



ChatGPT や Amazon もつまずいた、「AI 活用による人権侵害」をどう防ぐか

国連や EU も警鐘を鳴らす、社会を激変させる AI の「人権リスク」

株式会社オウルズコンサルティンググループ プリンシパル
矢守 亜夕美

※2023年3月14日付のJBpressの記事を一部変更して掲載しています。

昨今、企業によるSDGsやESGへの取り組み状況が評価されるようになり、脱炭素などの環境対応に加えて、「ビジネスにおける人権尊重」が重要な経営アジェンダとなっている。サプライチェーン上での強制労働・児童労働や差別、ハラスメントなどの「人権リスク」への対応を怠る企業には、厳しい視線が注がれているのが現状だ。

日本政府も昨年秋にサプライチェーン上の人権尊重に関するガイドラインを発表し、国内企業に取り組み強化を促している。こうした潮流の中で、にわかに注目を集めているのが「AI 活用に伴う人権リスク」の問題だ。一見、その便利さでビジネスの生産性や効率性を大幅に底上げしてくれる「救世主」にすら見える AI だが、実はその裏で深刻な人権リスクが顕在化してしまうことがある。

例えば最近、連日ニュースを賑わせている「ChatGPT」。米国の AI 研究機関 Open AI が開発したチャットボットであり、質問に対してかなり自然な文章で回答できるため注目を集めている。簡単な会話から識者顔負けの論文(内容は正確でないこともあるが)まで、幅広くアウトプットできるのが特徴だ。

この話題の AI について、開発過程での人権問題が指摘されているのをご存知だろうか。

ChatGPT が開発過程で起こした人権トラブル

AI に善悪を含めた正しい判断をさせたり、有害なコンテンツを自動的に排除させたりするためには、当然「悪い」情報もあらかじめ学習させる必要がある。OpenAI は、有害コンテンツを識別するためのラベリング作業を米国の Sama 社に委託していた。Sama 社は過去に Facebook 等にもコンテンツのラベリング業務を提供していたことで知られ、本拠は米国だがケニア、ウガンダ、インド等で多くの労働者を雇用している。ラベリング作業の担当者は、PC の前で長時間にわたり大量のコンテンツを確認してラベルを付けていくのだが、その中には児童虐待、性的虐待、殺人、自殺、拷問といった情報が多分に含まれていた。この過酷な作業を、同社が雇ったケニアの労働者に時給 2 ドル以下で行わせていたことが明らかになったのだ。

同社はカウンセラーとの面談機会などを設けていたが、従業員の中には「この仕事を通じて精神的な傷を負った」「拷問だった」とメディアに訴える人が複数現れた。この事実が明るみに出たことで、OpenAI は各所からの批判にさらされた。現在は OpenAI と Sama 社の契約は解除され、Sama 社は今後センシティブなコンテンツを扱う業務をすべて中止する旨を発表しているが、発注者であった OpenAI には引き続き責任を問う声がある。

ChatGPT の例は、過酷な業務を安価に発注していたことが問題視されたケースだが、他にも AI 活用に関わる人権トラブルは枚挙に暇がない。例えば、2018 年に米 Amazon が開発を試みた AI 採用システムの事例が有名だ。

米 Amazon が生んでしまった「女性差別 AI」

履歴書審査のアルゴリズムを構築するため、同社は過去 10 年分の履歴書を AI に学習させた。その結果、AI が女性よりも男性を優遇した評価を下すようになってしまったという。技術職などでは過去の応募・採用実績ともに圧倒的に男性が多かったため、「この職には男性の方が向いている」と AI が誤った前提を「学習」してしまい、性差別的な判断を下すようになったのだ。「女性」「女子大」などのワードが含まれる履歴書に一律で低評価を下してしまったこの AI システムは、開発中止を余儀なくされた。

米国の医療保険大手ユナイテッドヘルスも、過去に「AI による差別」を巡る同様の問題に直面している。同社は、患者のリスク度合いに応じた医療措置を AI が自動的に判断する「ハイリスク患者判定システム」を開発し、医療機関に提供していた。当初、現場では有用と見られていたが、AI 判定の結果、白人患者が黒人患者よりも優遇されていることが発覚した。症状が同程度の患者でも、白人の方が黒人より「ハイリスク」と判定され、手厚い医療的ケアを受けられる状況になっていたのだ。この AI のアルゴリズム自体に人種のデータが含まれていたわけではない。「患者が使う医療費」の過去データを学習させたところ、同じ症状でも白人の方が平均的により多くの医療費を支出していたため、AI が「医療費が高いということは、症状が重く重要度が高いだろう」と判断していたの

だ。その結果、米国に依然として存在する白人と黒人の経済格差が反映され、白人を優遇する評価につながってしまった。

多くの AI は大量の過去データを基に「学習」することで精度を高めるが、それはつまり、過去から存在してきた格差や不平等、バイアスなどをそのまま反映し、助長してしまうリスクが高いということでもある。便利な技術だから、と深く考えずに活用していると、意図せぬ差別や人権侵害を引き起こす危険性がある。

EU はハイリスク AI を規制する新ルールを検討

いわゆる「顔認識」を行う AI についても、同様の問題が多く指摘されている。米国では警察などによる顔認識 AI の利用が「有色人種への差別や不当な扱いにつながる」と強く批判され、ソフトウェアを提供していた IBM は 2020 年に事業撤退を宣言した。こうした事例や指摘を背景に、国連をはじめとする国際機関も、AI のもたらす人権リスクに警鐘を鳴らしている。

今年 2 月には、国連人権高等弁務官が「近年の AI 技術の進歩が害をもたらす可能性に心を痛めている」とのコメントを発表した。「人間の主体性、人間の尊厳、そしてすべての人権が深刻な危険にさらされている」とまで述べており、強い懸念を表明した形だ。この危機感を共有し、AI 技術への警戒を強めていると思われるのが欧州連合(EU)だ。EU では、現在「AI 規則(The AI Act)」の策定が進められている。2024 年の施行を目指して審議されており、もし正式に制定されれば世界各国の企業に影響が及ぶ見込みだ。今回の AI 規則案の最大の特徴は、AI を人権リスクのレベルに応じて分類し、レベルごとに規制内容を変える「リスクベースアプローチ」を採用している点だ。最も重大な「許容できないリスク」のある AI は EU 内でのサービス提供を禁止され、それに次ぐ「ハイリスク」AI には品質管理などが厳しく義務付けられる。

具体的に、どういったものが「許容できないリスク」と見なされるのか。例えば、子どもや障害者などの脆弱性につけ込んで精神・身体に害を及ぼすようなものや、公的機関によるソーシャルスコアリング(属性や過去の行動に基づいて個人の信用度を評価すること)がこれに該当する。人を対象に生体識別を行うものや、重要なインフラの管理・運営に用いられるもの、雇用・労働者管理に用いられるもの等も、用途によっては「ハイリスク AI」と判断される可能性がある。

AI 活用に必須となる「人権視点」と「倫理観」

AI 規則は EU 全体の「統一ルール」という位置付けになり、違反した際にはペナルティが課せられる。まだ議論中の段階だが、本規則案が通れば、制裁金として最大 3000 万ユーロ(約 43 億円)、もしくは全世界売上高の 6%のいずれか高い金額を課される可能性がある。さらに、違反企業には市場からの取り下げやリコール等の措置が下される可能性もあり、EU 域内でビジネスができなくなる恐れすらある。

かなり厳しい罰則だが、これは日本企業にとっても他人事ではない。EU 域内に居住する人をターゲットに AI システムやサービスを提供する場合は、本規則の適用対象となる。つまり、日本から EU の顧客向けに IT サービスなどを行っている企業もその範疇に入りうる。グローバル化が加速する中、日本を含むあらゆる国の企業が注視すべきルールと言える。

本規則案について、米 Google、IBM などのテック企業からは、規制対象となる AI の定義の明確化などを求める声が上がっている。AI 活用が規制されることで、今後の事業発展やイノベーションが阻害される事態を懸念

するためだ。他方、AI のリスクを深刻視する識者からはさらなる規制強化を求める声もあり、今後の動向に注目が集まっている。

消費者・NGO・投資家なども、人権尊重の視点を欠いた「倫理なき AI」には極めて厳しい目を向けるようになりつつある。国内でも「AI 倫理」の専門組織を立ち上げる企業が増えており、今後 AI を用いたサービス開発などを検討している企業は取り組みが必須となるだろう。

今この瞬間も急速に進化している AI が、果たして人類にとっての「救世主」となるか、人権を踏みにじる「悪魔」となるのか。AI を活用する企業の倫理観こそが、その命運を握っている。

著者



株式会社オウルズコンサルティンググループ
プリンパル
矢守 亜夕美 / Yamori, Ayumi

A.T. カーニー(戦略コンサルティング)、Google、スタートアップ等を経て現職。東京大学法学部(公法コース)卒。

事業戦略立案・組織改革など多岐にわたるプロジェクトを経験すると共に、近年はサステナビリティや「ビジネスと人権」関連のプロジェクトを多数担当。

著書に『すべての企業人のためのビジネスと人権入門』(共著: 日経BP社)がある他、日本経済新聞、読売新聞、時事通信、CBCテレビなど、人権・社会課題に関するコメント掲載多数。経済産業省/中小企業庁「ビジネスと人権」セミナー講師(2021年)、東京都人権プラザ主催「サステナビリティと人権」セミナー講師(2022年)、サステナブル・ブランド国際会議『「守り×攻め」の人権対応を実現するために」(2023年)等、登壇実績多数。

労働・人権分野の国際規格「SA8000」基礎監査人コース修了。

ISO30414 リードコンサルタント/アセッサー。

本資料は一般的な情報提供を目的とするものであり、その性質上、特定の個人や事業体に具体的に適用される個別の事情に対応するものではありません。関連する法令等の解釈を行ったものではなく、利用者が本資料を利用したことによる結果について、株式会社オウルズコンサルティンググループは一切の責任を負うものではありません。

また、書面による株式会社オウルズコンサルティンググループの事前承認なしに、第三者への配布・引用・複製を行うことはお断りしております。

株式会社オウルズコンサルティンググループ

〒106-0046 東京都港区元麻布 3-1-6

<https://www.owls-cg.com/>