

ビッグデータの倫理学

アンドレイ・ツヴィッター¹

凡例・出典

以下傍点は原文中の強調を示すイタリック体、〈〉は紹介者による語句の区切り、[]は紹介者による補足を表す。また、各項目中とりわけ重要と思われた内容には下線を施してある。

Andrej Zwitter, “Big Data ethics”, in *Big Data & Society*, July-December 2014, pp.1-6.

キーワード

- ・ 道徳的責任 (moral responsibility)
- ・ 社会的影響 (societal impact)
- ・ 教育 (education)
- ・ ビッグデータ・ガバナンス (Big Data governance)

論文の背景・目的・構成

ビッグデータとソーシャルメディア等の関連事象の発展スピードは、一般的な消費者たちが自らの行動とその波及効果を理解する能力を凌駕している。今や倫理学の主たる対象は、特定の見えやすい結果を伴う個人的な決定から、気づかぬ間にとってしまい、他者に意図せぬ結果をもたらす行動へと、移行しつつある。

ツヴィッターは、本論文において、ビッグデータが倫理学の諸概念、特に〈個人の道徳的行為者性 individual moral agency〉という、18世紀以来倫理学において自明とされてきた前提の再考が迫られていること、それが社会・政治・研究において再考を要請することを明らかにしようとしている。本論文の最も重要なテーゼは、〈力 power〉という概念が個人のものからネットワーク化されたものへと変容したという主張である。

本論文の構成は以下の通りである。(1) 道徳的責任に関する伝統的な倫理学的原理が略述される。(2) ビッグデータが有する倫理的に重要な4つの性質(データの多量性・有機性・グローバル性・相関性の優先)がまとめられる。(3) ビッグデータが力 powerの本質を変えるものだという発想、および、高度にネットワーク化された倫理学 hyper-networked ethics の登場について、立ち入って洞察される。(4) これらの変化によって、

¹ アンドレイ・ツヴィッター (Andrej Zwitter) 氏は、フローニンゲン大学教授。専門は国際関係論、倫理学、国際法・ヨーロッパ法。

社会・政治・研究においてどのような倫理的問題が生じるかが明らかにされる。

1. 伝統的倫理学

啓蒙主義の時代以来、義務論、功利主義などの倫理学は伝統的に個人の道徳的責任、つまり〈道徳的行為者性 moral agency〉に重きを置いてきた。この考えは個人主義と自由意志という前提に由来している。ある一つの存在 entity の責任は、それが道徳的行為者性をもつ程度に応じて決まる。道徳的行為者性は一般に、ある存在に内在的な条件によって規定されるが、この条件として広く受け入れられてきたのは次の3点、すなわち① 因果関係（行為者は、自らの行為が倫理的に問題となる帰結をもたらした場合、責任を負わされうる）、② 知識（行為者が自らの行為の帰結を知っていた/知っているべきであった場合、行為者はその帰結について非難されうる）、③ 選択（行為者が自己自身に対するより大きな危害を伴うことなく別の選択肢を選ぶ自由を有していた場合、行為者は帰結について非難されうる）だ。

しかし、こうした前提および条件が、現代技術とりわけビッグデータの発展によって異議を唱えられている。〈道徳的行為者性〉なしに〈帰結 outcome〉を道徳的に重要なものとする新たな路線が、コンピュータ・情報倫理（バイナム）やネットワーク倫理（フロリディ）、ソーシャルネットワークキングの倫理（ヴァラー）、企業体の分散的な道徳的責任（アースキン）の論議において進展してきており、ここに、ビッグデータが更に「多数の手 many hands」（多くの行為主体が分散的道徳性 distributed morality という形をとって一つの行為に参加するということをもたらす影響）の問題をはじめとする様々な変化を招いている。ビッグデータがもたらすそのような変化に注目することが必要だ。

2. ビッグデータの4つの性質

ビッグデータの倫理学には〈個人 person の道徳的行為者性という概念から遠ざかっていく〉方向と〈ビッグデータを操る主体（人間であれシステムであれ）の道徳的有責任を増大させる〉方向とがあるが、やはり一般的には前者つまり〈個人に基づかない倫理学 impersonal ethics〉に向かう傾向にある。そこで、前者の関心にとって重要なビッグデータの核心的性質を確認しておこう。それは（1）史上類を見ないほどにデータ量が多いこと（2011年に2日あたり50億ギガバイトだったのが、2015年には10秒あたり50億ギガバイトにまで増えている）、（2）有機的（organic）であること（無差別的なデータ集積により、統計的データよりもはるかに自然に現実を再現する）、（3）潜在的にグローバルであること（例えばグーグル）、（4）ビッグデータ分析が因果関係よりも相関関係（correlations）を重視すること、の4つだ。[この4性質が、1. でみた道徳的行為者性の3条件の成立を困難にするのである。]

そうしたビッグデータの具体例として倫理的に着目すべきは人間に関わるものだ。たしかにビッグデータに使われる全てのデータが人間によって生み出されるわけではない（例えば天文学的調査のための大型シノプティック・サーベイ望遠鏡なども大量のデータを生み出している）。とはいえ、人々とその相互作用に関係した巨大なデータ群（ソーシャルネットワークデータ、健康追跡データ、メールやメッセージ、グーグルの検索エンジンの使用など）が存在し、倫理的にはこちらの方が大きな問題となり得る。

3. 新たな力の分配

以上のようなビッグデータによる伝統的倫理学の前提の転換においてさらに問題なのが、ハイパーコネクト時代のなかでは、倫理学および道徳的責任論にとって決定的に重要な 〈力 power〉（他の行為主体への影響力）の概念が、よりネットワーク化されたものへと転換しつつあるということだ。個人の道徳的行為者性、つまり知識と行為能力 [1.で確認した条件] を保持することはもはや困難なのだ（ビッグデータのガバナンスが直面する困難もここにある）。

ビッグデータには「ビッグデータの収集者」「ビッグデータの利用者」「ビッグデータの生成者」の3種のステイクホルダーがある。そこでは3者の行使する〈力〉の相互作用が生じ、これが「依存的行為者性 dependent agency」（ある行為主体の行為能力が他の行為主体に完全に依存している状態）を呈示する。ネットワーク化された社会では、この依存的行為者性が、ビッグデータに関わる行為者の道徳的責任を判断する際つねに一要素となる。すると [伝統的倫理学ではあまり注目されてこなかった] 行為の 波及効果・副次被害 が重大なものとなるのだ。かくして、ネットワーク化という社会本性の変化は、力の実態を個人的/独立的なものからよりネットワーク的/相互依存的なものへと転換し、非意図的な帰結を増やし、私たちの自由意志概念の範囲を変化させるのである。

4. いくつかのビッグデータの倫理的難題

ビッグデータの4性質に基づく次のような倫理的難題が明らかになる。まず(1) データの多量性と (2) 有機性に関して、大量データの集積によりデータ化された現実世界では、実際にはどのようなデータが集められているか/それが何のために使われるかについての知識が欠如し、「データ生成者」（オンライン・カスタマーや携帯電話所持者など）は知識と自由意志の点で不利な立場に立たされる。次に (3) グローバルなデータはステイクホルダーどうしの力の不均衡を生む。そこでは情報から知識を生み出すノウハウをもつ企業体が利を得る。そして (4) 因果関係より相関関係を重視するビッグデータ分析のもとでは、私たちは〈なぜ〉を知ることなしに、見えるものを信じな

ければならない場合が増える。

4.1 プライバシー

私たちの生活がサイバーリアリティに反映され記録されるにしたがって、私たち一人ひとりの現在および過去は、技術とアクセスをもつ行為主体にとってほぼ完全に筒抜けになる。例えばアメリカでは、ソーシャルネットワークからアクセス可能なデータや IP アドレスに紐づけられたデータ等を自在に利用して、一人の人間をプロフィールし、その日常の行動を完全に筒抜けにする環視情報高速処理技術 (RIOT) というソフトウェアが開発された。

4.2 集団プライバシー

データ分析者たちは、人々の買い物の好みや健康状態などを知るためにビッグデータを用いているのであって、こうした情報が〈個人〉を特定するのは、稀なケースでしかない。しかし、非個人化 *de-individualization* (データがある一人の人間に結びつくのを可能にする要素を排除すること) は、匿名化のほんの一局面でしかない。ある集団への所属に関して重要で、それゆえ統計分析に有益な情報 (居住地、ジェンダー、年齢その他) は 〈集団プライバシー group privacy〉 の問題に関わってくる。データが非個人化されているという意味で匿名であっても、〈集団〉はますます筒抜けになりつつあるのだ。

統計情報やビッグデータから集められるそうした〈集団〉の情報は、ターゲットマーケティング等の手法に利用されうる。さらに、特定の集団がもつ選好と条件に関する様々な情報は、特定の行動を促進/抑制する動機づけ *incentive* を採り入れるために利用されかねない。例えば、集団 A がある選好 α (アイスクリーム等) を有し、かつこの同じ集団内の多くの人が条件 β (投票する政党を決めかねている等) を有することがわかった場合、条件づけ (政党 B に投票するとアイスクリームがもらえる等) を作り出すことで、この集団が β の領域内で特定の仕方でするまうよう α を提供することができる。これだけなら標準的な党派戦略だが、ここにビッグデータが絡むと、隠れた相関関係を発見する能力が増大し、この能力が更に、目的不明の動機づけを生み出す能力を増大させる。

一方でハイパーコネクト状態は他の戦略も招き入れる。例えば、Twitter に潜入するボットが、ある政党について虚偽の草の根議論を作り出す、という例が挙げられる (実際、2010 年マサチューセッツ州で行われた特別選挙の結果はまさにこのようなボットによって左右された可能性がある)。また 2013 年の報告によると、実はウェブ閲覧全体の 61.5% がボットによってなされたもので、その半数はサーチエンジンやその他サービスのための所謂「善良ボット」だったが、もう半数はマーケットの詮索や操作を目的とした悪意あるボットであったという。

4.3 傾向 propensity

ロサンゼルス等の都市では、既にビッグデータにより犯罪を犯しそうだと特定された町の住人や暴力団、個人を特別監視下におく「犯罪予測 predictive policing」が行われているが、これは政策上問題である。例えば殺人を犯す確率は、そのままでは多くの人からの批判を呼び起こさずにはおかない。また犯罪予測ほどには介入しなくとも、例えば、もしビッグデータ分析によってある人が 95%の確率で家庭内暴力に巻き込まれるなどという情報を手に入れば、どんな社会福祉機関もこの情報に基づいて政策的行動をとらない訳にはいかない。その結果として、本人やその家族、友人たちに対するステイグマが引き起こされてしまいかねないのである。

すると更に、データサイエンティストの役割の倫理性も問いに付されることになる。ビッグデータ研究における重要な変化の一つは、彼らの手によってアルゴリズムが単力で相関関係を検出できるようになったことだ。ビッグデータはデータマイニングが見つかる共通性に基づいて、ランダムな連関を極端なまでに増やす（例えば「ハリケーンの前には苺タルトの売上が伸びる」等）。これがソーシャルネットワーク分析と結びつくと、単に個人のプライバシーを大きく侵害する情報だけではなく、たまたま同時に起こったというだけで任意の連関をでっちあげることを可能にする情報を生み出しかねない。RIOT 等の調査技術については、この観点からも批判的検討がなされるべきだ。

4.4 研究倫理

研究倫理に関する議論は、上述のテーマに比べ遅れをとっている。ソーシャルメディアを研究目的で使用することについての十分な議論がなされていないのだ。①多くの議論で（従来型の）個人プライバシーは考慮されているものの、〈集団プライバシー〉の問題は考慮されていない。例えばフェイスブックはプライバシーが問題になるといつも槍玉に上がるのだが、一方でその議論は、非個人情報もその多くが非常に詳細に諸集団の地理的関係を暴露できることを見えなくしてしまう。つまり集団プライバシーの問題は、研究倫理についても当てはまるのだ。②多くの研究で未だに〈プライバシーに関係しない倫理的影響 non-privacy-related ethical effect〉が考慮されていない。例えば、ツイッターを使って研究する者は、特定の間人集団について不都合な事実を伝え、研究対象の集団に望ましくない影響を与えかねない。集団について不都合な情報を暴露してしまう研究上の発見というものが、研究倫理の次なる主要課題になるだろう。③「インフォームドコンセント」の問題もある。データがそれ自体すでに公共的であるとしても、自分は突然ツイッターやフェイスブックの研究のデータ主体になると本気で考えている人はいないのだ。しかし、社会現象の説明や分析のためと言って、インフォームドコンセントの欠如が他の研究形態（心理学や医学）であれば大きな研究倫理違反

となることを考慮せずにソーシャルメディアからデータを収集してくる研究者がいる。

5. 結論

本試論が試みたのは、ビッグデータが、個人性、自由意志、力という倫理学の伝統的前提に一定の変容をもたらすかもしれないということを指摘することだ。こうした変容の影響は、私たちが非常に長いあいだ目を向けてこなかった様々な領域に及ぶことだろう。

教育の領域では、大人子どもを問わず、自分のデジタル・フットプリント（ネット空間における自分の活動）の意図しない帰結についての（デジタル・リテラシーを超えた）教育が必要である。社会科学の研究は、この教育の欠如を考慮して、匿名的でありながら集団について多くを明かしてしまう社会的ビッグデータの使用に関する倫理的含意についての結論を引き出さねばならないだろう。法学および政治学の領域では、次の3つの発展が期待される。それは（イ）政治運動のオブザーバー、シンクタンクの研究者その他の調査者が、新種のデジタルな世論操作を調査するためのデータ犯罪科学の専門家となること、（ロ）弁護士や法学研究者と同様、法の執行や公共事業を担う機関もまた個人犯罪、犯罪確率、犯罪予防について考え直すこと、（ハ）国家が、地域の専門家や主観的判断よりもむしろグローバルなデータとアルゴリズムに基づいた、グローバルな戦略を開発する方法を漸進的に再設計すること、である。

ビッグデータは、個人の責任および力の分配をめぐる諸前提に大きな影響を及ぼすと言っても過言ではないように思われる。倫理学者たちは、データ化された世界で私たちはいかに生き得るのか/生きたいのか、いかにビッグデータの濫用を防ぐことができるのかについて、議論し続けていかねばならないだろう。

（林航平）