

倫理学とビッグデータ

リチャード・ハーシェル¹、バージニア・M・ミオーリ²

出典：

Herschel, Richard and Virginia M. Miori, "Ethics & big data," *Technology in Society* 49 (2017): 31-36.

キーワード：

- ・倫理学
- ・ビッグデータ
- ・カント主義
- ・功利主義
- ・社会契約論
- ・徳倫理

1. はじめに (pp.31-32)

- ・ビッグデータとは何か

ビッグデータは近年（論文が出版されたのは2017年）急激な速さで増加しているが、それは、コンピュータベースの技術やネットワークを用いることで作られ広まった情報を、獲得し、整理し、共有し、評価し、またそれに基づいて振舞うこと、これら全てに関わる。ビッグデータのソースには、気象情報、ソーシャルメディアの投稿、デジタル写真や動画、取引の履歴、RFID装置³、携帯電話のGPSなどが含まれ、以上のような多量のデータが集合

¹ セント・ジョゼフ大学ビジネス学部意思決定・システム科学科教授

² セント・ジョゼフ大学ビジネス学部意思決定・システム科学科教授

³ 株式会社デンソーウェーブのウェブサイトにおける説明は下記の通り。「RFIDとは、電波を用いてRFタグのデータを非接触で読み書きするシステムです。バーコードでの運用では、レーザなどでタグを1枚1枚スキャンするのに対し、RFIDの運用では、電波でタグを複数一気にスキャンすることができます。電波が届く範囲であれば、タグが遠くにあっても読み取りが可能です。」

<https://www.denso-wave.com/ja/adcd/fundamental/rfid/rfid/index.html>, 2020年4月22日アクセス

的にビッグデータと呼ばれる。

・ビッグデータの特徴

(1) 容量 (volume)、(2) 多様さ (variety)、(3) 速度 (velocity)、(4) 正確さ (veracity)。株式会社 IBM の特徴づけによる。(1) から (3) まではデータの質に関しては無関係だが、収集、配布されるデータが不完全で誤りを含むことが少なくないために、(4) に関してビッグデータは問題となる。

データ・アナリティクスに関する企業である SAS は、さらに下記の二つの特徴を付け足す。つまり、(5) ばらつき (variability)、(6) 複雑さ (complexity) である。(5) は、データフローが周期的なデータのピークと不整合になること、(6) はシステム間でデータのリンクや照合、転送などが困難になる点を踏まえて挙げられた特性である。

・ビッグデータに期待される効果

International Data Corporation (IDC) によれば、ビッグデータ技術やサービスには高い経済的収益が見込まれている。また、データマイニング⁴や分析によって、消費者の動向を把握し、予想し、さらには、新たな消費者の獲得を可能にするといったことが期待されている。

・ビッグデータのリスク

ビッグデータは扱いが難しく、不完全で間違いを含むことすらある。そのデータの中には人々のプライバシーを侵害する情報が含まれているかもしれない。ビッグデータ分析の不適切な使用によって生じる倫理違反行為の増加が予想される。

(1) Herold によれば、ビッグデータとビッグデータ分析に伴って、重大なプライバシーに対するリスクが生じるという。

匿名化されたデータファイルの使用を規制する確立した規則がない限り、大量のデータと強力なデータ分析によって、個人を特定する能力を排除できない。例えば、匿名化されたデータセットに完全に別のデータセットを組み合わせる場合、匿名性を保護するという目的に基づき、まず初めに、いくつかの削除されるべきデータ項目を決めておかないのならば、個人の再特定化が可能になってしまうのである⁵。

ス

⁴ 「データマイニングとは、未知の結果を予測するために、大量のデータセットに含まれている異常値、パターン、相関を発見するプロセスです」

SAS, https://www.sas.com/ja_jp/insights/analytics/data-mining.html

また、ビッグデータは個人の行為に影響を与えるために利用されうる。そのため、ビッグデータは、それを扱う組織が、個人のプライバシーを考慮に入れずに、様々な判断を下すために用いられる可能性がある。個人の取引（買い物）履歴の収集といったビッグデータの収集によって、企業は個人の同意なしに個人に関する知識を拡大させることができるのであって、その結果、「意思決定」とは、多くの場合、「予測可能な帰結を伴った個人による決定（individual decisions that have knowable outcomes）」であることをやめ、「おそらく多くの個人にとって意図しない帰結を持つデータ分析によって引き出された行為（actions derived from Big Data analytics which may have unintended consequences for many）」に、取って代わられることになるのだ。

(2) Zolan によれば、意思決定に用いられるビッグデータとは、多くの場合、常に正しいものであるとは限らず、バイアスのかかった文脈、あるいは、失われた文脈に基づいた不完全なものである場合が多い。このことが問題とされるのは、特に、ビッグデータを用いたアルゴリズムによって、欠陥のあるデータに基づいて、顧客のアイデンティティや行動に関する不正確な帰結が引き出されてしまう場合である。

また、Zolan によれば、ビッグデータを使用する際には、統計学的な偽陽性や不正確な発見を避けるよう、懐疑的であることや注意深くあることが、必然的に要求されるという。

(3) Zwitter によれば、ビッグデータによって、倫理学の中心は、説明責任を果たしたり道徳的判断を下したりする個人の能力から離れて、ビッグデータをコントロールするものの考察に移ったのだという。なぜならば、ビッグデータは、人々をターゲットにして操作して、人々をある一定の仕方で消費させたり行為させたりするために用いられるからである。このように、ビッグデータは、力（power）とは何であるのか、その力はどこにあるのか、そして、人々の行為を自由意志は実際どの程度決定しているのか、といったことを再定義することによって、倫理学上の議論の本質を根本的に変化させるのである。

2. ビッグデータと倫理的問題（pp.32-33）

・個人の権利保護のためのビッグデータ倫理の必要性

Richard と King によると、ビッグデータとその分析による予測によって、交際、買い物、医療、教育、投票、法の施行、テロ防止といった人のあらゆる行動や決定が、影響を被り始めている一方で、個人は、どのデータが集められているのか、どのデータが第三者に共有されているのかということについて、ほとんど把握していない。従って、ビッグデータの二次使用によって、プライバシーや機密性やアイデンティティが、過去、現在、さ

らには未来においてさえ、遡及的に侵害されうることを踏まえるならば、そもそも個人を同定している情報を取り扱うことに照準を絞った既存のプライバシー保護は、不十分なのである。さらに、最も個人を暴露しやすいセンシティブな情報の多くは、すでに政府や企業の手の内にあり、さらに増加しつつあるのであって、こうした組織から個人の権利を守るためのビッグデータ倫理が必要となるのだ。

・情報の再利用と不正アクセスが引き起こすプライバシー問題

Culnan と Williams は、「情報の再利用」と「不正なデータ」が乱用されてしまう可能性について注意を促す。彼らによれば、一方で、「情報の再利用」によって、組織は、収集した個人情報の新たな利用に関して、合法的な (legal) 決定を下すことができる。他方で、「不正なデータアクセス」によって、従業者たちは、本来見る権限を持たない個人情報を目にすることができる。この両者とも、個人が自分の情報へのアクセスを制限する能力を脅かし、個人に危害を加え、次いで、組織が消費者、株主、規制者に関与する正当性を毀損する可能性がある。また、「不正なアクセス」から帰結するプライバシーへの危害のうちには、機密性や信頼の侵害、個人のなりすましや身元詐称による経済的被害が含まれる。不幸なことに、ビッグデータは、このような危害への可能性を増強してしまうのである。

・複数のデータセット間で生じうる問題及びデータ収集の自動化に伴う問題

Nunan と DiDomenico は、市場調査に基づいて、ビッグデータが持つ含意について考察を行い、ビッグデータの利用によって作り出されるプライバシー問題を同定した。

第一の問題とは、一見プライバシー面で問題のなさそうなデータセットが他のデータセットと結びつくとき意図せずしてプライバシーを侵害するという問題であり、「意図しない使用のパラドックス (unintended use paradox)」と呼ばれている。例えば、Facebook から得られる公的に利用できる情報や写真のデータを利用し、この情報を顔認証アプリを通して照合することで、マッチングサイト上にある匿名の個人を特定する、という例が挙げられる。

第二のプライバシー問題は、データの収集が、人の活動とは独立に、自動的に行われるようになってきているという事実に関係している。つまり、電気や水道から飛行機に至るすべてのものに搭載されている、ネットワークに依存するセンサーの出現によって、収集された膨大なデータが、個人の同意に対する配慮なしに、自動的に分析されるという事態が出現している。

また、Mandinach によると、ビッグデータの処理は自動化されており、またアルゴリズム分析を用いた機器によって処理されるようになってきており、人間がこの過

程に参加しようとしても、このような努力は無駄なものになってしまうのである。

・ビッグデータの倫理問題に関する政府の立場や役割

いくつかの政府はビッグデータの使用と収集に内在的なプライバシーに関する問題を緩和しようと努めている。Terry によれば、欧州司法裁判所は、データの操作者(data controller)に対して、個人のデータを消去させるための合理的なあらゆるステップをとることを要求する、削除の権利(忘れられる権利)を認めている。それに対して、現在の合衆国政府の、ビッグデータの規制や、医療情報に対する危害への対応に関する見解は、一貫していないし、不十分であると Terry は考える。Terry は、連邦取引委員会(US federal trade commission / FTC)が勧告するように、データブローカーによって情報が収集・共有される前に、オプトインで明確に同意が表明される必要があることを支持する。

また King と Richards は、政府の介入を防ぐために、少なくとも組織はビッグデータ倫理に関する話し合いに参加すべきだと主張する。彼らは、FTC の勧告に同意して、大規模な第三者によるデータ分析のシステムの中において、人々は自身のプライベートな情報の流れを管理する能力を持つべきである、と主張する。従って、ビッグデータが倫理的に用いられるためには、データの所有者(データが使用される者)に対して、最低限、彼らのデータがどのように使用・売買されるのかということに関する、透明性のある見解が提供されるべきであるのだ。ここで重要となる点とは、あらゆる情報を、完全に公的、あるいは、完全に私的のいずれかとして考えることではなく、データは他のなんらかの目的のために利用されうるとする「暗黙の同意」など存在しない、ということなのである。

3. 倫理的観点 (pp.33-34)

ビッグデータ倫理の議論は採用する倫理学理論に左右される。以下4つの規範倫理学の理論を概観する。

3.1 カント主義

行為の結果よりも、行為の背後にある道德規則を重視する。イマヌエル・カント(1724-1804)の倫理学に由来する。カント主義では、われわれが何をなすべきかに着目する。我々のなすべきことは、義務(dutifulness)を反映したものであり、義務は善意志、すなわち誰もが従うべき規則に基づく正しいことを行ないたいという欲求を反映している。こうした規則の一つである定言命法とは、無条件に従うべき規則であり、他人を目的ではなく手段と

して扱ってはならない、などがその事例である。

3.2 功利主義

ものごとの善し悪しを行為や規則の結果に基づいて検討する理論であり、ジェレミー・ベンサム（1748-1832）とジョン・スチュアート・ミル（1806-1873）がその起源である。正しい行為は、共同体の幸福を最大化する行為であり、間違った行為は、その影響をうける集団の幸福の総量を減少させる行為である。正しい道徳的規則は、それが全員に採用された場合、関連するメンバーすべての幸福をもっとも増大させるものであり、功利主義では、行為や規則が最善の結果を達成するか計算が必要である。

3.3 社会契約論

人の道徳的または政治的責務はその人が住む社会を形成するために成立している同意に左右されると考える立場であり、トマス・ホッブス（1588-1679）、ジョン・ロック（1632-1704）、ジャン＝ジャック・ルソー（1712-1778）、ジョン・ロールズ（1921-2002）といった哲学者の議論に基づく。社会契約論では、人は社会生活の維持に必要な一定のガイドラインに従うことに同意する理性的能力をもつことが前提とされ、この同意をもとに、市民相互の関係を支配する道徳規則の確立の正当化を与える立場である。

3.4 徳倫理

徳倫理は義務や規則や行為の結果よりも、徳すなわち道徳的性格を強調する理論であり、アリストテレス（384 BCE-322 BCE）とプラトン（428 BCE-347 BCE）が起源である。道徳的に善い人は、有徳な人が正しいと理解することを実践し徳を発揮することで幸福を実現している。道徳的な徳は、一定期間を通じて繰り返される有徳な行為の習慣や傾向性によって形成される。

4. 倫理学理論のビッグデータ問題への応用（pp.34-35）

3節で確認した倫理学上の観点をを用いることにより、倫理学がビッグデータによるプライバシーへの懸念といった問題に取り組むことに寄与する理由とその仕方を、よりよく理解することができる。

4.1 カント主義とビッグデータ

カント主義的分析によれば、人々を単なる手段ではなく常に目的として扱わなければならない、人々の自律は尊重されなければならない、また、あなた以外の全ての人々が従うと想像することのできる道德規則にのみ従うべきである。

しかしながら、ビッグデータの本質とは、一般的に言って、人々が、自身の同意などを明示することによってオプトインでデータ収集やその利用に賛成するということが典型的ではない、という点にある。従って、人々のプライバシーはデフォルトで他の者の利益のために棄損されている、ということになる。また、ビッグデータは膨大で、人々はプライベートな情報を積極的に監督しコントロールするメカニズムも持たないために、人々は自身のデジタルなデータの共有に関して、盲目的な立場に置かれてしまっている。さらに、ビッグデータによって、人々は単なる「データの点 (data points)」としてしばしば取り扱われ、そして、その「データの点」は「その人が未来において目にする事」を操作するために利用されることになる。つまり、アルゴリズムに基づく操作によって、明示的な同意なしに、その人の未来における意志が前提とされるのである。

カント主義的観点は、比較的直接的に、ビッグデータの倫理を論じる手段を与えてくれる。ビッグデータに結びついた以上のような諸行為 (actions) は、個人の権利と平等な取り扱いに対する重大な危険となるために、ビッグデータは問題含みであるということになるのである。

4.2 功利主義とビッグデータ

カント主義と異なり、功利主義的観点からビッグデータの倫理を査定することは、困難に満ちている。行為功利主義の立場に立つ場合、功利計算をする際には、快苦の経験の強度、持続性、可能性、等々の多くの要素を考慮に入れなければならないし、規則功利主義の立場に立つ場合でも、幾分単純化されるとはいえ、ビッグデータの社会に対する利益と害悪を考量して、ビッグデータの使用に関する規則を査定しなければならない。さらに、功利計算は、その計算をなすものによってある事柄が他の事柄よりもいっそう重要視される場合があるといったように、内在的なバイアスを抱えているかもしれない。つまり、このアプローチの欠点とは、利益と害悪の両者を同定し量化するよう努める際に、あいまいさやバイアスが内在的に伴われる、という点にあるのであって、ビッグデータの価値を査定するための功利主義的手続きとは、複雑に過ぎ、内在的に欠陥を抱えており、また、ビッグデータとは何か、ビッグデータの使用は人々の生活にどの程度影響を与えているのかということに関する理解が、社会的にみて一般的に欠けているために、その手続きとは挫折を余儀なくされるのである。

4.3 社会契約論とビッグデータ

社会契約論によれば、合理的な人々が自身に利益を与えるということを理由にして受け入れることに賛成した、規則やルール形成が強調される。また同時に、この理論は、異なった社会が、同じ道徳的権利を、違った仕方で思い描き明確化し施行することも可能にする。従って、このような違いによって、諸社会の間での協調を可能にする交渉が必要となるような問題も生み出される。例えば、確かに、アメリカとヨーロッパにおいてデータ・プライバシー権について違いがあったものの、近年、両者の間で、データが共有された際に諸個人に認められるプライバシーのレベルに関して合意が取り交わされた。このように、人々はビッグデータを作り出す技術を利用したいと思いつつ、ビッグデータが人々に与える影響をコントロールすることも望んでいるのであって、社会契約論の立場に立てば、合理的な人々は、社会に与えられる利益に基づいて、ビッグデータのどのような側面が道徳的に正当なものであるのかについて、必然的に、集団的に決定を下すことになるだろうと、考えることができる。

4.4 徳倫理とビッグデータ

徳倫理の観点からは、ビッグデータの使用者の意図や使い方が、有徳な人の行為に整合的かどうか問題になる。ビッグデータの問題に照らすなら、ビッグデータを用いる人々の行為が他者のプライバシーを傷つけていないかどうかを知ることを通して、その人々が道徳的知恵 (*mora wisdom*) を持つかどうか評価されることになる。また、個人の権利を無視する人々は、不誠実さや貪欲といった悪徳を表出しているとみなされるだろう。しかし、徳倫理学は、採用された行為が有徳者に特有のものであると決定するために、その行為を調査する際に注意深くある (*scrupulous*) ことを要求するため、困難なアプローチとなりうる。結局、このアプローチとは、有徳な人ならば採用することになるようなアプローチなのだ。

5. 結論

倫理学上の理論を調査することによって、我々は、ビッグデータに関する道徳的状況に関する異なったパースペクティブをよりよく把握することができるし、提示されている議論の文脈や論理をよりよく理解することができるし、また、そうすることで、どのような意図された行為の方針が正当化されるのか、あるいは、正当化されるべきなのか、よりよく評価することができるのだ。

6. 今後の研究

ビッグデータを倫理学上の理論に則して検討した後になすべき重要な課題とは、ビッグデータの利用者の状態を示す、質的データの収集である。例えば、ビッグデータを使用する企業や、ビッグデータを提供する個人の状態に関する調査である。このような調査によって、企業や個人の現実の声に基づいたビッグデータ倫理学の構築が可能になるだろう。

(西條玲奈＋三上航志)