

EUが「AI倫理指針」を公表、日本とも連携へ

◆欧州委員会が、AI倫理指針を発表、今後は試験段階、国際的合意を目指す

AI（人工知能）は、医療、教育、自動車安全、農業、金融リスク管理など、多くの分野で利益をもたらす。また、詐欺やサイバーセキュリティ上の脅威の検出に役立ち、警察などがより効率的に犯罪に対処することを可能にする。そうしたポジティブな面の一方で、AIは、将来の仕事のあり方に新たな課題をもたらし、また、差別増長など倫理上の懸念など、ネガティブな面もあると指摘されている。

このため、欧州連合（EU）の欧州委員会は、個人データ保護や透明性などの基本原則を考慮したAI倫理指針を策定するため、産学官の「AIに関するハイレベル専門家グループ」による検討を進め、2019年4月8日、AIの開発と利用に関する「人間中心で信頼できるAI」の倫理指針を発表した。下記に7つの要件を示す。

欧州委員会が示した AI 倫理指針

1. 人間の代理機能と人間による監視 Human agency and oversight	AI は、人間の活動と基本的人権を支援することで公平な社会を可能にすべきで、人間の主体性を低下させたり、制限、誤誘導したりすべきではない
2. 堅牢性と安全性 Robustness and safety	信頼できる AI には、全過程を通じて、エラーや矛盾に対処し得る安全かつ確実、堅牢なアルゴリズムが必要
3. プライバシーとデータのガバナンス Privacy and data governance	市民が自身に関するデータを完全に管理し、これらのデータが市民を害し、差別するために用いられないことがないようにすべき
4. 透明性 Transparency	AI システムのデータの処理のされ方などのトレーサビリティ（追跡可能性）を確保すべき
5. 多様性・非差別・公平性 Diversity, non-discrimination and fairness	AI は、人間の能力・技能・要求の全分野を考慮し、アクセスしやすいものとすべき
6. 社会・環境福祉 Societal and environmental well-being	AI は、社会をより良くし、持続可能性と環境に対する責任を向上するために利用すべき
7. 説明責任 Accountability	AI システムと AI により得られる結果について、責任と説明責任を果たす仕組みを導入すべき

EUは、今回の「信頼できるAIのための主要な要件の設定」、「関係者からの意見などを得るための大規模な試験段階」開始の発表、そして「人間を中心としたAIに向けた国際的なコンセンサス作り」と3つのステップを踏んでいく計画だ。

倫理指針には、これらの要件を満たしているか確認するための「評価リスト」が含まれている。同委員会は19年夏から、倫理指針適用の試行期間を設定し、そのフィードバックに基づき、20年初めに評価リストを見直す予定で、産業界、研究機関、公的機関に対し、評価リストを試すよう要請し、参加機関へは、「欧州AI連合 (European AI Alliance)」への登録を呼びかけている。

今回の動きは、この分野への官民投資を今後10年で毎年最低200億ユーロ増加させることを目的とした、18年4月発表の「AI戦略」計画の一環である。

◆EUは日本など各国との連携も視野に

EUが指針を公表した背景には、日本、カナダ、シンガポールなど意思を共有するパートナー国との協力強化、G7やG20での議論を通じてEUがルールづくりで主導権を握りたい狙いもあるとみられている。19年6月の日本で開催のG20でAIが議題に上ることも考えられる。

日本政府も「人間中心のAI社会原則」を掲げ、夏には「AI戦略」をまとめ、AIやデータを活用して人手不足など社会課題の解決をめざすとしている。4月のイベント「アイサム (AI/SUM)」の基調講演で、世耕経済産業相は「人とAIが対立ではなく共に進化し、課題を解決していくことがめざす姿だ」と語っている。

日本と欧州の国レベルでは既に連携が進みつつある。4月16日には、日独仏3国間でAI研究を強化する基本合意書を締結した。6月には、フランス国立研究機構 (ANR)、ドイツ研究振興協会 (DFG)、日本の科学技術振興機構 (JST) の3つの研究支援機関による、AIに関する研究提案の募集が開始される。日独仏の研究者チームが対象で、3機関による評価を経て、先端技術の開発などが支援される。

この3カ国は18年11月に2日間にわたり、第1回「人工知能に関する日独仏合同シンポジウム」を開催しており、専門家350人以上を集めて盛況だったことが、今回の基本合意書の締結につながったとされている。シンポジウムでは、医療、交通、倫理、産業など幅広い分野でのセッションと、各国の専門家が意見を交わし、「共通の基準と人工知能の可能性についての共同理解を設定する人間中心のアプローチをすること」「人工知能の究極の目的は、人々に奉仕し、個人だけでなく社会全体の生活の質の向上に貢献すること」と共同声明で謳っている。

今後は、日本と欧州のAI分野での連携が、さらに進むだろう。 【赤山英子】