

生成AIという魔物の登場



月尾 嘉男

人間は古代ギリシャの時代から自動人形を製作するなど自身を神々の地位に接近させようと挑戦してきた。しかし刃物はメスにもドスにもなるという名言のように、技術には正負の両面がある。AIが浸透する情報社会を健全に維持していく制度を構築することが現代の重要かつ深刻な課題になっている。

2022年3月にイギリスのウォレス国防大臣が相手はウクライナのシュミハリ首相の偽物と気付かないまま15分間のテレビ会議をしたという事件が発生した。

さらに同年6月にはベルリンのギファイ市長がキウのクリチコ市長とテレビ会議をしたが、映像のクリチコ市長が偽物と気付いたのは何分か対談が経過してからであった。

このような事態を75年前に予想した天才が存在する。ケンブリッジ大学を優秀な成績で1934年に卒業したA・チューリングは第二次世界大戦中に敵国ドイツの暗号を解読するためにイギリス政府がロンドンの郊外に開設した組織で仕事をし、難攻不落とされて

いたドイツのエニグマという暗号装置の解析に貢献した学者である。

このチューリングが戦後の1950年にチューリングテストという概念を発表した。密閉された部屋に人間とコンピュータが存在し、外部から通信回線を経由して文章で質問すると、どちらかが返信してくる。その内容が人間からか、コンピュータからか判断できなければコンピュータは人間と同等の知能を保有しているとする理論である。

まだ登場したばかりのコンピュータは音声認識も音声合成もできなかった時代であるし、人工知能の研究が本格開始されたのは1956年に何人かのパイオニアが開催したダートマス会議からであるから、

1950年にはチューリングテストは思考実験でしかなかった。これが実体のあるシステムとして登場したのが生成AIである。

今3月、アメリカのカリフォルニア大学の研究者達が最新の生成AI「GPT4.5」を使用してチューリングテストを実施した。約280名の学生などが人工知能か人間かは事前に通知されないままキーボードから質問を送信し、その返答が二者のどちらであるかを判断する実験である。結果は73%の人々が見破れなかった。

人間は古代ギリシャの時代から自動人形を製作すること、すなわち自分の分身を創造して自身を神々の地位に接近させようと挑戦してきた。とりわけ時計の技

術が急速に発達した18世紀のヨーロッパではさまざまな機械仕掛の自動人形が製作され、日本でも江戸時代には木製の機巧人形が発明され、一部は現在まで保存されている。

これらは技術の進歩の一部であるが、刃物はメスにもドスにもなるという名言のように、技術には正負の両面がある。

トランプ大統領の一期目にオバマ前大統領がトランプ大統領を中傷する偽造動画が出回り、バイデン大統領が当選した2020年の選挙ではバイデン候補が幼児愛や同性愛の性癖があるという合成写真が大量に出回った。

これらは技術が稚拙で簡単に偽物だと見破ることができる水準であったが、生成AIは桁違いの精巧な文章や画像を生成することができ、チューリングテストは問題なく突破する能力を内蔵している。この魔物のようにもなる技術が浸透する情報社会を健全に維持していく制度を構築していくことが21世紀前半の重要かつ深刻な課題となる。

つきお・よしお 1942年生まれ。東京大学工学部卒業、東京大学教授、総務省総務審議官などを経て、現在は東京大学名誉教授。