

報告書2017

— 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7月28日

AIネットワーク社会推進会議

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 – 』

はじめに

序章

1. 「A I ネットワーク化」と「智連社会」
2. 本推進会議の目的

第1章 A I ネットワーク化をめぐる最近の動向

1. 国内の動向
2. 海外の動向

第2章 A I 開発ガイドライン

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A I 開発ガイドライン案の検討の方向性
3. A I 開発ガイドライン案の概要
4. 今後の展開

第3章 A I ネットワーク化が社会・経済にもたらす影響

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

結びに代えて

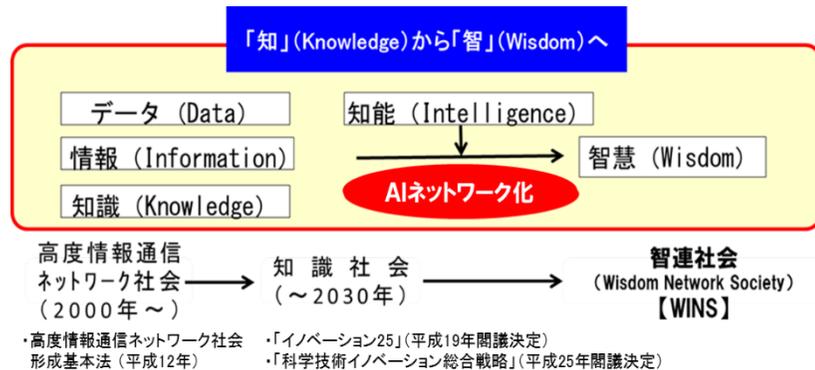
- <別紙1> 国際的な議論のためのA I 開発ガイドライン案
- <別紙2-1> 整理された論点に関する意見募集に寄せられた主な意見に対する考え方
- <別紙2-2> 報告書2017（案）に関する意見募集に寄せられた主な意見に対する考え方
- <別紙3> A I ネットワーク化が社会・経済にもたらす影響 ～先行的評価～
- <別紙4> A I ネットワーク化が社会・経済にもたらす影響 ～分野別評価～
- <別紙5> A I ネットワーク社会推進会議及び分科会構成員一覧
- <別紙6> 開催経緯
- <参考資料> 「A I ネットワーク社会推進フォーラム」関連資料

序章 AIネットワーク化と目指すべき社会像「智連社会」

AIネットワーク化

「AIネットワーク化」とは、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と接続され、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が形成されるようになること。

(注)「AIネットワーク」(Networked AI System)とは、インターネット等情報通信ネットワークと接続されたAIシステムを構成要素として含むシステム。
(備考) 従来の資料においては、「AIネットワークシステム」(“AI Network System”)と呼称。



「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)	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

智連社会(Wisdom Network Society:WINS[ウインズ])

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。

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」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

創造的かつ活力ある発展が可能となる社会

人機共生

総智連環

協調遍在

＜智連社会の基本理念＞

- ・ すべての人々による恵沢の享受
- ・ 人間の尊厳と個人の自律
- ・ イノベティブな研究開発と公正な競争
- ・ 制御可能性と透明性
- ・ ステークホルダーの参画
- ・ 物理空間とサイバー空間の調和
- ・ 空間を越えた協調による活力ある地域社会の実現
- ・ 分散協調による地球規模の課題の解決

第1章 AIネットワーク化をめぐる最近の動向

【国内】

○ 国際シンポジウム「AIネットワーク社会推進フォーラム」(平成29年3月13日・14日)

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、日米欧等の産学民官のステークホルダが意見交換。

< 主な議論の方向性 >

- ・ グローバルな議論を通じた合意形成：継続的に多様なステークホルダー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。
- ・ 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。教育、人材育成が重要。
- ・ 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。開発者は、説明責任を果たすことが重要。

○ 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 閣議決定)

- ・ 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。

【海外】

○ ホワイトハウス【2016年5月～】

大学やNPO等とともに、4回のワークショップを共催し、AIが社会にもたらす便益とリスクに関し検討。意見募集を経て、2016年10月に報告書「人工知能の未来に備えて」を公表しAIに関連する政策を提言するとともに、「米人工知能研究開発戦略」を策定。また、同年12月には、経済・雇用にもたらす影響に関する報告書を公表。

○ Partnership on AI (Amazon、DeepMind/Google、Facebook、IBM、Microsoft)【2016年9月～】

AI技術のベストプラクティスを研究して形成し、AIに関する公衆の理解を向上させ、AI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するために設立。(2017年1月にApple、5月にソニー等が企業構成員として参加)

○ 米国電気電子学会(IEEE)「倫理的に調整された設計」【2016年12月】

道徳的価値及び倫理的原則に即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を提示。

○ Future of Life Institute(FLI)【Jaan Tallinn(Skype共同開発者、DeepMind共同出資者)等、2014年3月～】

研究者、起業家等からなる非営利団体。頑健かつ有益(Robust and Beneficial)なAIを実現するための研究等を実施。2016年2月に、AIの安全性や倫理について指針を定める『アシロマAI原則』を公表。

第2章 国際的な議論のためのAI開発ガイドライン案①

目的

- ・ AIネットワーク化(AIシステムがインターネット等と接続され、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が形成されるようになること)の進展により、人間及びその社会や経済に多大な便益がもたらされる一方、不透明化等のリスクに対する懸念も存在
- ・ オープンな議論を通じ、国際的なコンセンサスを醸成し、非規制的で非拘束的なソフトローたるガイドラインやそのベストプラクティスをステークホルダ間で国際的に共有することが必要
- ・ 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とリスクの抑制により、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を実現

基本理念

- ・ 人間がAIネットワークと共生することにより、その恵沢がすべての人によってあまねく享受され、人間の尊厳と個人の自律が尊重される人間中心の社会を実現
- ・ 非拘束的なソフトローたる指針やそのベストプラクティスをステークホルダ間で国際的に共有
- ・ イノベーティブでオープンな研究開発と公正な競争、学問の自由や表現の自由といった民主主義社会の価値を尊重するとともに、便益とリスクの適正なバランスを確保
- ・ 技術的中立性の確保、開発者に過度の負担を課さないよう配慮
- ・ 不断の見直し、必要に応じた柔軟な改定、広範で柔軟な議論

用語の定義及び対象範囲

- ・ 「AI」: AIソフト及びAIシステムを総称する概念
 - 「AIソフト」: データ・情報・知識の学習等により、利活用の過程を通じて自らの出力やプログラムを変化させる機能を有するソフトウェア(例: 機械学習ソフトウェア)
 - 「AIシステム」: AIソフトを構成要素として含むシステム(例: AIソフトを実装したロボットやクラウドシステム)
- ・ AIシステムの「開発者」及び「利用者」は以下のとおり定義 ※ただし、「開発者」及び「利用者」は場面に応じて個別に決まる相対的な概念
 - 「開発者」: 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)
 - 「利用者」: AI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AI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)
- ・ 対象とするAIシステムの範囲→ネットワーク化され得るAIシステム(ネットワークに接続可能なAIシステム)
- ・ 対象とする開発者の範囲→定義された開発者すべて
- ・ 対象とする開発の範囲→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に限定

関係者に期待される役割

- ・ 各国政府及び国際機関による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、利用者等ステークホルダによる対話やベストプラクティスの共有、AIの便益及びリスクに関する認識の共有
- ・ 標準化団体等による推奨モデルの作成・公表
- ・ 各国政府によるAIの開発者コミュニティの支援、AIに関する研究開発を支援する政策の積極的な推進

第2章 国際的な議論のためのAI開発ガイドライン案②（AI開発原則案の解説）

開発原則	開発原則の解説（留意することが期待される事項等）
<p>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。</p>	<ul style="list-style-type: none"> 相互接続性と相互運用性を確保するため、①有効な関連情報の共有に向けた協力、②国際的な標準や規格への準拠、③データ形式の標準化及びインターフェイスやプロトコルのオープン化への対応、④標準必須特許等のライセンス契約及びその条件についてのオープン・公平な取扱い、などに留意することが望ましい。
<p>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。</p>	<ul style="list-style-type: none"> 採用する技術の特性や用途に照らし合理的な範囲で、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ことが望ましい。 (※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。)
<p>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。</p>	<ul style="list-style-type: none"> AIシステムの制御可能性について、あらかじめ検証及び妥当性の確認をするよう努めるとともに、採用する技術の特性に照らし合理的な可能な範囲において、人間や信頼できる他のAIによる監督・対処の実効性に留意することが望ましい。
<p>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。</p>	<ul style="list-style-type: none"> AIシステムの安全性について、あらかじめ検証及び妥当性の確認をするよう努めるとともに、AIシステムの本質安全や機能安全に資するよう、開発の過程を通じて、採用する技術との特性に照らし可能な範囲で措置を講ずるよう努めることが望ましい。 利用者及び第三者の生命・身体・財産の安全に関する判断を行うAIシステムについては、利用者等ステークホルダに対して設計の趣旨などを説明するよう努めることが望ましい。
<p>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。</p>	<ul style="list-style-type: none"> AIシステムのセキュリティについて、あらかじめ検証及び妥当性の確認をするよう努めるとともに、開発の過程を通じて、採用する技術の特性に照らし可能な範囲で対策を講ずるよう努めることが望ましい(セキュリティ・バイ・デザイン)。
<p>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。</p>	<ul style="list-style-type: none"> AIシステムのプライバシー侵害のリスクについて、あらかじめ影響評価を行うよう努めるとともに、開発の過程を通じて、採用する技術の特性に照らし可能な範囲で措置を講ずるよう努めることが望ましい(プライバシー・バイ・デザイン)。
<p>⑦ 倫理の原則 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。</p>	<ul style="list-style-type: none"> AIシステムの学習データに含まれる偏見などに起因して不当な差別が生じないよう、採用する技術の特性に照らし可能な範囲で所要の措置を講ずるよう努めることが望ましい。 国際人権法や国際人道法を踏まえ、AIシステムが人間性の価値を不当に毀損することがないよう留意することが望ましい。
<p>⑧ 利用者支援の原則 開発者は、AI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。</p>	<ul style="list-style-type: none"> AIシステムの利用者のために、①利用者の判断に資する情報を適時適切に提供し、かつ利用者にとって操作しやすいインターフェイスが利用可能となるよう配慮、②利用者を選択の機会を適時適切に提供する機能が利用可能となるよう配慮、③社会的弱者の利用を容易にするための取組、に留意することが望ましい。
<p>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。</p>	<ul style="list-style-type: none"> 開発原則①～⑧の趣旨に鑑み、利用者等に対しAIシステムの技術的特性について情報提供や説明を行うほか、ステークホルダとの対話を通じた意見聴取や、ステークホルダの積極的な関与を得るよう努めることが望ましい。 AIシステムによりサービスを提供するプロバイダなどとの情報共有・協力を努めることが望ましい。

第3章 AIネットワーク化が社会・経済にもたらす影響①

先行的評価・分野別評価

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、便益)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)を想定したシナリオに基づき評価(シナリオ分析)を実施

先行的評価

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評価

＜対象とするユースケース＞

- ① 災害対応 ② 移動(車両) ③ 健康 ④ 教育・人材育成
- ⑤ 小売・物流 ⑥ 製造・保守 ⑦ 農業 ⑧ 金融(融資)
- ⑨ 公共・インフラ ⑩ 生活

分野別評価

AIシステムを利活用する分野ごとに利用者の視点から評価

＜対象とするユースケース＞

- ★ 公共: まちづくり、パブリック・ガバメント、危機管理
- ★ 個人: 健康、移動、居住、教育・学び、仕事、財産、趣味・娯楽
- ★ 産業: モノ、カネ

※ 本報告書においては、下線の3分野について評価を実施

○ シナリオ分析から得られた示唆

- ・ 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。
- ・ 雇用が減少することが見込まれる業務もあるものの、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。
- ・ 各ユースケースにおいて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。 等

雇用・働き方への影響

- 情報通信分野の発展による雇用への影響の受け皿となっていたサービス業や運輸業務についても影響を受け、大規模な失業が発生するおそれがあるとの指摘がある。
 - ← この指摘については、主観的な予測をベースとしており信頼性は必ずしも高くないこと、新たな雇用創出の可能性が考慮されていないこと等に留意が必要
- 日本の労働市場においては、ルーティンタスクの集約度が高く、AIシステムの普及による影響を受け易い、非正規雇用の労働者が大きな影響を受ける可能性があるとの見方がある。
- 新しく創出される雇用への迅速な適応、円滑な移行のための教育や人材育成が重要である。

第3章 AIネットワーク化が社会・経済にもたらす影響②

留意すべき事項

- 様々なAIシステムが混在することに伴う問題
 - 例えば、性質や性能が異なる多様なAIシステム、新旧のAIシステム、目的が競合するAIシステムがネットワーク上に混在することに伴い、相互間の交渉・調整が必要
- AIシステムを利活用する者と利活用しない(できない)者が混在することに伴う問題
 - 例えば、AIシステムを利活用する者とAIシステムを利活用しない(できない)者とで、受けられるサービスに差異が生ずる可能性
- 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、個人情報への漏洩によるプライバシー侵害が生じた場合等における責任の分配の在り方が問題
- 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、価値の高いデータ(カメラが撮影した画像等)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
- AIシステムが取り扱うデータ・情報の性質
 - 例えば、プライバシー性や機密性が高いデータ・情報については、特に慎重な取扱いが必要
 - 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
- 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、有力なAIネットワークとの接続の可否・条件、データ等へのアクセスの可否・条件が事業活動に影響を及ぼす可能性
 - 例えば、競争関係になる事業者が、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、協調的価格設定等による競争の減殺・制限の可能性
- 人間とAI(ロボット)との役割分担
 - 例えば、対話型ロボットや介護ロボット等による対応で満足できるか、拒絶されないかなど人間とAI(ロボット)との役割分担が問題
 - 例えば、重要な判断(診断、採用、人事評価など)に当たり、AIシステムの機能に委ねてもよい事項の範囲等が問題
- AIシステムに関するリテラシーの向上
 - 例えば、AIシステムの便益を享受するためにも、特に高齢者のAIシステムに関するリテラシーの向上が重要
- 地域間の格差・偏在
 - 例えば、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システムの普及に地域間で差異が生ずる可能性
 - 関連して、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ことが必要
- AIシステムの導入・利活用のためのコスト
 - 例えば、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
- 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、情報公開請求や顧客への説明、監督官庁の検査・監査等に適切に対応できない可能性

第4章 今後の課題

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

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の評価に関する事項

- 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
- 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
- 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