

研究報告

AI ネットワーク化に関する 社会的・経済的・倫理的・法的課題

福 田 雅 樹

1. はじめに

(1) 報告の構成

御紹介にあずかりました総務省情報通信政策研究所の福田雅樹でございます。私からは、AI ネットワーク化に関する社会的・経済的・倫理的・法的課題に関し報告いたします。ただし、本日の報告のうち意見にわたる部分は私見であることをあらかじめ付言いたします¹⁾。

この報告の構成は、次のとおりであります。

第一に、緒論といたしまして、関連する議論の沿革を御紹介いたしますとともに、この報告の題名においても言及いたしております「AI ネットワーク化」の意義及び進展段階について御説明いたします。

第二に AI ネットワーク化の進展が社会・経済にもたらし得るインパクトについて御説明し、第三に AI ネットワーク化が社会・経済にもたらし得るリスクについて御説明いたします。

第四に、AI ネットワーク化の進展が社会・経済にもたらすインパクト及びリスクを踏まえ、AI ネットワーク化の進展をめぐる規範又は政策に関し、その基本的な方向性、形成の手法の在り方及び検討の方法の在り方についての問題提起をいたします。

第五に、AI ネットワーク化の進展をめぐる規範又は政策を形成するに当たっての指導理念となるものとして、AI ネットワーク化の進展を通じて目指すべきトゥルー・ノース (true north) としての社会像及びその基本

1) 本稿は、脚注を含むその全体を通じて、意見にわたる部分は、筆者の個人的な見解に基づくものである。

理念について御説明いたします。

第六に、AI ネットワーク化に関する社会的・経済的・倫理的・法的課題について御説明いたします。

最後に、この報告の結びに代えて、関連する国内外の動向を御紹介いたします。

(2) 関連する議論の沿革

総務省情報通信政策研究所においては、平成 27 年 2 月から、AI の高度化や普及に伴って進展することが期待される事象たる「AI ネットワーク化」に関する検討を進めております。同月から同年 6 月までにかけて開催した「インテリジェント化が加速する ICT の未来像に関する研究会」が一連の検討の嚆矢であります。この研究会においては、萌芽的な問題意識を整理し、今後の検討の基本的な方向性に関する提言を賜りました。この研究会及びその「報告書 2015」²⁾においては、AI の高度化や普及をめぐる動向と IoT、ビッグデータ等 AI 以外の情報通信分野の発展の動向をめぐる社会の変化を俯瞰的に捉えて「ICT インテリジェント化」と総称いたしておりました。「報告書 2015」においては、ICT インテリジェント化の進展を展望した上で、その社会や経済への便益その他のインパクト及び社会や経済にもたらすリスクを評価すべきこと等の提言を賜りました。

この提言を踏まえ、平成 28 年の 2 月から 6 月までにかけて、理工学系及び人文・社会科学系の研究者 37 人の御参加を得て、「AI ネットワーク化検討会議」を開催いたしました（開催時点の名称は「ICT インテリジェント化影響評価検討会議」でありましたが、第 2 回会合において改称いたしました。）。検討会議においては、ICT インテリジェント化について、これを牽引する AI をめぐる変化に着目して「AI ネットワーク化」として捉えなおした上で、AI ネットワーク化の進展を展望し、AI ネットワーク化の進展を通じて目指すべき社会像及びその基本理念を整理するとともに、AI ネットワーク化の進展が社会・経済にもたらす影響及びリスクの基礎的な評価を行い、これらの成果を踏まえて社会的・経済的・倫理的・法的

2) インテリジェント化が加速する ICT の未来像に関する研究会「報告書 2015」(総務省情報通信政策研究所、平 27)。

課題を整理いたしました。検討会議は、平成 28 年 4 月には中間報告書³⁾、そして 6 月には同検討会議の報告書といたしましては最終報告書となる「報告書 2016」⁴⁾を公表いたしております。

AI ネットワーク化検討会議における提言を踏まえ、御議論を更に進めていくために、平成 28 年 10 月から、検討会議を発展的に改組した会議体として「AI ネットワーク社会推進会議」を開催いたしております。推進会議は、多様なステークホルダにより学術的にも実務的にも高水準で多角的かつ実効的な議論を展開していただくべく、産学民の有識者及び利害関係者七十人余りの構成員並びに関係行政機関等からの十人余りのオブザーバの御参加を得た体制といたしております⁵⁾。

(3) 「AI ネットワーク化」の意義

AI の中には他のシステムと連携させずに単独で利活用してもその便益をその利用者にもたらし得るものもありますが、AI がその入力されたデータ等を操作して得られるデータ等に基づく出力をする機能を有することをその通有性とするシステムである⁶⁾ ことに加え、人工知能学会によると AI のほとんどの分野で利用されているとされる機械学習⁷⁾ がその入力されるデータ等が増えるにつれてその性能が向上するものであることから、AI の多くは、そのデータ等の入力元又は出力先を得るため、インターネット

3) AI ネットワーク化検討会議「中間報告書 AI ネットワーク化が拓く智連社会（WINS）—第四次産業革命を超えた社会に向けて—」（総務省情報通信政策研究所、平 28）。

4) AI ネットワーク化検討会議「報告書 2016 AI ネットワーク化の影響とリスク—智連社会（WINS）の実現に向けた課題—」（総務省情報通信政策研究所、平 28）。

5) AI ネットワーク社会推進会議は、平成 29 年 7 月までの議論の成果を報告書として公表している（AI ネットワーク社会推進会議「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—」（総務省情報通信政策研究所、平 29））。

6) 福田雅樹「『AI ネットワーク化』およびそのガバナンス『智連社会』に向けた法・政策の視座」福田雅樹＝林秀弥＝成原慧編『AI がつなげる社会—AI ネットワーク時代の法・政策』7 頁（弘文堂、平 29）参照。

なお、この報告においては AI が社会にもたらす便益やリスクに着目していることから、この報告において「AI」としては、特に断らない限り、システムとして機能し得る状態にあるものを想定している。

7) 機械学習が人工知能の他のほとんどの分野において利用されていることについて、人工知能学会「人工知能研究」（<http://www.ai-gakkai.or.jp/whatsai/AIresearch.html>）〔平成 30 年 2 月 1 日最終閲覧〕参照。

等情報通信ネットワークと接続され、当該情報通信ネットワークを介して他のシステムと連携させて利活用されるようになるものと見込まれます。また、後程御説明いたしますが、AIがこのような形で利活用されれば、当該連携がなければ得られない便益を得ることができます。

AIがこのような形で利活用される場合には、当該AI、当該情報通信ネットワーク及び当該他のシステムからなるネットワークの形成を観念することができます。このネットワークのことを「AIネットワーク」といいます⁸⁾。社会においてAIがインターネット等情報通信ネットワークと接続され、AIネットワークが形成されるようになること及び形成されたAIネットワーク相互間の連携等AIネットワークの高度化が進むことを「AIネットワーク化」と総称いたしております⁹⁾。

「AIネットワーク化」という概念は、AIネットワーク化検討会議において、AIを起点に置いて見た社会に生ずる変化を一連の事象として把握するために定式化した概念であります。翻って、AIのみならず様々なシステムがインターネット等情報通信ネットワークを介して互いにつながって様々なデータ等を流通させて連携するという事象が進んでいる社会を起点とすれば、「AIネットワーク化」については、そのような社会において情報通信ネットワークを介して他のシステムとつながってデータ等を流通させて連携し得る存在（すなわち、そのような社会において形成が進んでいく様々なシステムが複雑につながるネットワークの構成要素となり得る存在）としてAIが加わっていく事象をいうものとして捉え直すこともできるものと考えられます。

この報告においては、「利用者」に幾度も言及いたします。以下この報告において「利用者」とは、AIネットワークの構成要素たるAIの機能を利活用する¹⁰⁾個人又は団体をいいます。自ら又は他人が開発したAIを自

8) この定義は、AIネットワーク社会推進会議・前掲注（5）3頁に掲げる「AIネットワーク」の定義を筆者が敷衍したものである。AIネットワーク化検討会議・前掲注（3）及び同・前掲注（3）においては「AIネットワーク」のことを「AIネットワークシステム」と呼んでいたが、この報告においては、AIネットワーク社会推進会議・前掲注（5）における用語法に合わせるよう、その全体を通じて「AIネットワーク」と呼んでいる。

9) この定義は、AIネットワーク化検討会議・前掲注（3）5頁及びAIネットワーク社会推進会議・前掲注（5）3頁に掲げる「AIネットワーク化」の定義を筆者が敷衍したものである。

10) この報告において「利活用」とは、利用又は活用をいう。「利用」とはその客

ら設置して AI ネットワークを形成し、自ら当該 AI の機能を活用する個人又は団体は典型的な利用者ではありますが、AI の機能を情報通信ネットワークを介して他人の用に供する役務（以下「AI ネットワークサービス」といいます。）を他人に有料又は無料で提供する個人又は団体（以下「プロバイダ」といいます。）から AI ネットワークサービスの提供を受けてその AI の機能を活用する個人又は団体（以下「AI ネットワークサービスの利用者」という場合があります。）が多くなると見込まれます¹¹⁾。

「利用者」は、最終利用者（利用者であって、自らが活用する AI の機能を他人に提供しない者をいいます。）とプロバイダに大別されます。最終利用者は、自らが設置する AI の機能を活用する最終利用者と、AI ネットワークサービスの利用者たる最終利用者（以下「AI ネットワークサービスの最終利用者」という場合があります。）に大別されます。プロバイダは、自らが設置する AI を用いる AI ネットワークサービスのプロバイダ（以下「第一次プロバイダ」といいます。）と、他のプロバイダから提供を受けた AI ネットワークサービスを他人に有料又は無料で再販するプロバイダ（以下「再販プロバイダ」といいます。）¹²⁾ に大別されます。

AI ネットワーク化をめぐる諸問題の主たる関係者としては、利用者（自らが設置する AI を構成要素とする AI ネットワークの最終利用者、AI ネットワークサービスのプロバイダ（第一次プロバイダ、再販プロバイダ）、AI ネットワークサービスの最終利用者）のほか、開発者（情報通信ネットワークに接続され得る AI の研究又は開発を行う個人又は団体をいいます。）

（この文の場合には、AI ネットワークの構成要素たる AI の機能）を自らのために用いる行為自体をいい、「活用」とはその客体を用いることにより得られる便益その他の効用を自らの何らかの目的のために活かすことをいう。AI 及び AI ネットワークに関する限り利用と活用は一体的に行われることが多いことに鑑み、この報告においては、「利用」と「活用」とを一体的に捉えたものを「利活用」という。

- 11) 管轄権の問題を掛けば、AI ネットワークサービスは電気通信事業法 2 条 3 号に規定する電気通信役務に該当し、AI ネットワークサービスを他人の需要に応ずるために提供する事業は同条 4 号に規定する電気通信事業に該当するものと解される。ただし、当該事業が同法 164 条 1 項各号に掲げる電気通信事業である場合には、当該事業については、同条 3 項に定める例外を除き、同法の規定は適用されない。
- 12) この「再販プロバイダ」は、当該他のプロバイダ（第一次プロバイダ又は他の再販プロバイダ）との関係では当該 AI ネットワークサービスに用いられる AI ネットワークの利用者であり、当該他人との関係では当該 AI ネットワークサービスを用いて自らが提供する AI ネットワークサービスのプロバイダである。

す。)¹³⁾、開発者による研究若しくは開発のための作業に自ら従事し又は当該作業を補助する者（開発者の受託者、使用人等）、データ等を利用者又は開発者に対し提供する者、情報通信ネットワークに接続され得る AI を他人に販売等する者（販売店等）、他人のために AI ネットワークを構築する作業を行う者（システムインテグレータ等）、AI ネットワークの用に供し又は AI ネットワークと接続する情報通信ネットワークを運営する者（電気通信事業者等）、最終利用者から AI ネットワークに係る AI の機能を中間投入要素とする製品又は役務の供給を受ける者、AI ネットワーク又は AI ネットワークサービスの利活用をめぐる事故、犯罪等の被害者等が挙げられます。ただし、ある者が利用者、開発者等のいずれに当てはまるのかは、その着目する具体的な場面に応じて異なり得ます¹⁴⁾。

(4) AI ネットワーク化の進展段階

AI ネットワーク化検討会議及び AI ネットワーク社会推進会議においては、AI ネットワーク化がどのように進展した状況において、どのような影響やリスクが見込まれ、どのような課題が考えられるのかということを検討するための物差しといたしまして、あくまでも理念的な整理ではありますが、AI ネットワーク化の進展を次の四段階からなるものとして整理いたしております¹⁵⁾。

第一段階は、AI が、インターネット等情報通信ネットワークと接続され、当該 AI が当該情報通信ネットワークを介して AI 以外のモノからなるシステム（例えば、AI を実装しないクラウドやエッジのシステム、IoT のデバイス、最終利用者が操作する端末機器等）と連携する AI ネットワークを形成することにより、AI について見れば他の AI とは連携せず機能するという段階であります。この段階は、今まさに国内外の様々な分野において急速に進展しております。

13) この報告において「開発者」としては、研究又は開発をめぐる権利が帰属する個人又は団体を想定している。

14) 例えば、自らが開発者たる AI を自ら設置して AI ネットワークを形成し、自ら当該 AI の機能を利活用する個人又は団体は、当該 AI の開発者であると同時に当該 AI ネットワークの利用者でもある。福田・前掲注 (6)37 頁参照。

15) AI ネットワーク化検討会議・前掲注 (3)6-10 頁及び AI ネットワーク社会推進会議・前掲注 (5)4 頁を参照。

第二段階は、複数の AI がインターネット等情報通信ネットワークを介して相互に連携する AI ネットワークを形成する段階であります。この段階においては、相互に連携する複数の AI 又は AI ネットワークを取りまとめる AI 又は AI ネットワークも出現するようになり、一の利用者の下で、又は複数の利用者の下で、様々な AI が様々な情報通信ネットワークを介して垂直的又は水平的な関係を形成して様々な形で立体的に連携するようになり、また、解決すべき問題に応じて連携の内容や組み合わせを動的に変更するようになるものと展望されております。

第三段階は、第二段階の後でなければならないというのではなく、AI ネットワークが利活用される分野によっては、第一段階の次に第三段階が実現した後に第二段階が実現するということもあり得ます。この第三段階は、センサやアクチュエータを実装する AI ネットワークと人体の器官（手足、目、耳、脳等）とが連携する段階であります。例えば、センサを実装する AI ネットワークとの連携を通じて、遠隔地の状況を知覚できるようにする等人間の感覚器官の能力を拡張することができるようになり、また、ロボットの手足等アクチュエータを実装する AI ネットワークとの連携を通じて、遠隔地にあるロボットを操作できるようにする等人間の身体機能を拡張することができるようになるものと展望されております。

これら AI と AI 以外のモノからなるシステムとの間の連携、複数の AI 相互間の連携及び AI ネットワークと人体の器官との間の連携がそれぞれ社会のあらゆるところで進展し、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、社会のあらゆる場面において、AI、モノ及びヒトの相互間における様々な連携を通じて、ヒト相互間の AI ネットワークを介した連携が実現する段階に至るものと展望されております。

2. AI ネットワーク化の進展が社会・経済にもたらし得るインパクト

(1) 総論

AI をインターネット等情報通信ネットワークと接続し、他のシステムと連携させて当該 AI の機能を利活用する場合には、その連携の内容に応

じて、当該他のシステムから出力されるデータ等を用いて当該 AI の機能を利活用し、又は当該 AI から出力されるデータ等を当該他のシステムの動作のために提供することのほか、当該 AI を介して当該他のシステムを操作すること、当該他のシステムからの操作を受けて当該 AI の機能を利活用すること等により、当該連携がなければ得られない便益が得られます。

特に複数の AI を相互に連携される場合には、個々の AI の機能のみに基づく便益に加えて、当該連携がなければ得られない便益として、複数の AI が共通の問題に取り組めるようになること¹⁶⁾、個々の AI をモジュールとして組み合わせることにより高度な機能を実現し得るようになること¹⁷⁾、連携に係る AI を利活用する業務全体を状況に応じて動的に調整し得るようになること¹⁸⁾、連携に係る AI を利活用する役務を連動させることにより可能となる新たな役務が提供されるようになること、連携に係る AI を利活用する業務自体の融合が進むこと¹⁹⁾等の便益がもたらされる得るものと展望されております。

これらの便益は、連携の相手となるシステムが増えるにつれて高まり得るものであります。特に、機械学習の機能を有する AI 等利活用の過程において入力されるデータ等が増えるにつれて性能が向上する AI の便益は、その連携の相手となるシステムが増えることに伴って入力が増えるにつれてその性能自体が向上することによっても更に高まり得ます²⁰⁾。しかも、これらの便益は、時には実空間とサイバー空間の境界を越え、時には国境を越え、即座にもたらされ、波及し得るものであります²¹⁾。

AI ネットワーク化が社会の様々な場面において進展していけば、これら便益の発生、増大及び波及により、人間の活動の効率化や高度化、人間

16) 中西崇文『シンギュラリティは怖くない—ちょっと落ちついて人工知能について考えよう』172-175 頁(草思社、平 29) 参照。

17) 中西・前掲注 (16) 176-178 頁参照 (モジュールとする AI の組み合わせを動的に変更することにより、必要となる機能をリアルタイムで実現し得るようになる可能性をも指摘している。)

18) AI ネットワーク社会推進会議・前掲注 (5) 34-37 頁及び別紙 3 参照。

19) AI ネットワーク社会推進会議・前掲注 (5) 38-42 頁及び別紙 4 参照。

20) 関連する指摘として、ケヴィン・ケリー (服部桂訳)『＜インターネット＞の次に来るもの—未来を決める 12 の法則』42-43 頁 (NHK 出版、平 28) を参照。

21) 便益の具体例については AI ネットワーク化検討会議・前掲注 (3) 12-27 頁並びに AI ネットワーク社会推進会議・前掲注 (5) 別紙 3 及び別紙 4 を、便益の波及の可能性については同書 3 頁、29 頁及び 49 頁を参照。

が活動し得る領域の拡大等様々な恵沢が人間の社会・経済に幅広く連鎖的にもたらされ、その結果として社会・経済の構造も幅広く連鎖的に変化していくものと展望されます。

特にAI ネットワーク化の進展が産業にもたらすインパクトについては、AI ネットワーク化の進展を通じて隣接業界への進出や異業種間の連携による新たな製品・役務やビジネスモデルの創出が進み、ひいては製造業のサービス業化に代表される異業種間の融合等産業構造の変革が進んでいくとともに、あらゆる産業において情報通信産業化が進展していくという評価も可能であろうとの見方が示されております²²⁾。

このようにAI ネットワーク化が進展することにより様々な恵沢が社会・経済に幅広くもたらされ得るものと展望されている一方で、その副作用として、これまで人間が自らの知能や身体機能を用いて行ってきた活動がAI ネットワークに代替されることによる技術的失業の増加等負のインパクトの可能性も懸念されております。AI ネットワーク化が社会・経済にもたらす恵沢を豊かなものとするとともに負のインパクトを抑制することが今後の課題となります。

(2) 社会・経済の分野ごとのインパクト

AI ネットワーク化検討会議においては、AI ネットワーク化の進展により社会・経済の各分野においてどのようなインパクトがもたらされ得るのかについて、社会・経済の分野を大きくは公共（まち）、生活（ひと）及び産業（しごと）に区分し、細かくは16に区分いたしまして、その時点における国内外の先行研究、構成員の御意見等を踏まえ、その展望を整理いたしました。次のスライドは、社会・経済の分野ごとのインパクトに関する展望の例を掲げるものであります²³⁾。

22) AI ネットワーク化検討会議・前掲注 (3)35-36 頁及び同・前掲注 (4)10-11 頁を参照。

23) このスライドは、AI ネットワーク化検討会議・前掲注 (3)18-27 頁に基づき筆者が整理したものである。

AIネットワーク化の進展に伴い社会・経済の各分野にもたらし得るインパクトの例	
分野	インパクトの具体例
公共	
公共インフラ	-公共インフラに関する需要と供給のリアルタイムでのデータの収集・分析により、異常気象、災害など急な環境変化に即時に対応 -メンテナンスのオートメーション化により、効率化を実現
防災	-災害の影響のリアルタイムでの予測の高度化及びこれと連動した避難誘導の実現
スマートシティ	-街全体において街頭カメラの活用やエネルギーマネジメントの実現により、効率的で、快適かつ安全な街の実現
行政	-AIシステムによるオープンデータの分析結果の活用により、行政の水準の向上 -個人や企業から発信されるデータ、情報等を活用した将来予測の実現により、政策の立案が精緻化
生活	
生活支援 (パーソナルアシスト)	-身体、室内のセンサーやロボットを活用することにより、各人の生活パターンに沿った家事等雑務支援が実現 -人間との自然な会話が可能でAIシステムが出現
個人の活動	-3Dプリンター等を利用したパーソナルファブリケーションが普及することと相俟って、製品・サービスの利用者によるカスタマイズが一般化 -センサやメディアの発達により出会いの支援や体験共有が高度化することにより、他者とのつながり方が質的に変化
産業	
領域共通	-バックオフィス業務等単純作業の自動化が進み、個人適応された自動化（自分代行秘書サービス等）の実現 -自動栽培や農業用ドローン、インテリジェントファームিং等の実現による、生産効率の向上や収穫量の拡大
農林水産	
製造	-製造プロセスとサプライチェーンのスマート化により、動的な需給バランスに対応した生産最適化や高度な多品種少量生産（マスカスタマイゼーション）が実現 -利用者の稼働データの分析により、デジタルマーケティングや、付加価値の高いアフターサービス・メンテナンスサービスが実現 -製品の設計段階からの自動化の実現により、開発作業が効率化・高速化
運輸・物流	-自動運転レベルの向上により、事故の減少、渋滞の解消、環境負荷の低減、地方や高齢者等の交通難民の解消が進展
卸売・小売	-インテリジェントコマースや購買レコメンデーション等個々の顧客のデータのきめ細かい分析結果の活用が進み、消費が喚起
金融・保険	-リスク評価の精緻化等により、商品・サービスの高度化・多様化 -トレーディング、ローン審査、与信管理の自動化が普及
医療・介護	-患者のバイタルデータによる発病予測や遠隔予備情報による健康管理等の実現により、健康寿命が延伸 -研究論文の自動分析の実現により、研究や新薬開発が加速
教育・研究	-教科の学習からキャリアの設計に至るまで、個人に応じたきめ細かい教育が進展 -優れた実演家や熟練技術者、クリエイター等の「暗黙知」の「形式知」化により、教育内容のアーカイブ化が実現し、教育の質が向上
サービス業	-営業業務や、店舗におけるバックヤードの作業、コールセンターにおける応答の業務等のうち、比較的単純な作業について、ロボット等による自動化の進展 -不動産の適正価格の自動評価等の実現により、不動産取引の円滑化
建設	-危険作業や密着作業へのロボット技術の導入等により、建設現場が女性、高齢者等にとって従事しやすいものに変化 -建造物の劣化度がわかるセンサや、データの高度な解析による新しい機能性材料の開発により、建築物の安全性が向上

これら社会・経済の分野ごとのインパクトの展望は、平成 28 年 4 月に整理したものであり、既に陳腐化しているかもしれません。今後、AI ネットワーク社会推進会議において、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を逐次形成し、分析を進めていく予定であります²⁴⁾。

3. AI ネットワーク化の進展が社会・経済にもたらし得るリスク

(1) 総論

ここまでの御説明のとおり、AI ネットワーク化の進展は、利用者に様々な便益をもたらし、もって便益に基づく恵沢を社会・経済に幅広く連鎖的にもたらし得るものと見られておりますが、その一方で、AI ネットワーク化の進展が社会・経済にもたらし得るリスクについても指摘されており

24) その成果の一部として、AI ネットワーク社会推進会議・前掲注 (5) 別紙 3 及び別紙 4 を参照。

ます。ここで「リスク」とは、将来生ずるかもしれない望ましくない事象と人々が認識するものをもってリスクとする構築主義的リスク観の立場におけるリスクをいいます²⁵⁾。

機械学習の機能を有する AI のように、利活用の過程において入力されるデータ等からの学習等により自らの出力やプログラムを将来に向けて継続的に変化させる機能を有する AI については、利活用の過程における学習等による変化に起因して予見、理解、検証又は制御が困難な事象が生ずるおそれがあるため、従来の情報通信技術とは異なる特別の留意が求められることが指摘されております²⁶⁾。

また、AI ネットワーク化が社会において広く進展していけば、用途、機能、性能、新旧、学習に用いたデータ等様々な面を異にする様々な AI がインターネット等を介して直接又は間接につながる情報通信ネットワーク上に混在するようになり²⁷⁾、社会の様々な分野に属する様々な属性の個人又は団体が様々な意図に基づきこれら様々な AI を他の様々なシステムと連携させる AI ネットワークを形成し、様々な場面で様々なデータ等を入出力させて利活用し、相互に影響を及ぼし得るようになるという複雑な状況が連鎖的に変化し続けていくようになるものと考えられます。

さらに、将来的には AI と他のシステムとの連携の組み合わせを自動的に変更することができるようになる可能性も指摘されています²⁸⁾。そのような可能性が現実化すれば、その結果として AI が自律的な成長を遂げる可能性も生ずるという展望も示されております²⁹⁾。AI が自律的に成長を遂げるとなると、AI ネットワークをめぐる状況は、その複雑さが一層高まるものと考えられます。

AI ネットワーク化の進展に伴いこの複雑な状況が連鎖的に変化し続けていくと、AI がその利用者にもたらし得る便益及び社会に対する恵沢が

25) 「リスク」については、「望ましくない事象の発生確率×予測される損害規模」をもってリスクとする客観主義的リスク観及び「将来生じるかもしれない望ましくない事象（と人々が認識するもの）」に焦点を当てる構築主義的リスク観があるとされる（中山竜一「互恵性と責任の政治学—リスク現実化の「前」と「後」—」立命館言語文化研究 26 巻 4 号 145-146 頁（平 27）参照。）。

26) AI ネットワーク社会推進会議・前掲注（5）25-26 頁参照。

27) ネットワーク上に様々な AI が混在することに伴う問題に留意すべきことの指摘として、AI ネットワーク社会推進会議・前掲注（5）45-46 頁参照。

28) 中西・前掲注（16）176-178 頁参照。

29) 中西・前掲注（16）178 頁参照。

高まり得る一方で、あらかじめ適切な手立てを講じておくことを怠れば、複数の AI ネットワーク相互間の連鎖反応等を通じた AI の出力やプログラムの変化に伴い、AI の出力の予測、理解、検証、制御等が困難となり、しかし、利用者又は第三者が不測の不利益を被るおそれが高まっていきかねないものと考えられます³⁰⁾。

このほか、AI ネットワークの構成要素に不具合が生じた場合、AI ネットワーク上を流通するデータ等に誤りや偏りがあった場合、複数の AI ネットワーク相互間の目的の競合や対立があった場合³¹⁾、AI ネットワーク又は AI ネットワーク上のデータ等を取り扱う者に過失又は他人への害意に基づく行為があった場合その他他人の権利利益を侵害する結果を招く態様での利活用があった場合等には、あらかじめ適切な手立てを講じておくことを怠れば、これまでに御説明した AI 固有の特徴や AI ネットワークをめぐる状況の複雑さと相俟って、利用者又は第三者が不測の不利益を被るおそれがあると考えられます。

これら利用者又は第三者が不測の不利益を被るおそれは、リスクとして捉えられ得るものであります。これらのリスクも、AI ネットワークの便益と同様に、時に実空間とサイバー空間の境界を越え、時に国境を越え、即座に波及し得るものであり³²⁾、また、連携の相手が増えるにつれて高まり得るものであります。

(2) リスクの例

AI ネットワーク化検討会議は、AI ネットワーク化の進展が社会・経済にもたらし得るリスクを AI ネットワークの機能に関するリスクと法制度・権利利益に関するリスクに大別しております³³⁾。AI ネットワークの機能に

30) 中川裕志「シンギュラリティ以前—人工知能と社会—改訂版」26-28 頁（平成 28 年 4 月 4 日）<<https://www.slideshare.net/hirsoshnakagawa3/ss-64699644>>〔平成 30 年 2 月 1 日最終閲覧〕参照。

31) 複数の AI ネットワーク相互間の目的の競合又は対立に関することについて、長尾真「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配付資料 3 頁（平成 29 年 3 月 13 日）を参照。

32) AI ネットワーク社会推進会議・前掲注 (5) 3 頁、29 頁、32 頁及び 49 頁を参照。

33) AI ネットワーク化検討会議・前掲注 (3) 44-48 頁及び同・前掲注 (4) 36-39 頁を参照。これらのリスクについては、両報告書のほか、AI ネットワーク社会推進会議・前掲注 (5) 別紙 3 及び別紙 4、堀浩一「人工知能の研究開発をどう進め

関するリスクとはAI ネットワークの機能上の理由によって利用者が意図しない事象が生ずるリスクであります。法制度・権利利益に関するリスクとはAI ネットワークの利活用の仕方に伴い利用者若しくは第三者の権利若しくは利益又は社会の法的な仕組みが害されるリスクであります。これらのリスクの具体例は、次のスライドのとおりであります³⁴⁾。

AIネットワーク化の進展が社会・経済にもたらし得るリスクの例

リスクの種類及び例	リスクの具体例
AIネットワークの機能に関するリスク	
セキュリティに関するリスク	-AIネットワークがハッキングやサイバー攻撃を受けて不正に操作されること -AIネットワークに対する攻撃が秘かに行われ、被害に気付かないこと
情報通信ネットワークに関するリスク	-情報通信ネットワーク上に多種多様なAIシステムが連在することにより、AIネットワークが正常に動作せず、意図しない事象が発生 -情報通信ネットワークの構成要素の不具合によりAIネットワークが正常に動作せず、意図しない事象が発生 -クラウド等におけるデータ漏洩・消失やシステム障害
不透明化のリスク	-AIネットワークが複雑に形成される中で、入出力の予見、理解、検証又は説明が困難化
制御喪失のリスク	-高頻度取引(High Frequency Trading)におけるフラッシュクラッシュ -ファームウェアの乗っ取り等によりAIネットワークが暴走し、人間による制御が困難化ないし不可能化
法制度・権利利益に関するリスク	
事故のリスク	-自動運転車やロボット等の自律的判断に基づく動作による事故 -AIネットワークが入出力するデータ等の誤りによる事故
犯罪のリスク	-AIネットワークを悪用したマルウェアによる犯罪
消費者等の権利利益に関するリスク	-アップデートの懈怠に伴う不具合により消費者等の権利利益が毀損
プライバシー・個人情報に関するリスク	-AIネットワークによる個人情報の収集・利用が不透明化することにより、個人情報のコントロールが困難化 -AIネットワークを用いたプロファイリングによるプライバシーの侵害
人間の尊厳と個人の自律に関するリスク	-人間の意思決定過程を見えない形で操作 -現実の環境における人種、性別等の偏りの存在を反映したデータ等への入力に応じて出力されたデータ等に依拠することに起因する意図せざる不当な差別
民主主義と統治機構に関するリスク	-投票等国民の行動の操作 -AIネットワークを国家の統治に活用する場合における意思決定過程の不透明化や責任の所在の曖昧化

これらのリスクのうちには、不透明化のリスク、制御喪失のリスク等利活用の過程において入力されるデータ等からの学習等により自らの出力やプログラムを将来に向けて継続的に変化させるというAI固有の特徴に関連するリスクのみならず、セキュリティに関するリスク、プライバシー・個人情報に関するリスク、消費者等の権利利益に関するリスク等従来の情報通信技術の利活用においても広く指摘されているリスクも含まれます。そのようなリスクといえども、専らAI以外の従来の情報通信技術を利用する場合と比べ、AIの性能の向上やAI ネットワーク化の進展に伴うシ

るか」情報管理 58 巻 4 号 254-257 頁（平 27）、中川・前掲注（30）並びに福田＝林＝成原編・前掲注（6）所収の各論文及び座談会をも参照。

なお、これらのリスクのうちには、事故のリスク等両者の側面を有するリスクもある。

34) このスライドは、AI ネットワーク化検討会議・前掲注（3）44-48 頁及び同・前掲注（4）36-39 頁を筆者が整理したものである。

システム間の連携の複雑化等に伴って量的に増大し、その不利益の抑制が困難となりかねないものであることから、AI 固有の特徴に関連するリスクと同じく、AI ネットワーク化の進展がもたらし得るリスクと捉え、その対処の在り方を検討することが必要となります。

(3) リスクへの対処の在り方

リスクに対処するための枠組みとしては、「リスク分析」が知られております。「リスク分析」は、リスクの所在を把握し、被害の発生時期、生起確率、規模等を評価するプロセスたる「リスク評価」、リスク評価の結果を踏まえてリスクへの対処を決定し、実施するプロセスたる「リスク管理」及びステークホルダーが情報、意見等を交換するプロセスたる「リスク・コミュニケーション」の三点から構成されます³⁵⁾。

AI ネットワーク化に関するリスク評価については、新たなリスクに関する経験の蓄積が追いつかないことから、被害の生起確率や規模等の予測が不可能いし困難なリスクの顕在化に対処できるよう、リスク認知のバイアス³⁶⁾に留意しつつ、複数のシナリオを想定して準備を進めるべきことが指摘されております³⁷⁾。AI の技術の加速度的な発展及び AI の利活用の急速な拡大に鑑みると、シナリオの作成及びシナリオに基づく検討は、AI ネットワーク化をめぐる現実の動向及び将来の展望の変化に必ずべく、不断に見直していくことを前提として継続的に進められるべきものと考えられます。

AI ネットワーク化に関するリスク管理については、不確実なリスクの管理については「予防原則」³⁸⁾が環境分野等において国際的な支持を集め

35) AI ネットワーク化検討会議・前掲注 (3)42 頁参照。

36) 「リスク認知のバイアス」とは、非専門家は思い浮かべやすいリスクや顕著に認識できるリスクを過大に評価しがちなことをいう(キャス・サンスティーン(角松生史=内野美穂監訳・神戸大学 ELS プログラム訳)『恐怖の法則』47-65 頁(勁草書房、平 27) 参照)。

37) AI ネットワーク化検討会議・前掲注 (3)42-43 頁参照。

38) 「予防原則」とは、不可逆的又は甚大な損害が発生するおそれがある場合には、科学的不確実性があるとしても、費用対効果の高い何らかの予防的措置(差止めとは限らない。)をとることを要請する法原理をいう(AI ネットワーク化検討会議・前掲注 (3)43 頁、中山・前掲注 (25)145-146 頁等を参照)。

ている³⁹⁾ことを踏まえた上で、予防原則にはリスク・トレードオフ⁴⁰⁾に伴う限界があること、不確実なリスクの顕在化を予見することが困難であること等にも留意しつつ、規制的手法のほか、経済的手法、情報的手法、自主規制等多様なソフトな手法をも組み合わせた対処の在り方を検討すべきことが指摘されております⁴¹⁾。

AI ネットワーク化に関するリスク・コミュニケーションについては、そのための場を作り、多様なステークホルダ相互間でリスク・コミュニケーションを行うことにより、AI ネットワーク化のリスクに対する懸念を緩和するとともに、リスクの適切な対処の在り方について検討を進めていくべきことが指摘されております⁴²⁾。

AI ネットワーク社会推進会議においては、AI ネットワークの具体的な利活用の場面を想定して、そのインパクト及びリスクに関するシナリオを作成することにより、インパクトの展望を得るとともに、リスク評価を行ってリスク管理及びリスク・コミュニケーションの在り方の検討を進めております⁴³⁾。

4. AI ネットワーク化の進展をめぐる規範・政策の在り方

今後、AI が利活用される際には AI ネットワークとして利活用されることが多くなり、しかも、AI が AI ネットワークとして利活用される場合には、個々の AI が単独で利活用される場合には得られない便益が得られ、当該便益による恵沢が社会・経済に幅広く連鎖的にもたらされ得る一方で、個々の AI が単独で利活用される場合よりも多岐にわたるリスクがもたらされ、当該リスクに伴う不利益が利用者又は第三者に幅広く連鎖的にもたらされるおそれがあるものと見られております。しかも、これら便益及びリスクは、時には国境を越えて即座に波及し得るものであります。

39) 予防原則は 21 世紀の新たな法の一般原理として定着しつつあり、新しい科学技術の潜在的なリスクが大惨事をもたらす前に、ひとまず予防原則を選択肢のひとつとして真剣に顧慮することは許されないとするものとして、中山・前掲注 (25) 146 頁を参照。

40) サンステイーン・前掲注 (36) 17-44 頁参照（「リスク・トレードオフ」とは、あるリスクを予防することにより別のリスクが生ずるおそれをいう。）。

41) AI ネットワーク化検討会議・前掲注 (3) 43 頁参照。

42) AI ネットワーク化検討会議・前掲注 (3) 44 頁参照。

43) その成果の一部として、AI ネットワーク社会推進会議・前掲注 (5) 別紙 3 及び別紙 4 を参照。

また、AI がもたらし得るリスクに関し、今後開発される AI に透明性や制御可能性を求めることが容易でないことも指摘されますが、単独の AI に着目するだけでは対処しにくいリスクをめぐる問題を AI ネットワークの利活用をめぐるリスクをめぐる問題として捉えることとすれば、AI の動作をネットワークを介して人間又は他の AI が監視し又は制御したり、リスクを顕在化させた AI をネットワークから切り離したりする等直接又は間接につながる AI ネットワーク全体を視野に入れた対処を検討し得るようになるものと考えられます。

これらのことに鑑みると、AI がもたらし得る便益及びリスクをめぐる問題については、これを単に個々の AI をめぐる問題として捉えるのではなく、広く AI ネットワーク化の進展をめぐる問題として捉えることが適切であると考えられます⁴⁴⁾。そこで、AI ネットワーク化の進展をめぐる規範又は政策に関し、その基本的な方向性、形成の手法の在り方及び検討の方法の在り方についての問題提起のため、AI ネットワーク社会推進会議等における御議論を踏まえつつ、私見として一試論を御説明いたします⁴⁵⁾。

(1) 基本的な方向性

AI ネットワーク化の進展をめぐる規範・政策については、これを産業、金融、交通、教育、医療・介護、地域等社会の構成員が活動する個々の分野における個別的な事情に対応するための規範・政策と、社会全体における AI ネットワーク化の進展の在り方に関する分野共通のないし分野横断的な規範・政策とに大別できましょう。そのうち、個々の分野における個別的な事情に対応するための規範・政策については、当該個別的な事情に対応するための規範・政策である以上、当該分野自体に関する規範・政策の一部として検討されるべきものと考えられます。そこで、ここでは、AI ネットワーク化の進展をめぐる規範・政策のうち、社会全体における AI

44) 福田雅樹「AI ネットワーク化の展望と課題」情報通信学会誌 35 巻 2 号 65 頁(平 29)のほか、AI ネットワーク化検討会議・前掲注 (3)4-5 頁及び AI ネットワーク社会推進会議・前掲注 (5)3 頁を参照。

45) 私見については、その全体を通じて、福田・前掲注 (6)18-21 頁及び福田・前掲注 (44)65 頁を参照。

ネットワーク化の進展の在り方に関する分野共通的ないし分野横断的な規範・政策の基本的な方向性に関する一試論を御説明することといたします。

社会全体における AI ネットワーク化の進展の在り方に関する分野共通的ないし分野横断的な規範・政策を通じて究極的に達成すべき目的としては、第一に AI ネットワーク化の進展が社会・経済にもたらす恵沢を全体として豊かなものとしていくこと、第二に社会のすべての構成員が AI ネットワーク化の進展状況に応じてその恵沢を享受できるよう、社会におけるその構成員の「包摂」を AI ネットワーク化の進展状況に応じて確保することが挙げられましょう。

AI ネットワーク化の進展が社会・経済にもたらす恵沢を全体として豊かなものとしていくことのためには、AI ネットワークの円滑かつ健全な形成及び利活用を確保していくことが必要となるほか、AI ネットワーク化の進展が社会・経済にもたらすリスクに伴う不利益や恵沢の副作用として生ずる負のインパクトを忌避して社会において AI ネットワークが受容されなくなることを回避することが必要となるものと考えられます。

また、社会におけるその構成員の「包摂」を AI ネットワーク化の進展状況に応じて確保していくことのためには、社会の構成員の誰しもがその希望する AI ネットワークをその機能に応じた手頃かつ公平な条件で安心して安全に利活用できる状況を AI ネットワーク化の進展状況に応じて確保していくことが必要となると考えられます⁴⁶⁾。これは、社会における AI ネットワーク化が進展していくにつれて、AI ネットワークを利活用できることを前提として様々な製品や役務が供給されたり、社会の仕組みが形成されたりすることが増えていき、社会の構成員が活動するためには利用者となることを余儀なくするようになるものと見込まれるからであります⁴⁷⁾。

これらのことを整理すると、AI ネットワーク化の進展が社会・経済にもたらす恵沢を全体として豊かなものとしていくとともに社会におけるその構成員の「包摂」を AI ネットワーク化の進展状況に応じて確保するためには、AI ネットワークの円滑かつ健全な形成及び利活用を確保するこ

46) AI ネットワーク社会推進会議・前掲注 (5)52-53 頁参照。

47) 今日の社会においてその構成員たる個人又は団体が活動するためには電話、インターネット等の利用者となることを余儀なくしていることと同様である。

とに加えて、社会の構成員の誰しもがその希望する AI ネットワークをその機能に応じた手頃かつ公平な条件で安心して安全に利活用できる状況を AI ネットワーク化の進展状況に応じて確保すること⁴⁸⁾、端的に言えば、利用者となることのできることの利益を含むものとしての利用者の利益を保護することが必要となる⁴⁹⁾と同時に、社会における AI ネットワークの受容を確保することが必要となるものと考えられます。

AI ネットワークの円滑かつ健全な形成及び利活用を確保するとともに利用者の利益を保護することのために必要となることとしては、AI ネットワークサービスの提供を含む AI ネットワークの形成及び利活用に関する営為の適正かつ合理的な運営及び当該営為をめぐる公正な競争を確保すること、AI ネットワークにおけるデータ等の適正かつ円滑な流通及び利活用を確保すること並びに AI ネットワークを安心して安全に利活用できるようリスクの顕在化又は波及に伴う利用者の不利益を抑制することが考えられます。

他方、社会における AI ネットワークの受容を確保することのために必要となることとしては、リスクの顕在化又は波及に伴う第三者の不利益を抑制すること、AI ネットワーク化の進展が社会・経済にもたらす恵沢の副作用として生ずる負のインパクトを抑制すること、そして、これまでの社会において AI ネットワーク化の進展を念頭に置かずに形成されてきた法制度その他の社会の仕組みを AI ネットワーク化の進展と整合するものとするのが考えられます。

以上のことから、AI ネットワーク化の進展をめぐる規範・政策のうち、社会全体における AI ネットワーク化の進展の在り方に関する分野共通的ないし分野横断的な規範・政策の基本的な方向性に関する一試論としては、AI ネットワーク化の進展が社会・経済にもたらす恵沢を全体として豊かなものとしていくとともに社会におけるその構成員の「包摂」を AI ネットワーク化の進展状況に応じて確保することを究極的な目的とし、AI ネットワークの円滑かつ健全な形成及び利活用を確保するとともに利用者の利益を保

48) AI ネットワークサービスの料金が無料であるとしても、その提供を受けるに当たり利用者のデータ等をプロバイダの事業活動の用に供せしめることが条件とされる場合には、当該条件が手頃かつ公平なものであるのか否かに加えて、当該条件が安全で安心し得るものであるのか否かも問題となり得る。

49) AI ネットワーク社会推進会議・前掲注 (5)52-53 頁参照。

護すること並びに社会における AI ネットワークの受容の確保を図ることを直接的な目的として⁵⁰⁾、AI ネットワークの形成及び利活用に関する営為の適正かつ合理的な運営及び公正な競争を確保すること、AI ネットワークにおけるデータ等の適正かつ円滑な流通及び利活用を確保すること、リスクの顕在化又は波及に伴う利用者及び第三者の不利益を抑制すること並びに AI ネットワーク化の負のインパクトを抑制することに関する規範・政策を形成していくとともに、AI ネットワーク化の進展を念頭に置かず形成されてきた法制度その他の社会の仕組みを AI ネットワーク化の進展と整合するものとなるよう見直していくという方向性が考えられます。

AI ネットワークについてはネットワークならではの特徴としてネットワーク効果⁵¹⁾やネットワークの拘束性⁵²⁾が見られる可能性があるほか、利活用の過程において入力されるデータ等が増えるにつれて性能が向上する AI システムをその構成要素とする AI ネットワークについてはデータネットワーク効果⁵³⁾が見られる可能性があります。AI ネットワークの便益及びリスクは、ネットワーク効果やデータネットワーク効果により増大し、しかも、情報通信ネットワークを介することにより、時として国境を越え、即座に波及し得るものであります。また、AI ネットワークの形成及び利活用並びにこれらをめぐる競争や紛争は、技術や標準化の動向、利用者のニーズの状況、AI ネットワークやその基盤となる情報通信ネットワークの状況、データ等の創造・流通・蓄積の状況等市場の状況のほか、有力な AI ネットワークとの接続の可否やその条件、有力な AI 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といった事情により大きく左右され得るものと考えられます⁵⁴⁾。その一方で、データ等の中には、プライバシー等個人の権利利益を理由として慎重な取扱いを要するものがある

50) このことは、例えば、AI ネットワークサービスの提供に関し公正な競争を確保するための規範・政策について検討する場合においても、一般競争法の基準に照らし検討すれば足りるのではなく、これら固有の目的を達成するための必要性や合理性に照らし検討すべきであることをその含意とする。

51) 差し当たり、黒田敏史「固定ブロードバンドの実証分析」依田高典＝根岸哲＝林敏彦編著『情報通信の政策分析—ブロードバンド・メディア・コンテンツ』54-55 頁（NTT 出版、平 21）を参照。

52) 差し当たり、福田雅樹＝林秀弥「情報通信プラットフォームに関する競争法的考察（2・完）」名古屋大学法政論集 253 号 246 頁（平 26）を参照。

53) 差し当たり、AI ネットワーク社会推進会議・前掲注（5）47 頁を参照。

54) AI ネットワーク社会推進会議・前掲注（5）46-47 頁参照。

ります。したがって、AI ネットワーク化の進展をめぐる規範・政策を検討していく際には、市場の状況や動向を注視することのほか、ネットワーク効果等 AI ネットワークの特徴やデータ等の性質を踏まえ、個人の権利利益と公共の利益との調和の確保に留意することが必要であると考えられます。

(2) AI ネットワーク化の進展をめぐる規範・政策の形成の手法の在り方

AI の研究・開発及び利活用並びに AI ネットワークの形成及び利活用は、少なくとも現時点においては揺籃期にあることから、AI ネットワーク化の進展をめぐる規範・政策のうち、少なくとも社会全体における AI ネットワーク化の進展の在り方に関する分野共通のないし分野横断的な規範・政策については、柔軟性を重視し、非規制的かつ非拘束的なソフトローによるガバナンスの形成を基本とするアプローチによるべきものと考えられます⁵⁵⁾。なお、ソフトローによるガバナンスはステークホルダの自主的な尊重を待みとするものであることから、AI システムや AI ネットワークについて通曉するステークホルダたる開発者やプロバイダのアカウントビリティの確保が重要となってくるものと考えられます⁵⁶⁾。

また、AI ネットワーク化がインターネット等情報通信ネットワークを介して国際的に進展していくことが展望されるものであり、AI ネットワーク化がもたらす便益及びリスクが国境を越えて即座に広く波及し得るものであることに加え、ガラパゴス化は避けるべきことに鑑みると、社会全体における AI ネットワーク化の進展の在り方に関する分野共通のないし分野横断的な規範・政策については、国際的な枠組みの形成を先行させることを志向すべきものと考えられます⁵⁷⁾。

55) AI ネットワーク社会推進会議・前掲注 (5) 10 頁及び 23 頁を参照。

なお、社会の構成員が活動する個々の分野における個別的な事情に対応するための規範・政策については、その手法についても、当該個別的な事情に対応するために当該分野自体に関する規範・政策の一部として検討されるべきものと考えられる。

56) 開発者のアカウントビリティの重要性に関する指摘として、AI ネットワーク社会推進会議・前掲注 (5) 10 頁及び 23 頁を参照。

57) 社会の構成員が活動する個々の分野における個別的な事情に対応するための規範・政策については、国際的な枠組みの形成を先行させるのか各国の独自の規範・

AI ネットワーク化の進展が社会・経済にもたらすインパクト及びリスクは、AI ネットワークの構成要素となり得る AI システム自体の研究及び開発の在り方並びに AI ネットワークの形成及び利活用の在り方の双方に左右されるものであります。しかるに、AI システムの研究又は開発の主体と AI ネットワークの形成又は利活用の主体は AI システム及び AI ネットワーク並びに他人に対する関わり方においてその性格を異にするものであり、観念的にはその立場を区別し得るものでありますが、実際には両者が重なることもあります。したがって、AI ネットワーク化の進展の在り方に関する分野共通のないし分野横断的なガバナンスは、これを AI ネットワークの構成要素となり得る AI システムの研究及び開発のガバナンスと AI ネットワークの形成及び利活用のガバナンスとに大別し、両者を相互に補完的なものとして関連付けて形成すべきものと考えられます⁵⁸⁾。

(3) AI ネットワーク化の進展をめぐる規範・政策の検討の方法の在り方

AI ネットワーク化の進展をめぐる規範・政策について、非規制的かつ非拘束的なソフトローによるガバナンスの形成を基本とするアプローチによる場合においては、ソフトローの実効性を確保するため、当該ソフトローを開発者、プロバイダ、最終利用者等 AI の研究・開発及び AI ネットワークの利活用の関係者自身の自主的な尊重が得られるものとする必要があります。

また、AI ネットワーク化が社会・経済にもたらし得るインパクト及びリスクは、その波及も含め、社会・経済のあらゆる分野に及び得るものであることに加え、その内容も多岐にわたり得るものと見られることから、このソフトローによるガバナンスの形成に向けての検討に当たっては、当該ソフトローを自主的に尊重することが期待される者の知見に加えて、AI や情報通信に関する理工系の知見のほか、人文・社会科学の幅広い知見や関係する実務家の知見等をも結集し、多角的に検討することが必要となつてまいります。

政策の形成を先行させるのかについても、当該分野自体に関する規範・政策の一部として検討されるべきものと考えられる。

58) AI ネットワーク社会推進会議・前掲注 (5)28 頁及び 49-50 頁を参照。

しかも、AIの研究・開発及び利活用並びにAIネットワークの形成及び利活用は、現時点では揺籃期にあり、特に研究・開発はAIの技術の専門家によると、加速度的に進んでおり、その速さたるや専門家による予測を次々と塗り替えざるを得ない程のものであるとのことです⁵⁹⁾。このため、ある時点における知見に基づいて検討した結果も、その前提となる知見を次々と塗り替えざるを得ず、不断の見直しが必要となるものと考えられます。

揺籃期にある現時点から検討を進めていくこと自体が時期尚早であるとする批判があり得ようものと想定されます。しかしながら、研究・開発の発展の速さに鑑みますと、時期尚早であることを理由として検討を進めずにいると、気が付けば既成事実が急速に形成されてしまい、手遅れとなってしまうかねません。したがって、拙速を避けるべきであるとしても、手遅れになることがないように、注意深く漸進的に、不断の見直しを前提として検討を早々から着実に進めていくべきものと考えられます⁶⁰⁾。

これらのことから、このソフトローによるガバナンスの形成に向けての検討については、手遅れと拙速の双方を避けるべく、当該ソフトローを自主的に尊重することが期待される者たる開発者、プロバイダ、最終利用者、関係行政機関等AIの研究又は開発及びAIネットワークの利活用の関係者自身に加え、理工系及び人文・社会科学系の有識者、関係する実務家、関係行政機関等多様なステークホルダの参画を得て、多角的かつ継続的に議論し、不断の見直しを前提とする漸進的な合意に立脚しつつ着実に進めていくべきものと考えられます⁶¹⁾。

ここでは、ソフトローの尊重が期待される関係者自身を含む多様なステークホルダの参画及び議論を通じた合意の必要性について、ガバナンスの実効性を確保するという実務的な見地から御説明いたしました。ガバナンスの民主的正統性という規範的な見地からも説明できます。ソフトローによるガバナンスについては、その形成の過程、内容及び運用につい

59) 松尾豊「人工知能の未来—ディープラーニングの先にあるもの」ICTインテリジェント化影響評価検討会議第1回会合資料8の17-19頁（平成28年2月2日）参照。

60) AIネットワーク化検討会議・前掲注（4）68頁及びAIネットワーク社会推進会議・前掲注（5）63頁を参照。

61) AIネットワーク社会推進会議・前掲注（5）10頁及び23頁を参照。

て、立法府における討議を通じた民主的な統制が予定されておりません。AI ネットワーク化の進展がもたらすインパクト及びリスクが社会のあらゆる分野に幅広く及び得るものであることに鑑みると、そのガバナンスの形成の過程、内容及び運用について、立法府における討議を通じた民主的な統制に代わる民主的な正統性を何らかの形で担保することが期待されます。このような規範的な見地からも、このソフトローによるガバナンスについて、ソフトローの尊重が期待される関係者自身を含む多様なステークホルダの参画及び議論を通じた合意が重要となってくるものと考えられます。

5. 目指すべき社会像及びその基本理念⁶²⁾

(1) 目指すべき社会像「智連社会」

AI ネットワーク化の進展をめぐる規範・政策を形成するに当たっての指導理念となるものとして、AI ネットワーク化の進展を通じて実現することを目指すべきトゥルー・ノース（true north）として AI ネットワーク化検討会議及び AI ネットワーク社会推進会議が構想する社会像たる「智連社会」及びその基本理念について御説明いたします。

AI ネットワーク化が円滑かつ健全に進展した場合における社会の変化の方向性としては、社会の主たる牽引力が「データ」・「情報」からなる「知識」（Knowledge）⁶³⁾ という「知」から、これら「知」に基づき「知能」（Intelligence）⁶⁴⁾ を活用する人間の能力としての「智慧」（Wisdom）⁶⁵⁾ という「智」へと変化していくことを指摘することができます。ここで「知能」としては人間の脳、AI その他の情報処理システムが有する機能を想定し

62) この章については、その全体を通じて、AI ネットワーク化検討会議・前掲注 (3) 12-16 頁及び AI ネットワーク社会推進会議・前掲注 (5) 4-8 頁を参照。

63) 「データ」としては断片的な事実、数字や文字が、「情報」としてはデータの組み合わせに意味を付与したものが、「知識」としてはデータや情報の体系的集積がそれぞれ念頭に置かれている（AI ネットワーク化検討会議・前掲注 (3) 13 頁及び AI ネットワーク社会推進会議・前掲注 (5) 5 頁参照。）。

64) 「知能」としてはデータ・情報・知識から新たにデータ・情報・知識を創造する機能が念頭に置かれている（AI ネットワーク化検討会議・前掲注 (3) 13 頁及び AI ネットワーク社会推進会議・前掲注 (5) 5 頁参照。）。

65) 「智慧」については、注 (66) に対応する本文及び注 (66) に掲げる文献を参照。

ているのに対し、「智慧」としては「知能」という機能を活用することにより得られる人間の能力を想定しております。

平成12年に制定された高度情報通信ネットワーク社会形成基本法（平成12年法律144号）2条においては、同法を通じて目指すべき社会像として「高度情報通信ネットワーク社会」が規定されております。「高度情報通信ネットワーク社会」とは、「インターネットその他の高度情報通信ネットワークを通じて自由かつ安全に多様な情報又は知識を世界的規模で入手し、共有し、又は発信することにより、あらゆる分野における創造的かつ活力ある発展が可能となる社会」をいいます。情報及び知識という「知」の流通及び利活用に着目した社会像であります。

このような社会においてAIネットワーク化が円滑かつ健全に進展すれば、入力されるデータ・情報・知識という「知」を操作して新たにデータ・情報・知識という「知」を出力する機能を有するAIの「知能」が高まるとともに、AIの「知能」とヒトの「知能」との連携が進み、さらにはヒトの「知能」相互間のAIネットワークを介した連携が進み、その結果として、「データ・情報・知識に基づき、知能を活用することにより、人間や社会の在り方を構想し、その実現に向けた課題を解決するための人間の能力」としての「智慧」（智）^{66）}を連結する「智のネットワーク」（Wisdom Network）が形成され、人間の「智慧」の連携・協調が進み得るようになるものと展望されております。

このような展望の下、AIネットワーク化検討会議及びAIネットワーク社会推進会議においては、「データ」・「情報」・「知識」という「知」に着目する社会像たる「高度情報通信ネットワーク社会」が実現した後に、AIネットワーク化が円滑かつ健全に進展していくことを通じて実現することを目指すべき社会像として、「知」に基づき「知能」を活用する人間の能力としての「智慧」（智）に着目する「智連社会」（Wisdom Network Society ウイんズ）が構想されております。

AIネットワーク化の進展を通じて目指すべき社会像たる「智連社会」は、人間がAIネットワークと主体的に共生し、すなわち、人間がAIネットワー

66）AIネットワーク化検討会議・前掲注（3）13頁及びAIネットワーク社会推進会議・前掲注（5）5頁参照。併せて、濱田純一「AIネットワーク化と智連社会」AIネットワーク社会推進フォーラム配付資料2-5頁（平29）を参照。

クを使いこなし、データ・情報・知識を自由かつ安全に創造・流通・連結して、人間の能力としての「智慧」の連結からなる「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⁶⁷⁾であります。

ここでAIネットワークがAI、情報通信ネットワーク及び当該情報通信ネットワークを介しての連携先たるシステムからなる物理的な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な存在の機能を利活用することにより得られる能力たる「智慧」が相互に連携し合い、相互に協調し合う様子を捉え、これを観念的に描写する概念であります。「智のネットワーク」自体は観念的な概念であります。が、「智のネットワーク」を形成することにより実現されるヒト・モノ・コト相互間の空間を越えた協調には、国境等地理的な空間を越えた協調のほか、実空間とサイバー空間との境界を越えた協調⁶⁸⁾も含まれます。

(2) 智連社会の基本理念

AI ネットワーク化検討会議は、AI ネットワーク化の進展を通じて目指すべき社会像たる「智連社会」の基本理念として8項目を掲げております。この基本理念は、高度情報通信ネットワーク社会形成基本法3条から9条までに定める高度情報通信ネットワーク社会の形成についての基本理念⁶⁹⁾を参考にして、AIの普及及び高度化並びにAIのネットワーク化の影響を踏まえて、「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

67) 人間中心の社会を構築することの重要性について、AI ネットワーク社会推進会議・前掲注(5)10-11頁及び23頁を参照。目指すべき社会像を人間中心の社会像としていることは、人間とAIとの役割分担を検討する際の指導理念として特に重要となるものと考えられる。

68) 実空間とサイバー空間との境界を越えた協調とは、CPS (Cyber-Physical System) にほかならない。CPSとは、実空間から収集されサイバー空間にあるデータをAI等が解析した結果が実空間に自動制御等の形でフィードバックされるという連関をいう。CPSとAIとの関係については、See WORLD ECONOMIC FORUM, EXTREME AUTOMATION AND CONNECTIVITY: THE GLOBAL, REGIONAL, AND INVESTMENT IMPLICATIONS OF THE FOURTH INDUSTRIAL REVOLUTION 13 (2016)。

69) 高度情報通信ネットワーク社会形成基本法3条から9条までについては、「高度情報通信ネットワーク社会の形成についての基本理念」を「定める」ものと規定されている(同法10条)。

あり、AI ネットワーク化の進展をめぐる規範・政策を検討するに当たって立脚すべきものと考えられます。

智連社会の基本理念の第一の項目としては、すべての人々が、AI ネットワークを容易かつ主体的に利用する機会を有し、個々の能力を創造的かつ最大限に発揮し又は拡張することが可能となり、もって AI ネットワークの恵沢をあまねく享受できることが掲げられており、第二の項目としては、個人が人間としての尊厳をもった自律的な主体として AI 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が掲げられております。これら第一及び第二の項目は、先程御説明いたしました社会全体における AI ネットワーク化の進展の在り方に関する分野共通のないし分野横断的な規範・政策の基本的な方向性の一試論において掲げた規範・政策の究極的な目的に直接対応するものであり、智連社会が備えるべき性質の中でも必然的かつ本質的なものであります。

智連社会の基本理念の第三の項目としては、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 AI ネットワークが提供されることが掲げられており、第四の項目としては、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が掲げられております。第三及び第四の項目は、第一及び第二の項目にのっとり形成される規範・政策について、その内容を検討するに当たっての主たる要素を示すものとして捉え得るものと考えられます。

智連社会の基本理念の第五の項目としては、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が掲げられております。この項目は、先程 AI ネットワーク化の進展をめぐる規範・政策の検討の方法の在り方に関し御説明した内容に関連するものであります。

智連社会の基本理念の第六から第八までの項目においては、AI ネットワークを基盤として形成される「智のネットワーク」によるヒト・モノ・コト相互間の協調及びその恵沢の空間的な拡がりの在り方に関する理念が掲げられております。第六の項目には、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が掲げられ、第七の項目には、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が掲げられ、第八の項目には人類が、AI ネットワークを基盤と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が掲げられております。

6. AI ネットワーク化に関する社会的・経済的・倫理的・法的課題

AI ネットワーク化検討会議の「報告書 2016」においては、AI ネットワーク化の進展が社会・経済にもたらすインパクト及びリスクを踏まえ、今後注視し、又は検討すべき社会的・経済的・倫理的・法的課題として多岐にわたる課題が掲げられております⁷⁰⁾。時間の都合上、この研究会との関係が深い課題たる「AI 開発原則・指針の策定」、「AI ネットワーク化の進展に向けた協調の円滑化」、「競争的なエコシステムの確保」及び「利用者の保護」について御説明いたします。

まず、「AI 開発原則・指針の策定」について御説明いたします。

AI ネットワーク化検討会議においては、AI ネットワーク化の進展が社会にもたらすインパクト及びリスクに鑑み、OECD プライバシーガイドライン等を参考に、AI ネットワークの構成要素となり得る AI に関し、その研究開発に当たり留意すべき事項を「AI 開発原則」（仮称）として整理し、AI 開発原則自体及びその内容の説明からなる指針たる「AI 開発ガイドライン」（仮称）を国際的に参照される枠組みとして策定して共有することが AI ネットワークの社会における受容の向上及び智連社会への円滑な移行のために必要かつ効果的であるとの考え方に立脚し、AI 開発ガイドラインの策定に向け、ステークホルダの参画を得つつ国内外において検討及び議論を進めていくべき旨が提言されております。

AI ネットワーク化検討会議においては、AI 開発原則に盛り込むべき事項の素案として次の 8 項目が掲げられております⁷¹⁾。

70) AI ネットワーク化検討会議・前掲注 (4)40-67 頁参照。なお、AI ネットワーク社会推進会議が整理した今後の課題について、AI ネットワーク社会推進会議・前掲注 (5)49-59 頁を参照。

71) AI ネットワーク社会推進会議が平成 29 年 7 月に取りまとめた「AI 開発ガイドライン案」（後掲・注 (74) 参照。）においては、これら 8 項目のほか、「連携の原則」が加えられた。本文に掲げた 8 項目の原則の内容及び順序は、「AI 開発ガイドライン案」における原則の内容及び順序とは異なるものであることに注意されたい。

- ① 透明性の原則 AI ネットワークの動作の説明可能性及び検証可能性を確保すること。
- ② 利用者支援の原則 AI ネットワークが利用者を支援し、利用者を選択の機会を適切に提供するように配慮すること。
- ③ 制御可能性の原則 人間による AI ネットワークの制御可能性を確保すること。
- ④ セキュリティ確保の原則 AI ネットワークの頑健性及び信頼性を確保すること。
- ⑤ 安全保護の原則 AI ネットワークが利用者及び第三者の生命・身体の安全に危害を及ぼさないように配慮すること。
- ⑥ プライバシーの原則 AI ネットワークが利用者及び第三者のプライバシーを侵害しないように配慮すること。
- ⑦ 倫理の原則 AI ネットワークの研究開発において、人間の尊厳と個人の自律を尊重すること。
- ⑧ アカウンタビリティの原則 AI ネットワークの研究開発者が利用者等関係するステークホルダに対しアカウンタビリティを果たすこと。

平成 28 年 4 月 29 日・30 日に開催された G7 香川・高松情報通信大臣会合においては、高市早苗総務大臣から、これら 8 項目を示しつつ、AI 開発ガイドラインの策定等に関し OECD 等国際機関の協力も得て国際的な議論を進めるよう提案した⁷²⁾ ところ、参加各国から賛同が得られました。また、同年 11 月 17 日に OECD が開催した「技術予測フォーラム」においても、我が国からの出席者が G7 香川・高松情報通信大臣会合の結果、我が国における検討の状況等に関し説明したところ、加盟各国の参加者や OECD 事務局らから、AI 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ました⁷³⁾。

現在、AI ネットワーク社会推進会議において、AI 開発ガイドラインの策定に向けての今後の OECD、G7 等における国際的な議論に付するため

72) See Minister for Internal Affairs and Communications, *Proposal of Discussion toward Formulation of AI R&D Guideline* (G7 香川・高松情報通信大臣会合配付資料、平成 28) <http://www.soumu.go.jp/joho_kokusai/g7ict/english/main_content/ai.pdf> (last visited Feb. 1, 2018) ; AI ネットワーク社会推進会議・前掲注 (5)23 頁。

73) AI ネットワーク社会推進会議・前掲注 (5)21 頁参照。

の案の検討が進められております⁷⁴⁾。

ここまでの御説明から明らかなように、AI 開発ガイドラインとしては、我が国における規制若しくは拘束的な規範又は国際的な規制若しくは拘束的な規範の導入を目指すものではなく、OECD プライバシーガイドライン等と同様に、非規制的かつ拘束的なソフトローとしてステークホルダー間で国際的に共有される指針として国際場裏において形成されるものが想定されております。

次に、「AI ネットワーク化の進展に向けた協調の円滑化」、「競争的なエコシステムの確保」及び「利用者の保護」について御説明いたします。これらの課題は、いずれも AI ネットワークの円滑かつ健全な形成及び利活用の確保と直接関連する課題であることから、性質上互いに関連し合い、重なり合います。

先程御説明いたしましたように AI ネットワークについてはネットワーク効果、ネットワークの拘束性及びデータネットワーク効果が見られる可能性があります。ネットワーク効果やデータネットワーク効果は、ネットワークの拡大に伴いその効用を高め得る一方で、有力な AI ネットワークをより有力なものとする傾向をもたらし得ます。ネットワークの拘束性も、利用者によるプロバイダの変更のしやすさ等に影響を及ぼし得るおそれが考えられます。これらの事情は、AI ネットワークの形成及び利活用の進展並びにこれらの営為をめぐるエコシステムにおける競争及び利用者の利益に影響を及ぼし得るものと考えられます。

もっとも、AI ネットワークは揺籃期にあり、そのエコシステムの変遷については予断を抱き難いことから、規制の創設等の介入は、少なくとも当分の間は謙抑的たるべきと考えられます。

これらの事情の下、「AI ネットワーク化の進展に向けた協調の円滑化」

74) AI ネットワーク社会推進会議は、平成 29 年 7 月までの検討の結果として、「AI 開発ガイドライン案」を国際的な議論のための案として取りまとめた（AI ネットワーク社会推進会議・前掲注 (5)23-30 頁並びに別紙 1 及び福田・前掲注 (44)66-68 頁を参照。）。この「AI 開発ガイドライン案」については、同年 9 月 25 日・26 日に伊国において開催された G7 情報通信・産業大臣会合の成果文書において紹介されている（*See G7 ICT and Industry Ministerial Meeting, G7 MULTISTAKEHOLDER EXCHANGE ON HUMAN CENTRIC AI FOR OUR SOCIETIES* (Annex 2 to the G7 ICT and Industry Ministerial Meeting, *G7 ICT AND INDUSTRY MINISTERS' DECLARATION: MAKING THE NEXT PRODUCTION REVOLUTION INCLUSIVE, OPEN AND SECURE*), at 1 (Sep. 26, 2017).）。

については、特に相互接続性・相互運用性の確保が課題となります。また、「競争的なエコシステムの確保」及び「利用者の保護」については、AI ネットワークの形成及び利活用の動向、データ・情報・知識の集積の動向、市場における当事者間の紛争及びその影響の動向等の継続的注視の在り方、注視のための情報の収集の在り方、AI ネットワークの形成に向けての当事者間の協議の円滑化の在り方、AI ネットワークの形成又は利活用の関係者間の紛争の円滑な解決の在り方等が今後検討すべき課題となり、しかも、これらの課題は、「AI ネットワーク化の進展に向けた協調の円滑化」という見地からも課題となるものと考えられます。

7. 結びに代えて

結びに代えて、国内外の動向について御説明いたします。

AI のもたらし得る便益及びリスクについては、欧米の産学民官においても議論が見られます。G7 香川・高松情報通信大臣会合の数日後たる平成 28 年 5 月 3 日に、ホワイトハウスが AI がもたらし得る便益及びリスクに関し検討する旨を表明しました⁷⁵⁾。ホワイトハウスは、同年 5 月から 4 回にわたりワークショップを開催し、同年 10 月に報告書を公表しました⁷⁶⁾。ワークショップの初回には総務省情報通信政策研究所が招聘されたので、AI ネットワーク化検討会議の御議論の状況等に関し情報を提供しました。

このほか、欧米の産学民官において関連する議論がなされております。その動向及び議論の要旨を整理したのがこのスライドではありますが、AI ネットワーク化検討会議等の議論と通底するところが多いことが御覧いただけましょう。

我が国の政府においては、世界最先端 IT 国家創造宣言（平成 28 年 5 月 20 日改定・閣議決定）において「AI の普及及びネットワーク化に伴う社

75) Ed Felten, *Preparing for the Future of Artificial Intelligence*, WHITE HOUSE BLOG < <https://obamawhitehouse.archives.gov/blog/2016/05/03/preparing-future-artificial-intelligence> > (last visited Feb. 1, 2018).

76) NAT'L SCI. & TECH. COUNCIL, EXEC. OFFICE OF THE PRESIDENT, *PREPARING FOR THE FUTURE OF ARTIFICIAL INTELLIGENCE* (2016).

会的・倫理的課題に関し、AIの研究開発に当たっての留意事項に関する国際的に参照すべき原則の策定やその見直し等に関する国内外における継続的な議論の推進、その他所要の措置を講ずる」ことが決定されており、AI ネットワーク社会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の検討等を進めております。今後、OECD、G7 等の協力を得て、国内外における継続的な議論の推進に取り組んでまいります。

私からの報告は、以上でございます。御清聴くださり、ありがとうございます。

欧米における議論の動向

	米国	欧州
政府機関	<ul style="list-style-type: none"> ● 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、10月に報告書「人工知能の未来に備えて」を公表するとともに、「米国人工知能研究開発戦略」を策定。 	<ul style="list-style-type: none"> ● 欧州議会【2016年4月～】 法務委員会が2016年4月にロボット・AIの法的・倫理的問題に関し公聴会を開催。5月には、ロボットの民事法的ルールに関し報告書の草稿を公表。10月には、関連するワークショップを開催。 ● 英国下院科学技術委「ロボティクスと人工知能」【2016年10月】 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、AIの開発・利用に関する原則の策定やAIを巡る規制に関する政府への助言等を司る委員会をアラン・チューリング研究所の下に設置するよう提言。
企業・大学・NPO等	<ul style="list-style-type: none"> ● (Amazon、DeepMind/Google、Facebook、IBM、Microsoft) Partnership on AI 【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。 ● (スタンフォード大学) One hundred year study of Artificial Intelligence (Stanford AI 100) 【2014年12月～】 AIが法制度、経済、社会等にもたらす長期的影響を調査。 ● Future of Life Institute (FLI) 【Joan Tallinn (Skype共同開発者、DeepMind共同出資者) 等、2014年3月～】 研究者、起業家等からなるボランティアベースの団体。強健かつ有益(Robust and Beneficial)なAIを実現するための研究等を実施。 ● Open AI 【Elon Musk(テスラモーターズCEO) 等、2015年12月～】 安全なAIを構築して、AIの便益が可能な限り広範に等しく配分されることを目指して設立された非営利団体。 ● We Robot 【Michael Fromkin(マイアミ大学教授) 等、2012年4月～2016年4月までに5回】 事故時の責任、プライバシー、人種差別等、ロボットに関する法的問題について議論。 	<ul style="list-style-type: none"> ● (サンタナ大学(伊国)等) RoboLaw: ロボット技術に対する法と倫理 【2012年3月～2014年3月】 ロボット技術が現行の法的枠組みに適合し得るものであるのかを広範に調査・考察し、ロボット技術の発展が今後の社会規範、人々の価値観、社会的行動プロセスに与える影響を予測。 ● 英国規格協会「ロボット及びロボットシステムの倫理的設計・利用のためのガイド」 【2016年4月】 ロボットの開発者が遵守すべき倫理的基準を提示。 ● (オックスフォード大学哲学部) Future of Humanity Institute (FHI) 【2015年～】 技術が人類の未来に与える影響に注目する視点から、AIの制御可能性や安全性等について学際的に研究。 ● (ケンブリッジ大学) 未来のインテリジェンスのためのレバ・ヒューム・センター(LCfI) 【2015年12月～】 人類に利益をもたらすためのAIの在り方について、コンピュータ科学、認知学、哲学、社会学等多分野の研究を推進。

欧米における議論の要旨

ホワイトハウス報告書「人工知能の未来に備えて」(2016年10月)「結論」

・実務家(practitioner)は、AI対応システム(AI-enabled systems)が以下のとおりであることを確保しなければならない。

- ① 統御可能(governable)であること ② オープンで、透明で、理解可能であること ③ 人々と効果的に機能し得ること
④ 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

ホワイトハウス「米国人工知能研究開発戦略」(2016年10月) * 連邦政府の予算によるAI研究の方針を策定

(4) AIシステムの安全性及びセキュリティを確保する

AIシステムが広く利用されるようになる前に、当該AIシステムが安全(safe)かつ堅牢(secure)で、制御され(controlled)、十分に定義され、かつ十分に理解された方法で動作するものであることをあらかじめ確保しておくことが必要。

Partnership on AI「信条」(Tenets)(2016年9月) * ステークホルダーの関与と対話を重視

⑥ AIの便益を最大化し、AI技術の潜在的な課題に対処するため

- a. 個人のプライバシーとセキュリティを保護、b. AIの発展により影響を受け得るすべての当事者の利益を理解し、尊重
c. AIの研究と技術のコミュニティが、AI技術の社会への潜在的影響について社会的責任を負い、敏感であり、直接関与
d. AIの研究と技術を頑健で、信頼し得て、信用に値するものとともに、堅牢な制約の範囲内とすることを確保
e. 国際条約や人権に反するAI技術の開発と利用に反対するとともに、危害を与えないセーフガードと技術を推進
⑦ AI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るものであることが重要である

AI100 (One Hundred Year Study on Artificial Intelligence)「2030年における人工知能と生活」(2016年9月)

・将来への指針:現状のデジタル技術について成功の見込みのある規制を導く原則が出发点を提供する。適切な規制は、透明性、説明責任、専門性の向上等の取組の好循環を強化する可能性がある。

欧州議会法務委員会「ロボティクスを巡る民事法ルールに関する欧州委員会への提言に伴う報告書案」(2016年5月)

・倫理原則: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。倫理指針の枠組みは、恩恵、無害及び自律の原則並びに人間の尊厳等欧州基本権憲章に定められた原則に依拠すべき。

英国下院科学技術委員会報告書「ロボティクスと人工知能」(2016年10月)「結論と提言」

4. 我々の検討は、真摯な考慮を必要とする幾となる倫理的論点(検証及び妥当性確認、意思決定の透明性、偏見の最小化、アカウンタビリティの向上、プライバシー及び安全)を照らし出した。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、これらの論点について継続的に注視することが必要。

5. 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にAIに関する常設の委員会を設置するよう提言。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。