

日本語仮訳

AI における倫理及びデータ保護についての宣言  
第 40 回データ保護プライバシー・コミッショナー国際会議  
2018 年 10 月 23 日（火）ブリュッセル

第 40 回データ保護プライバシー・コミッショナー国際会議は、

マラケシュでの第 38 回データ保護プライバシー・コミッショナー国際会議における AI、ロボット工学、プライバシー及びデータ保護に関する当初の議論を考慮し、

AI システムが、使用者及び社会に大きな利益をもたらし得ること、これらの利益には、プロセスの迅速性向上及び意思決定支援、民主的プロセスへの参加のための新しい方法の創設、公的部門及び産業における効率性の改善、資源及び機会のより公平な配分の達成、国民の健康・医療・セキュリティ・持続可能な発展・農業及び運輸などの様々な分野における新たな方策及び解決策の提供、科学的研究及び教育における新しい機会の提供、個人に対するよりパーソナライズされたサービスの提供が含まれることを認識し、

AI における特定の分野、特に、大量の情報の処理、人の行動及び性格の分析及び予測に関する著しい進展、そして、近い将来において著しい進展が見込まれるロボット工学、コンピュータービジョン及び自立型システムなどの関連分野を考慮し、

ビッグデータ及び AI の急速な進歩、特に、機械学習、特に、深層学習技術の進歩に伴い、潜在的な決定を導く複雑な操作をアルゴリズムが解決できるようになるが、そのプロセスはより不透明になることを強調し、

プライバシー及びデータ保護に関する権利の尊重は、AI の進展によってますます課題になってきていること、またこの進展は、倫理的かつ人権的配慮によって補完されるべきであることを確認し、

特に機械学習の技術は、また一般に AI システムは、その発展を大量の個人データの処理に依存しており、データ保護及びプライバシーに影響を与える可能性があることを

考慮し、また、AI の分野における現在の市場集中の傾向によって引き起こされる潜在的リスクについても考慮し、

個人情報の収集・使用及び開示—プライバシー及びデータ保護の従来の範囲—と、特に差別や表現及び情報の自由に関するより広範な人権への直接的な影響との相関関係を認識し、その結果として、データ保護及びプライバシー機関がより幅広く人権を考慮し、人権を扱う他の機関と協力する必要性を認め、

機械学習ベースのシステム及び AI システムの訓練に使用されるいくつかのデータセットには、特定の個人又はグループを不公平に差別しうる決定につながり、また、潜在的に特定のサービスやコンテンツの利用を制限し、これにより、表現及び情報の自由など個人の権利を侵害し、又は、個人的、社会的、職業的な生活における特定の側面から人々を排除する結果をもたらすこととなる、固有のバイアスが含まれていることが明らかとなっていることを指摘し、

その決定につき説明が不可能な AI 動力システムは、プライバシー及びデータ保護法における説明責任のみならず、エラー及び損害が生じた場合の責任に関する基本的な疑問を惹起させることを強調し、

AI の分野における多くの利害関係者は、プライバシー、データ保護及び人間の尊厳に関連するリスクと同様に、AI の悪用のリスクについても懸念を表明しており、例えば、大衆監視と組み合わせた AI の進展は、基本的な権利及び自由の抑制に使用される可能性についての懸念を惹起すると指摘していることに留意し、

これらのリスク及び諸課題は、個人及び社会に影響を与えるであろうこと、また、将来的な影響の程度及び性質は、現在のところ不明確であることを強調し、

強力なデータ保護及びプライバシー保護措置は、データ処理過程における個人の信頼の構築に資し、データ共有を推進することによりイノベーションを促進することから、信頼の重要性を強調し、

AI 及び機械学習システムによって引き起こされた現在の諸課題は、国際的なレベルでの全てのデジタル開発における人権の促進及び保護を確保するため、国際的なアプロ

ーチ及び基準の採択の必要性を高めているとの見解を採り、

データ保護機関及びデータ保護プライバシー・コミッショナー国際会議が行った、この変わりゆく環境への適応に当たり、データ保護及びプライバシー原則を維持すること、特に、将来の変化に備えるべく資源を投入し、新しいスキルを開発するというコミットメントを再確認する。

第40回データ保護プライバシー・コミッショナー国際会議は、AIシステムに関するすべての創作、開発及び使用においては、人間の尊厳、無差別及び基本的価値と同様に、人権、特に、個人データの保護及びプライバシーの権利が尊重されなければならない、かつ、個人がAIシステムの管理及び理解を維持できる解決策が提供されなければならないと考える。

したがって、当会議は、AIの開発における人権保護のための本質的価値として、次の指導原則を支持する。

1. AI及び機械学習の技術は、特に次に示すところによって、基本的人権を尊重し、公正の原則に従ってデザイン、開発及び使用されるべきである。
  - a. AIシステムの使用につきその当初目的との一貫性を確保すること、そして、データがその当初の収集目的と矛盾しない形で使用されることを確保することにより、個人の合理的な期待を考慮すること。
  - b. AIの使用が個人に与える影響の程度のみならず、グループ及び社会全体に与える集団的影響についても考慮すること。
  - c. AIシステムは、人間の発達を促進し、それを妨害又は危険にさらすことのないよう開発されることを確保すること、したがって、特定の使用において説明と線引きの必要性を認識すること。

2. AI システムは、特に次に示すところによって、説明責任と同様に、その潜在的な影響及び結果に対する継続的な注意及び警戒が確保されるべきである。
  - a. すべての関連する利害関係者の個人、監督機関及び必要に応じてその他の第三者に対する説明責任を促進すること。これには、AI システムに対する監査の実施、継続的監視及び影響評価、監督メカニズムの定期的な見直しといった手段を含む。
  - b. 行為者及び利害関係者全体の繋がりを含めた共同かつ連帯の責任を、例えば、共同基準の作成及びベストプラクティスの共有により培うこと。
  - c. AI 及び社会におけるその潜在的影響に関する及びその理解のための十分な情報を確保するため、注意喚起、教育、研究及び研修に投資すること、そして、
  - d. 信頼できる第三者への依頼や独立の倫理委員会の立ち上げなどにより、すべての関連行為者に対し証明可能な管理プロセスを確立すること。
3. AI システムの透明性及び明瞭性は、効果的な執行の観点から、特に次に示すところによって、改善されるべきである。
  - a. 説明可能な AI に関する公的及び私的な科学研究への投資
  - b. 各関連する聴衆のために求められるそれぞれ異なったレベルの透明性及び情報を考慮し、例えば、革新的な通信手段の開発を通じて、透明性、明瞭性及び到達可能性を促進すること。
  - c. 組織の慣行について、特に、提供された情報の有意性を確保しつつ、システムのアルゴリズム的透明性及び監査可能性の促進により、さらに透明化すること。そして、
  - d. 情報に基づく自己決定の権利について、特に、個人が AI システムと直接交流しているとき又は彼らがそのシステムによって処理される個人データを提

供するとき、常に適切な方法で知らされていることを確保することにより、保証すること。

- e. 個人の期待に沿った継続的な協力を立証するために、またそのようなシステムに対する人間による全体的な制御を可能とするために、AI システムの目的及び効果に関する十分な情報を提供すること。
4. 「エシックス・バイ・デザイン」の全体アプローチの一部として、AI システムは、特に次に示すところによって、プライバシー・バイ・デフォルト及びプライバシー・バイ・デザインの原則の適用により、責任をもってデザインされ開発されるべきである。
- a. データ主体のプライバシー及び個人データが、処理方法の決定時点及びその処理の時点の双方において尊重されることを確保するため、開発されたシステムの形態に比例した、技術的かつ組織的な手段及び手続を実施すること。
  - b. AI プロジェクトの開始時点において予想される個人や社会に対する影響及びそのライフサイクル全体にわたる関連する開発について、評価し文書化すること。そして、
  - c. システムの倫理的かつ公正な使用のため、及び、すべての AI システムの開発・運用の一部としての人権の尊重に係る具体的な要件を明らかにすること。
5. 特に次に示すところによって、すべての個人の権限強化が推進され、広範な参画の機会の創設と同様に、個人の権利行使が促進されるべきである。
- a. データ保護及びプライバシー権—これには、適用できる場合には、情報に対する権利、アクセス権、処理に反対する権利及び削除権が含まれる—を尊重すること、また、教育及び注意喚起キャンペーンを通じてこれらの権利を促進すること。
  - b. 無差別の権利と同様に、表現及び情報の自由を含む関連する権利を尊重すること。

- c. 個人の成長や見解に影響を与える技術への適用に異議を申し立てる又は抗議する権利を認めること、及び、適用できる場合には、もしそれが個人に重大な影響を与える場合には自動化された処理のみに基づく決定には従わない個人の権利を保証すること、また、適用できない場合にはその決定に異議を申し立てる個人の権利を保証すること。
  - d. 平等な権利強化を育成し、例えば、適応できるインターフェイスやアクセスが可能なツールを通じて、広範な参画を促進する AI システムの能力を活用すること。
6. AI におけるデータの使用の結果もたらされるであろう違法なバイアス又は差別は軽減及び緩和されるべきである。これには、次の方法が含まれる。
- a. 人権及び無差別に関する国際的な法的手段の尊重を確実にすること。
  - b. バイアスを特定し、対処し、緩和するための技術的方法の研究に投資すること。
  - c. 自動化された意思決定に使用される個人データ及び情報が、正確、最新かつ可能な限り完全であることを確保するための合理的なステップを踏むこと。
  - d. バイアス及び差別への対処に際しての具体的な指針及び原則を詳細に定め、個人及び利害関係者の意識を促進すること。

上記の原則を考慮して、第 40 回データ保護プライバシー・コミッショナー国際会議は、その開発及び使用が倫理及び人間の価値に即して行われ、人間の尊厳を尊重することを確保するため、AI に関する一般的なガバナンス原則の確立及びこの分野における協調した国際的な取組の発展を求める。これらの一般的なガバナンス原則は、すべての重要な横断的問題に対処するため、マルチステークホルダーアプローチに基づき、AI 技術の急速な革新により生ずる課題に取り組むことができるものでなければならない。AI の開発は国境を越えた現象であり、すべての人類に影響を与えることから、これらは国際的なレベルで行われなければならない。本会議は、この国際的な取組に関与し、競争、市場及び消費者規制など他の分野における一般及び部門の当局との共働及び支援をす

べきである。

したがって、第 40 回データ保護プライバシー・コミッショナー国際会議は、国際的なレベルにおける将来の一般的なガバナンスへの貢献として、また、AI における倫理及びデータ保護に関する原則に沿った指針をより詳細に定めるため、AI の開発に係る課題に取り組む常設作業部会を立ち上げる。この AI における倫理及びデータ保護に関する作業部会は、AI システムの開発に関与するすべての関係者（政府及び公的機関、規格統一団体、AI システムデザイナー、プロバイダー及び研究者、企業、市民及び AI システムのエンドユーザーを含む）による現在の決議の原則に対する理解及び尊重の促進を担当することになるであろう。AI における倫理及びデータ保護に関する作業部会は、本会議の他の作業部会が行う業務を考慮するものとし、その活動を定期的に本会議に報告するものとする。したがって、本会議は、この分野における力強い倫理文化及び個人の認識の創設を目的としたデジタル倫理に関する活発な議論を積極的に支援するよう努める。