

「AIと倫理」の議論はなぜ噛み合 わないのか？ — 法律家の視点から —

第16回ITSシンポジウム2018(同志社大学)

名古屋大学大学院法学研究科教授
林 秀 弥

shuya.hayashi@law.nagoya-u.ac.jp

自己紹介



京都大学大学院法学研究科博士課程単位認定退学、京都大学助手、神戸市外国語大学講師を経て、2005年4月より名古屋大学助教授。2007年4月より同准教授。2013年4月より同教授。専門は経済法、情報法

【公職・学会等】

- ・総務省・電波監理審議会委員
- ・総務省・電気通信市場検証会議 座長代理
- ・総務省 情報通信政策研究所 特別研究員
- ・林野庁直轄事業契約監視等委員会 外部委員
- ・国土交通省中部地方整備局・入札監視委員会委員
- ・日本学術会議 連携会員
- ・日本学術会議 若手アカデミー委員会 委員（
- ・日本経済法学会 日本国際経済法学会 理事
- ・American Antitrust Institute Advisory Board Member
- ・情報通信学会 常任理事 編集委員長
- ・地区防災計画学会 理事

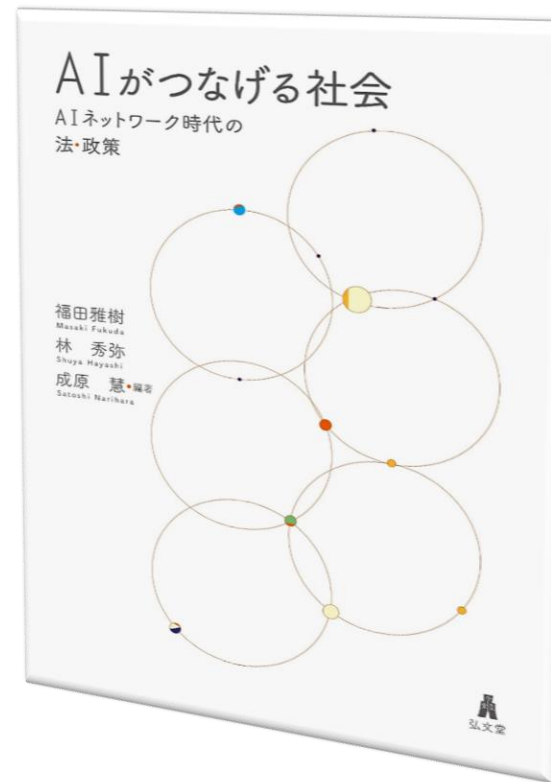
最近のAIに関する研究成果

『AIが繋げる社会- AIネットワーク時代の法・政策』(弘文堂、2017年)

AIが繋げる近未来の社会設計を考える

本書は、AIの発展による雇用環境の激変といった注目を集める論点から、AIにより日常的に情報収集が行われることによるプライバシーへの影響、AIを利用した選挙がもたらす民主主義への影響など統治機構に関する問題、さらにはAI・ロボットに関する権利や責任の分配まで、近未来の社会において生じ得るさまざまな問題を想定。来るべき社会に備えた法の役割や政策の課題を、将来の具体的なユースケースを想定したシナリオを示しつつ、第一線の執筆者たちが提言します。

- 第Ⅰ部 総論
- 第Ⅱ部 研究開発
- 第Ⅲ部 データ・情報・知識の流通と利活用
- 第Ⅳ部 プライバシーとセキュリティ
- 第Ⅴ部 AIネットワーク時代における社会の基本ルール
- 第Ⅵ部 AIネットワーク時代における人間



問題の所在：課題解決における利用者選好の検討の欠落

これまでの自動運転をはじめとするAIに関する課題の多くは、技術的側面からの解決アプローチのみであった。例えば、Loukissas and Mindell (2014) 等における技術論の視点は、新技術への移行に関する課題が提示されており、有益な示唆を含んでいるが、**そこには需要に関連する意思決定、サービスのスイッチに係る障壁、といった重要な心理的・法的課題が未検討のままである。**たとえ技術的に自動運転が実現したとしても、利用者の選好が新たなサービスとしての自動運転に向かわなければ普及は難しい。

(参考文献) Yanni Alexander Loukissas, David Mindell (2014)

“Visual Apollo: A Graphical Exploration of Computer–Human Relationships” Design Issues, Volume 30, Number 2, pp.4–16.

「AIと倫理」の議論はなぜ噛み合っていないのか

1. 自動運転の技術レベルと利用者選好の関係性の研究の希薄性

自動化のレベルによって利用者が感じる効用、懸念は異なる。運転補助のような運転機能の一部のみの自動化であれば、得られる効用も小さいかもしれないが、抵抗感も小さいことが予想される。逆に、完全自動走行は、リスクを過大に評価するおそれがある。自動運転のレベルごとに利用者が感じる効用および懸念を、定量的側面、定性的側面の両面から、まだ明らかにできていないのではないか。

2. 自動運転普及に資する法制度的対応

自動運転の普及は、これまで出現していなかった新たなリスクを生起させる可能性があり、それらのリスクには、技術と同時に法制度的対応が必要となる。例えば、自動運転によりネットワークの遅延や停止により自動運転車が動作しなくなったり想定外の動作をしたりするリスクが懸念されるが、このリスクには技術的解決とともに、リスク発生時の責任の所在を明確にする法整備が必要となる。1. の課題との融合を図ることによって、特に利用者¹に生じる新たな懸念に対応するために必要な法制度的対応を明らかにすることが必要ではないか。

法律家はまずリスクから考える

AIネットワーク化の進展がもたらすリスク(1/2)

1. AIネットワークの機能に関するリスク:

AIネットワークに期待される機能が適正に発揮されないリスク

2. 法制度・権利利益に関するリスク

AIネットワークにより権利利益等が侵害されるリスク

出典:総務省「AIネットワーク社会推進会議」の資料より抜粋

リスクの種類	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システムに対するハッキングやサイバー攻撃 -AIシステムに対する攻撃が秘かに行われ、被害に気づかないこと
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システムが混在することにより、AIシステムが正常に動作せず、意図しない事象が発生 -情報通信ネットワークの不具合によりAIシステムが正常に動作せず、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のアルゴリズム等が不透明化し、人間にAIシステムの適正な制御が困難化ないし不可能化
制御喪失のリスク	-AIシステムが暴走し、人間による制御が困難化ないし不可能化

AIネットワーク化の進展がもたらすリスク(2/2)

リスクの種類	例
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
犯罪のリスク	-AIシステムを悪用したマルウェアによる犯罪 -自律型兵器のテロ等犯罪への悪用
消費者等の権利利益に関するリスク	-AIシステムが適正に利活用されないことにより消費者、青少年、高齢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システムによる個人情報収集・利活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システムが人々の信念、健康、将来の行動等を推論することにより、プライバシーが侵害
人間の尊厳と個人の自律に関するリスク	-AIシステムが人間の意思決定過程を見えない形で操作することにより、個人の自律が侵害 -AIシステムが一定の分野で人間を上回る知能を持つことにより、人間中心主義的な価値体系が動揺
民主主義と統治機構に関するリスク	- AIシステムによる投票等国民の行動の操作 -AIシステムを国家の統治に利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

倫理の問題も「AIネットワーク化」の 進展段階に応じて

倫理の問題も「AIネットワーク化」の進展段階に応じて

1. AIシステムが、他のAIシステムとは連携せずに、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単独で機能。

2. 複数のAIシステム相互間のネットワークが形成され、ネットワーク上のAIシステムが相互に連携して協調。

3. 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身体又は脳と連携することを通じて、人間の潜在的な能力が拡張。

4. 人間とAIネットワークが共生し、人間社会のあらゆる場面においてシームレスに連携。

AIガバナンスにおけるステークホルダ 参加の重要性

AIガバナンスにおけるステークホルダの参加の重要性

- AIネットワーク化に関するステークホルダとは？
 - AIのソフトウェアの研究者・プログラマ
 - ネットワーク化され得るAIシステムの研究者・開発者
 - システムインテグレータ、サービスプロバイダ
 - 最終利用者(個人、団体)
 - 関連する第三者
 - (例 AIネットワークの最終利用者がその機能を用いて提供するサービスの顧客)
 - 政府
 - 国際機関
 - 市民社会(例 非営利団体)
- AIネットワーク化のガバナンスの過程においては、関係するステークホルダの参加及び当該ステークホルダ相互間のコミュニケーションを確保して、コンセンサスを形成するとともにベストプラクティスを共有することが強く期待されよう。

さいごに：今後の主な議論の方向性

さいごに: 今後の主な議論の方向性

○ 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
- ・ 継続的に議論を行っていくことが重要

○ 人間中心の社会の構築

- ・ A I 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

○ A I 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（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
- ・ 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

説明責任

- AI・PFが反差別ポリシーを構築し、最低限これをユーザーに明示することが必要となろう。
- 第三者による監視の機会等を設ける必要があるのではないか。そのためには、監査に必要な制度整備を設ける必要があるかもしれない。

御静聴くださり、ありがとうございます。