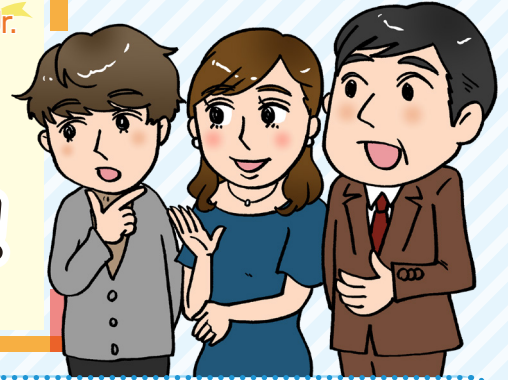


連載

Jr.

先生、質問です!



みなさんの身近にも急速に増えてきた人工知能。ジュニア会員からの素朴な疑問にさまざまな角度からお答えいただきました。



人工知能が悪い方に発達して私たちの暮らしを脅かす可能性はないのでしょうか。

匿名希望
ジュニア会員

Q

結論からいうと、人工知能が自ら人間を脅かすことはできません。なぜなら、現在の人工知能技術では、人間が機械にさせたいことを示さなくてはならないからです。これは童話に登場する「オズの魔法使い」にもたとえられます。この物語では、魔法使いは背後で人に動かされていました。同様に、人間が目標を与えてやらなければ、人工知能システムは何もできません。

けれども、問題がないわけではありません。1つは、誰かが、あなたにとって悪いことを人工知能にさせようとする問題があります。たとえば、あなたから泥棒をするように命令したりといったことです。法律を作ったりすることも対策の1つですが、法律があっても人間の泥棒がいるように防ぎきれません。そこで、人間の泥棒に対して警備員がいるように、AI泥棒に対してはAI警備員が活躍することになるでしょう。

もう1つは、人間の指示の仕方が良くない問題です。たとえば、そこにあるカバンを取ってくるように人に頼んだとしましょう。このとき、頼まれた人が、頼んだ人が思っていたのとは違うカバンを取ってきてしまうことがあります。同様に、人工知能も、頼んだ人が思っていたとおりのことをしないことがあります。この問題に対しては、曖昧な指示に対しては確認をとるなど、頼む人が意思を的確に伝達できるように人工知能を改良していくことになるでしょう。

将来あなたも、こうした人工知能を作ってゆくことに参加してみてもはどうでしょうか？



神鷹敏弘
産業技術総合研究所

A

A



松原 仁

[正会員]

東京大学次世代知能科学
研究センター

いまの人工知能はデータから学習する機械学習が主流になっています。人工知能を教育するのは人間の子どもを教育するのに似ています。人間の子どもが親や先生や本などから学習して育っていくように、人工知能はデータを人間が与えることによって学習して育っていきます。いい子になるように育てようとしても環境などの影響で悪い人になってしまうことが残念ながらときどきあるように、悪いデータを与えてしまうと人工知能も悪いことをするように学習してしまう可能性があります。我々人間はいわば人工知能の親に相当します。親として子どもである人工知能がいい子に育つようにどういうデータを与えればいいかをよく考えないといけません。また悪い人がAIを悪い子に育てようとするのを防がないといけません。人間にとって教育がとても重要であるように、人工知能にとってもとても重要なのです。

その可能性はあります。この懸念は人工知能に限ったことではなく、科学技術が持つ宿命です。技術を良いことに使うのも、悪用するのも使う人次第、というのが1つの考え方ですが、人工知能のような高い能力を持つ技術は、人に大きな恩恵をもたらす一方、悪用されたときの影響も甚大なものとなります。人工知能を使って本物にしか見えない嘘の画像を生成し、誤った情報（デマ）を拡散させることで、人々を混乱させるフェイク画像問題などが顕著な例です。また、人工知能は大規模で複雑なシステムとなりますので、人が入念に設計しても想定外の動作をしてしまう可能性もあります。そこで、人工知能を開発する際のガイドラインの策定が欧州や日本においても進められています。重要なのが、人工知能を人が確実に制御できることと、人工知能が医療診断をする場合などにおいて、どのような理由で診断したのかを、人が理解できるようにすることです。しかし、人工知能にまったく問題がなくても、人工知能が学習するためのデータに偏りがあれば、人工知能も偏った判断をしてしまう課題も指摘されています。

そして、人類にとっての人工知能の最悪の用途が兵器への転用です。これを国際的に禁止しようという議論も盛んに行われていますが、残念ながら世界的合意には至っていないのが現状です。まさに今、人類の賢さが試されているのです。

A



栗原 聡

[正会員]

慶應義塾大学理工学部

「先生、質問です！」
への質問はこちら



<https://www.ipsj.or.jp/magazine/sensei-q.html>