

人工知能と人間社会に関する検討すべき論点

【論点整理】

人工知能¹は人間社会に多大な便益をもたらし、健全に利活用されれば持続可能社会の強力な推進力になることが期待できる。その実現のため、人工知能と人間社会の適切な関係のために今現在検討すべき論点をまとめた²。

倫理的論点

人は様々な道具や機械を利用して、状況に応じた選択や判断を行ってきた。人工知能の進展によって、多大なデータに基づいた正確・迅速な判断、半自動的な操作、統計的に的確な選択が可能となる事例が増えてきている。人工知能が人の選択や判断を支援することで、正確さや迅速さの向上がもたらされ、人が犯しやすい認知バイアスや偏見³の影響を受けないなど多くの便益がもたらされる。ただし、状況や対象に応じて、人による判断と人工知能による判断のバランスを考慮する必要がある。人工知能技術の進展に伴い、人工知能・機械と人間の関係性に徐々に変化がみられる可能性があり、将来的にはその新たな関係性に基づく新たな倫理観が形成されるものと想定される。**(人工知能技術の進展に伴って生じる、人と人工知能・機械の関係性の変化と倫理観の変化)**

人工知能は、人にしかできないと思われてきた行為や判断を補助・代替できるようになりつつある。その一方で、人工知能によって人の心や行動が操作・誘導されたり、評価・順位づけされること、感情、愛情、信条に働きかけることに対する不安や懸念が生じる可能性がある。特に、本人が気づかないところでそれらが行われることには倫理的検討が必要ではないか。将来的には、人工知能によって拡張される人の時空間感覚と身体感覚に加えて、能力概念や感情の問題の相互作用や影響があり、その受容のための人間観の捉え直しが行われていく可能性がある。**(人工知能に知らぬ間に感情や信条、行動が操作された**

¹ 人工知能の定義は様々あるが、ここでは人の知的活動（認知、推論、学習、それらに基づいた行動）を代替しうる技術を主たる対象とし、さらに人工知能に関わるデジタルイゼーション技術も対象として検討を行った。

² 論点や事例の全てを検討することや全ての事例に共通する論点を抽出することは論理的に不可能であること、状況は常に変わっていることから見落としがある危険性については常に認識しなければならない。

³ 確証バイアス（自分の仮説に都合のよい情報を過剰評価し、確証に重きを置く認知的傾向）や根本的帰属エラー（人の行動の結果を状況・環境よりも個人特性のせいにする認知的傾向）などの認知バイアス、および個人的背景や文化的背景に基づく偏見など。

り、順位づけ・選別されたりすることへの懸念。感情を含む人間観の捉え直し)

人工知能の利活用によって、生産性が量的に向上するのみならず生産物の質が上がり、これまでに創れなかったものや膨大な時間とコストが必要だったものが容易・安価に生産され、誰もが利用できるようになることが予想される。そのとき、人が行う行為・創造と人工知能が行う行為・創造、そして人と人工知能が協働しての行為・創造の価値（有用性、オリジナリティ、芸術性など）がそれぞれどのように評価されるか、社会的に受容されるかを検討する必要性が想定される。特に、人と人工知能が協働することは人間能力の拡張とも言え、新しい価値観の基盤となる可能性がある。また、人によって異なる価値観やビジョンを尊重し、様々な選択肢や価値の多様性を担保することが大切であろう。（人工知能が関与する行為・創造に対する価値の検討、価値観やビジョンの多様性の確保）

法的論点

人工知能がもたらす便益や成果等についても、あるいはリスクや事故、権利侵害等についても、責任分配を明確にすることで、ユーザも事業者も人工知能の便益を享受しやすくなると思われる。技術進展のレベル（たとえば、自動運転のレベル0~4のそれぞれなど）に対応した責任分配を明確にし⁴、不確実で確率的に生じるようなリスクに対しては保険で対応することが、人工知能技術が社会に受容され、その便益を享受するために有効だろう。産業界における萎縮効果やレピュテーションリスクへの過反応を防ぐためにも責任分配の明確化は重要であろう。人工知能を利用することによるリスクのみならず、利用しないことで便益を失うリスクや責任も意識して人工知能の利活用を検討することが大切である。（人工知能による事故等の責任分配の明確化。人工知能を使うリスク、使わないリスクの考慮）

人工知能はビッグデータの活用でより有益なものとなる。ただし、その利便性と個人情報保護（プライバシー）はトレードオフの関係になる可能性がある。それらを両立し、萎縮効果を生まないための制度（法律、契約、ガイドライン）の検討が必要である。個人情報についてのデータアクセス権、データポータビリティ、忘れられる権利、そしてそのセキュリティなどについて日本としての考え方を整理しておく必要があるのではないか。これらを実際に適用する1つの対象として行政サービスへの人工知能活用を政府が検討することが期

⁴自動運転のレベル2までについては、交通事故の責任は基本的にドライバーにあるとされている。

待される。(ビッグデータを活用した人工知能の利便性確保と個人情報保護の両立)

人工知能の活用によって高付加価値な創作物が容易に生産されるようになるが⁵、人工知能による創作物や人工知能と協働した創作物、計算結果などの権利は誰のものか(帰属割合)の検討が必要であろう。また、人工知能の開発と利活用を促進するために、適切な契約とガイドラインによって、人工知能の開発者や活用者、元データ所有者等への適切な権利(インセンティブ)の配分方法を適時検討することが一般的になることも想定される。(人工知能を活用した創作物の権利とインセンティブの検討)

人工知能による仕事や雇用の変化に対して、従来法(道路関連法、業務関係法、薬事法、労働関連法⁶など)の解釈で対応が十分か、法律の修正が必要か、新しい法律が必要か、あるいは法律の概念自体を変えて対応する必要があるかを検討して、議論を続けていく必要がある。(法律概念の再検討の可能性)

経済的論点

情報社会においてインターネット上のビッグデータを駆使できた少数の企業が台頭したように、所謂人工知能覇者企業が出現して、既存のビジネス勢力図が抜本的に変化する可能性がある。また、人工知能を活用することで多大な労働力を必要としなくても大規模な企業活動が可能となり、多くの企業にとってはコストメリットが高く、機動力も上がることが予想される。

被雇用者にとっては、現在の仕事・業務内容(タスク)が人間から人工知能・機械に代わり⁷、人間は単純労働から解放され、より創造的な業務を主に行うようになることが想定される。そのような新しい業務に対応する能力として、労働者は個人の能力を最大限に発揮するために仕事を変える転職力や創造的労働に必要な能力、人工知能を活用する能力などの獲得が必要であろう。人工知能を活用した起業なども増加し、個人事業主化が進むことが想定される。

(人工知能による働き方の変化：個人対象)

人工知能の利活用によって単純労働・長期労働・重労働が減少し、高付加価値な労働と企業に従属しない自由な働き方への変化が予想される。それに合わせて、企業は持続的な経済成長を目指して経営判断の迅速化、雇用の再配置の

⁵ 熟練工と同等かそれ以上の作業を行う産業用ロボットや著名な芸術家の作風を学び創作する人工知能技術などが実用化されつつある。

⁶ 経済的論点で後述するように個人事業主化が進むと、被雇用者を対象とする労働法について捉え直しが必要となる可能性がある。

⁷ 技術の進化に応じて新しい業務(タスク)が出てくることから人が行う仕事自体がなくなることはないと思われる。

迅速化、テレワークなど空間と時間に制約されない働き方の促進を検討することが必要であろう。(人工知能の利活用による雇用形態と企業の変化：企業対象)

人工知能を活用して経済成長を促し、個人に適した多様な労働形態の確保のためには、まずは労働移動⁸を可能とする能力を教育し、学習する機会の提供が国の政策として必要である。それに加えて、何らかのマクロ経済政策やセーフティネットについてそれらが必要かを検討することが大切であろう。人工知能による生産効率向上や経済活性化、予測可能性の向上などの恩恵・利益をどのように社会的に公平に配分し、経済格差をなくすかを検討する必要がある。

(人工知能の利活用を促進するための政策：国対象)

社会的論点

安全安心な社会の実現や、少子高齢化が進む中でも少ない労働力で高い生産性が上がるなど、人工知能がもたらす社会的便益は大変大きい。しかし、他の多くの道具や技術と同様に人工知能も1つの技術であり、その使用が社会的に強制されるものではない。個人の信条として人工知能との関わりの自由を確保し、人工知能の利用を強制されないことおよび自由に利用できることを保証すること、そして人工知能を利用する者と利用しない者との間に社会的対立が生じないような配慮が必要ではないか。そのためにも、異なるビジョンや考え方を持つ者同士が議論する場をつくり、専門家の意見も含めて可視化し、継続的に検討を続ける必要がある。また、人工知能の恩恵を得るために個人情報を提供していた者も利用を停止した際に情報を消去できる仕組みなどオプトアウト・オプトインの方法と制度を検討する必要がある。(人工知能との関わりの自由、忘れられる権利)

人工知能の便益を最大限に享受するには人工知能に関する正しい知識に加えて、プライバシーに関するデータの知識、デジタル機器に関するリテラシーなどがユーザに必要となる。社会のすべての構成員がこれらの知識やリテラシーを有していれば、社会的便益は最大化する。しかし、これらの知識やリテラシーが無い人が居る場合には、いわゆる人工知能格差が出現する可能性がある。例えば、本来移動弱者に便益となるべきライドシェアが知識不足により利用できず、代わりに利用する通常のタクシーが高額になってしまうような事態が想定される。それゆえ、人工知能に関するリテラシーや知識、資産の有無によって新たな格差や社会コスト負担の不均衡が生じないような配慮・政策が必要であろう。(人工知能による格差、デバイド。人工知能に関連する社会的コスト)

⁸ 労働者個人が、労働市場において企業間、産業間、職業間などを移動すること。

の不均衡)

人工知能が社会で活用される場面が多くなるため、将来的には人工知能に対する依存や過信・過剰な拒絶など新たな社会問題や社会的病理が生じる可能性もある。正しい情報の公開、議論の場の提供、教育施策などを通じた対処の必要性も検討事項として想定される。(新たな社会的病理の可能性、対立、依存)

教育的論点

過去の社会においても新しい道具・技術の出現によってユーザがそれを使いこなす訓練を行い、その利活用から便益を得てきた。人工知能についても、その便益とリスクを理解し、責任の所在を把握し、人工知能によってどのような選択、判断、操作がなされているかを把握するリテラシーや検証するスキルを身に着けることが望ましい。また人工知能に関するリテラシーのみならず、それを適切に使いこなすための法律に関するリテラシーを身に着けること、それを実現するための法律に関するコミュニケーションも重要であろう⁹。つまり、ユーザが人工知能を主体的に利用し、協働・協調して創造的活動ができる能力の育成が必要である。(人工知能を利活用するための個人の能力の育成)

人工知能が補助・代替できることが増える中で、人工知能には何ができて何ができないのかを研究し、そのエビデンスに基づいた人工知能利活用社会に適した教育カリキュラムを検討する必要がある。特に子どもに対しての教育は急務である。子どもの人工知能利用に関しては、親がどこまで監督するべきかも含めて検討の必要性が想定される。また、人工知能で代替可能としても、人に残すべき能力とは何かを検討し、その育成と保護の検討の必要性も想定される。(人にしかできない能力の育成)

学校教育や独学可能な教育環境の整備によって個人の人工知能に関するリテラシーやスキルを高める政策が必要である。ただし、それでも残るリテラシー・スキル保有者と未保有者の格差があれば、社会政策的な対応が求められる。(教育格差に対する政策)

⁹ 人工知能を利用したサービスなどを利用する際にプライバシーなどに関する使用許諾書への同意が必要であるが、既に現状において契約に関する項目や記述が多岐であり、一般ユーザが正しくそれらを理解した上で判断することが困難になっている。

研究開発的論点

研究開発する側は、高い倫理観を持って研究開発に従事し¹⁰、各種学会や所属機関の倫理規定¹¹やガイドライン等を順守して価値創出に努めること、およびそれらに関して説明責任（アカウンタビリティ）を持つことが求められるだろう。サイバーセキュリティや安全性を確保してユーザが人工知能を安心して利用できる環境を整備する必要がある。特に個人情報（プライバシー）の保護、それをどこまで利用可能とするかの選択を安全に可能とする技術の開発が必要である。（倫理観、アカウンタビリティ、セキュリティ確保、プライバシー保護）

人が人工知能を制御できることを担保する技術（制御可能性）、人と人工知能の制御権の切り替えをスムーズにするインタフェース、人工知能の推論・計算の過程・論理を説明できる技術（透明性）、人工知能の利用度合い情報の埋め込み技術などの開発が必要と想定される。（制御可能性、透明性）

社会に対して新技术を語る際には、技術によって得られる便益とリスクを恣意性なく的確に表現することに務め、ユーザは知り得た新技术である人工知能を自らの判断で利用するか否かを最終的に判断することが求められるだろう。（人工知能に関する適切な情報公開と、それに基づく責任ある使用判断）

¹⁰ 総務省「AI ネットワーク化検討会議」では、「研究開発の原則」の策定として、(1)透明性の原則、(2)利用者支援の原則、(3)制御可能性の原則、(4)セキュリティ確保の原則、(5)安全保護の原則、(6)プライバシー保護の原則、(7)倫理の原則、(8)アカウンタビリティの原則を含むことを発表している（2016年4月中間報告書、同年6月報告書2016）。

¹¹ 人工知能学会は、2016年6月に「人工知能研究者の倫理綱領（案）」を提案している。今後、研究機関・大学・企業内に人工知能に関する研究倫理審査委員会（Institutional Review Board: IRB）が設置されることが予想される。