

最近、家電製品や金融商品にもAI (Artificial Intelligence:人工知能) が取り入れられ、高度なサービスが提供され始めています。このコラムでも、取り上げられているSociety5.0を実現させるためには、AIの活用は必要不可欠です。私たちの生活を快適にしてくれているAIですが、危険とも隣り合わせであるということを認識していますか？

皆さんは、AIが引き起こしたGoogleフォトの人種差別問題やAmazonの人材採用における性差別問題などを聞いたことがあるでしょうか？

Googleは、2015年、画像認識ソフトウェアによる自動タグ付け機能において、黒人の写真に「ゴリラ」というタグが付けられたことで問題になりました。これは、学習データの人種の構成の偏りが要因だったといわれています。Amazonは、2014年頃、採用を効率化するための人工知能システムを開発していました。しかし、開発から1年が経過した頃、AIは経歴が10年以上ある男性求職者にのみフォーカスして、雇用すべき候補として選び出し始めたそうです。この原因も、学習データとして用いていた履歴書の大半が男性であり、学習データの構成が偏っていたからではないかと言われています。このシステム開発は、2017年に中止しています。

このように、学習に用いられるデータセットによって、人種差別や性差別を生み出してしまうことがあるのです。また、AIが学習を行うためのデータセットを用意するためには、人が画像のタグ付けを行ったり、適切な人を選び出すためのアルゴリズムの調整を行ったりします。この作業の際、人間の潜在意識にある偏りが反映してしまうことがあり、アルゴリズムの偏りを排除することはとても難しい問題と言われています。

また、最近では、AIによる自然言語処理が発展し、AIにツイートさせたり、フェイクニュースを生成させたりすることも可能となりました。人間のものと区別ができないAIチャットボットの開発も進んでいます。現在のAIを用いれば、他人の名誉を毀損するような内容を自動生成して発信したり、ネット上で公開されている他人の著作物を収集して別のコンテンツへ無断で使用したりすることも可能です。更には、悪意あるツイートやニュースを大量に生成し、社会を混乱させることも不可能ではありません。

このような状況下で、より公正なデータセットやAIを目指し、研究者らはデータの見直しや、バイアス検知アルゴリズムの開発などを行っています。また、各国の政府や企業はAIの倫理に関するガイドラインの整備を急速に行っています。日本でも、2018年5月、内閣府において「人間中心のAI社会原則検討会議」が設置され、「人間中心のAI社会原則」(<https://www8.cao.go.jp/cstp/aigensoku.pdf>)が定められました。

ジュニア会員の皆さんは、AIの仕組みやプログラミングを勉強すると同時に、是非、AI倫理についても勉強してください。AIは、私たちの生活に無くてはならない存在になると思います。皆さんには、単なる精度や快適さを求めたAI開発ではなく、いかに安全で人と共生可能であるかを考慮したAI開発に携わって頂ければと思います。

(日本女子大学 小川賀代)

