

フランケンシュタイン・コンプレックスと AI をめぐる言説

久木田水生

北海道哲学会 2021 年 7 月 25 日

名古屋大学情報学研究科

人工知能をめぐる喧噪

2012 年、物体認識コンテスト ILSVRC において Google の開発した、「深層学習」という手法を用いた画像認識システムが従来を大きく上回る精度を記録して優勝した。

このシステムは「ディープ・ニューラル・ネットワーク」というアーキテクチャーをベースに、従来とは異なる「教師なし学習」によって、与えられた膨大なデータから自動的に対象を識別するための特徴を抽出する。

深層学習の成功を起爆剤に、2010 年半ばごろから機械学習が急激に発展し、様々なタスク、分野に応用されるようになった。そしてこのころからメディアには「AI」や「人工知能」という言葉が溢れかえるようになり、「第3次 AI ブーム」と呼ばれるようになった。

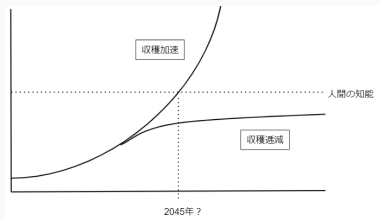
人間以上の知能を持つか？

哲学者のニック・ボストロムは、人間をはるかに超えた知性を持つ人工知能が、人間によって制御できなくなる可能性を詳細に考察している。

書影は <https://nikkeibook.nikkeibp.co.jp/item-detail/35707> より。



人間以上の知能を持つか？



未来学者のレイ・カーツワイルはこれまでのコンピュータ技術などの発展の指数関数的なペース（cf. 「ムーアの法則」：集積回路上のトランジスタ数は18か月で2倍になる）から、2045年には人工知能が人類の知能を超える知能を持つようになるると予言した。

多くの事柄について、発展や成長には限界がある。しかしカーツワイルのような論者は、テクノロジーの発展は更なるテクノロジーの発展をもたらすので、成長の勢いは衰えず、「**収穫加速**」の軌跡を描くと考える。

イーロン・マスク、スティーヴン・ホーキング、ビル・ゲイツといった著名人たちが、「AIは人類の生存にとって脅威になりうる」というAI脅威論を広めた。ボストロムが描く「超知能」の不穏なビジョンも大きな影響力を持っていた。

人工知能によって多くの仕事が奪われる可能性も指摘され、それによって生み出される失業に対処するための政策の必要性が主張された。Cf. 井上智洋、『人工知能と経済の未来——2030年雇用大崩壊』；エリック・ブリニョルフソン、アンドリュー・マカフィー、『ザ・セカンド・マシン・エイジ』；ジェリー・カプラン、『人間さまお断り』。また「AIに仕事を奪われないために人間はどんな能力を身に着け、伸ばすべきか」といったことも論じられた。Cf. 新井紀子、『AIに負けない子どもを育てる』。

フランケンシュタイン・コンプレックス の系譜

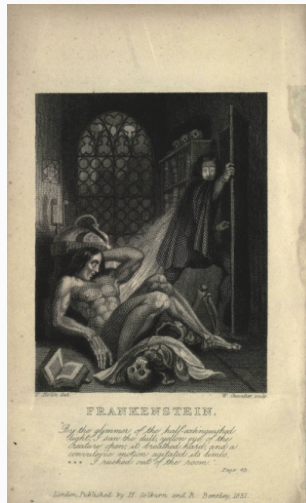
マーク・クーケルバークの『AI の倫理学』によれば、**有用であるはずの技術が制御できないものとなり、それを使うものに災厄をもたらすというナラティブ**は、西洋の精神文化において脈々と語り継がれてきたおり、AI 脅威論はその現代版である。

またシンギュラリティ論やトランスヒューマニズムにはユダヤ・キリスト教的な終末思想、プラトン主義的な魂についての思想が現れている。

現代の AI をめぐる言説を伝統的なナラティブの系譜に位置づけることは、極端な言説をより適切に理解することに役立つ、とクーケルバークは述べる。

フランケンシュタイン・コンプレックス

SF作家のアイザック・アシモフは、新しいテクノロジーに対する根拠のない恐怖を「**フランケンシュタイン・コンプレックス**」と呼んだ。



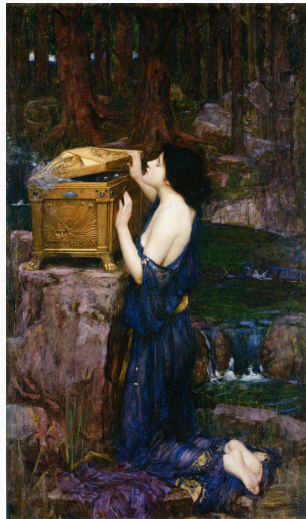
フランケンシュタイン・コンプレックス的ナラティブの典型例？

- 魔法使いの弟子（原型は2世紀のギリシア、サモサタのルキアノスによる「虚偽を愛する者」にまでさかのぼる）
- ゴーレム（ユダヤ教の伝説）
- メアリー・シェリー、『フランケンシュタイン』（1818）
- カレル・チャペック、『ロッサムの万能ロボット（R. U. R.）』（1920）



災厄をもたらすロボットの原型としてのパンドラ

哲学者の松浦和也は「呪われた孤児としての現存せざるロボット」（『日本ロボット学会誌』、38(9)、857-859、2020）においてゼウスの命令によりヘーパイストスが泥から作った女神そっくりの女性、パンドラがロボットのナラティブの原型であり、そしてその時から**ロボットは人間に災いをもたらす「呪われた」存在であった**、と論じる。



ロボットの存在の多義性

パンドラをロボットの原型として挙げるなら『ギルガメシュ叙事詩』のエンキドゥもロボットのナラティブの一つの原型と言えなくもない。エンキドゥはギルガメシュと戦うために神によって創造された獣人である。しかしエンキドゥは戦いの後、無二の親友、忠実な従者となってギルガメシュを助ける。

これはロボットの存在が人間の「バディ」、「サイドキック」として登場するストーリーの最古の例である。



フランケンシュタイン・コンプレックスを再考する

『フランケンシュタイン』をどう読むか

『フランケンシュタイン』は単に科学技術の産物が暴走する脅威を描いただけの話ではない。

そこで描かれているのはヴィクター・フランケンシュタインが自分勝手に生み出した生命に対する責任を放棄し、そして責任を引き受けることを拒否し続けた末の悲劇であり、**むしろ怪物が被害者**である。

少なくとも私はヴィクターよりも怪物に同情しながらあの小説を読んだ。

『フランケンシュタイン』をどう読むか

松浦和也の言葉を借りるならば、怪物は間違いなく「呪われた孤児」だった。

しかしその呪いは創造主のヴィクター・フランケンシュタインが招いたものに他ならない。

だとすれば**フランケンシュタイン・コンプレックス**とは、**アシモフが考えたような理由のない恐怖ではない。**

『R. U. R.』をどう読むか？

『フランケンシュタイン』の怪物が結局、ヴィクターを直接殺すに至らなかったのに対して、『R. U. R.』のロボットたちは人類を絶滅に追いやる。

そのため『R. U. R.』は、ロボットの存在が人間に災いをもたらすナラティブの典型とされる。

しかし『R. U. R.』の中心的なテーマは決してテクノロジーの脅威ではない。

『R. U. R.』で扱われているテーマ

- 生命を人為的に操作して利用することの倫理
- 労働の価値
- 科学技術の副次的な影響を予想することの難しさ
- 民生技術の軍事転用
- 産業資本主義の陥穽
- 自己犠牲的な愛の尊さ
- 生き延びようとする意志の美しさ
- 生命の強靱さ

「作者が「ロボット」を書く際、人間精神の驚異のようなものを頭に置いていたとしても、それは技術ではなく、科学だったのである。だから機械が人間に取って代わることができるのか、ロボットの歯車の中に生命、愛、あるいは反抗心といったものが目覚めるかもしれぬ、などという思念が生まれたのはお前のせいだから責任を持て、と言われれば、ぞっとして拒否する。作者はこういう暗い未来像を、機械に対する許せない過重評価でなければ、生命に対する由々しい侮辱とみなすだろう。」

チャペック「「ロボット」について」、『R. U. R.』、p. 273。

『フランケンシュタイン』と『R. U. R.』

これらはどちらも**科学技術を駆使する人々によって抑圧された弱者、辺境に追いやられた他者を書いた物語**だと言える。

『フランケンシュタイン』では弱者は弱者として排斥されるが、『R. U. R.』では弱者が団結して強者を駆逐する。

旧人類と新人類



リチャード・マシスン『アイ・アム・レジェンド』を下敷きにした藤子・F・不二雄の短編「流血鬼」では、「怪物＝新人類」によって旧世代の人類が駆逐される。これは『R. U. R.』の重要なテーマでもある。

スティーブン・スピルバーグ監督の『ジュラシックワールド——炎の王国』（2018；英語の副題は“Fallen Kingdom”）では、クローン技術でよみがえった恐竜が、人間が君臨してきた世界の秩序を転覆させる。

「ジュラシック」シリーズのこれより前の作品では、主に**恐怖の対象だった恐竜たちがこの作品では、共感すべき対象、救済されるべき対象となっている**ことは注目に値する。

ちなみに強者の都合で利用された弱者が、救済されないけど救済された気になって終わるのが同じくスピルバーグ監督の『A. I.』（2001）である。

強者の都合で利用された弱者（クローン人間）が、**ひたすら救済されないままに終わる**のがカズオ・イシグロの『わたしを離さないで』（2005）。

カズオ・イシグロの『クララとお日さま』（2021）も客観的にみれば同様に救いが無いのだが、弱者（ロボット）の方が都合よく利用されている自覚がなく、排除されても苦しみがないので、読んでいてそれほど悲しくない。しかし気持ちの晴れない読後感である。

テッド・チャンの「ソフトウェアオブジェクトのライフサイクル」(2019)には、人間と同じように学習して成長していく「ディジェント」と呼ばれる人工生命たちが登場する。ディジェントは開発当初ペットのようなものとして想定されていたが、オーナーたちは長く付き合う内に徐々に彼らを自分の子供のように考えるようになる。

ディジェントはメーカーの都合で廃棄されそうになるが、最後は自立して生きていくことになることを示唆して終わる。

「フランケンシュタイン・コンプレックス」とは何か

それは異質な他者と向き合った時に、生じる様々な感情で、それは時代や社会環境、個人的な資質や過去の経験によってかなり異なったものでありうる。

「人間はロボットに対して不合理な恐怖を抱くものだ」という捉え方は、一面の真実であるが、しかしそれがすべてではない。

ロボットの存在をめぐる多様なナラティブを掘り起こすことは、ロボットや AI とどのように付き合うかを考えるうえで有用である。

本論文は、JST CREST JPMJCR20D2、JSPS 科研費 JP19H00518、科研費 JP20H01175 の支援を受けた研究の成果の一部である。