

〈書評〉

『AI時代の「自律性」—未来の礎となる概念を再構築する』

河島茂生編著，勁草書房，2019年10月，209p.

ISBN 978-4-326-00047-0 本体 3,500円＋税

本書は、AI時代の人間の尊厳を基礎から考えるために「自律性」(autonomy)という概念に着目しこれを精査することによって、人間の持つ自律性とAIの自律性が混然一体となって論じられている現状に一石を投ずる試みである。

機械が自律性を備えて人間の領域に侵入し、権利をさえ主張する可能性を持つ一方、人間が大きな社会システムの歯車となり、その自律性が損なわれるという事態がAI時代の人間の尊厳の危機として喧伝される。このように、自律性こそが問題の鍵となる概念であるにも関わらず、これを体系的に整理した研究がないという。

本書は三部構成、全六章からなる。序章「なぜ、いま自律性を問わなければならないか」では、まず人間の持つ自律性の特質を明らかにする手がかりとして、生物の持つ自律性を検討する。生物は、「自分で自分を作る」ことで内部と外部を形成する。外部の環境は生物の動きを完全に決定することはできない。この生物が持つ根本的な自律性、ラディカル・オートノミーを根底に、人間の自律性の特質に迫ろうとする。

人間の場合はその上に理性の働きがあり、自由意志としてその自律性が顕著に現れる。しかし人間が自由意志を持つことの強調は、人々に過度の心理的負担を強いる際限なき自己責任論にも転化しうる。また無意識の働きと意識的な理性との関係という問題や、

自律的であるつもりで実は無自覚に他者の方向付けに従っている場合もあることが指摘される。

他方、AIやロボットが人間の介入なしに自力で行動できる場合に、このような類の自律性を持つ人工物を道徳的行為者ないし「電子的人間」として権利や義務を認める議論が主に欧州で起こっていることに言及するとともに、機械の場合には見かけ上の自律であり、本質は「自動」(automated)であるとの可能性も示唆されている。

その上で、本書の理論的支柱であるネオ・サイバネティクスに議論が進む。サイバネティクスは、生物をモデルに体温を一定に保つなどのホメオスタシス(恒常性)の機能を機械に持たせ、エアコンの温度を一定に保つようなフィードバック制御の理論として知られているが、ネオ・サイバネティクスは、生物の必要かつ十分な条件として、自分で自分(auto)を制作(poiesis)する内閉したシステム、つまりオートポイエーシス(自己制作)にあると考えることで、生物と機械との線引きをした理論である。

そしてオートポイエーシスという特徴は生物及び人間の個体だけでなく、社会にも見られるものであり、デジタル環境にあつてはAIを含むコンピュータ技術がそれを介在している状況、すなわち「人間＝機械」複合系(西垣通)といえるシステムが生成していることが指摘される。

第Ⅰ部「自律性とはいったいなにか」の第一章「生命の自律性と機械の自律性」(西田洋平)では、生物つまりオートポイエーシスのシステムの「観察者」概念を立てたフェルスターの議論を通じて、機械の見かけ上の自律性と、生物の持つ「制御の制御による閉鎖系が持つ自律性」(ラディカル・オートノミー)とが根本的に異なる点が示される。続く第二章「生きられた意味と価値の自己形成と自律性の偶然」(原島大輔)では、グレーザーズフェルドのラディカル構成主義やヴァレラの自律系の議論の要点を追いながら、自律性の概念をとりまく議論の系譜を整理した上で、記号接地(メイクセンス)及び自己形成の知能としての知識と行為の合一(リアライズ)に論が進む。ここで意味や価値と呼ばれているものを論じる視点は、単なる生物的な生存レベルのそれではなく、生きる意味や美的な価値にも通じる可能性を含んでいる点が、今後のネオ・サイバネティクスの展開の鍵として注目される。

第Ⅱ部「情報技術と心の自律性」の第三章「ロボットの自律性概念」(谷口忠大)では一転して、ロボティクス(ロボット工学)の観点から、記号創発システムを突き詰めた自律ロボットが人間の自律性(ラディカル・オートノミー)に原理的に限りなく近づき得る予想が示され、それまでの議論に対し挑戦的な議論が展開される。続く第四章の「擬自律性はいかに生じるか」(棕本輔)では、視覚情報(画像データ)のパターン処理を題材に、観察者である人間がAIを擬人化する解釈を加えることによって、あたかも意味を持つように見えるという点を指摘し、「機械の自律性=イリュージョン」(幻影)という論を展開している。

第Ⅲ部「AI社会に組み込まれる個人」の第五章「他者と依存し合いながら生起する社会的自律性」(ドミニク・チェン)では、独我性と他者性の問題を取り上げ、他者や環境との共依存的な生起、縁起的な社会的自律性という観点から、強すぎる近代的な個人観を克服して「力と衝撃」(プレローマ)ではなく「生命的な共進化」(クレアトゥラ)の社会モデルを模索する方向性が示される。第六章「組織構成員の自律的思考とAIをめぐる実証的分析」(辻本篤)では、他者と構成する組織において、構成員が組織的な目的の下で拘束されつつも自律的である「創発的思考」を生む条件について考察し、基礎情報学の「タイプⅢアプリケーション」(有機的機械)の議論を引き合いに出し、AIが構成員の問題発見及び自律的思考を支援する可能性とともに意思決定におけるAIへの依存の危険性が顧慮されている。

本書の意義は、ネオ・サイバネティクスを中心とした自律性に関する学問的な議論を丁寧に進めながら、AI時代の人間の尊厳を論じる前提として議論の土台を提供しようとしたことにある。少なくともAIと人間の自律性を論じる際に欠かせない論点を示し得た点でその試みは成功していると言える。

本書の課題を挙げるとすれば、「むすびにかえて」で编者自身が述べてもいるように、生命システムと心的システムとの関係について生物の自律性を人の心も有しているという議論にとどまっている点がある。人とその他の生物は単なる延長線上にあるのか。生物と人間を連続的にとらえる見方こそがオートポイエーシスの特徴であるとともに、その違いも今後精緻化すべき課題であろう。

(評者：竹之内 禎)