

倫理的な人工知能のための 10 大原則



UNI Global Union

Nyon, Switzerland

World Executive Board, 10-11 October 2017

倫理的な人工知能のための 10 大原則

「労働の未来」に関する UNI 動議の原則 6「人々と地球を優先する AI 倫理の保証」は、以下を要求する。

人工知能 (AI) は人々と地球を優先すべきである。これは国際規模の AI 倫理の議論が必須である理由である。倫理的 AI に関して全てを網羅する国際条約は、人類の生存にとって最も実現可能な保証となる。

AI が人類の生活のあらゆる側面に入り込んでいる今こそ、AI の適切な使用について議論し決定する時であるという点で、有識者達の意見は一致している。UNI は、労働者や社会にとっての AI の便益を強調すると同時に、AI の意図しない負の結果を予防し、これに対処する上で有益な、倫理的 AI に関する国際条約を求めてきた。我々は、人類と企業はこれについて責任を負う主体であるということを強調する。

本文書は、この重要な要求を運用可能なものにし、組合や職場代表、グローバル労組同盟が行動を取れるよう、10 の項目に分けた。

目次

1 AI システムの透明性を要求する	16
2 AI システムに「倫理的ブラックボックス」を装備する	16
3 AI を人々と地球に役立つものにする	16
4 人間が主導権を握るアプローチを採用する	17
5 ジェンダーレスで偏見のない AI を確保する	17
6 AI システムによる恩恵を共有する	17
7 「公正な移行」を確保し、基本的自由および権利の支持を確かなものとする	17
8 グローバルなガバナンス構造を確立する	17
9 ロボットへの責任帰属を禁止する	18
10 AI の軍拡競争を禁止する	18

1 AI システムの透明性を要求する

透明性のある AI/AS であってこそ、システムが「どのようにして」「なぜ」、特定の決定を下したのかについて、我々は知ることができ、またロボットの場合には、なぜそのように行動したのかを知ることができるのである。特に、

- a) オープンソース・コードは、透明性の確保にとって必要条件でも十分条件でもないことを強調する。明瞭性は、複雑性によって曖昧にすることはできない。
- b) ユーザーにとって、透明性は重要である。なぜなら、システムが何をどのような理由で行うのかを理解するための簡潔な方法をユーザーに提供することで、システムの信頼と理解を構築するからである。
- c) AI/AS の妥当性確認と認証を行う上で、透明性は重要である。なぜなら、システムが精査の過程に晒されるからである。
- d) 事故が生じた場合、事故を招いた内部プロセスが明らかになるよう、AI/AS は事故調査員に対して透明性を確保し、説明責任を負う必要がある。
- e) 労働者は、AI システムの決定や結果、背景にあるアルゴリズムに関する透明性を要求する権利を持たねばならない（後述の原則 4 を参照）。これには、AI やアルゴリズムが下した決定に抗議し、人間による再調査を求める権利が含まれる。
- f) 労働者は、AI システムの実施や開発、配置について意見を求められなければならない。
- g) 事故が生じた場合、裁判に関わる裁判官、陪審員、法律家、鑑定人は、証拠を得て決定を下すため、透明性と説明責任を要求する。

透明性の原則は、残りの原則が遵守される上での必須条件である。

運用上の解決策については、以下の原則 2 を参照。

2 AI システムに「倫理的ブラックボックス」を装備する

AI システムにおける十分な透明性は、システムに関する情報を「倫理的ブラックボックス」の形式で記録する装置によって促進されるべきである。この形式は、システムの透明性や説明責任を確保する関連データを含むだけでなく、倫理的配慮に関してシステムに組込まれた明確なデータと情報も含む。

ロボットに適用された場合、倫理的ブラックボックスは、全ての決定とその根拠、動き、ロボット操作者用のセンサーデータを記録する。ブラックボックスが提供するデータは、人間のユーザーにとって理解可能な言語でロボットの動きを説明し、人間とロボットの関係性やユーザー体験を改善する上でも役立つだろう。倫理的ブラックボックスからの情報出力は、簡潔かつ迅速なものでなければならない。

3 AI を人々と地球に役立つものにする

ここには、AI の開発、適用、使用に関する倫理規定が含まれる。AI システムが、その全運用過程を通じて、人間の尊厳、人格、自由、プライバシー、文化やジェンダーの多様性といった原則や基本的人権と矛盾なく存在し、これらを向上させるためである。加えて AI システムは、地球の生態系や生物多様性を守り、さらには改善させるものでなければならない。

4 人間が主導権を握るアプローチを採用する

絶対的な前提条件は、機械が合法的であり、法人が常に機械を管理しこれに対して責任を有する時、AIの開発は、責任ある安全で有益なものでなければならない、ということである。

そのためには AI システムが、プライバシー法を含め、既存の法に合致するように設計・運用される必要がある。人々は、システムがデータを分析し活用する能力を持つことを前提に、AI システムが生み出したデータにアクセスし、これを管理・制御する権利を持つべきである。（「労働者データのプライバシーと保護のための 10 大原則」の原則 1 を参照）

5 ジェンダーレスで偏見のない AI を確保する

AI/AS の設計および保守において、システムが否定的で有害な人間の偏見に対して統制を持ち、その偏見がジェンダー、人種、性的指向、年齢など何であれ、偏見は識別され、システムによって偏見が広まらないことが不可欠である。

6 AI システムによる恩恵を共有する

AI 技術によって、可能な限り多くの人々が恩恵を受け、エンパワーされるべきである。AI がもたらす経済的繁栄は、全人類が恩恵を受けられるよう、広く公平に分配されるべきである。そのため、経済的・技術的・社会的なデジタル面での格差解消を目的とした、グローバルおよび国家レベルの政策が不可欠である。

7 「公正な移行」を確保し、基本的自由および権利の支持を確かなものとする

AI システムが発展し、拡張現実が形成されるにつれ、労働者や業務がこれに置き換えられていくようになるだろう。そうした失業者の求職を支援する特定の行政措置など、デジタル・リアリティへの公正な移行を確保するための方針の導入は不可欠である。

デジタル経済への広範な移行を伴う AI システムは、あらゆる職業の全階層の労働者が、社会保障制度ならびに雇用適性を維持するための継続的生涯学習の機会を利用できるようになることを必要とする。あらゆる雇用形態の全ての労働者が、この 2 つに対する権利を持ち、利用できるよう解決策を見出すことは、国家と企業の責任である。

加えて、労働の臨時雇用化と個別化が高まる社会においては、あらゆる雇用形態の全労働者が、同じ強力な社会権と基本的権利を持たねばならない。全ての AI システムは、AI の配置・拡大が、人権法や ILO 条約、団体協約に規定された労働者の権利と協調して進むよう、抑制と均衡を図らねばならない。ILO の中核的条約を反映する、システムに組み込まれた「アルゴリズム 8798」は、まさにこの目的にかなうものだ。労働者の権利が守られなかった時は、システムはシャットダウンされなければならない。

8 グローバルなガバナンス構造を確立する

UNI は、多くの利害関係者が加わるディーセント・ワークと倫理的 AI に関するガバナンス構造を、グローバルおよび地域レベルで確立することを提言する。ここには、AI 設計者、メーカー、オーナー、開発者、研究者、雇用主、法律家、市民社会団体、労働組合が含まれるべきだろう。倫理的 AI への移行と導入を確保するための、内部通報制度およびモニタリング手続きが確立されねばならない。このガバナンス構造は、コンプライアンス・プロセスや手続きについて勧告する権限を持つ。

9 ロボットへの責任帰属を禁止する

ロボットは、既存の法やプライバシーを含む基本的権利と自由に合致する限りにおいて、設計・運用されるべきである。これは、法的責任の問題と関わる。ブリソン他の議論（2011）を踏まえて、UNI は、ロボットに対して法的責任のある人物に責任が帰属することを強く主張する。ロボットは、法の下で責任主体とはなりえない。

10 AIの軍拡競争を禁止する

サイバー戦争を含む自律型致死性兵器は禁止すべきである。