

報告書2018（案）

— A I の利活用の促進及びA I ネットワーク化の健全な進展に向けて —

平成30年●月●日

A I ネットワーク社会推進会議

報告書2018の構成

A I ネットワーク社会推進会議 『報告書2018 - A I の利活用の促進及びA I ネットワーク化の健全な進展に向けて - 』

はじめに

序章

1. 「A I ネットワーク化」と「智連社会」
2. 本推進会議における取組

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
3. 国際的な議論の動向

第2章 A I ネットワーク化が社会・経済にもたらす影響

1. A I ネットワーク化するシナリオ分析（分野別評価）
2. 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
3. 教育・人材育成
4. A I ネットワーク化の進展に関する指標

第3章 A I 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
2. 関係する主体の整理
3. A I 利活用原則案
4. A I 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

結びに代えて

<別紙1> A I ネットワーク化するシナリオ分析（分野別評価）

<別紙2> 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

<別紙3> 教育・人材育成に関する主な意見

<別紙4> 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5> 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6> 開催経緯

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。

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

人機共生

総智連環

協調遍在

創造的かつ活力ある発展が可能となる社会

<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
- ・ イノベー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

(注) 主なものを記載

国内の動向

○ 「人間中心のAI社会原則検討会議」開催（平成30年5月8日～）

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、G7や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・利活用などについて、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、政府は、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。

海外の動向

○ 米国電気電子学会（IEEE） 「倫理的に調整された設計 第2版」公表（平成29年12月12日）

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、悪用への警戒を掲げる。第1版での検討テーマに加え、感情コンピューティング、政策、AIにおける古典倫理、ICTにおける複合現実、幸福をテーマとする章を新たに追加。

○ 欧州委員会独立諮問機関 the Group on Ethics in Science and New Technologies (EGE) 声明公表（平成30年3月9日）

AI・ロボティクス・自律システムの構築、利活用及びガバナンスなどに関する国際的な理解の確保を欧州委員会に対し要求。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倫理原則（人間の尊厳、自律性、責任、正義・公平・連帯、民主主義、法の支配とアカウントビリティ、セキュリティ・安全性・心身の整合性、データ保護とプライバシー、サステナビリティ）を提案。欧州委員会は、上記声明を踏まえたAI倫理ガイドラインを策定、2018年末までに公開予定。

国際的な議論の動向

○ OECD・総務省共催による国際カンファレンス「AI : Intelligent Machines, Smart Policies」開催（平成29年10月26～27日）

世界各国の産学官の有識者が一堂に会し、AIの発展・普及が社会経済にもたらす様々な機会と課題、政策の役割と国際協調の在り方などについて、議論が行われた。AI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AI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。

○ OECDデジタル経済政策委員会（CDEP）開催（平成29年11月21～22日、平成30年5月16日～18日）

AIに関する今後の取組について議論が行われ、平成31年以降理事会勧告作成に向けた作業に着手することについて加盟国の合意が得られた。また、事務局から分析レポートが報告され、平成30年11月の会合で最終化の予定。

○ G7情報通信・産業大臣会合（イタリア、平成29年9月25～26日）

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AIというビジョンを共有。閣僚宣言サマリー「附属書2:我々の社会のための人間中心のAIに関するG7マルチステークホルダー交流」において、AI開発ガイドライン案に言及。

○ G7イノベーション大臣会合（カナダ、平成30年3月27～28日）

「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。議長サマリー「附属書B:AIに関するイノベーション会合声明」において、AI開発ガイドライン案に言及。

第2章 AIネットワーク化が社会・経済にもたらす影響①

シナリオ分析(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

➤ 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(本報告書においては、下線の3分野について評価を実施)

＜対象とするユースケース＞

※「まちづくり」、「健康」、「モノ」については、報告書2017で評価を実施

★ 公共：まちづくり、**行政**、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

○ シナリオ分析から得られた示唆

- ・ 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。
- ・ 特に高齢者に有益なAIシステムの利活用と考えられるため、高齢者のリテラシーを向上させる方策が求められる。 等

エコシステムの展望

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を展望し、AIの利活用における便益や課題を整理

＜利用シーンにおけるAIの利活用の展望＞

- 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、豊かな人生
- 企業における利用：ヒト、モノ、カネ、情報

＜ケーススタディ＞

- 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

＜ネットワーク化の観点からの課題＞

- 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

第2章 AIネットワーク化が社会・経済にもたらす影響②

<エコシステムの展望において整理された課題>

【主として生命・身体の安全、権利・利益等を守るための課題】

- ・ 生命・身体・財産の安全に関する課題(事故の防止など) : 安全
- ・ AI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
- ・ 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AIとの関係等に関する課題】

- ・ 人間とAI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
- ・ AI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】

- ・ AI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
- ・ 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
- ・ AI間の連携に関する課題(AI間の交渉・調整など) : 連携

【主としてデータに関する課題】

- ・ AI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

これらの課題を踏まえ、
「AI利活用原則案」
を取りまとめ

教育・人材育成

AIネットワーク社会に向けた教育・人材育成の在り方等を検討(主な意見を紹介)

- 人間とAIとの関係・役割分担 : 人間中心を意識する必要がある。価値判断やコミュニケーション、創造性がポイント。
- 社会の変化及び人間の技能・能力 : 幅広く基礎を押さえることが重要。教養、コミュニケーション能力、数学、哲学が重要。
- 学校教育 : 現状の学校教育に強い危機感。「自分で学ぶ」という意識改革が重要。
- 職業スキルの習得 : 「学び続ける」ことが重要。今後は企業内での人材育成が困難。
- 生涯教育 : リカレント教育が重要。教育全体をデザインする必要がある。早く議論を始めるべき。
- AIの利活用 : 個人に応じた教育パターンを作れる。学ぶ・自己実現のための時間を確保のためにAIを利活用。

指標

AIネットワーク化の進展に関する指標を検討

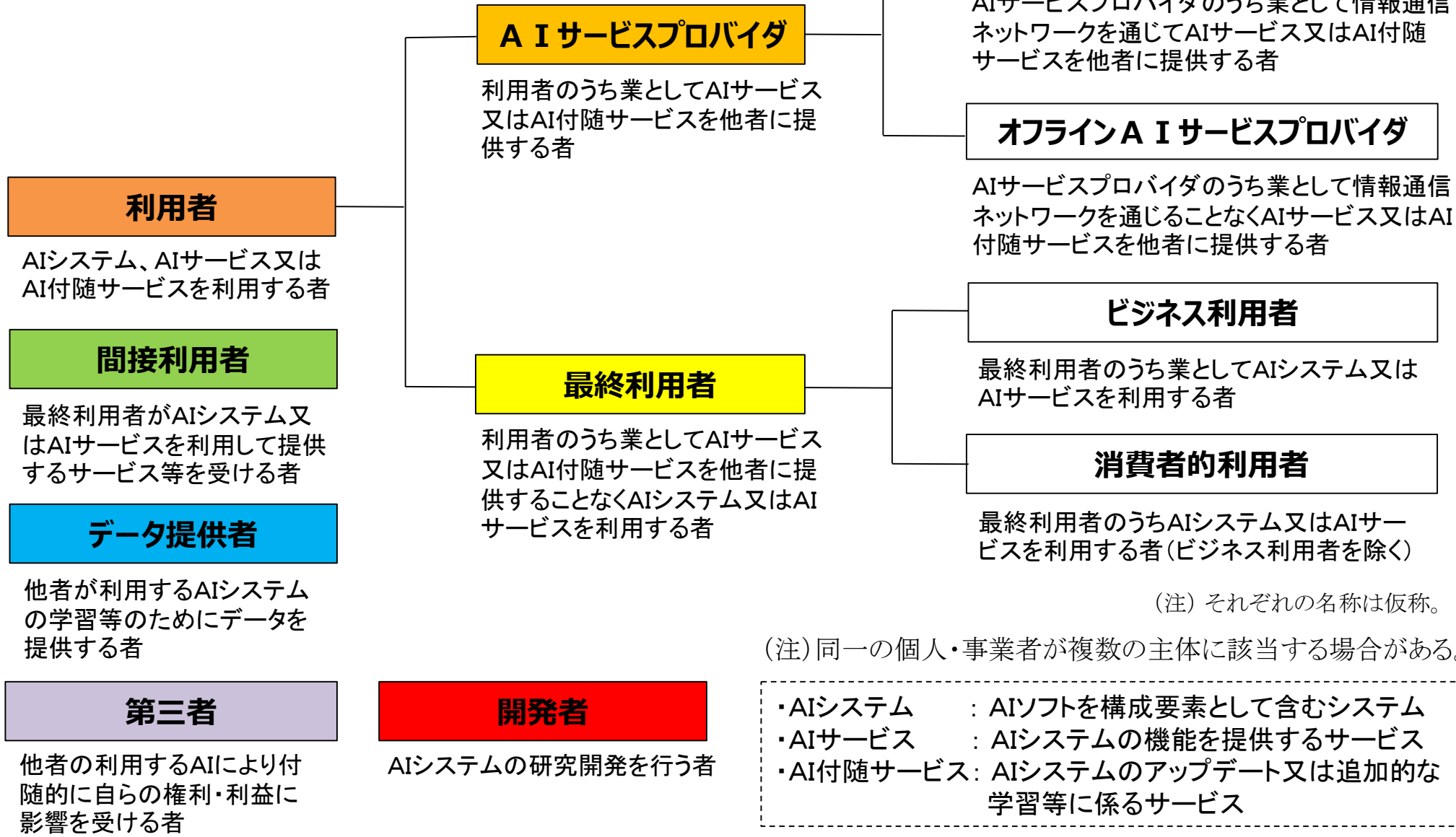
- AIネットワーク化事前指標(AIネットワーク化の前提となる環境整備に関するもの)
- AIネットワーク化初期指標(AIネットワーク化の初期段階に関するもの)



引き続き、統計データの収集やアンケート調査における有用な項目の検討を実施

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

関係する主体の整理



(注) それぞれの名称は仮称。

(注) 同一の個人・事業者が複数の主体に該当する場合がある。

- ・AIシステム : AIソフトを構成要素として含むシステム
- ・AIサービス : AIシステムの機能を提供するサービス
- ・AI付随サービス: AI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

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②

AI利活用原則案

国際的な議論のためのものとして、また、非規制的かつ非拘束的なものとして取りまとめ

① 適正利用の原則 [安全][役割分担]

利用者は、人間とAI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AIシステム又はAIサービスを利用するよう努める。

② 適正学習の原則 [データ][正当性・公平性]

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。

③ 連携の原則 [連携]

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。

④ 安全の原則 [安全]

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。

⑤ セキュリティの原則 [セキュリティ]

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。

⑥ プライバシーの原則 [プライバシー]

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 [正当性・公平性]

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。

⑧ 公平性の原則 [正当性・公平性]

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。

⑨ 透明性の原則 [ブラックボックス化]

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。

⑩ アカウンタビリティの原則 [受容性]

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。

主に
便益の増進
に関係

主に
リスクの抑制
に関係

主に
信頼の醸成
に関係

(注) AIの開発において留意することが期待される事項については、本推進会議において「国際的な議論のためのAI開発ガイドライン案」を取りまとめた(『報告書2017』)。関係するステークホルダ(政府、業界団体等)が取り組む環境整備に関する課題については、第4章「今後の課題」において整理している。

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③

AI利活用原則案の各原則の内容に関する主な論点

AI利活用原則案	各原則の内容に関する主な論点(例)
① 適正利用の原則	<ul style="list-style-type: none"> ・ 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、AI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、AI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AIの利活用が適正な範囲・方法で行われているか定期的に確認することが期待されるのではないか。 ・ 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。
② 適正学習の原則	<ul style="list-style-type: none"> ・ 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)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質に留意することを期待することが適当か。
③ 連携の原則	<ul style="list-style-type: none"> ・ 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。
④ 安全の原則	<ul style="list-style-type: none"> ・ 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。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。
⑤ セキュリティの原則	<ul style="list-style-type: none"> ・ 利用者は、AIのセキュリティに留意し、その時点で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。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。
⑥ プライバシーの原則	<ul style="list-style-type: none"> ・ 利用者は、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、AI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。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。
⑦ 尊厳・自律の原則	<ul style="list-style-type: none"> ・ 利用者は、AIを利活用する際の社会的文脈を踏まえ、他者の尊厳と自律を尊重することが期待されるのではないか。 ・ 利用者には、AIにより意思決定や感情が操作されるリスクや、AIに過度に依存するリスクに留意することを期待することは適当か否か。
⑧ 公平性の原則	<ul style="list-style-type: none"> ・ 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、AIの学習等に用いられるデータの代表性やデータに内在する社会的なバイアスに留意することが期待されるのではないか。 ・ 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当か。
⑨ 透明性の原則	<ul style="list-style-type: none"> ・ 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。 ・ 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。
⑩ アカウンタビリティの原則	<ul style="list-style-type: none"> ・ 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。

第4章 今後の課題

AIの開発及び利活用並びに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の利活用に関する指針の策定: 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た最終アウトプットの取りまとめに向けた検討
- 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: ステークホルダ間の協力・ベストプラクティスの共有、法制度等の在り方の検討等
- AIシステム又はAIサービス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)の在り方の検討等
- 技術開発に関する課題: AIネットワークのセキュリティを高めるための技術やセキュリティ攻撃の際の早期復旧のための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、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、AI生成物の知的財産制度上の在り方の検討等

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・国際的な共有等
- 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のリスクが顕在化した場合の責任の分配の在り等々の検討
- 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