員に就いている。

³⁶ 次に掲げる URL のウェブサイトに掲載。

<http://www.kantei.go.jp/jp/singi/titeki2/tyousakai/kensho_hyoka_kikaku/2017/johoza_i/houkokusho.pdf>

³⁷ シアトルで開催されたホワイトハウスとワシントン大学の共催によるワークショップには、主催者側からの招へいを受けて、総務省情報通信政策研究所調査研究部主任研究官が参加し、非公式会合において検討会議における A I 開発ガイドライン案に関する検討を紹介するとともに、参加者らと意見交換を行った。

³⁸ <https://obamawhitehouse.archives.gov/blog/2016/05/03/preparing-future-artificial-intelligence>

と

- ・ 利用の安全性及び公正性を注視し、規制の枠組みを公衆を守りながらイノベーションを促進するものに適合させること
- ・ 熟練した多様な労働力の開発に加えて、基礎研究及びA Iの公益への利用を支援すること
- ・ より早く、より効果的に、より低いコストで公衆に奉仕するため、政府が自らA Iを利用すること
- ・ A Iによる公共政策の各分野への変化を理解し適応するため、その能力を涵養し続けること

が掲げられるとともに、実務家 (practitioner) が確保すべき事項として、A I対応システムが、

- ・ 統御可能であること
- ・ オープンで、透明で、理解可能であること
- ・ 人々と効果的に機能し得ること
- ・ その操作は人間の価値及び願望と一致し続けるであろうこと

が掲げられている³⁹。

また、同月 13 日付けで、N S T C ネットワーキング・情報技術研究開発小委員会は、連邦政府の予算によるA I研究（連邦政府自身のA I研究のみならず、連邦政府の助成を受けた大学等のA I研究を含む。）の方針として「米国人工知能研究開発戦略」を策定した。戦略においては、連邦政府の予算によるA I研究の目標として、社会に便益をもたらす新たなA Iに関する知識及び技術を生み出しつつ、ネガティブな影響を最小化することが掲げられ、この目標を実現するために優先的に取り組むべき事項が設定されている⁴⁰。

さらに、ホワイトハウスは、同年12月20日付けで、報告書「人工知能、自動化、そして経済」を公表した。報告書においては、A Iによる自動化が経済や雇用に与える影響が分析され、A Iの開発への投資、A Iがもたらす変化に対応した教育改革、労働者の支援などの政策的対応が提言されている⁴¹。⁴²

³⁹ White House, *Preparing for the Future of Artificial Intelligence* (2016).

⁴⁰ National Science and Technology Council Networking and Information Technology Research and Development Subcommittee, *The National Artificial Intelligence Research and Development Strategic Plan* (2016).

⁴¹ White House, *Artificial Intelligence, Automation, and the Economy* (2016).

⁴² ホワイトハウスにおけるA Iに関する検討について中心的な役割を担ったエドワード・フェルテン氏（前・ホワイトハウス科学技術政策局C T O補佐官、プリンストン大学教授）がフォーラムに参加し、基調講演及びA Iネットワーク化のガバナンスに関するパネルディスカッションに登壇し、ホワイトハウスにおけるA Iに関する検討を紹介するとともに、本推進会議の構成員らと意見交換を行った。

(2) Partnership on A I

2016年(平成28年)9月28日、A I技術のベストプラクティスを研究して形成し、A Iに関する公衆の理解を向上させ、A I及びその社会的影響に関する議論と関与のためのオープンなプラットフォームとして、Partnership on Artificial Intelligence to Benefit People and Society (Partnership on A I) が設立された⁴³。

Partnership on A Iは、設立に当たり次に掲げる事項を「信条 (Tenets)」として公表した。

- ・ A I技術ができるだけ多くの人々に裨益し、能力を与えるよう努める。
- ・ 一般市民を啓蒙し、その声を傾聴するほか、積極的にステークホルダを関与させて、焦点に対する意見を求め、我々の活動を周知させ、疑問に対処する。
- ・ A Iの倫理的・社会的・経済的・法的影響に関するオープンな研究と対話を約束する。
- ・ A Iの研究開発の努力は、広範なステークホルダの積極的な関与を得るとともに、彼らに説明可能なものとする必要がある。
- ・ 領域固有の懸念及び機会が理解され対処されることの確保に資するよう、ビジネスコミュニティのステークホルダに関与させ、その代表者の参画を得る。
- ・ A Iの便益を最大化し、A I技術の潜在的な課題に対処するため、個人のプライバシーとセキュリティの保護、すべての当事者の利益の理解と尊重、A Iの潜在的影響に関する社会的責任の確保、A Iの研究・技術の頑健性、信頼性、信用性、堅牢性の確保、国際条約や人権に反するA I技術の開発・利用への反対、危害を与えないセーフガードと技術の推進に取り組む。
- ・ A Iシステムの動作は、その技術を説明するため、人々の理解と解釈が可能なものであることが重要である。
- ・ 上記の目標の達成に資するよう、A Iの科学者や技術者間の協力、信用及び開放性の文化を創造するよう努める⁴⁴。⁴⁵

⁴³ 設立当初の企業構成員は、Amazon、DeepMind/Google、Facebook、IBM、Microsoftの5社で、暫定共同議長は、エリック・ホロヴィッツ氏 (Microsoft 技術フェロー) 及びムスタファ・スレイマン氏 (DeepMind 共同創業者) である (<<https://www.partnershiponai.org/>>)。2017年(平成29年)1月21日には、Appleが企業構成員に、米国人工知能学会、米国自由人権協会、オープンAIが非営利構成員に新たに加わるとともに、評議員に、企業構成員各社を代表する6名に加え、研究者や人権擁護団体の役職者ら6名が独立評議員として就任したことが公表された

(<<https://www.partnershiponai.org/2017/01/partnership-ai-update/>>)。また、同年5月16日には、ソニー、イーベイ、インテル、マッキンゼー・アンド・カンパニーなど8社が新たに企業構成員に、ユニセフ、電子自由財団、ヒューチャー・オブ・ヒュ머니ティ・インスティテュート (F H I) など14の組織・機関が新たに非営利構成員に加わったことが公表された (<https://www.partnershiponai.org/2017/05/pai-announces-new-partners-and-initiatives/>)。

⁴⁴ <https://www.partnershiponai.org/tenets/>

⁴⁵ Partnership on A Iから、暫定共同議長のエリック・ホロヴィッツ氏 (Microsoft 技術フェ

(3) 米国電気電子学会 (IEEE)

電気電子学会 (IEEE) の人工知能及び自律的システムにおける倫理的考慮のためのグローバル・イニシアティブ⁴⁶は、2016年(平成28年)12月13日に報告書「倫理的に調整された設計—人工知能及び自律的システムにより人間の福利を優先させるビジョン 第1版—公共の議論のために」を公表した。

報告書においては、人工知能(AI)／自律的システム(AS)の潜在能力から十分に便益を得るためには、それらの技術が我々の道徳的価値及び倫理的原則に照らして人間と調和するように確保することが必要であるとの問題意識に基づいて、道徳的価値及び倫理的原則に照らして人間と調和するAIを設計する方法について技術者の議論を促すための論点が提示されている。

また、あらゆる種類のAI／ASに適用される高度な倫理的配慮を定式化し、

- ・ 人権に関する最も高度な理念の具現化
- ・ 人間及び自然環境への最大限の便益の優先
- ・ AI／ASが社会技術システムとして発展することに伴うリスクと負の影響の緩和

という3つの原則が掲げられている。

さらに、AI／ASに人間の規範や価値を実装するための方法論を検討した上で、AI／AS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⁴⁷。

報告書については、公表されてから2017年(平成29年)3月まで意見募集が行われた。意見募集の結果やIEEEにおける議論を踏まえ、報告書の第2版が同年秋頃に公表される予定である。⁴⁸

ロー)及び評議員のグレッグ・コラード氏(Google Inc. 主席研究員)がフォーラムに参加した。ホロヴィッツ氏は特別遠隔講演・特別遠隔インタビュー(インタビュアーは本推進会議の堀幹事・開発原則分科会技術顧問)に登壇し、コラード氏は講演及びAIネットワーク化のガバナンスに関するパネルディスカッションに登壇し、Partnership on AIや企業のAIに関する取組を紹介するとともに、本推進会議の構成員らと意見交換を行った。

⁴⁶ イニシアティブは、その目標として、技術者コミュニティの検討作業のための重要な参照点となる認識及び提言(案)を提示すること並びにIEEEの標準(standard)策定のための提言を行うことを掲げている。

⁴⁷ 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).

⁴⁸ IEEEグローバル・イニシアティブに参加しているサイラス・ホーディス氏(ハーバード大学ケネディ行政大学院ザ・フューチャー・ソサイエティAIイニシアティブ共同創設者)がフォーラムに参加し、AIネットワーク化がもたらす豊かさと幸せに関するパネルディスカッションに登壇し、AIの倫理的課題等に関する議論を紹介するとともに、本推進会議の構成員らと意見交換を行った

(4) 英国下院

英国下院は、2016年（平成28年）10月12日付けで、科学技術委員会が同年9月13日付けで取りまとめた報告書「ロボティクスと人工知能」を公表した。

報告書においては、ロボット・AIに関する社会的・倫理的・法的問題を検討した上で、政府に対して提言が行われている。AIのガバナンスについては、AIという生成途上の領域に包括的な規制を設けることは時期尚早である一方、AIシステムの倫理的、法的及び社会的次元に関する注意深い精査を開始することが不可欠であるとした上で、実効的なガバナンスが継続的に評価され実行されるよう、検証と妥当性確認、意思決定の透明性、偏見の最小化、アカウントビリティの向上、プライバシー、安全等の倫理的論点について継続的に注視することが必要であるとされている。また、AIの発展に関する社会的、倫理的及び法的含意について検討するため、アラン・チューリング研究所⁴⁹にAIに関する常設の委員会を設置するよう提言し、この委員会は、AI技術の開発及び利用を統治するための原則の策定等に注力すべきと勧告している⁵⁰。

(5) 欧州議会

欧州議会法務委員会は、2016年（平成28年）4月にロボット・AIの法的・倫理的問題に関する公聴会を開催するなどロボットに関する法的問題の検討を進めてきており、同年5月にロボットの民事法的ルールに関する報告書の草稿を公表し、同年10月には関連するワークショップを開催した。

一連の検討を踏まえ、欧州議会は、2017年（平成29年）2月16日付けで、「ロボティクスに係る民事法的規則に関する欧州委員会への提言」を採択した。提言では、EUはロボット・AIの開発・設計・利用において尊重されるべき基本的な倫理原則を策定する上で不可欠な役割を果たせるとの認識が示されるとともに、倫理指針の枠組みは、恩恵、無害、自律及び正義の原則並びに人間の尊厳など欧州連合条約第2条及び欧州連合基本権憲章に定められた原則・価値に依拠すべきであるとされている。その上で、ロボット・AIに関する民事法上の問題を中心に検討を行い、ロボット・AIを所管するEU機関の設置、スマート・ロボットの登録、危害に関する損害賠償責任、知的財産権、相互運用性等の確保等について定めるEU法の策定を欧州委員会に求める提言がなされている。また、基本権の尊重、予防原則、包摂性、アカウントビリティ、安全性、可逆性、プライバシー、便益の最大化と危害の最小化などを内容とするロボット開発者向

⁴⁹ データサイエンスに関する英国の国立研究機関である。ケンブリッジ大学など5つの大学と英国工学・物理科学研究会議により設立された研究所であり、大英図書館に本部を置いている。

⁵⁰ House of Commons Science and Technology Committee, *Robotics and artificial Intelligence: Fifth Report of Session 2016-17* (2016).

けの倫理的行動規範の策定についても提言されている⁵¹。⁵²

(6) Future of Life Institute (FLI)

Future of Life Institute (FLI) は、ジャン・タリン氏 (Skype 共同開発者、DeepMind 共同出資者)、イーロン・マスク氏 (SpaceX 創業者、Tesla Motors 共同創業者)、スティーブン・ホーキング氏 (宇宙物理学者) ら起業家や研究者等により 2014 年 (平成 26 年) 3 月に設立された非営利団体であり、頑健かつ有益な AI を実現するための研究等を実施している⁵³。

FLI は、2017 年 (平成 29 年) 1 月に米国アシロマにおいて学界や産業界等の関係者の参加を得て AI の研究開発に関する会合を開催し、その成果として、同年 2 月に「アシロマ AI 原則」を公表した。原則は、次に掲げる 23 の項目から構成されている。

- ・ 有益な知性の開発という研究の目標
- ・ AI の有益な利用のための研究への投資
- ・ 科学と政策の連携
- ・ AI 研究における協調、信頼、透明性の文化の促進
- ・ AI システムの安全基準を低下させる競争の回避
- ・ AI システムの安全性とセキュリティ
- ・ AI システムの事故時の透明性
- ・ 司法の意思決定に関与する自律的システムの説明可能性
- ・ AI システムの利用、悪用、動作に関する設計者及び製造者の責任
- ・ 高度に自律的な AI システムの人間の価値への調和
- ・ AI システムの設計及び操作における人間の尊厳、権利、自由、文化的多様性の理念への適合
- ・ パーソナルデータに関する個人のアクセス、管理、コントロールの権利
- ・ パーソナルデータの利活用時における自由とプライバシーの尊重
- ・ 可能な限り多くの人々の便益の享受と能力の向上
- ・ AI により生み出される経済的便益の人類による共有
- ・ 人間による AI システムの制御
- ・ 社会的・市民的プロセスの破壊の禁止
- ・ 自律型兵器による軍拡競争の回避
- ・ 将来の AI の能力の上限の仮定の回避

⁵¹ European Parliament, *European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics* (2015/2103(INL)).

⁵² 欧州議会から、ロバート・ブレイ氏 (法務委員会事務局課長) がフォーラムに参加し、講演及び AI ネットワーク化のガバナンスに関するパネルディスカッションに登壇し、欧州議会における AI に関する検討を紹介するとともに、本推進会議の構成員らと意見交換を行った。

⁵³ <https://futureoflife.org/>

- ・ 高度なA I がもたらす根本的な変化の重大性への注意
- ・ A I システムがもたらすリスクの軽減
- ・ 再帰的に自己改良を行うA I システムの安全性と制御
- ・ 超知能 (Superintelligence) の開発における倫理的理念と人類の便益の追求

F L I は、インターネット上でこの原則への署名を募っており、現在、世界各国からA I ・ロボット研究者を中心に3,600名以上の個人から賛同を得ている⁵⁴。⁵⁵

(7) O E C D (経済協力開発機構)

O E C D は、2016年(平成28年)11月17日に人工知能をテーマとする技術予測フォーラム (Technology Foresight Forum : T F F) を開催した。技術予測フォーラムには、本推進会議から、平野幹事・開発原則分科会長とクロサカタツヤ開発原則分科会・影響評価分科会構成員が参加し、A I 開発ガイドライン等に関する本推進会議の検討状況を紹介した。加盟各国の参加者やO E C D 事務局らから、A I 開発ガイドラインの策定に向けて国際的な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)5月17日から19日にかけて、デジタル経済政策委員会 (Committee on Digital Economy Policy : C D E P) が開催され、総務省から3月に開催されたフォーラムの成果等を報告したところ、加盟各国の参加者やO E C D 事務局らから、A I の開発等に関するガイドラインの策定に向けてO E C D においても議論と検討を進めていくことに賛同する姿勢が示された。

今後、同年10月にパリで総務省とO E C D の共催により開催される国際シンポジウム、同年11月に開催されるO E C D デジタル経済政策委員会 (C D E P) 等において、A I 開発ガイドラインの策定などA I ネットワーク化の推進に関する国際的な議論が進められることが期待される⁵⁶。

(8) G 7

2016年(平成28年)4月29日及び30日に開催されたG 7香川・高松情報通信大臣会合(以下「G 7会合」という。)において、高市早苗総務大臣から、G 7各国が中心となり、O E C D 等国際機関の協力も得て、A I ネットワーク化が社会・経済に与える

⁵⁴ <https://futureoflife.org/ai-principles/>

⁵⁵ F L I から、ジャン・タリン氏 (F L I 共同創設者) がフォーラムに参加し、講演及びA I ネットワーク化のガバナンスに関するパネルディスカッションに登壇し、F L I におけるA I に関する議論を紹介するとともに、本推進会議の構成員らと意見交換を行った。

⁵⁶ O E C D から、ダグラス・フランツ氏 (事務次長)、ウォンキ・ミン氏 (デジタル経済政策委員会議長) 及びアン・カブラン氏 (科学技術イノベーション局デジタル経済政策課長) がフォーラムに参加した。フランツ氏はビデオ講演を行い、ミン氏は講演に登壇し、カブラン氏は講演及びA I ネットワーク化のガバナンスに関するパネルディスカッションに登壇し、A I 開発ガイドライン策定への期待を表明するとともに、本推進会議の構成員らと意見交換を行った。

影響やA Iの開発原則の策定などA I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。

2017年（平成29年）5月26日及び27日にタオルミナにおいてG7首脳会議が開催され、その成果の一つとして付属文書「イノベーション、技能及び労働に関するG7人間中心の行動計画」が取りまとめられた。付属文書においては、包摂的な経済成長及び経済発展のための不可欠な牽引役として、人工知能の開発及び利用を前進することが掲げられた。これを受けて、今年9月にトリノで開催されるG7情報通信・産業大臣会合において、人工知能の開発及び利用の前進に向けた議論が深められる予定である⁵⁷。

⁵⁷ 2017年（平成29年）のG7議長国であるイタリアから、ベネデッタ・アレーゼ・ルチーニ氏（イタリア共和国経済財政大臣顧問）及びマルコ・ジリ氏（トリノ工科大学長）がフォーラムに参加した。ルチーニ氏はイタリア経済財政大臣特別メッセージ及びA Iネットワーク化のガバナンスに関するパネルディスカッションに登壇し、ジリ氏は基調講演及びA Iネットワーク化と人間社会に関する座談会に登壇し、G7におけるA Iネットワーク化に関する議論の推進を表明等するとともに、本推進会議の構成員らと意見交換を行った。

第2章 AI開発ガイドライン

1. 背景・経緯

平成28年4月29日及び30日に開催されたG7会合において、高市早苗総務大臣から、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、このうち主に開発原則分科会において、国際的な議論のためのAI開発ガイドライン案の作成に向けた検討を行ってきた。同年12月28日から平成29年1月31日にかけて、本推進会議において整理したAI開発ガイドライン案の論点に関する意見募集を行い、産学民の多様なステークホルダから意見が寄せられた。

その後、平成29年3月13日及び14日に開催されたフォーラム（国際シンポジウム）⁵⁸において、AIネットワーク化のガバナンスについて、人間中心の社会の実現を目的とすべきこと、非規制的かつ非拘束的なアプローチ（いわゆるソフトロー）が適切であること、開発者は説明責任を果たすことが重要であること、連携、透明性、制御可能性、安全性、プライバシー、セキュリティ、公正などの観点が重要であること、といった議論がなされ、多様なステークホルダによるオープンでグローバルな議論を通じた合意形成が重要であり継続的に議論を行っていくことが重要であることに関し概ね意見の一致を見ることができたものと考えられる。

なお、第1章で詳述しているとおり、この間、欧米各国においても政府や民間においてAIに関する社会的・経済的・倫理的・法的課題に関する検討が加速し、活発な議論が行われた。フォーラム等における意見交換等を通じて、こうした海外の検討や議論についても情報収集を行い、本推進会議におけるAI開発ガイドライン案の検討に当たり参考にしている。

2. 整理された論点に関する意見募集の結果等を踏まえたAI開発ガイドライン案の検討の方向性

意見募集に寄せられた意見等を踏まえ、AI開発ガイドライン案について、次のような方向で検討を進めてきた⁵⁹。

⁵⁸ 詳細については、第1章1. (1)及び参考資料参照。

⁵⁹ 意見募集に寄せられた主な意見の概要とそれらに対する考え方については、別紙2参照。提

- ・ AI開発ガイドラインを、開発者が遵守すべき基準を定めるものとしてではなく、開発者が留意して対応し、対応状況について利用者等に対しアカウンタビリティを果たすことが期待される事項に関し国際的に共有されることが期待される非拘束的な指針を示すものとして策定することを明確にした。
- ・ AI開発ガイドラインの策定及び運用に当たっては、関連する多様なステークホルダ間のコミュニケーションを図り、ベストプラクティスの国際的な共有に資するものとするを明確にした。
- ・ AI開発ガイドラインにおける用語としての「AI」を
 - (ア) ソフトウェアたる「AIソフト」
 - (イ) 「AIソフト」を構成要素として含むシステムである「AIシステム」を総称する概念とし、「AIソフト」及び「AIシステム」の定義を示した上で、「AIソフト」及び「AIシステム」の関係を整理することとした。
- ・ 学習等により自らの出力やプログラムを変化させる機能に着目してAIソフトを定義することにより、AI開発ガイドラインにおける用語としてのAIの定義が、現在既に実用化されている機械学習や深層学習等に基づく特化型AIを主たる想定対象としつつ、今後開発される多種多様なAIについても、学習等により自らの出力やプログラムを変化させる機能を有するものである場合には、含み得るものとするを明確にした。
- ・ AI開発原則の対象範囲とするAIシステムは、ネットワーク化され得るAIシステム、すなわち、ネットワークに接続され得るAIシステムとした。
- ・ 「開発者」の定義については、AIシステムの研究開発（AIシステムを利用しながら行う研究開発を含む。）を行う者（自らが開発したAIシステムを用いてAIネットワークサービスを他者に提供するプロバイダを含む。）とした。
- ・ AI開発ガイドラインの策定に当たっては、技術的中立性に鑑み、特定の技術（例：深層学習等）や方法（例：アジャイル型の開発）に基づくAIの研究開発を阻害しないよう、個々のAIに係る技術や方法の特性に応じてAIシステムの開発者に期待される開発原則への対応の在り方が異なり得ることを明示するなど十分留意することとした。
- ・ 開発者と利用者との責任の適切な分担に留意しつつ、AI開発ガイドラインの検討を進めるとともに、利用者が留意して対応すべき事項に関する指針の在り方についても検討を進めることとした。

3. AI開発ガイドライン案の概要

本推進会議における検討を踏まえ、国際的な議論のためのAI開発ガイドライン案を

出された意見の全部について、次に掲げるURLのウェブサイトに掲載。
 <http://www.soumu.go.jp/main_content/000490102.pdf>

作成した（別紙1）⁶⁰。A I 開発ガイドライン案は、A I 開発において留意されることが期待される事項のうち各分野に共通する項目を整理したものである。以下、A I 開発ガイドライン案の概要を記すとともに、必要に応じて、趣旨を解説する。

(1) 目的

A I ネットワーク化による国境を越えた便益及びリスクの波及などガイドライン策定に当たっての背景となる認識及び問題意識を述べた上で、本ガイドラインの目的として「A I ネットワーク化の健全な進展を通じてA I システムの便益の増進とリスクの抑制を図ることにより、利用者の利益を保護するとともにリスクの波及を抑止し、人間中心の智連社会⁶¹を実現すること」を掲げている。

(2) 基本理念

上記(1)の目的に鑑み、次に掲げる理念を一体的に基本理念として掲げている（概要のみ）。

1. 人間中心の社会の実現
2. 指針とそのベストプラクティスのステークホルダ間での国際的な共有
3. 便益とリスクの適正なバランスの確保
4. 技術的中立性の確保と開発者の負担の軽減
5. 本ガイドラインの継続的な見直しと柔軟な改定

(3) 用語の定義及び対象範囲

① 定義

本ガイドラインにおける「A I」の概念を「A I ソフト及びA I システムを総称する概念」と位置付けた上で、「A I ソフト」を学習等により自らの出力やプログラムを変化させるソフトウェアの機能に着目して定義し、その例として、機械学習ソフトウェアを挙げる一方で、「A I システム」を「A I ソフトを構成要素として含むシステム」と定義し、その例として、A I ソフトを実装したロボットやクラウドシステムを挙げている。学習等により自らの出力やプログラムを変化させる機能に着目してA I ソフトを定義する理由は、このように定義されたA I ソフト及びそれを構成要

⁶⁰ 平成29年6月14日から同年7月7日までにかけて、「国際的な議論のためのA I 開発ガイドライン案」を含め本報告書（案）に関する意見募集が行われた。意見募集に寄せられた意見等を踏まえ、推進会議において検討が行われた結果、本ガイドライン案について、説明を補足するなど、若干の修正が行われた。

⁶¹ 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）を構築することにより、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、序章及び中間報告書12-14頁参照。

素として含むA I システムについては、学習等を通じた変化によって開発者が予見し抑制することが困難なリスクが生ずるおそれがあることなどから、従来のICTとは異なる特別の留意が求められることによる。

本ガイドラインにおけるA I の定義は、現在既に実用化されている特化型A I を主たる対象として想定しているが、自律性を有するA I や汎用A I (Artificial General Intelligence) の開発など今後予想されるA I に関する急速な技術発展を見据え、今後開発される多種多様なA I ⁶²についても、学習等により自らの出力やプログラムを変化させる機能を有するものである場合には、含み得るものとしている⁶³。なお、A I の技術発展はスピードが早くその方向性も多様かつ不確実であると指摘されていること、専門家の間でも汎用A I の実現時期については予測が分かれていること、特化型A I と汎用A I は連続的なものであるとの見方も有力であり、両者を峻別することは必ずしも容易ではないこと、本ガイドラインは非拘束的な枠組みと

⁶² ホワイトハウスの報告書「人工知能の未来に備えて」においては、民間部門の専門家の共通認識及びこれに対するNSTC技術委員会の同意として、汎用A I は少なくとも数十年間は実現しないだろうとの見方を紹介した上で、慎重な立場からは超知能 (superintelligence) がいつの日にか誕生する可能性に注意を払うことが求められるが、このような関心がA I に関する公共政策の主要な推進力となるべきではない旨が述べられている一方で、汎用A I の実現プロセスについては専門家の間でも意見が分かれているものの、A I 研究者のうち人間レベルの汎用A I が2040年までに実現する確率が50%以上であろうと予測する者が半数にのぼることを指摘した上で、不確実性を踏まえ、今後のA I の発展への注視が必要である旨が提言されている (White House, *Preparing for the Future of Artificial Intelligence* 7-8, 23-24 (2016))。

一方、我が国の代表的なA I 研究者らにより設立された特定非営利活動法人全脳アーキテクチャ・イニシアティブは、脳全体のアーキテクチャをモデルにして研究を行うことにより、今後十数年以内に人間のような汎用A I が開発されることになるとの見通しを示している (<https://wba-initiative.org/2357/>)。

また、Future of Life Institute (FLI) の「アシロマA I 原則」においては、原則(19)において、「将来のA I の能力についてのコンセンサスが存在しない以上、将来のA I の能力の上限について強い仮定を置くことは避けるべきである」と述べられている (<https://futureoflife.org/ai-principles/>)。

⁶³ 本推進会議の構成員の意見及び意見募集に寄せられた意見においても、A I の定義に関しては、人間と同程度の自律性を有するA I に限定すべきであるという意見から、自律性を有しない特化型A I に限定すべきであるという意見まで、多様な意見が提示された。本ガイドライン案においては、このような見解の多様性を踏まえ、学習等により自らの出力やプログラムを変化させる機能に着目してA I ソフトを定義することにより、学習等により自らの出力やプログラムを変化させる機能を有する多種多様なA I を包摂し得る一般的なA I の定義を試みている。本ガイドラインにおけるA I の定義の在り方については、A I の技術発展の動向等を踏まえ、今後継続的に議論を行っていくことが期待される。また、一部の構成員からは、A I ソフト及びハードウェアやバイオなどソフトウェア以外の方法によりA I ソフトの機能を実現するものを包含して「A I モジュール」と定義し、A I ソフトの概念に代えて用いるべきであるとの意見も提示されたが、現段階においては、「A I モジュール」の概念について定まった理解が確立していないことなどから、本ガイドライン案においては、差し当たりA I ソフトを基本概念として採用しているが、定義の在り方も含め、A I モジュールを構成要素とするA I システムを本ガイドラインの対象とするか否かについては、将来の検討課題とする。

して国際的に共有されることを目指していること、そして、国内外におけるA Iの開発の在り方に関する検討においては、特化型A Iのみならず、自律性を有するA Iや汎用A Iも射程に入れて検討を行うものが有力となっていること⁶⁴などに鑑みると、本ガイドラインにおけるA Iの定義から汎用A Iを明示的に除外することは適当ではないものと考えられる。

また、本ガイドラインに関係する主体として、A Iシステムの「開発者」及び「利用者」を定義している。「開発者」は、自らの開発したA Iシステムを用いてA Iネットワークサービスを他者に提供するプロバイダを含め、A Iシステムの研究開発を行う者として定義している。「利用者」は、最終利用者（エンドユーザ）のほか、他者が開発したA Iシステムを用いてA Iネットワークサービス（情報通信ネットワークを通じてA Iシステムの機能を提供するサービスをいう。以下同じ。）を第三者に提供するプロバイダを含む形で定義している。

なお、A I開発ガイドラインは、主に開発者がA Iの開発において留意することが期待される事項から構成されたものであるが、A Iは利活用時の学習等によりその

⁶⁴ 欧州議会が2017年2月に採択した報告書においては、自律性を有するスマート・ロボットが主たる検討対象とされている（European Parliament, European Parliament resolution of 16 February 2017 with recommendations to the Commission on Civil Law Rules on Robotics (2015/2103(INL))）。

Future of Life Institute (FLI)の「アシロマA I原則」においては、A I一般に適用される原則に加え、高度な自律性を有するA Iや汎用A Iを念頭に、それらの安全性や倫理的問題に関する原則を策定している（<https://futureoflife.org/ai-principles/>）。

米国電気電子学会（IEEE）グローバル・イニシアティブが公衆の議論及び意見募集に供するため公表した報告書「倫理的に調整された設計」においては、あらゆる種類のA I／自律型システム（AS）に適用される高度な倫理的配慮を定式化するための論点が提示されるとともに、汎用A Iに特有の倫理的問題等についても検討が行われている（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)）。

OECDのアン・カブラン氏（科学技術イノベーション局デジタル経済政策課長）も、フォーラムにおいて、A I開発ガイドラインは、現に開発されているA Iに加え、将来開発が見込まれるA Iも見据え、策定される必要がある旨を発言している。

我が国の人工知能学会の策定した「倫理指針」も、第9条において「人工知能が社会の構成員またはそれに準じるものとなるためには、上に定めた人工知能学会員と同等に倫理指針を遵守できなければならない」と定め、将来開発される可能性のある人間と同等の高度な自律性を有するA Iを念頭に置いた原則を盛り込んでいる（<http://ai-elsi.org/wp-content/uploads/2017/02/%E4%BA%BA%E5%B7%A5%E7%9F%A5%E8%83BD%E5%AD%A6%E4%BC%9A%E5%80%AB%E7%90%86%E6%8C%87%E9%87%9D.pdf>）。

一方、内閣府の人工知能と人間社会に関する懇談会の報告書は、汎用A Iが近未来に実現される可能性は低いとの見解などを紹介した上で、「現存する人工知能技術、あるいは近未来に実現することが確実な人工知能技術」を対象として、A Iの倫理的・法的・経済的・社会的論点等について検討を行っている（人工知能と人間社会に関する懇談会「報告書」8-10頁（平成29年））。

出力やプログラムを変化させる可能性があることなどから、利用者がA Iの利活用において留意することが期待される事項から構成されるA I利活用ガイドラインの策定などA Iの利活用のガバナンスの在り方についても、国際的な議論に向けて、今後検討を進めていくことが期待される。

② 対象範囲

本ガイドラインの対象とするA Iシステムの範囲は、A Iシステムがネットワークを通じて国境を越えて利用され、広く人間及び社会に便益やリスクをもたらす可能性があることから、ネットワーク化され得るA Iシステム（ネットワークに接続可能なA Iシステム）としている。ネットワーク化され得るA Iシステムを本ガイドラインの対象とする理由は、ネットワーク化され得るA Iシステムについては、ネットワークを通じて国境を越えて即時に便益及びリスクが波及する蓋然性が高い一方で、ネットワーク化され得ないA Iシステムについては、国境を越えて即時に便益及びリスクが波及する蓋然性は低いと考えられることから、前者については、後者と異なり、開発者が留意することが期待される事項を国際的に共有する必要性が高いものと考えられることによる。

本ガイドラインの対象とする開発者の範囲は、本ガイドラインが拘束的ではないソフトローとしての指針であることに鑑み、定義された開発者すべてとしている。

本ガイドラインの対象とする開発の範囲は、学問の自由の尊重、社会に与える影響の大きさ等に鑑み、閉鎖された空間（実験室、セキュリティが十分に確保されたサンドボックス等）内での開発は対象とせず、ネットワークに接続して行う段階としている。

(4) A I 開発原則

A I 開発原則として、次の9原則を掲げている。

(主にA Iネットワーク化の健全な進展及び便益の増進に関する原則)⁶⁵

① **連携の原則**-----開発者は、A Iシステムの相互接続性と相互運用性に留意する。

(主にリスクの抑制に関する原則)

② **透明性の原則**-----開発者は、A Iシステム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。

③ **制御可能性の原則**-----開発者は、A Iシステムの制御可能性に留意する。

④ **安全の原則**-----開発者は、A Iシステムがアクチュエータ等を通じて利用者及

⁶⁵ 連携の原則は、A Iネットワーク化の健全な進展を通じてA Iシステムの便益を増進することを主たる目的とするものであるが、リスクの抑制に関する複数の開発者の取組が相互に調和して有効に機能するようにすることも目的として含むものである。

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**-----開発者は、A Iシステムのセキュリティに留意する。

⑥ **プライバシーの原則**-----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。

⑦ **倫理の原則**-----開発者は、A I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。

(利用者等の受容性の向上に関する原則)

⑧ **利用者支援の原則**-----開発者は、A Iシステムが利用者を支援し、利用者を選択の機会を適切に提供することが可能となるよう配慮する。

⑨ **アカウントビリティの原則**-----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。

(5) A I 開発原則の解説

各原則の留意事項について解説を付している。各原則の解説に共通するおおよその方向性として、次の事項が挙げられる。

- ・ A Iシステム相互間の連携やA Iシステムと他のシステムとの連携を円滑化すること。
- ・ 開発者間で情報を共有し協力すること。
- ・ 国際的に共有された指針や標準・規格が確立されている場合には、当該指針や当該標準・規格を参照すること。
- ・ 技術的中立性に鑑み、各々のA Iシステムにおいて用いられる技術の特性や用途に照らし可能な又は合理的な範囲で各原則の留意事項に対応すること。
- ・ A Iシステムの学習等による出力又はプログラムの変化の可能性に留意して利用者が適切な対応を取ることが可能となるよう、利用者に対し情報提供すること。
- ・ A Iシステムについて、リスク評価のための検証及び妥当性の確認⁶⁶を行うとともに、開発の過程を通じて適切に措置を講ずること。
- ・ A I開発原則への対応状況に関し、利用者等関連するステークホルダに対しアカウントビリティを果たすこと。

⁶⁶ 欧米の政府や民間におけるA Iシステムの制御可能性、安全性、セキュリティ等に関する検討においても、リスクを評価し抑制するための手法として、検証 (verification) [※形式的な整合性の確認]及び妥当性の確認 (validation) [※実質的な妥当性の確認]を求めるものが有力になっている (The Future of Life Institute (F L I), *Research Priorities for Robust and Beneficial Artificial Intelligence (2015)*, National Science and Technology Council, *The National Artificial Intelligence Research and Development Strategic Plan (2016)*, House of Commons Science and Technology Committee, *Robotics and artificial intelligence, Fifth Report of Session 2016-17 (2016)*)。

(6) 関係するステークホルダの役割

別添では、本ガイドラインを踏まえ、関係する産学民官のステークホルダに期待される役割の例を掲げている。概要は、次のとおりである。

- ・ 【各国政府及び国際機関】 多様なステークホルダ間の対話の促進に向けた環境整備
- ・ 【開発者、市民社会を含む利用者など関係するステークホルダ】 対話によるベストプラクティスの共有とA Iの便益及びリスクに関する認識の共有・協力⁶⁷
- ・ 【標準化団体等】 本ガイドラインの趣旨に適う推奨モデルの作成・公表⁶⁸
- ・ 【各国政府】 A Iの開発者コミュニティ及びA Iに関する研究開発の支援

4. 今後の展開

本推進会議が作成したA I開発ガイドライン案を踏まえ、今後、平成29年9月にトリノで開催されるG7情報通信・産業大臣会合、同年10月にパリで総務省とOECDの共催により開催される国際シンポジウム、同年11月に開催されるOECDデジタル経済政策委員会(CDEP)等において、国際的に共有される非拘束的なA I開発ガイドラインの策定に向けて、国際的な議論が進められることが期待される。また、我が国においても、A I開発ガイドラインの策定に向けた国際的な議論をリードすることができるよう、本推進会議が作成したA I開発ガイドライン案を踏まえ、産学民官の関係するステークホルダにおいて継続的に議論が進められ、合意形成が図られることが期待される。

なお、A Iシステムの研究開発は多様な分野に及び、分野ごとにA Iシステムのもたらす便益やリスクは異なる可能性があることに鑑みると、分野ごとの事情に応じて留意することが期待される事項については、A Iの開発の在り方のうち分野共通的な事項について定めるA I開発ガイドラインの上部に接合される分野別ガイドラインとして、策定の可否を含め、各分野の国際機関を含む関係するステークホルダによる議論が行われることが期待される。

また、A Iシステムはその利活用の段階において学習等により出力やプログラムが継続的に変化すること等から、開発段階のみならず利活用段階においても留意することが期待される事項が想定されることに鑑みると、A I開発ガイドラインとは別に、利用者（最終利用者（エンドユーザ）のほか、他者が開発したA Iシステムを用いて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。）が留意することが期待される事項等により構成されるA I利活用ガイドラインの策定などA Iの利活用に関するガバナ

⁶⁷ Partnership on A Iも、A I技術に関するステークホルダ間の対話やベストプラクティスの共有に積極的な姿勢を示している (<https://www.partnershiponai.org/>)。

⁶⁸ 倫理的考慮を組み込んだA Iの設計に関する標準の在り方につき検討するものとして、see e.g., The IEEE Global Initiative for Ethical Considerations in Artificial Intelligence and Autonomous Systems, *supra* note.

ンスの在り方についても、国際的な議論が行われることが期待される。

第3章 AIネットワーク化が社会・経済にもたらす影響

1. 背景・経緯

(1) AIネットワーク化に関するリスク・シナリオ分析

報告書 2016 において、近い将来に社会の様々な場面において利活用の急速な拡大が見込まれるロボット⁶⁹を題材として、AIネットワークを利活用する具体的な場面を想定したシナリオの作成を試み⁷⁰、シナリオに即したリスク対処の在り方を検討した結果が紹介されている。

そこでは、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、

- ・ AIネットワークの利活用の様々な場面における局所的な制御喪失のリスクへの対応が速やかに求められること
- ・ 事故等のリスクが顕在化した場合には、ネットワークを通じて他のAIシステム等にリスクが波及する可能性があり、AIネットワーク全体のレベルでリスクの顕在化の連鎖に対処するための取組が必要であること

等の知見を得ることができた。

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⁷¹。

また、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
- ・ AIネットワーク化の進展等に応じたシナリオの継続的な見直し
- ・ シナリオを踏まえた政府の取組の在り方の検討

⁶⁹ 自律性を有するAIを実装するロボットだけではなく、AIネットワークを通じて機能するロボットも対象としている。

⁷⁰ 機能に関するリスク（セキュリティに関するリスク、情報通信ネットワークシステムに関するリスク、不透明化のリスク、制御喪失のリスク）について8とおり、法制度・権利利益に関するリスク（事故のリスク、犯罪のリスク、消費者等の権利利益に関するリスク、プライバシー・個人情報に関するリスク、人間の尊厳と個人の自律に関するリスク、民主主義と統治機構に関するリスク）について12とおり、計20とおりのシナリオを作成した。これらのシナリオは、その作成に当たり、検討会議の構成員のうち、特に赤坂構成員、板倉構成員、クロサカ構成員及び中西構成員の協力を得たものである。

⁷¹ 詳細については、報告書 2016 第4章参照。

が提言されている⁷²。

(2) 本推進会議における検討

検討会議を発展的に改組した本推進会議においては、親会の下に開発原則分科会及び影響評価分科会が置かれ、このうち主に影響評価分科会において、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、A I 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、A I 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、A I 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A I 開発ガイドライン案の検討に向けた先行的評価
- ・ A I システムを利活用する分野ごとの分野別評価

を実施した⁷³。

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）については、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⁷⁴を基礎として本推進会議が整理したユースケース、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ることとした。

⁷² 詳細については、報告書 2016 第 5 章 11. 参照。

⁷³ 先行的評価は、A I システム相互間の連携によりインパクトが増大することの具体的な可能性を展望するとともに、セキュリティやプライバシー等に関するリスクを具体的に想定することにより、その結果から得られる示唆を国際的な議論のためのA I 開発ガイドライン案の検討に役立てることを主たる目的として実施したものである。これは、同ガイドライン案の検討に当たっては、A I システムの利活用の場面におけるインパクトやリスクに関する具体的な展望を得た上で各原則の内容等を検討すべきとの指摘があったことによるものである。

分野別評価は、その結果から得られる示唆を今後のA I システムの利活用のガバナンスの在り方（後述（第 4 章(2)参照））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が、得られる示唆の内容によっては、国際的な議論のためのA I 開発ガイドライン案の更なる見直し等にも役立てる場合があり得る。本報告書においては、先んじて実施した 3 分野の評価の結果を掲げているが、今後はこの 3 分野以外の 9 分野について評価を進める予定である。

⁷⁴ スタンフォード大学「A I 1 0 0」〈<https://ai100.stanford.edu/>〉、人工知能と人間社会に関する懇談会〈<http://www8.cao.go.jp/cstp/tyousakai/ai/>〉、一般社団法人産業競争力懇談会（COCN）〈<http://www.cocn.jp/>〉が掲げるユースケース、構成員から提供されたユースケース等。

2. 先行的評価

(1) 評価の枠組み

先行的評価については、「A Iシステム相互間の連携前の段階（A Iシステムが他のA Iシステムとは連携せず、インターネット等を介して単独で機能し、利用者を支援する段階）」と「A Iシステム相互間の連携後の段階（A Iシステム相互間のネットワークが形成され、利用者の便益が飛躍的に増大する段階）」とでは、A Iネットワーク化のインパクトやリスクが異なり得るものと考えられることから、この2段階に分けて評価を行うこととした。

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、A I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、A Iシステムの利活用に伴い新たに生ずるリスク（A Iネットワーク化固有のリスク）だけではなく、A I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A I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⁷⁵。

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⁷⁶。

⁷⁵ 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。

⁷⁶ 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。

(2) 先行的評価

先行的評価については、国内外における既存のユースケースを基礎として整理するユースケースを用いることから、ここでは、A I システムの利活用の分野を主に社会・経済における製品・サービスの供給面の領域の別に応じて、次に掲げるユースケースを対象として評価を行った。

- ・ 災害対応に関するユースケース
- ・ 移動（車両）に関するユースケース
- ・ 健康に関するユースケース
- ・ 教育・人材育成に関するユースケース
- ・ 小売・物流に関するユースケース
- ・ 製造・保守に関するユースケース
- ・ 農業に関するユースケース
- ・ 金融（融資）に関するユースケース
- ・ 公共・インフラに関するユースケース
- ・ 生活に関するユースケース

先行的評価においては、これらのユースケースごとに、次に掲げる災害対応に関する例のように、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3のとおりである⁷⁷。

⁷⁷ 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。

1. 災害対応に関するユースケース

災害が発生した際にAIシステムを利活用することにより、救助、支援業務等が効率化、高度化。
AIシステム相互間の連携後には、救助や支援業務、救急搬送、支援物資の運搬及び分配について、連携の対象となる業務全体としてのリアルタイムでの最適化に向けた提案又はその実施が可能。

【AIシステム相互間の連携前の段階】

(AIシステムが他のAIシステムとは連携せず、インターネット等を介して単独で機能し、利用者を支援。)

<シナリオ上想定される利活用(主なものの例示)>

- 被災者(要救助者)のスマートフォンやウェアラブル端末から得られる様々な情報(位置情報やバイタルデータ等)をAIシステムが解析し、要救助者の容体の変化に応じて、救助を要請する。
- カメラを搭載したドローンが空からの目視が困難な被災者の探索を行い、救助を要請する。
- 避難所で生活する被災者に対し避難所生活や復旧に関する情報を提供する。また、健康情報や生活情報等をもとに健康状態を推定しアドバイスを送る。

【AIシステム相互間の連携後の段階】

(AIシステム相互間のネットワークが形成され、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)>

- AIシステム相互間の連携前の段階における利活用に加えて、時々刻々と変化する要救助者の状況や被災状況、被害予測に基づき、リソースや物資の需給バランスのリアルタイムでのシミュレーションを行い、
- 救助や支援業務に要するリソースの最適な配分を提示して、その作業全体のリアルタイムでの最適化に向けた提案をする。
 - 救急搬送や物資運搬のための最適なルートを設定し、車両等の運用全体のリアルタイムでの最適化に向けた提案をする。
 - 各避難所における支援物資の需給及び運搬車の運用状況のマッチングにより、物資の運搬及び分配全体のリアルタイムでの最適化に向けた提案をする。

(注) AIシステムに決定まで委ねる場合には、提案内容のリアルタイムでの実施が可能。

【AIシステム相互間の連携前の段階】

<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)	リスク評価(例)
被災者(要救助者)からの救助要請	・要救助者の早期発見、早期救助につながる。	・誤った推定・判断により、本来必要とされる救助が要請されないおそれがある。	生命身体に危害が及ぶリスクの評価
ドローンによる被災者の探索	・人間では探索が困難な被災者の発見、救助が可能となる。	・通信の遅延や情報通信ネットワークの遮断、ハッキング等により機能不全に陥ったりするおそれがある。	情報通信ネットワークのオペレーション評価
避難所生活に関する助言や健康状態の推定	・被災者の不安を取り除き、復旧に対する意欲をかきたて、健康維持管理や病気の予防や感染症の拡大防止に役立つ。	・被災者に関する重要な情報がインプットされていなかったこと等により、被災者に誤ったアドバイスを行い被災者の健康状態を害するおそれがある。	生命身体を害する影響等の評価

【AIシステム相互間の連携後の段階】

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)	リスク評価(例)
要救助者の状況や被災状況等に応じた救助や支援業務全体のリアルタイムでの最適化に向けた提案又はその実施	・時々刻々と変化する状況に即応して救助や支援業務全体の最適化を図ることが可能となる。	・AIシステムを利活用しない者が救助や支援の対象から外され、又は不当に順番が劣後されるおそれがある。	生命身体に危害が及ぶリスクの評価
救急搬送や物資運搬のための最適なルート設定、車両等の運用全体のリアルタイムでの最適化に向けた提案又はその実施	・被災状況や活用できるリソースに応じ、車両等の運用全体の最適化を図ることが可能となる。	・一部のAIシステムが個別最適を目指して過剰な要求を行った結果、全体の最適化が実現しないおそれがある。	支援物資やリソースの過不足に及ぼす影響の評価
支援物資の分配全体のリアルタイムでの最適化に向けた提案又はその実施	・避難所ごと、自治体ごとに支援物資の分配・運搬全体の最適化を図ることが可能となる。	・誤った情報(デマ等)が溢れる中、正しい情報に基づく適切な判断、提案ができないうおそれがある	誤って認識・推測する確率の評価

リスク管理(例)	リスク・コミュニケーション(例)
<ul style="list-style-type: none"> ・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 ・判断又は決定をAIシステムに委ねる範囲及び条件の設定 ・複数のAIシステム間の連携に関する仕組みの確立(インターフェース標準化やプロトコルの変換等) 	<ul style="list-style-type: none"> ・災害時におけるステークホルダー間の平常時からの双方向的な対話・協働の場の形成 ・緊急時のコミュニケーション(クライシス・コミュニケーション)を平常時のリスク・コミュニケーションとは区別して実施できる体制の整備

(注) 想定される利活用のうち、主なものについて記載

(3) シナリオ分析から得られた示唆

先行的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⁷⁸。

- ・ 同じ領域内においても様々なA Iシステム相互間の連携が可能となることにより、A I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A I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⁷⁹。

<A Iシステム相互間の連携の例>

- 災害対応： 救急搬送や物資運搬のための最適なルートを設定し、車両等の運用全体のリアルタイムでの最適化に向けた提案をする。
- 健康： 医療機関と消防機関等との間において、救急車、病院の受入態勢等についてリアルタイムで調整を行い、搬送及び治療全体のリアルタイムでの最適化に向けた提案をする。
- 製造・保守： 複数の企業が保有するリソース・技術等がA Iシステムを通じて連携し、生産計画の策定や生産ラインを企業横断的に調整し、リソース活用等の全体としての最適化に向けた提案をする。

※ それぞれの利活用において、A Iシステムに決定まで委ねる場合には、提案内容のリアルタイムでの実施が可能となる。

- ・ 複数の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。これらのリスクは、今回対象としたユースケース以外のA Iシステムの利活用においても広く想定され得るものであることから、特に留意が必要となるものと考えられる。

<共通して想定されるリスクの例>

- A Iシステムやロボット等がハッキングされること等により、不正に操作されたり、機能不全に陥ったりするおそれ、個人情報（センシティブ情報（病歴や健康状態等）を含む。）や機密情報が流出するおそれがある。
 - 一部のA Iシステムが個別最適を目指して過剰な要求を行い続けるなどネットワークを介してつながるA Iシステム相互間の調整に関する技術的又は法的な仕組みが有効に機能しない結果、全体としての最適化の実現が妨げられるおそれがある。
 - なりすましや偽装による不正取引・契約、不正注文が行われるおそれがある。
- ・ これらのリスクへの対処として、リスク管理、リスク・コミュニケーション等が重要となるものと考えられる。また、これらのリスクへの対処は、便益の増進とも密接に関連し得るものである。

⁷⁸ A Iネットワーク化の健全な進展に向けた今後の課題の検討に当たって留意すべき事項については、分野別評価におけるシナリオ分析から得られた示唆と併せて本章5. に記載。

⁷⁹ 長尾・前掲注（8）2頁参照。

<リスク管理、リスク・コミュニケーションの例>

- 専門家による学習データやパラメ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
- 認証等のセキュリティ対策の実装、データの匿名化・暗号化の実施、個人情報の運用・管理ガイドラインの策定
- ステークホルダ間の双方向的な対話・協働の場の形成

3. 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、AIネットワーク化の進展に伴いAI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。

- ・ AIネットワーク化が主としてAI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し得るのかを重視し、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、AI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AI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。
- ・ AI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。

先行的評価においては主として社会・経済における製品・サービスの供給面の領域ごとに評価を実施したことを踏まえ、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たことに鑑み、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た。

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、先行的評価における枠組みと概ね同様である⁸⁰。

⁸⁰ 本章2. (1)参照。

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
	パブリック・ガバナンス		■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	腐敗対策、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	■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	◇時間軸	■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

(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⁸¹。

- ・ 公共：まちづくりに関するユースケース
- ・ 個人：健康に関するユースケース
- ・ 産業：モノに関するユースケース

分野別評価においては、これらのユースケースについて、次に掲げるまちづくりに関する例のように、AI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4のとおりである⁸²。

⁸¹ 他の分野の評価については、本報告書の公表後、本推進会議において行う予定である。

⁸² 前掲注(77)同様、各ユースケースに記載されているAI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。

まちづくりに関するユースケース（概要）

A I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、インフラの異常検知・故障予測とメンテナンスロボットの連携、交通量・流の把握に基づく円滑な交通の実現、カメラによる迷子等の特定と対話型ロボットによる対応など、それぞれのA Iシステムが単独で機能する場合に比べて、時々刻々と変化する状況に即応した**安心・安全かつ快適なまちづくりを進めることができるようになる**。

さらに、A Iシステム相互間のネットワークが公共インフラ、行政、スマートシティ、小売、移動などの領域を横断して形成されることにより、地域の特性に応じた都市計画の策定、在・不在に応じた自動パトロールの実施などの行政サービスが提供されるとともに、パーク&ライドの実現による環境負荷の軽減（CO2の削減）なども図られ、領域が横断する前に比べて、**まちの諸機能が領域を越えて連携し、高度化するようになる**。

【領域横断前の段階】（領域内でA I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拡大）

＜シナリオ上想定される利活用（主なものの例示）＞

- ① 各種インフラに設置されたセンサや衛星写真等から異常検知、故障予測を行い、メンテナンスロボットが即応して補修等を行う。
- ② リアルタイムで把握できる交通量・流に基づいて信号を制御し円滑な交通を実現するとともに、道路状況に即応して渋滞予測やその回避ルートの自動調整を行う。
- ③ カメラの映像から、迷子や困っている人を捜し、即時に対話型ロボットが対応して、困りごとを解決するとともに、不審者を見つけて警備を強化する。

【領域横断後の段階】（A I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）＞

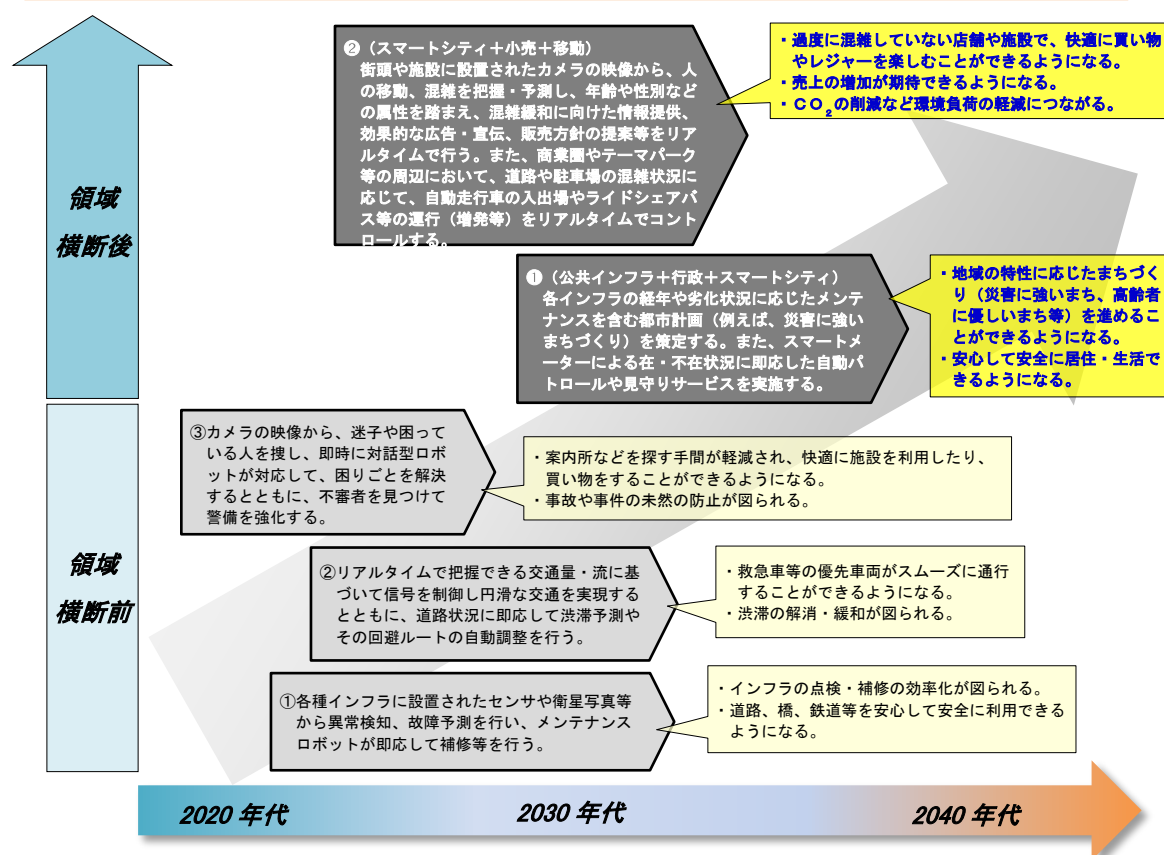
①（公共インフラ+行政+スマートシティ）

各インフラの経年や劣化状況に応じたメンテナンスを含む都市計画（例えば、災害に強いまちづくり）を策定する。また、スマートメーターによる在・不在状況に即応した自動パトロールや見守りサービスを実施する。

②（スマートシティ+小売+移動）

街頭や施設に設置されたカメラの映像から、人の移動や混雑を把握・予測し、年齢や性別などの属性を踏まえ、混雑緩和に向けた情報提供、効果的な広告・宣伝、販売方針の提案等をリアルタイムで行う。また、商業圏やテーマパーク等の周辺において、道路や駐車場の混雑状況に応じて、自動走行車の入出場やライドシェアバス等の運行（増発等）をリアルタイムでコントロールする。

まちづくりに関するユースケース（将来ビジョン）



まちづくりに関するユースケース（評価サマリ）

【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 各種インフラに設置されたセンサーや衛星写真等から異常検知、故障予測を行い、メンテナンスロボットが即応して補修等を行う。	・インフラの点検・補修の効率化が図られる。 ・人間の介入を減らすことができ、作業の安全性の向上が図られる。 ・道路、橋、鉄道等を安心して安全に利用できるようになる。	・異常検知の精度が十分でないことや誤った判断を行うこと等により、異常の見逃し等が生ずるおそれがある。また、異常検知の閾値が低すぎると、過剰反応が発生するおそれがある。 ・ハッキング等により、各種インフラの衛星写真やセンサーから得られた情報が流出し、悪用されるおそれがある。	・見逃された又は放置された異常箇所を車両等が通行することにより、事故が発生するおそれがある。 ・過剰反応によって、無駄なメンテナンスや誤ったメンテナンスが行われるおそれがある。 ・流出した情報がインフラ等に対する物理的な攻撃に利用されるおそれがある。	生命身体に危害が及ぶリスクの評価 AIシステムによる誤った判断が及ぼす影響範囲の評価
② リアルタイムで把握できる交通量・流に基づいて信号を制御し円滑な交通を実現するとともに、道路状況に即応して渋滞予測やその回避ルートの自動調整を行う。	・リアルタイムで高精度の情報を提供することができるため、情報サービス事業の収益機会が増大する。 ・救急車等の優先車両がスムーズに通行することができるようになる。 ・渋滞の解消・緩和が図られる。	・民間のデータ提供者、信号制御を行う警察、交通情報を提供する公的機関・事業者、自動運転車間で、意思疎通が十分に図られず、機能不全に陥るおそれがある。 ・交通情報を所有する関係者間で情報が共有されず、渋滞予測・信号制御の精度が向上しないおそれがある。	・AIシステムが機能不全に陥ること等により、信号制御ができなくなるおそれがある。また、誤った渋滞予測や回避ルートを設定するおそれがある。 ・その結果、交通事故が発生したり、新たに渋滞が発生するおそれがある。	社会的な受容度に影響を及ぼすリスクの評価 情報流出時の影響の評価
③ カメラの映像から、迷子や困っている人を探し、即時に対話型ロボットが対応して、困りごとを解決するとともに、不審者を見つけて警備を強化する。	・案内所などを探す手間が軽減され、快適に施設を利用したり、買い物をするようになる。 ・事故や事件の未然の防止が図られる。	・画像認識の精度が十分でない場合、人物を誤って特定するおそれがある。 ・対話型ロボットの学習不足等により、迷子や困りごとに対応できないおそれがある。 ・カメラの映像（個人情報）が本人同意がなく、第三者に提供、共有されるおそれがある。	・迷子や困りごとが即時に解決されないおそれがある。 ・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。	プライバシー侵害の評価 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)		
<ul style="list-style-type: none"> ・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交換等） ・データの暗号化・匿名化の実施 		<ul style="list-style-type: none"> ・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 ・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 		
雇用・働き方への影響(例)		その他留意すべき事項(例)		
<ul style="list-style-type: none"> ・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。 ・交通情報サービス事業者やデータ提供者の収益増加に伴い、新たな雇用が創出される。 ・案内所や迷子対応の業務から、付加価値の高い業務に配置転換することができる。 		<ul style="list-style-type: none"> ・事故が発生した場合の責任の分配の在り方が問題となる。 ・人間とロボット（AI）との役割分担に留意する必要がある。 ・各地域の利用者の財政状況や利活用に対するマインドにより、収集・提供される情報の質・量に偏在が生じ、地域間でバラつきが生ずるおそれがある。 		

【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 (公共インフラ+行政+スマートシティ) 各インフラの経年劣化状況に応じたメンテナンスを含む都市計画(例えば、災害に強いまちづくり)を策定する。また、スマートメーターによる在・不在状況に即応した自動パトロールや見守りサービスを実施する。	・地域の特性に応じたまちづくり(災害に強いまち、高齢者に優しいまち等)を進めることができるようになる。 ・安心して安全に居住・生活できるようになる。	・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、地域の特性に応じた適切な都市計画が策定されないおそれがある。 ・ハッキング等により、在・不在の情報や見守りサービスに活用する映像が流出し、悪用されるおそれがある。 ・スマートメーターと自動パトロール車におけるAIシステム間の連携が十分でなく、不在者宅付近のパトロールが手薄になったり、同じ場所に複数のパトロール車が駆けつけるなどのおそれがある。	・計画を作り直すなどの無駄なコストが発生するおそれがある。 ・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。 ・安心感が薄れていくおそれがある。	AIシステムによる誤った判断や差別的な判断が及ぼす影響範囲の評価 AIシステムへの信頼への影響の評価
② (スマートシティ+小売+移動) 街頭や施設に設置されたカメラの映像から、人の移動、混雑を把握・予測し、年齢や性別などの属性を踏まえ、混雑緩和に向けた情報提供、効果的な広告・宣伝、販売方針の提案等をリアルタイムで行う。また、商業圏やテーマパーク等の周辺において、道路や駐車場の混雑状況に応じて、自動走行車の入出場やライドシェアバス等の運行(増発等)をリアルタイムでコントロールする。	・過度に混雑していない店舗や施設で、快適に買い物やレジャーを楽しむことができるようになる。 ・売上の増加が期待できるようになる。 ・CO2の削減など環境負荷の軽減につながる。	・画像認識の精度が十分でない場合、適切に属性(年齢や性別)が分類できないおそれがある。 ・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、差別的な広告・宣伝を行うおそれがある。 ・AIシステム間の連携が十分でなく、機能不全に陥り、適切に車の入出場のコントロールができなくなる。混雑が深刻化するおそれがある。 ・カメラの映像（個人情報）が本人同意がなく、第三者に提供、共有されるおそれがある。	・効果的でない広告・宣伝を行うことで無駄なコストが発生するおそれがある。 ・購買意欲等が低下し、売上の増加が見込まれなくなるおそれがある。 ・人権が侵害されるおそれがある。 ・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。	プライバシー侵害の評価 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)		
<ul style="list-style-type: none"> ・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 ・個人情報取得又は活用の際の本人同意の確保、名寄せの制限 		<ul style="list-style-type: none"> ・同種のAIシステムを活用する事業者間やAIシステムを活用しない者との間における対話・協働の場の形成 ・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 		
雇用・働き方への影響(例)		その他留意すべき事項(例)		
<ul style="list-style-type: none"> ・商業施設やテーマパークの売上・消費が増え、雇用が創出される。 ・駐車場等の警備員・整理係の業務に関連する雇用が減少する。 ・自治体の都市計画の策定に関連する雇用が減少する。 ・パトロールや見守りの業務に関連する雇用が減少する。 		<ul style="list-style-type: none"> ・利用者間の調整に当たっては性質や性能が異なる多様なAIシステムがネットワーク上に混在したり、新旧のAIシステムがネットワーク上に混在したりすることに留意する必要がある。 ・AIシステムの判断がブラックボックス化された場合、当該AIを所有・管理する公的機関(自治体等)への情報公開請求に対して適切に対応できない可能性がある。 ・パブリックなスペース(公道など)とプライベートなスペース(店舗内など)が混在することを踏まえて、データの価値とプライバシーとのバランスをとることが重要である。 		

(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⁸³。

- ・ 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⁸⁴。

<領域横断・融合の例>

- 公共インフラ+行政+スマートシティ、スマートシティ+小売+移動
- 医療・介護+農林水産+小売+生活支援、医療・介護+行政+金融・保険
- 製造+運輸・物流+小売、製造+運輸・物流+金融・保険

- ・ 雇用・働き方への影響について、雇用が減少することが見込まれる業務もあるものの、より付加価値の高い業務への配置転換や新たな雇用の創出の可能性も見込まれる。

<雇用・働き方への影響の例>

- 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、家電メーカー、ロボットメーカーに関連する雇用等の創出が見込まれる。
- 行政（自治体）の窓口業務から政策立案等の業務への配置転換、生産ラインの調整等の業務から商品・サービスの企画・開発等の業務への配置転換が可能となることが見込まれる。
- パトロールや見守りの業務に関連する雇用、在庫管理や商品の発送・納品等に係る調整等の業務に関連する雇用等の減少が見込まれる。

- ・ 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⁸⁵が重要である。

<共通して想定されるリスクの例>

- 誤ったデータや一部の領域の事情に偏ったデータに基く学習や学習不足の結果、領域を横断した最適な調整等が実現しないおそれがある。
- 画像認識の精度が十分でない場合、属性（性別、年齢等）、体調の急変等を正しく認識できず、期待された便益を享受できないおそれがある。
- 個人情報（センシティブ情報（病歴や健康状態等）を含む。）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。

⁸³ 先行的評価におけるシナリオ分析から得られた示唆と共通するものについては、本章2.

(3)参照。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。なお、前提注(73)と同じく、A Iネットワーク化の健全な進展に向けた今後の課題の検討に当たって留意すべき事項については、先行的評価におけるシナリオ分析から得られた示唆と併せて本章5. に記載。

⁸⁴ 長尾・前掲注(8)2頁参照。

⁸⁵ リスク管理、リスク・コミュニケーション等については、先行的評価におけるものと共通するものが考えられる。本報告書36頁参照。

なお、分野別評価については、本報告書において対象とした3分野以外の分野の評価を今後進めることとしており、その結果として得られる示唆は、主として今後のAIシステムの利活用のガバナンスの在り方の検討に役立てるほか⁸⁶、示唆の内容によっては、国際的な議論のためのAI開発ガイドライン案の更なる見直し等にも役立てる場合があり得る。

4. 雇用・働き方への影響

報告書2016においては、AI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⁸⁷。

1990年代以降、情報通信分野の発展が雇用に与える影響については、ルーティンタスクに従事していた中間層の雇用が代替されたが、サービス業や運輸業が受け皿となったことにより、社会全体としては雇用が確保されていたとの見方がある⁸⁸。しかしながら、AIシステムの普及により、受け皿となっていたサービス業や運輸業等これまでは人間にしかできないとされていた業務についてもAIシステムへの代替が進むようになることから、受け皿がなくなり大規模な失業が発生するのではないかと指摘がある⁸⁹。また、AIシステムに関する技術革新のスピードが非常に速いため、補完的イノベーション⁹⁰が技術革新に追いつかず、AIシステムがもたらす生産性向上・経済成長よりも先に格差の拡大が生じてしまうのではないかと指摘もある⁹¹。

ただし、これらの指摘に関しては、

- ・ AIシステムによる代替が進むとされている職種は、主観的な予測をベースとして挙げられており、必ずしも信頼性が高いとは言い難いこと。
- ・ 新しい技術による雇用の代替可能性は、その価格に依存するが、その点が考慮されていないこと。
- ・ 新しい技術による生産性向上・経済成長がもたらす雇用創出の可能性が考慮されてい

⁸⁶ AIシステムの利活用のガバナンスについては、後述（第4章(2)参照）。

⁸⁷ 報告書2016第2章2. 参照。

⁸⁸ 山本影響評価分科会構成員「労働経済学研究に基づくAIネットワーク化の労働市場への影響」（提出資料）参照。

⁸⁹ Frey, C., & Osborne, M., *The future of employment* (2013)参照。

⁹⁰ 一般的に、技術革新が起きた場合、それをうまく活用するように機械の導入や組織改革等が行われることにより、生産性向上・経済成長が生じ、新しい雇用が創出されると言われている。ただし、その間にはタイムラグがあり、すぐに新しい雇用が創出されるわけではないこと（生産性のパラドックス）に留意することが必要である。

⁹¹ エリック・ブリニユルフソン＝アンドリュウ・マカフィー（村井章子訳）『ザ・セカンド・マシン・エイジ』（日経BP社、平成27年）参照。また、経済成長と格差の拡大が同時に進行する可能性があるとの指摘もある。

ないこと⁹²。

- ・ 射程とするタイムスパンにより、影響の在り方や大きさが異なること。
- ・ 失業以外への影響について、A Iシステムのサポートにより労働者のストレスの軽減、女性・高齢者・障がい者などの活躍の促進など良い影響がもたらされる可能性があること。

等に留意する必要がある⁹³。

ここで、我が国の労働市場に目を向けてみると⁹⁴、ルーティンタスクの集約度が比較的高い⁹⁵（ルーティンタスクが他の先進国に比べ多く残っている）状況にあることから、A Iシステムの普及による影響を受け易いとの見方がある⁹⁶。また、従来は、正規雇用を中心とする日本的雇用慣行により情報通信分野の発展による雇用の代替が生じにくかったが⁹⁷、A Iシステムの普及が進んでいくことにより、日本的雇用慣行そのものがなくなり、中長期的には代替が進むのではないかとの見方もある⁹⁸。さらに、非正規雇用の業務はルーティンタスクが中心になっていることから、A Iシステムの普及により、非正規雇用の

⁹² 必ずしも新しい技術による生産性向上・経済成長が雇用創出をもたらすとは限らず、失業の増大と経済成長が同時に進行する（雇用なき経済成長が継続する）可能性があるとの指摘もある。

⁹³ 山本・前掲注(88)参照。

⁹⁴ 我が国においては、少子高齢化等に伴い中長期的に労働人口が減少するとの見通しがある中で、A Iシステムの普及による雇用の代替が人間の労働力不足を補い得るという側面もあることに留意が必要である。また、現状においても、人手不足と言われる産業や業界もあり、特に、そのような産業や業界においては、A Iシステムの普及により人間の負担の軽減が図られるものと考えられる。

⁹⁵ 国際比較において22箇國中4番目に高い（1位：韓国、2位：イタリア、3位：ロシア）。DeLaRica, S. & Gortazar, L. *Differences in Job De-Routinization in OECD Countries* (2016)参照。

⁹⁶ 山本・前掲注(88)参照。

⁹⁷ 従来は、企業特殊的な人的投資により人材育成を行っていたことから、雇用の代替に当たっての調整費用が大きくなるため、情報通信分野の発展による雇用の代替が進まなかったものと考えられる。このことに関連して、大内伸哉（影響評価分科会構成員）『A I時代の働き方と法 2035年の労働法を考える』第1章及び第4章（弘文堂、平成29年）参照（日本的経営の柱とされていた終身雇用、年功型処遇及び企業別組合があてはまる正社員こそが日本の労働者のメインストリームであるとして、日本型雇用システムは、長期雇用の下に、年功型の職能給という安定的な処遇体系をもち、その反面、職務や教育訓練についての企業の権限（人事権）を広く認めるという柔らかな労働組織があったため、技術革新に対する高いアダプタビリティ（適応力）をもっていたと指摘している。）。

⁹⁸ 山本・前掲注(88)参照。このことに関連して、大内・前掲注(97)参照（企業内の再配置を可能とするアダプタビリティは、技術の発達スピードが比較的緩やかで、企業内での教育訓練によって、新たな技術が必要とする技能の習得が追いつくことが前提であったとして、現在の技術の発達スピードは格段に速まっており、企業内での再配置が困難になっていることや再配置するための雇用が創出されていないことを挙げ、ホワイトカラーの雇用の多くは、既に危機的な状況にあるとの警告に真剣に耳を傾けなければならないと指摘している。）。

労働者は大きな影響を受ける可能性があるとの見方もある⁹⁹。

このような雇用・働き方への影響を踏まえると、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¹⁰⁰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。

教育・人材育成の重要性については、国際的にも指摘されており、米国ホワイトハウスの報告書『人工知能、自動化、そして経済』（平成28年12月20日）において、AIによる自動化が経済や雇用に与える影響を分析し、AIの普及がもたらす変化に対応した教育改革、労働者の支援などの政策的対応を提言している。また、フォーラムにおいても、教育・人材育成の重要性に関する指摘が多数あった¹⁰¹。

いずれにしても、教育・人材育成の成果が現れるまでには長期の時間を要することから、一刻も早くアクションを起こすべきである。

5. 留意すべき事項

先行的評価及び分野別評価におけるシナリオ分析等から、AIネットワーク化が社会・経済にもたらす影響に関して、重要な示唆を得ることができた。

今後、AIネットワーク化の健全な進展に向けた今後の課題を検討するに当たっては、特に次の事項に留意することが必要である。

ア 様々なAIシステムが混在することに伴う問題

- 例えば、AIシステム相互間の連携を進めるに当たっては、性質や性能が異なる多様なAIシステムがネットワーク上に混在したり、新旧のAIシステムがネットワーク上に混在したりするようになることが問題となるとともに、目的が競合し又は対立するAIシステムがネットワーク上に混在してその相互間の交渉や調整が必要

⁹⁹ 山本・前掲注(88)参照。

¹⁰⁰ 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。

¹⁰¹ 例えば、

- ・ AIネットワーク化の便益を世界の多様な文化的背景を持つ人々が享受できるように、教育、学際的研究などの取組が必要である[マルコ・ジリ（トリノ工科大学長）]。
- ・ 教育は、トップ及びボトムの双方から進めることが必要である[ベネデッタ・アレーゼ・ルチーニ（イタリア共和国経済財政大臣顧問）]。

等。

となり得ること¹⁰²も問題となる。

イ AIシステムを利活用する者と利活用しない(できない)者との混在することに伴う問題¹⁰³

- 例えば、AIシステムを利活用することにより、自身に適した行政サービスや金融サービス(保険)を受けることが可能となるが、AIシステムを利活用しない(できない)者は、そのようなサービスを受けることができない可能性がある。

ウ 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

- 例えば、自動運転車による事故が発生した場合、商品の配達において遅延が生じた場合、個人情報漏洩してプライバシーが侵害された場合など、責任の分配の在り方が問題となる。
- 上記に関連して、責任の分配の実効性を確保する手段として、AIシステム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システム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関する検討が必要であるとの指摘があった。

エ データ・情報の流通と個人情報保護・プライバシーとのバランス

- 例えば、カメラで撮影した画像は、様々な場面において価値が高く非常に有益なデータとなり得るものであるが、その取得や利活用に当たっての本人同意の在り方等が問題となる可能性がある。また、パブリックなスペース(公道など)とプライベートなスペース(店舗内など)が混在することを踏まえて、データ・情報の価値と個人情報保護・プライバシーとのバランスをとることが重要である。

オ AIシステムが取り扱うデータ・情報・知識の性質

- 例えば、プライバシー性や機密性が高いデータ・情報(個人:病歴、遺伝情報等、法人:財務情報、営業秘密等)・知識については、プライバシー侵害等が生じないよう特に慎重な取扱いが求められる。
- 例えば、AI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。

カ 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における公正な競争の確保

- 例えば、AIネットワークサービスの提供に新規参入しようとするプロバイダに

¹⁰² 長尾・前掲注(8)3頁参照。

¹⁰³ 濱田・前掲注(10)7頁参照。

としては、自らのA Iシステム等と有力なA Iネットワークとの接続の可否やその条件、有力なA Iネットワークが集積するデータ等へのアクセスの可否やその条件が事業活動に影響を及ぼす可能性があり、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における公正な競争の確保の在り方が問題となる¹⁰⁴。

- 例えば、競争関係にある事業者が、それぞれ提供するサービス等に関する価格設定等に当たり、
 - ・ 同一のA Iシステム又は共通のアルゴリズムを用いるA Iシステムを使用
 - ・ 各事業者の価格設定等を委ねるA Iシステムを相互に連携させて使用等することを通じた協調的価格設定等¹⁰⁵による競争の減殺又は制限の可能性¹⁰⁶が考えられる。

キ 人間とA I（ロボット）との役割分担

- 例えば、対話型ロボットが迷子に対応する場面や介護ロボットが介護する場面において、A I（ロボット）による対応で満足できるか、拒絶されないかなど人間とA I（ロボット）との役割分担が問題となる。
- 例えば、専門職（医師、弁護士、会計士等）とA I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A Iシステムの機能に委ねてもよい事項の範囲が問題となる。

¹⁰⁴ A Iネットワークサービスのプロバイダは、その利用者から収集するデータを活用して自らのA Iのアルゴリズムを改良することにより、当該サービスの品質を向上させることができる。このように利用者から収集するデータを活用したサービスの品質向上については、利用者の便益が高まるという面がある一方で、利用者からのデータの収集、収集したデータによるサービスの品質向上、サービスの品質向上に伴う利用者数の増加、利用者数の増加により収集されるデータの増加という循環が確立されやすい（この傾向は、「データ主導ネットワーク効果（data-driven network effect）」又は「データネットワーク効果」（data network effect）と呼ばれる。）ことから、市場において有力なプロバイダの地位がより強固になって、新規参入が相対的に困難となり、その結果として競争が鈍化するおそれが懸念される旨が指摘されている（See Organization for Economic Cooperation and Development Directorate for Financial and Enterprise Affairs Competition Committee, Big Data: Bringing Competition Policy To The Digital Era - Background note by the Secretariat -, DAF/COMP (2016) 14 October 27 2016, at paras 22-25 [available at [https://one.oecd.org/document/DAF/COMP\(2016\)14/en/pdf](https://one.oecd.org/document/DAF/COMP(2016)14/en/pdf)]).

¹⁰⁵ デジタル・カルテルと呼ばれることがある。

¹⁰⁶ See Organization for Economic Cooperation and Development Directorate for Financial and Enterprise Affairs Competition Committee, *supra* note (102), , at paras 80-81; Ezrachi, Ariel and Stucke, Maurice E., *Two Artificial Neural Networks Meet in an Online Hub and Change the Future (Of Competition, Market Dynamics and Society)* (April 9, 2017) [Available at SSRN: <https://ssrn.com/abstract=2949434> or <http://dx.doi.org/10.2139/ssrn.2949434>].

ク AIシステムに関するリテラシーの向上¹⁰⁷

- 例えば、健康・介護の分野においては、特に高齢者向けのサービス提供が考えられることから、AIシステムの便益を享受するためにも、AIシステムに関するリテラシーの向上が重要である¹⁰⁸。

ケ 地域間の格差・偏在

- 例えば、各地域の利用者（自治体等）の財政状況や利活用に対するマインドにより、AIシステムの利活用の状況、収集・提供される情報に差異が生じ、AIシステムの利活用が進んでいる地域と進んでいない地域が存在することとなる（移動を伴う場合、突然、AIシステムが機能しなくなるなどの問題が生ずるおそれがある。）。
- 上記に関連して、地域間の格差だけではなく、所得格差やAIネットワーク化の進展に伴う所得の再分配にも留意する必要があるとの指摘があった。

コ AIシステムの導入・利活用のためのコスト

- 例えば、製造（モノ）や物流の分野において、多くの事業者がAIネットワークに参加した方が全体として便益が増大するものと考えられるが、中小・零細事業者におけるAIシステムの導入コストや利活用のためのコストが問題となる。

サ AIシステムの判断がブラックボックス化することに伴う問題

- 例えば、行政（自治体）がAIシステムを利活用する場合、AIシステムの判断がブラックボックス化すると、情報公開請求に対して適切に対応できない可能性がある。また、金融機関をはじめ事業者が利活用するAIシステムの判断がブラックボックス化すると、顧客への説明や監督官庁の検査・監査等に適切に対応できない可能性がある。

¹⁰⁷ 濱田・前掲注（10）7頁参照。

¹⁰⁸ 高齢者のリテラシーの向上を期待するのではなく、リテラシーが高くなくても利活用できる商品・サービスが提供されることが望ましいとの指摘があった。

第4章 今後の課題

報告書 2016 に掲げられている「今後の課題」¹⁰⁹並びに本推進会議におけるA I 開発ガイドライン案の策定に向けた検討及びA I ネットワーク化が社会・経済にもたらす影響の評価に関する検討の結果等を踏まえ、今後の課題を次のとおり整理する。

○ A I ネットワーク化の健全な進展に関する事項

(1) A I 開発ガイドライン（仮称）の策定

第2章において述べたように、本推進会議においては、A I 開発ガイドライン（仮称）の策定に向け、G 7やOECD等における国際的な議論の用に供するため、「国際的な議論のためのA I 開発ガイドライン案」（別紙1）を整理した。今後、G 7やOECD等において、このA I 開発ガイドライン案を踏まえて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A I 開発ガイドライン（仮称）が策定されることが期待される。

なお、A I 開発ガイドライン（仮称）については、その基本理念¹¹⁰に掲げるとおり、A I ネットワーク化の進展等に応じて、関係するステークホルダの参画を得るなど広範で柔軟な議論を通じて、継続的に見直し、必要に応じて改定することが期待される。本推進会議においても、A I 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A I 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。

(2) A I 利活用ガイドライン（仮称）の策定

A I システムは、特段の措置が講じられているものを除き、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。また、A I システムは、情報通信ネットワークを介して他のA I システムと連携することにより、システム間の連携の効果及びいわゆるネットワーク効果が相俟って、利用者全体にとっての便益及びリスクの双方が高まり得るものである¹¹¹。しかるに、ネットワークに接続されないA I システムの便益及びリスクは、基本的に、当該A I システムが所在する場所から即座に広く波及することはないが、ネットワークに接続されたA I システムの便益及びリスクは、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¹¹²。

¹⁰⁹ 本章末尾参照。詳細については、報告書 2016 第5章参照。

¹¹⁰ 別紙1 2. 参照。

¹¹¹ 須藤議長「開会の辞」（A I ネットワーク社会推進フォーラム 開会の辞 配付資料）7頁（平成29年）参照。

¹¹² 須藤・前掲注（111）7頁参照。併せて、宍戸構成員・開発原則分科会長代理「A I ネットワーク化のグローバルガバナンス—枠組み、戦略、手法、開発と利活用」（A I ネットワーク

これらのことから、A I ネットワーク化の健全な進展を通じたA I システムの便益の増進及びリスクの抑制については、ネットワークに接続され得るA I システムの開発を対象とするA I 開発ガイドライン（仮称）を国際的に共有すること等を通じた開発のガバナンスと併せて、利活用のガバナンスとして、ネットワークに接続されたA I システム又はA I ネットワークサービスの利活用に当たり、利用者（最終利用者たる個人又は団体のほか、他者が開発したA I システムを用いて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。）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な共有を図ることが必要であるものと考えられる¹¹³。

そこで、本推進会議においては、今後、A I 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と併せて、ネットワークに接続されたA I システム又はA I ネットワークサービスの利活用に関し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たる「A I 利活用原則（仮称）」及びその解説からなる国際的な指針たる「A I 利活用ガイドライン（仮称）」の策定に関し、国際的な議論の動向も踏まえながら多角的に検討を進めることが必要である。

なお、「A I 利活用ガイドライン（仮称）」の策定など利活用の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、A I 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、A I 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ものと考えられる。そこで、本推進会議においては、「A I 利活用ガイドライン（仮称）」の策定等に関する検討に向け、前章で紹介した分野別評価及びエコシステムの展望の検討並びにこれらを通じた論点の整理を進めていくことが課題となる。

(3) A I システム相互間の円滑な連携の確保

A I 開発原則¹¹⁴の第1原則とした「連携の原則」においては、開発者はA I システムの相互接続性と相互運用性に留意することを掲げている¹¹⁵。同原則の解説¹¹⁶においては、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

社会推進フォーラム パネルディスカッション「A I ネットワーク化のガバナンスの在り方」
配付資料）1頁（平成29年）参照。

¹¹³ なお、A I 開発ガイドライン案3-1においても指摘しているとおり、開発者と利用者は、場面に応じて個別に決まる相対的な概念であることに留意することが必要である。

¹¹⁴ A I 開発原則については、第2章3. (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。

¹¹⁵ 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。

¹¹⁶ A I 開発原則の解説については、第2章3. (5)参照。詳細については、別紙1-5参照。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。

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を掲げている。

これらのことを踏まえると、A Iシステム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ネットワークサービスのプロバイダが他の開発者やA Iネットワークサービスのプロバイダ、利用者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、前章で行った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（以下「シナリオ分析」という。）から得られた示唆（以下「シナリオ分析に基づく留意事項」という¹¹⁷。）において指摘したように、A Iシステム相互間の連携を進めるに当たっては、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ようになることが問題となるほか、目的が競合するA Iシステムがネットワーク上に混在して相互間の交渉や調整が必要となり得ること¹¹⁸も問題となる。したがって、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システムの混在の可能性に留意することが必要である。

(4) 競争的なエコシステムの確保

A Iネットワーク化の進展を通じて智連社会¹¹⁹を実現するためには、

- ・ 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の形成
- ・ A IシステムやA Iネットワークサービスの提供
- ・ ネットワーク化されたA Iシステム又はA Iネットワークサービスにより創造・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り、シナリオ分析に基づく留意事項においても、このような問題意識に基づく指摘をしている。

¹¹⁷ シナリオ分析に基づく留意事項については、前章5．参照。以下本章におけるシナリオ分析に基づく留意事項に関する記述において同じ。

¹¹⁸ 長尾・前掲注（8）3頁参照。

¹¹⁹ 序章1．参照。

これらのエコシステムは、今後のA I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A IシステムやA Iネットワークサービスに対するニーズの変化、A I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、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と考えられる¹²⁰。

(5) 利用者の利益の保護

A I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A Iシステム又はA Iネットワークサービス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A Iシステム又はA Iネットワークサービス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A Iシステム又はA Iネットワークサービスを手頃な条件で¹²¹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¹²²。

¹²⁰ 報告書2016においては、競争的なエコシステムを確保するという観点から、関係する市場の形成の進展に応じて、その動向の継続的注視を行うべきであるとし、その説明として、関連する市場の動向、A Iネットワーク化やデータ寡占等に着眼したデータ等の創造・流通・蓄積の状況、事業者間の競争状況その他市場の動向の注視・評価を挙げ、更にこれらの前提となる課題を次のように掲げている(報告書2016・50頁参照。)

- ・注視対象(A Iの範囲、データの範囲、市場等)の画定、注視の視点、評価基準等の在り方の検討

- ・注視・評価に必要となる情報の収集の在り方の検討

- ・A Iネットワークサービスの供給者による行為であって、公正な競争を阻害するおそれがあるものの類型化の検討

なお、報告書2016においては、競争的なエコシステムを確保するという観点からの今後の課題として、市場の動向の継続的注視と併せて、関係する市場の形成の進展に応じて、A I相互間のネットワークの形成に関する当事者間の協議の円滑化に取り組むべきであるとし、その内容を次のように敷衍している(報告書2016・50頁参照。)

- ・A I相互間のネットワークの形成に関する当事者間の協議をめぐる紛争の動向及び影響の継続的注視

- ・必要に応じ、当事者間の協議を円滑化する観点からの紛争処理の在り方の検討

¹²¹ 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。

¹²² A Iネットワーク化の進展に伴い、社会のあらゆる場面において、社会の構成員がネットワーク化されたA Iシステム又はA Iネットワークサービスを利用できることを前提として、様々な物、サービスや社会のシステムが提供されるようになるものと見込まれる。そのような社会において個人又は団体が当該社会の構成員として振る舞うためには、自らが関連するA Iシステム又はA Iネットワークサービスの利用者となることが不可避となろう。このような社会において人間の「包摂」(inclusiveness)を確保する見地からは、当該社会の個人又は団体に対し、関連するA Iシステム又はA Iネットワークサービスを手頃な条件で安心して安全に利用できる状況を確保することが必要となる(須藤・前掲注(111)6頁参照。併せて、宍

このため、A I ネットワーク化のガバナンスにおいては、A I 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¹²³。A I 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A I 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¹²⁴。

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 ネットワークサービスのプロバイダから利用者に対する
 - ・ A I システム又はA I ネットワークサービスの技術的特性に関する情報¹²⁵
 - ・ 継続的なアップデートへの対応の必要性に関する情報¹²⁶
 - ・ その他リスクの顕在化又は波及の抑止に資する情報

戸・前掲注(112)2頁参照。このことに関連し、シナリオ分析に基づく留意事項においては、例えば、A I システムを利活用することにより、自身に適した行政サービスや金融サービスを受けることが可能となるが、A I システムを利活用しない(できない)者は、そのようなサービスを受けることができない可能性があることを指摘している。)

また、ネットワーク化されたA I システム又はA I ネットワークサービスの利用が増えること自体により、ネットワーク効果として、利用者全体にとって当該A I システム又はA I ネットワークサービスの便益が高まり得る(須藤・前掲注(111)6頁参照。)。その一方で、A I システムについては、様々なリスクが指摘されていることに加えて、特段の措置が講じられているものを除き、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から、利用が増えること自体によって、利活用の仕方次第では、ネットワーク効果として、利用者全体にとって当該A I システム又はA I ネットワークサービスのリスクも高まり得る(須藤・前掲注(111)6頁参照。)。これらのことから、ネットワーク効果の見地からは、利用者の数が増えることを通じて便益を高めると同時に、効果的にリスクを抑制することが求められ、かくして、この見地からも、利用者が手頃な条件で安心して安全に利用できるようその利益が保護されることが求められる。

¹²³ 須藤・前掲注(111)8頁及び戸・前掲注(112)2頁参照。

例えば、A I システムが他のA I システム又は他の種類のシステムと連携することにより、システム間の連携の効果及びネットワーク効果の双方が相俟って、その便益及びリスクの双方が高まり得ることに鑑みると、(3)に掲げるA I システム相互間の円滑な連携の確保も、このネットワーク効果による利用者全体の便益の増進及びリスクの増大に伴う不利益の抑制に資することを通じて利用者の利益の保護に資するものと認められる。

また、(4)に掲げる競争的なエコシステムの確保も、これを通じた公正な競争の進展の結果として、利用者の利益の保護に資するものと認められる。

¹²⁴ 連携の原則は、連携に係るA I システムと他のA I システム等との円滑な連携を通じて、連携に係るA I 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。利用者支援の原則及びアカウントビリティの原則も、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。

¹²⁵ A I 開発原則に掲げる事項に関する情報を含む。

¹²⁶ 報告書2016・54頁参照。

の自発的な提供¹²⁷の在り方の検討

- ▶ AIネットワークサービスのプロバイダ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
- ▶ AI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¹²⁸

なお、(4)において述べたように、AIネットワーク化をめぐるエコシステムが今後の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ネットワーク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に鑑みると、利用者の利益の保護についても、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、利用者の利益の保護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、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と考えられる¹²⁹。

(6) 技術開発に関する課題

AI開発原則の個々の項目の実効性を高めるとともに、上記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。
 - ・ AI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AI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
 - ・ 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

¹²⁷ これらの情報提供のうち開発者によるものについては、AI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検討することが期待されよう。

¹²⁸ シナリオ分析に基づく留意事項においても、「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。

¹²⁹ 報告書2016においては、AIネットワークサービスの利用者の利益を保護するという観点から、関係する市場の形成の進展に応じて、その動向の注視・評価を行うべきであるとし、更にその前提となる課題を次のように掲げている（報告書2016・53頁—54頁参照。）。

- ・ 注視すべき市場の画定、評価基準等の在り方の検討
- ・ 注視・評価に必要な情報の収集の在り方の検討
- ・ AIネットワークサービスの供給者による行為であって、利用者の利益を阻害するおそれがあるものの類型化の検討

なお、報告書2016においては、AIネットワークサービスの利用者の利益を保護するという観点からの今後の課題として、市場の動向の注視・評価と併せて、関係する市場の形成の進展に応じて、AIネットワークサービスの供給者と利用者との間の紛争処理の在り方の検討を掲げている（報告書2016・53—54頁参照。）。

- ・ 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ネットワーク上を流通する情報・データに関する事項

(7) セキュリティ対策

AI開発原則の第5原則とした「セキュリティの原則」においては、開発者はAIシステムのセキュリティに留意することを掲げている。また、シナリオ分析においては、各ユースケースに共通して想定されるリスクとして、セキュリティに関するリスク（ハッキング、偽装・なりすまし等）が挙げられている¹³⁰。

今後、AIシステム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

(8) プライバシー及びパーソナルデータの保護

AI開発原則の第6原則とした「プライバシーの原則」においては、開発者はAI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を掲げている。このことに関連して、シナリオ分析に基づく留意事項においては、「AIシステムが取り扱うデータ・情報の性質」及び「データ・情報の流通と個人情報・プライバシーとのバランス」を挙げている。

今後、AIシステム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、AIシステムが取り扱うデータ・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
- AI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

¹³⁰ 先行的評価において挙げられたリスクについては、前章2. (3)参照。分野別評価において挙げられたリスクについては、同章3. (3)参照。

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

- 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

(9) コンテンツに関する制度的課題

シナリオ分析に基づく留意事項においては、「AIシステムが取り扱うデータ・情報の性質」を挙げている。また、新たな情報財の保護・利活用の在り方に関し、AI生成物の知財制度上の在り方について引き続き検討すべきとされている¹³¹。

今後、AI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、AI生成物の知財制度上の在り方の検討¹³²
- AI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プンデータの推進
- 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の評価に関する事項

(10) 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 2016 においては、今後の課題として、「影響・リスクに関するシナリオの作成・共有」が掲げられていたところ、本推進会議において、AIネットワーク化が社会・経済にもたらす影響の評価に関する検討を進め、シナリオ分析（先行的評価及び分野別評価）を行った¹³³。

今後、継続して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。

- 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

(11) 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

シナリオ分析の検討において、一定の評価の軸を設定してインパクト及びリスクを

¹³¹ 第1章1.(3)⑥参照。

¹³² 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（AI）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。

¹³³ 先行評価については、前章2. 参照。分野別評価については、同章3. 参照。

評価すべきとの指摘があった。分野別評価において、一部取り入れているが、十分に反映できているとは言い難い。また、報告書 2016 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」が掲げられている。

今後、評価指標の設定について継続的に検討することが必要である。

- 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

(12) A I システムの利活用に関する社会的受容性の醸成

A I 開発原則の第 9 原則とした「アカウントビリティの原則」においては、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を掲げている。これは、A I 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、シナリオ分析において、社会的受容性が求められるものと考えられるような A I システムの利活用の事例も扱っている¹³⁴。

今後、A I 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 A I システムの利活用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

○ A I 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

(13) 人間と A I システムとの関係の在り方に関する検討

A I 開発原則の第 7 原則とした「倫理の原則」においては、開発者は A I 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を掲げている。また、シナリオ分析に基づく留意事項において、「人間と A I (ロボット) との役割分担」を挙げている。

今後、人間中心とする智連社会に向けた人間と A I システムとの役割分担、A I システムに委ねる判断の範囲の妥当性、A I 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 A I 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 A I 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
- A I ・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
- 「A I 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

¹³⁴ 例えば、健康情報（病歴や遺伝情報等）に応じて金融（保険）サービスをカスタマイズして提供する例など。

(14) ステークホルダ間の関係の在り方に関する検討

シナリオ分析に基づく留意事項においては、「事故の発生等A I 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」を挙げている。また、報告書 2016 において、今後の課題として、A I ネットワーク化に対応した社会の基本ルールの在り方について、A I ネットワークに関する権利義務及び責任の帰属の在り方の検討、A I ネットワークに関する事故時の責任の帰属の在り方の検討等が挙げられている。

今後、A I ネットワーク化の進展を見据え、上記(12) 人間とA I システムとの関係の在り方に関する検討とともに、A I システム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。

- A I システム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¹³⁵間の役割分担の在り方の検討
- A I システム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
- A I システム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¹³⁶
- 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
- A I システムの利活用をめぐる紛争の解決の在り方の検討

(15) 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

シナリオ分析においては、様々な業務に関連する雇用（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や家電メーカーやロボットメーカーに関連する雇用等）の創出や付加価値の高い業務への配置転換が可能となることを見込まれるとの結果を得ている。また、A I システムの普及による雇用・働き方への影響に関する検討においては、教育・人材育成の重要性が指摘されている¹³⁷。

今後、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

¹³⁵ ここでステークホルダとしては、開発者（A I のメーカー等のほか、自らが開発したA I システムを用いるA I ネットワークサービスのプロバイダを含む。）、A I システムの供給者（販売者等）、利用者、第三者等が挙げられる。このうち利用者には、A I システム又はA I ネットワークサービスの最終利用者（個人又は企業その他の団体）のほか、他者が開発したA I システムを用いるA I ネットワークサービスのプロバイダが含まれる。

¹³⁶ A I 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、A I システム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A I システム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。

¹³⁷ A I システムの普及による雇用・働き方への影響については、前章5. 参照。以下本章における雇用・働き方への影響に関する記述において同じ。

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
- AI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

(16) 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

シナリオ分析に基づく留意事項においては、「AIシステムに関するリテラシーの向上」、「AIシステムを利活用する者と利活用しない（できない）者とが混在することに伴う問題」、「地域間の格差・偏在」を挙げている。

今後、すべての人々がAIシステム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。

- 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システム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システム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システム利活用による便益を享受できるような環境の整備

(17) セーフティネットの整備

AIシステムの普及による雇用・働き方への影響において、大規模な失業が発生するのではないかと（特に我が国においては、中長期的にAIシステムへの雇用の代替が進む、非正規雇用の労働者への影響が大きいのではないかと）との指摘がある。

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

(18) その他の課題

報告書2016に掲げられている「今後の課題」のうち、上記以外の事項についても、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

(参考) 報告書 2016 に掲げられている「今後の課題」¹³⁸

- AI ネットワーク化の健全な進展に関する事項
 - ・ AI 開発原則・指針の策定
 - AI 開発原則及びその内容の説明から構成される指針(AI 開発ガイドライン)の策定に向けた議論の推進
 - ・ AI ネットワーク化の進展に向けた協調の円滑化
 - 相互接続性・相互運用性の確保(確保の対象や方法の検討等)等
 - ・ 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、ネットワーク形成に関する当事者間の協議の円滑化等
 - ・ 利用者の保護
 - 関連する市場の動向の継続的注視、消費者保護に関する国際的な制度調和の在り方の検討等
- AI ネットワーク上を流通する情報・データに関する事項
 - ・ セキュリティの確保
 - 情報セキュリティ等のAI ネットワークへ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に関する制度的課題
 - プライバシー・パーソナルデータの保護と利活用のバランスに留意した検討等
 - ・ コンテンツに関する制度的課題
 - 多種多様かつ大量のコンテンツの創造・流通に適した法制度の在り方の検討等
- AI ネットワーク化が社会・経済にもたらす影響・リスクの評価に関する事項
 - ・ 影響・リスクに関するシナリオの作成・共有
 - AI ネットワーク化の進展に応じたシナリオの継続的な作成・共有・見直し
 - ・ 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の継続
- 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
 - ・ 社会の基本ルールに関する検討
 - AI ネットワークに関する権利義務及び責任の帰属の在り方や事故時の責任の帰属の在り方等の検討
 - ・ AI ネットワーク化に対応した人材育成

¹³⁸ 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。

- 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等の育成等
- AIネットワークに関するリテラシーの涵養
 - リテラシー教育の在り方の検討、科学コミュニケーションの在り方の検討
- AIネットワーク・ディバイド形成の防止
 - AIネットワーク・ディバイドの要因となるデジタル・ディバイドの解消、高齢者等の利用環境整備等
- AIネットワーク化に対応した就労環境の整備
 - 労働法制の在り方の検討等
- 人間の在り方に関する検討
 - 「AI依存」等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、教育の改革等
- セーフティネットの整備
 - 所得の再分配の在り方の検討

結びに代えて

本報告書では、A I ネットワーク化をめぐる国内外の動向を概観した上で、A I 開発ガイドライン案を提示するとともに、A I ネットワーク化の影響（インパクト及びリスク）の評価を行った上で、それらの検討を踏まえて今後の課題を整理した。

第1章で見たとおり、「A I ネットワーク社会推進フォーラム」において、A I 開発ガイドラインの策定をはじめとするA I ネットワーク化のガバナンスの在り方に関し国際的な議論を継続していくことに対し国内外の参加者から広範な賛同を得られたところであり、今後、G7やOECD等の国際社会において議論が加速することが期待される。また、A I ネットワーク化のガバナンスの在り方に関する国際的な議論をリードすることができるよう、我が国においても、本報告書を踏まえ、産学民官において継続的に議論が進められ、合意形成が図ら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。