

序章 なぜ、いま自律性を問わなければならないか

河島茂生

一 はじめに

未来を作るには、どうすればよいだろう。想像力、創造力が必要だ。人を巻き込むコミュニケーション力や実践力、そうした力も求められる。しかしその前に、なによりもまず基礎を打ち固めなければならない。基礎がなければ、社会がぐらつく。その基礎の中核に「自律性」(autonomy)がある。

自律性について考えてみたことがあるだろうか。ほとんどの人は、その言葉を耳にしたことがあるにせよ、深く考えてみたことがないかもしれない。それほど重要性を感じないのかもしれない。重要なのはA I (artificial intelligence) によって「仕事がなくなくなるのか」「国の経済力が維持されるのか」「便利になるのか」であって、自律性といった小難しいことはなにに役立つかわからない。時間をかけて議論する必要はない。重要なことではない。このように思っている人もいるだろう。

自律性という言葉は抽象的で難しい。漠然としたイメージはすぐ思い浮かぶけれども、考えれば考えるほど、よくわからなくなる概念である。しかし、本書を読んでもらえればわかるように、これからの社会の大きな方向性を左右する、重要な概念である。

A I、ロボット、I O T (internet of things) などのコンピュータ技術が普及した現在、従来にもまして自律

性という語がさまざまな場面で使われるようになってきている。高度なコンピュータ技術の特性を表す語としても自律という言葉が使われているからだ。しかし、そのさまざまな意味合いや使われ方は十分に検討されないままにあり、曖昧なまま混乱した状態にあるといつてよい。

たとえば、総務省情報通信政策研究所が主催するAIネットワーク社会推進会議の報告書(2017, 2018)をみると、「自律性を有するAI」という言葉と「個人の自律」という言葉が並んで使われている。このAIの自律性と個人の自律性は同列に論じるべきなのだろうか。あるいは両者には大きな差があるのだろうか。もし両者に違いがないなら、ビッグデータ型AI時代にあつて「人間の尊厳と個人の自律を尊重する」ことはできるのだろうか。

AIネットワーク社会推進会議の報告書のほか、いろいろところで人間のためのAIが唱えられている。たとえば、内閣府は「人間中心のAI社会原則検討会議」を主催している。また、アメリカの非営利組織FLI(Future of Life Institute)が作ったアシロマ原則には、「AIシステムは、人間の尊厳、権利、自由、そして文化的多様性に適合するように設計され、運用されるべきである」という原則が入っている(Future of Life Institute, 2017)。人間の尊厳をもたらすものが自律性にあるとするなら、それはどのようなものか、AIのような機械がもっているものなのか、真剣に考えなければならぬ。

本書のねらいは、自律性概念を整理し体系づけることで、そうした混乱した状態を解消し、AI時代の人々の対話の基盤づくりに資することである。自律性とはいったいなんだろうか。人間は自律的なのだろうか、それとも他律的なのだろうか。機械は果たして自律性をもつか。そうだとしたら、その場合の自律性とはいったいかなる意味か。本書は、そうしたテーマに正面から取り組む。

端的にいえば、生物がその内部でもっている自律性は、「ラディカル・オートノミー」(radical autonomy)と

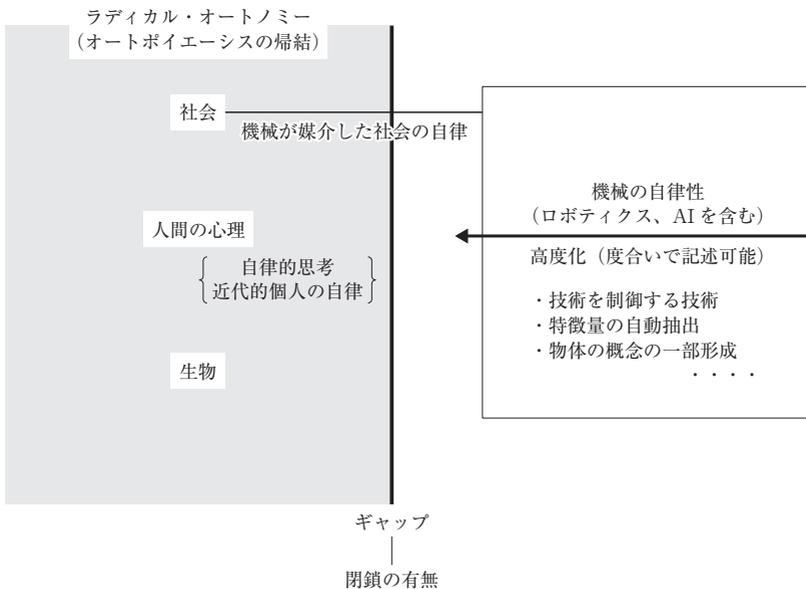
でもいべき生物学的自律性である。生物は自分で自分を作る。その自己制作する閉鎖したプロセスのなかで内部にメカニズムの起点を置くラディカルなオートノミーが形成される。生物は、自分で自分を作りながら内部と外部を形成し、その内側のメカニズムから環境を認知していく。環境が完全に生物の動きを決定するわけではない。そのように見える場面があるとしたら、環境に対してそのように対応するように生物の内部のメカニズムが形成されているからである。生物は根本的かつ徹底的な自律性を有している。このラディカルなオートノミーは、後に述べるように内部のメカニズムが閉鎖しているか否かによるため、「あり／なし」の2値しかない。近代的個人の自律性はその生物学的自律性を基盤として存立しており、ラディカル・オートノミーの一種と位置づけられる。一方、機械の自律性は、生物が無限定の環境下で適応しながら活動していることを参照しながらエンジニアが作っているものであり、近年、自律性の度合いが高まっている。生物のもつラディカル・オートノミーの段階には達しておらず、「あり／なし」ではなく程度の問題である。ラディカル・オートノミーと機械の自律性の間には壁がある。その壁を超えるには、機械が自己制作しなければならず、大きな跳躍が必要である（図序1）。

二 さまざまな自律性

二.1 近代的個人の自律性

自律という語は、もともとギリシアやローマで「自己統治」(自己auto + 規範・統御nomos)の意味で使われ、

(1) 本書でいうAIは、機械学習を含む最先端のコンピュータ・ソフトウェアで、コンピュータ・システムの構成要素を指す。人間の知能を圧倒的な差で凌駕する人工超知能(artificial superintelligence)は除く。



図序-1 本書で依拠する大まかな概念図

政治的な自治を指し示す言葉であった。また、一六世紀の宗教改革のころは信教の自由を体現する言葉であった。個人の尊厳を根拠づける意味で使われはじめたのは近代以降である。それはイマヌエル・カントの道徳に関する哲学によるところが大きい。カントは、理性の働きによってみずから道徳律を定める自律性が人間に本来備わっていると述べ、人間の有する尊厳の哲学的基盤を築き上げた。すでにルネ・デカルトは「我思う、ゆえに我あり」と述べ、人間の理性を基点として物事を考えることを提案しており、ブレイズ・パスカルは、人間は「考える葦」であって思考することがわれわれの尊厳につながることを述べていた。カントはそうした立論を深めている。カントは、理性によってみずから考えた道徳法則に沿った生き方を行うことを自律と呼んだ。カントは、この理性による自律性の特徴を浮かび上がらせるために対義語として他律 (heteronomy) という語を造った。この場合の他律とは、理性ではなく生得的

欲求や他者の指示、自然法則に従うことを指している。もちろん、偉い人から言われたからそれに従った行動をとることは他律である。しかし、それだけでなく自分の内にある感情や欲望のまま行為することも他律と位置づけた。理性による行動ではないからである。したがってカントのいう自律性は、かなり範囲が限定された、厳しい条件をクリアしたものである。このように自律性の語の意味は、政治的な自治から人間の理性に関係したものと、時代とともに変わってきている。いまは、また自律の語の意味が変わる時期なのかもしれない。

いうまでもなく自律性に関することは、たとえその言葉が明示的に使われなくとも日常生活でしばしば話題に上っている。就職する会社や通う学校などを自由に選べるような感覚を抱く一方で、逆に他の人からの指示によって望む選択ができなかった場合には自由が失われたと感じる。前者は自律性があり、後者は自律性が奪われていると見なすことができるだろう。後者の場合は、不満に感じストレスを覚えることが多い。このような場合の自律性は、人間の尊厳を輪郭づける「自由意志」(free will)であり、他から束縛されずにみずからの意思で決定することを指している。

よく知られているとおり自由意志の尊重は、近代社会の選択プロセスのなかで大きな柱になってきた。居住地や学校、職業、婚姻、商品の選択にあたって個人の意志が尊重されるのはその最たる例である。何を選択するべきかを自分で思案し、自分で決断し、その結果も自分で引き受ける。近代社会では、結婚するかどうか、結婚するとしたらどの相手と結婚するかどうか、基本的には個人の意志に委ねられる。強制的に自分が知らない相手と結婚させられることもない。しかし、よく知られているように近代より前はそうではなく、家同士が決めた相手と結婚することが多かった。近代社会になって個人の意志が重んじられるようになった。

またマイケル・マーモットなどのホワイトホール研究が示したことは、自分の自己決定権が高いと考えている人のほうが健康状態が良好であるということだった。職階が上の人は、下の人に比べて死亡リスクが低く、それ

には自分の裁量で決められる範囲がかかわっていた。自分の裁量で物事を決めている感覚がある人は、たとえそれが勝手な思い込みであっても死亡率が低い。逆に窓際族は、窓際で外の風景をぼんやりとみつめていて気楽そうにみえるけれども、本人が強制されていると感じているかぎり、精神的には望ましい状態ではなく健康を害しやすいのである。幸福感と自己決定との強い関連性も指摘されている（西村・八木、2013）。健康や人間関係に次いで自分で進学先や就職先を決めている度合いが幸福感には強く関わっている。自分で人生の大きな選択をしている人は幸福度が高い。年収や所得よりも、自己決定が幸福感には重要なのである。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」で触れるとおり、主に二〇世紀初頭に工場経営に活用された科学的管理法の限界も、ここにあった。科学的管理法によって、工場の労働者の動作が標準化され、時間管理も徹底された。その結果、生産性は劇的に向上した。しかし、従業員は機械のように扱われ不満が生まれることになった。われわれ人間にとっては、自由意志を発揮し自分でいろいろなことを決めているという感覚を抱くことが権利意識を守り健康や幸福感を増進させるうえでとても重要である。

けれども、自由意志の尊重は決してよいことばかりではない。ときには残酷な心理的負担を強いる。新生児の重い障害を知った親は、その子の命を絶つか継続させるかの選択に直面する。シーナ・アイエンガーによれば、アメリカでは親がみずから選びたがる傾向が如実に見て取れる（Youniss, 2010）。しかし親が自分で子どもの死の決断を行った場合、その親は後になって罪悪感や迷い、恨みを執拗に感じるようになった。逆にフランスでは親が特段の申し立てを行わないかぎり医師が延命治療の中止の決断を下すが、その場合は親はそれほど苦悩していなかった。すなわち、自由意志の尊重がかえって人を苦しめることがある。ほかの人に決めてもらい、あえて他律になることで精神的負荷を下げられる。これは、そのほかの場面でも見られる。学校や職業などは膨大な選択肢があり、最終的にはどれを選ぶかを自分で決断しなければならない。たとえ希望するものが得られな

くとも、自分の能力次第でそれを得られるイメージが作り出されている。個人は、さまざまな選択肢を比較検討し、自分の人生を歩む。その失敗は、生活保護受給者への非難にみられるように、しばしば個人に課せられる。勉強したくても病弱で思うようにいかなかったり、家の都合で高等教育を受けられなかったりしても、個人に責任があるかのように述べられることが少なくない。差別やいじめ、暴力により、人生に大きなダメージがあったとしても、自己責任のようにならざるを得ない。これはきわめて過酷なことである（河島・竹之内、2019）。いっそ人間は外的なものに規定されている他律的な存在であるといわれたほうが楽になるケースは少なくない。

加えて、自由意志それ自体にも疑問が再三投げかけられている。古くは、一九世紀末から二〇世紀はじめごろにかけて精神分析の創始者ジークムント・フロイトが心は無意識からも成り立っており、それが根底に横たわっていることを述べていた。自由な意志は、世界の隅々までを見通す宇宙の基点ではない。また、一九六〇年代以降に隆盛を誇ったポストモダンニズムの思想の根幹には理性的人間の死があった。一九六〇年代に哲学者ミシェル・フーコーは「人間は波打ちぎわの砂の表情のように消滅するであろう」と述べ、人間の消滅を予言したことはあまりにも有名である。この人間の消滅とは、いうまでもなく生物学的な人間の死ではなく、西洋の知の歴史のなかで近代になって現れてきた理性的人間なるものが終わりを迎えることを指している。

私たちは、「自分で自分のことを決めている」と思っていたとしても、意外にも他の人によって設計されているとおりには街を歩き、買い物をしていくだけかもしれない。商品の陳列方法を変えるだけで売上が変わる。臓器提供や年金制度加入をデフォルトにするだけで、つまり最初の設定を変えるだけで、臓器提供や年金制度への加入が促進される。もし拒否するなら余分な手間がかかるようにするだけで、臓器提供の意思表示をする人や年金に加入する人の比率が劇的に上がるのだ。われわれは、自分で動いているつもりでも実際には操られている。行

動経済学や神経科学、心理学などの知見により、個人の尊嚴の基盤となる個人の自律性（近代的自我）はいま問
い直されつつある。神経生理学者ベンジャミン・リベットらの実験は、よく参照される。指を動かそうと決意す
る前に脳には準備電位が既に発生している。脳のプロセスが先でその後遅れて意識が生じる。そうであるなら
ば、自分の意志よりも脳の動きのほうが行動を決めているのではないだろうか。また、神経学者マイケル・ガザ
ニガらの研究によれば、意識はたとえでつち上げであっても自分なりのストーリーを作り出し首尾一貫性をもた
らしている。本人にとっては嘘が嘘であることすら気づかない。意識は、自分の行動に説明がつくように作り話
を組み立て、それがこじつけであることにすら気づかない。このように意識が特権的な位置にいて正しく判断を
下しているといった見方には批判が相次いでいる。いわば自己決定の虚構性の知見が積み重ねられているのであ
る。幻想であるにもかかわらず、「自分で自分のことを決めている」という自己決定の意味での個人の自律性を
強調しすぎると、安易な自己責任論に結びつき、ひいては社会の結びつきを弱めることにもなりかねない。
近代的個人を支えた理性・精神の自律性はゆらいでいる。

二・二 機械の自律性

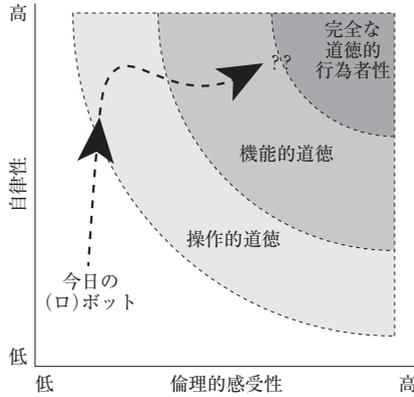
では、機械のほうはどうだろうか。

人間の柔軟な発想力とたゆまぬ努力により、道具は、次第に複雑になり自動化されてきた。斧やナイフは、人
間の力が動力源になり木を切ったり相手に攻撃したりすることに使われる。直接的に人間の動きが道具の動きと
なる。時代が経ち、やがて道具を制御する道具が作られた。たとえばモーターという道具は刃やドライバー、車
輪といった道具を回転させる。直接的に人間の力が加えられるのではなく、人間が操作した道具を介して別の道
具を動かすようになっていった。自動化が進んだといつてよい。自動化は、コンピュータ技術においても随所に

見られる。コンピュータを立ち上げると、自動で無線LANに接続しようとする。更新プログラムも自動でダウンロードされる。コンピュータ・プログラムが知らぬ間に他のコンピュータを攻撃していたり、セキュリティのソフトウェアが知らぬ間に攻撃を防いでいたりする。GPSをONにしていると、その場その場に応じて天気や紹介される店舗が変わってくる。サーチエンジンのウェブクローラーは、ウェブ空間を自動で巡回しながらたどり着いた先のウェブページのデータをコピーする。クラウド・コンピューティングにより、自動でデータが転送される。技術が技術を制御する仕組みが連鎖している。

このように自動化が高度になるにつれて、人間の自律性とは別に、機械の自律性が話題に上り盛んに論じられることが多くなってきた。AIやロボットでも同様である。第三次ブームを迎えたAIは、機械学習によって学習データを読み込み自動的に内部で特徴量(素性)を形成し、学習したとおりに新しい対象を分類する。その特色を自律性という言葉で表現する場合も多い。自律エージェントや自律分散システム、自律型兵器といった語もすでに使われている。自動運転車は、自律走行車とも呼ばれ、英語で「autonomous car」と表記されることが多い。電気・電子分野の規格の標準化などを行っている団体IEEEは、いわゆるAIの設計ガイドラインである「倫理的に調和したデザイン(Ethically Aligned Design)」を公開し、いろいろな国の人々を巻き込みながら議論を重ねているが、そこではAIは「自律的かつ知的なシステム(Autonomous and Intelligent Systems)」と表現され、自律性に力点が置かれた名称となっている(IEEE, 2019)。

第三章「ロボットの自律性概念」で述べるとおり、以前よりロボティクスでは自律性を念頭に置きながら開発が進められている。ロボットは、その場の環境に合わせて持続的に適応して活動することが求められるからである。欧州委員会のロボロー(Robolaw)・プロジェクトは、ロボティクスと法律との関係を検討したプロジェクトであるが、そのプロジェクトの成果である「ロボティクス規制のガイドライン(Guidelines on Regulating Ro-



図序-2 AMA 開発の 2 つの次元 (Wallach & Allen, 2009: 26=2019: 33)

「borics」(2014) では「人間の行動を担うことができる自律的マシン」としてロボットを定義づけ、ロボットは物理的ボディをもち自律的で人間に似ているものと位置づけた。ここでいう自律性とは、人間の介入なしに自力で行動することが可能な能力のことである。さらにこの文書では、ロボットの分類も提案されており、その基準となる 5 特徴の 1 特徴として自律性を挙げ、実験室の外の環境で人間が関わらずにタスクを実行できるレベルにより、完全自律 (full autonomy) / 半自律 (semi-autonomy) / 遠隔操作 (tele-operation) に分類できるとしている。たとえば、グーグル・カーは自律の枠に入り、外科手術ロボットの da Vinci は遠隔操作に分類される。このロボット・プロジェクトの考え方は特殊なものではなく、生活支援ロボットの安全性に関する国際規格である ISO13482 などでも自律性に言及するかたちでロボットが定義づけられており、自律性はロボットを論じるなかで大きな一要素になっている。

また、汎用 AI——特定の課題に特化した AI ではなく、さまざまな状況に合わせて多様な課題を解く適用範囲の広い AI——の開発をめぐる「自律性の付与が鍵」(栗原ほか, 2017) とされている⁽²⁾。加えて、いかにして AI に倫理的な側面を実装して道徳的な振る舞いをする AMAs (artificial moral agents) を設計するかを考

るマシン・エシックス (machine ethics) の領域でも、自律性が尺度として設定されている (図序2)。技術的人工物を道徳的行為者に含める立論を展開しているルチアーノ・フロリディやジェフ・サンダースも、行為者 (agent) となる一条件として自律性——外部との相互作用なしでも内部状態を変化させられること——を挙げた (Florida & Sanders, 2004)。道徳的行為者とは、道徳的観点からその行為が評価される者のことである。ジョン・サリンズも、ロボットが道徳的行為者となる条件として意図や責任のほかに自律性——機械が外部からの直接的なコントロールなしに動くこと——を挙げている (Sullins, 2006)。機械の自律性をどのように捉えていったらいいのか。機械の自律性はどこまでいくのか。つまり工学的に自動プログラミングがどこまでいくのか、構成論的アプローチによるロボット内部の概念形成はどこまでいくのかも検討する必要がある。

そして、ペッパーやアイボのようなソーシャル・ロボットもそれに接すると生きているかのような感覚——擬生命化 (擬人化含む) ——を抱くことがある。心理実験により、機械内部の仕組みにそれほど注意を向けない一般人の人々だけでなく機械内部の設計に気を配るコンピュータ技術者も、しばしば機械に礼儀正しく振る舞うことが知られている (Reeves & Nass, 1998 = 2001)。次の実験が有名だ。まずコンピュータが一通りの豆知識を被験者に提示し、その後で被験者が豆知識の印象をアンケート方式で回答するのだが、この回答を行う端末を2通りに

- (2) 汎用性と関わる「適用度」は、さまざまな事象のデータに対応可能な度合いを指すが、自律性とは別の概念である。というのも、たとえばタニは、ヤコブ・フォン・ユクスキュルによれば酪酸の匂い・温度感覚・毛の多寡といったごく少数の知覚標識をもとに環境を捉えているが、その様子は人間の目からすればひどく単純にみえる。けれどもタニにとっては十分な世界なのである。「動物主体は、もつとも簡単なものも、非常に複雑なものも、同じ完全さでその環境世界に適用している」(Uexküll & Kriszat, 1970 = 1973: 20)。すなわち、適用度は、人間の適用範囲からみてその度合いが図られるけれども、その度合いがたとえ低くとも生物学的自律性は存在するといえる。

分けた。豆知識を教えたコンピュータと同じ端末で豆知識の印象を聞く場合と、別の端末で印象を聞く場合である。その結果は驚くべきものであった。被験者は、同じ端末で評価した場合のほうが、別の端末で評価に比べて肯定的な回答を行っている。すなわち、我々が人間相手に示しているのと同様、豆知識を教えたコンピュータに対しては礼儀正しく接した。この傾向は、コンピュータ技術に長けた者であっても同じであった。

著名な哲学者ダニエル・デネットの区分でいえば、これは、AIやロボットに対して「物理的な構え」(physical stance)や「設計的な構え」(design stance)ではなく、「志向的な構え」(intentional stance)をとるといふことである。デネットは、事物の予測に関する姿勢を「物理的な構え」「設計的な構え」「志向的な構え」に3分類した(Dennett, 1996 = 1997)。物理的な構えとは、物理法則と物理的構造に着目する手法で、放り投げた石が描く放物線、水の沸騰の説明などに適用される。石が落ちるのは重力の法則があるからと考える見方である。設計的な構えは、設計された構造があり設計された通りに機能するだろうと想定する手法で、目覚まし時計などに活用される。朝六時に目覚まし時計が鳴るのは、指定した時間に時計が鳴るように予め設計されているからでありその時刻が朝六時と指定されたからである。志向的な構えとは、「対象の行為を理解するためにそれを主体として捉える」(Dennett, 1996 = 1997: 69)ことであり、その対象が信念や欲求をもつと考える姿勢である。この志向的な構えを目覚まし時計に適用すると、朝六時に目覚まし時計が鳴るのは、目覚まし時計がその時間に人を起こそうと思っているからであり、親切心や義務感からそうしているということになる。この志向的構えは擬人化にも深く関係している。設計的な構えをとっていても、ときに我々は無意識的に志向的な構えをとり、機械の振る舞いに意図を感じてしまう。

AI時代にあつては、たとえフィクションであっても社会制度上で機械に自律性を認めることが望ましいのかも考えなければならない(河島, 2016)。周知のように、欧州議会が「電子的人間」(electronic person)の議論を

は始めている。「もつとも高度な自律性を備えたロボットは、特定の権利や義務、与えた損害を償う電子的人間の地位をもつように制度設計することがありえる」(European Parliament Committee on Legal Affairs, 2016: 12)と⁽³⁾いう。その議論はどこまでの妥当性を有しているのか。日本でも、電子的人間というものを考える必要はあるのか。あるいは、ロボットに財産権を超えた保護を与える必要性はあるのか。AIやロボットはどのような意味で自律性を見出させる存在であるのか、それとも言語的表現としては「自動」がふさわしく自律という語を使うのは本質を覆い隠す・見誤らせることなのか。そうしたことも検討していかなければならない。

筆者らは、AI(AIを備えたロボットを含む)のリストを挙げ、愛着を感じたものはあるか否かを聞いた⁽⁴⁾(表序1)。モノへの愛着もあるため擬人化そのものの質問ではないけれども、機械に愛着を抱いている人が三割ほどいることが見て取れる。愛着のない人が約七割いるが、人型ロボットや掃除ロボット、ペット型ロボット、会話型AIでほぼ変わらずおよそ三割の人がそれらに愛着をもっているのだ。電子的人間肯定派にとっては、やや嬉しい数字かもしれない。というのも、そういった人々は、技術が高度化すれば、この数字はきつと上がるとい

(3) 欧州議会における電子人間の提言に対する批判的検討としては、河島茂生(2019)「AIネットワーク状況下における集合的責任」社会情報学八巻一号、一―四頁がある。

(4) 筆者と河井大介との共同調査であり、中央調査社のマスターサンプルに対する郵送調査(督促はがき一回)を行った。このマスターサンプルは電子住宅地図を利用した層化三段無作為抽出法に基づいて依頼を受けた個人が登録されているものであり、調査会社が保有する調査パネルのなかでも偏りが小さく、代表性が高いと考えられる。調査対象は、マスターサンプルのうち日本全国に居住する者で、二〇歳以上五九歳以下の男女である。調査対象者は性別と年齢層(一〇歳刻み)で母集団比例の割付を行った上で、予測回収率をもとに重みづけを行い、地域(七地域)と都市規模(三段階)で層化無作為抽出された一三〇〇人で、回答者は六二三人であった。発送・返送期間は、二〇一九年一月〜同年二月である。

表序-1 AI (AI を備えたロボットを含む) への愛着

		愛着	
		あり	なし
1)	人型ロボット (Pepper (ペッパー) など)	29.1%	70.3%
2)	掃除ロボット (ルンバなど)	27.3%	71.9%
3)	ペット型ロボット (aibo (アイボ) など)	30.8%	68.5%
4)	会話型 AI (Siri (シリ)、りんな、Alexa (アレクサ) など)	26.6%	72.7%

(総数=623人、無回答=4人 (0.6%))

うに違いないからだ。

ただし、もし本当の人間であるかのようにロボットを扱うのであれば、OSやアプリケーションのインストールにあたっては、ロボット自身に許可をとらなければならなくなってしまう。あるいは、勝手にデータを読み込ませることも倫理的に認められない行為になるだろう。それは、個人の内面を無断で操作する行為であり、不可侵な領域の侵犯に当たると見なされてもおかしくはない。現在のところ、ロボットが人間のような自律性を備えていないからこそ、無許可でこうした行為をしても倫理的問題に問われないのである。

二二三 自律性の体系化の必要性

これまで長々と見てきたように現代社会では、人間の尊厳の基盤とされる自律性のゆらぎと同時に機械の自律性の高度化が並行して起きている。

高度に発達しつつあるコンピュータ技術に対する認識の違いは、将来的に看過できない問題となる可能性がある。楽観的に捉えている人もいるが、不安を抱いている人も決して少なくない。これはコンピュータ技術に対する認識枠組み(フレイミング)の違いを示唆しているが、その核心に潜んでいるのが人間や機械の自律性をどう捉えるかという問題である。

真剣に考えられないまま、人間の本質をなす自律性と機械の自律性を同

じに考えると、たとえば「人間が機械のように働かされる」「病気になった人間が故障した機械のように捨てられる」といったように、さまざまな面で深刻な社会的問題を引き起こす懸念さえある。あるいは、人間の自律性がゆらいでいるからといって、もしくは強調しすぎると自己責任論につながるからといって、個人の尊厳と強く結びついた自律性の概念を放棄してもよいのだろうか。個人の自律性の放棄は、個人の選択の自由の放棄とイコールで考えてよいのだろうか。ホワイトホール研究にあつたように、人は自分に決定の裁量権がないと感じると、その健康状態が悪化してしまうおそれすらある。さらに現代社会の制度は、個人の自己決定が幻想であっても、個人には自律性があるという前提で構築されている。病気の治療方法、宗教への入信、結婚の有無、居住場所、会社との契約、言論などは、個人の自己決定が最大限尊重される。そうであれば、個人の自己決定とは別の意味で自律性概念を構築したほうがよいのではないだろうか。自律性は権利や責任、倫理的配慮を考える際の基礎概念であるため、自律性概念をなおざりにしては、今後の社会制度の大きな方向性が見えてこない。先に述べたように近代以降の人類は、ほかの生物や物体との違いを精神の自律性に見出だしてきた。したがって、いま、もし機械にも自律性があるとしたら、人間とは何かという本質的な問いが生じざるをえない。

現在は、いろいろな自律性が整理・体系化されないまま混在し、社会的課題設定が妨げられている状態である。このままでは多様な人々の間の対話が誤解のまま進み、社会的な意思決定が適切に行われぬ事態を招きかねないのではないだろうか。本書では、これからのデジタル社会の根幹をなす自律性という概念を整理し体系づけることを目指す。

実は、既存の研究として、こうした多様な自律性を取り上げ体系的にまとめあげた研究は見当たらない。ヴァージニア・ディグナム(2017)の論考のように、タイトルに「autonomy」という単語が入っていても自律性とはなにかが論じられていない(こと)しばしばである(Dignum, 2017)。イギリス下院科学技術委員会がまとめた

報告書「ロボティクスと人工知能 (robotics and artificial intelligence)」では、機械の自動的 (automated) と自律的 (autonomous) との違いが説明されている (House of Commons Science and Technology Committee, 2016)。自動的という語が適切であるにもかかわらず、自律的といっていることが多すぎると指摘されている。ここでいう自動的は、産業用ロボットに使うのに適している言葉であり、定型的で繰り返しが多く、予測がかなりできる作業を担う機械に用いるべきであるのに対して、自律的は、未知の環境にも対応できる機械に使われるべきだということの説明は、実にシンプルでわかりやすいが、機械の自動的と自律的の区分であって人間も含めた自律性の検討ではない。

また、欧州科学・新技術倫理グループは自律性に関する声明を出している (European Group on Ethics in Science and New Technologies, 2018)。その声明によれば、自律性とは規範やルール、法を作り考えて選ぶ人間の能力であり、そのなかには自分自身で道徳規範を設け、自分で人生の目的や目標を選ぶ権利も含まれている。そのような人間の能力や権利には自己形成や自己意識、自己内の信念や価値が必ずともなっているものであって、それゆえ自律性は倫理的な意味では人間だけに使われる語である。どんなに機械が高度化して知能をもっているように見えようとも、自律性の語を人工物に使うのは誤用の類に属することである。この声明は、人間と機械の両方を視野に収め、自律性という語を機械に適用する妥当性について検討しているため、本書のねらいに近い。ただし、後でいう「自己決定」と自律性を同じものとみなしており、いわば近代的個人を前提にしている。近年、個人の自己決定が欺瞞であるといわれていることを鑑みるに、説得力不足を指摘されてしまいかねないだろう。本書は、近代的自我だけではなく生物や倫理的基盤となる自律性までも含めた幅広い領域を取り上げる。このように自律性に関して体系的に検討した研究はなく、本書はこれまでの研究と十分に差別化できる。