

社会イノベーション事業に AI を活用するための AI 倫理原則

日立製作所
2021年2月22日

日立製作所(以下、「日立」)は、社会イノベーション事業に AI^{*1}を適切に活用することにより、社会価値・環境価値・経済価値を一層向上するとともに、人間の尊厳に立脚した快適で持続可能な社会の実現と、世界の人々の Quality of Life(QoL)の向上に貢献することをめざしています。そのために、日立は、100年以上にわたり取り組んできた高品質な OT(Operational Technology、制御・運用技術)に、50年以上の実績のある IT(Information Technology、情報技術)をベースとした AI 技術の知識と、脳科学研究などを通じて培った科学技術における倫理に関する知見を統合し、AI の利活用を推進します。日立は、AI 倫理原則として、AI の計画・社会実装・維持管理の3つのフェーズにおける行動規準と、全てのフェーズに共通する7つの実践項目を定め、この原則に基づき AI を利活用します。

行動規準

I 持続可能社会実現のために、AI の開発、利活用を計画します

AI の潜在的な倫理的リスクを抑制し、新たな価値の創出を図るには、計画の時点において、サービス、ソリューションやプロダクトに AI を利活用しようとする目的が適切であることが重要です。日立は、社会の多様な課題を解決し、快適で強靱な持続可能社会の実現や、世界中の人々の QoL 向上のために AI の開発と利活用を計画します。

II 人間中心の視点で、AI を社会実装します

AI の判断結果が、個人を尊重し、社会の利益に資するよう、責任をもって AI を社会実装し、AI と人間との共生を可能とすることが重要です。日立は自由、公平、公正の理念にのっとり、人間中心の視点で AI の社会実装を行うとともに、AI が想定通りに機能するよう検証に努めます。

III 提供価値が長期間にわたり持続するよう、AI を維持管理します

AI の社会実装後も、AI が長期かつ安定的に価値を提供し続けることが重要です。日立は、AI の提供価値が社会や環境の変化に応じ継続的に受け入れられるよう維持管理に努めます。

実践項目

1. 安全重視

AI および AI を活用したシステムやソリューションが想定通りの動作をするよう品質を検証し、利用者やそれに関係する者の生命や健康をはじめとする人権、財産、名誉、信頼、信用を守ることに努めます。また、同時に、地球環境の破壊あるいは悪化を防止し、人々に安全な暮らしを提供する AI の実現、運用に努めます。

2. プライバシー保護

AI の学習、評価、運用に利用する入力データおよび AI が出力するデータに関して、個人情報適切に扱い、プライバシーを含む権利を保護するよう AI の実現、運用に努めます。

3. 公平性実現

AI の判断結果が、多様なステークホルダーの利益に資するとともに、人種、性別、国籍などによる差別や偏見を発生させたり助長させたりすることがないよう AI の実現、運用に努めます。

4. 適正開発・利活用

AI が設計された用途や動作条件から逸脱した使い方をされないよう、AI の具体的な利活用シーン(ユースケース)におけるリスクポテンシャルを踏まえた開発を行い、AI の適正な利活用に努めます。

また、利用者や運用者等に対し利用方針、利用条件等を示すことで AI の適正な運用に努めます。さらに、AI が利用される動作環境の変化、AI やその判断結果に対する人々の意識や社会状況の変化などを随時確認し、AI の適切な維持管理に努めます。

5. 透明性・説明責任重視

AI の判断結果の根拠などを検証し説明できるよう AI の透明性確保に努めるとともに、AI やその判断結果に関して用途や状況に応じ、説明責任を果たすよう努めます。

6. セキュリティ重視

情報漏洩、改ざん、システムの破壊、サービスの妨害などを防止するよう、システムや運用レベルでの対策も含め、セキュリティを重視した AI の実現、運用に努めます。

7. 法令遵守

利用される国や地域の法令を遵守した、AI の実現、運用を行います。

*1 AI (Artificial Intelligence、人工知能): データ、情報、知識の学習により、自らの出力や処理方法を変化させる機能を持つ処理モジュール、および、このようなモジュールを構成要素として含むシステム

以上