



イノベーションを支える
データ倫理規範の形成

データ化の時代のデータガバナンスの新興モデル

出典・凡例

本稿は, Micheli, M., Ponti, M., Craglia, M., & Berti Suman, A. (2020). Emerging models of data governance in the age of datafication. *Big Data & Society*.

<https://doi.org/10.1177/2053951720948087> の要約である。

要約作成にあたり, 原語を付すために () を用いた。原文でイタリックにより強調されている箇所は太字にした。また, 本稿に付された注はすべて訳注である。



アブストラクト

現在、企業のプラットフォームが大量の個人データを収集し、搾取する支配的モデルが注目されている一方で、中小企業などの他の主体もデータガバナンスに参加しており、その中から新たなデータガバナンスモデルが生まれつつある。

本稿はデータガバナンスの社会科学に基づいた概念化を提案し、データインフラストラクチャ(data infrastructure)¹の概念から教訓を得て、利害関係者の役割、彼らの相互関係、価値の明確化、及びガバナンスの原則の機能として4つのモデルを特定し、データの力関係に対処するために、データの運用をめぐるこれらの主体の競争的な闘争を検討した。

この概念化は、企業主体による支配的モデルの中で出現したガバナンスモデル内の力関係と多面的な経済的・社会的相互作用を前面に押し出す。これらのモデルは、市民社会と公的機関が、データガバナンスを民主化し、データを通して生み出される価値を再分配するための重要な主体であることを強調している。

モデル、そしてその基礎となる原理と限界の議論を通じ、特にデータガバナンスを巡る政策論争が欧州で非常に活発に行われている今、データガバナンスのための社会技術的イマジナリー(imaginary)²の将来の調査に影響を与えることを本稿は目指している。

キーワード

データガバナンス、ビッグデータ、デジタルプラットフォーム、デジタルインフラストラクチャ、データ・ポリティクス³(data politics)、データ政策

¹ データの共有と利用を促進するデジタルな社会基盤。

² 「特定の社会集団および人々が社会全体を想像するための対応する社会に共通する一連の価値観、制度、法律、および象徴」といった一連の意味を持つ社会学の用語。

³ オープンデータといったデータの政治的側面に焦点をあてる分野。

背景

(本稿の方向性)

- デジタルプラットフォームを運営する一部企業による独占に注目が集まり、それに対応しようとする動きが欧州などで出現している。
- こうした今後の法改正を考慮し代わりに（ビッグ）データを運用するために実施され、理論化された社会的実践に焦点を当てる。
- それにより、現在のデータ環境から下意上達による主体の可能性に至るまで、科学的な注目の新たな方向づけを唱えている研究者や政策立案者に対応する。

(本稿の手法)

- 本稿は、主体間の力関係を強調するデータガバナンスについての**社会科学的な**手法を採用することでこの種の探究に貢献している。
- 科学技術研究(Science and Technology Studies、以下 STS)と批判的データ研究(Critical Data Studies、以下 CDS)に基づいて、データインフラストラクチャを分析レンズとして使用し、各データガバナンスモデルを、利害関係者の役割、相互関係、価値の明確化、各モデルにおけるガバナンスの原則、手段、メカニズムの組織の状況的、偶発的、関係的なインスタンス化⁴として概念化した。

(本稿の焦点)

- 現在の「プラットフォーム社会」におけるデータガバナンスの支配的モデルは、少数の企業の巨大テクノロジープラットフォームにより確立されたものだが、他の主体が、異なるモデルを通じて個人データを制御し、そこから価値を生み出すことに徐々に関与している。これら代替モデルが本稿の焦点である。
- 故に、我々はプラットフォームではなく社会的主体により開発されたデータへのアクセスと制御のための実践に注目する。これらは、将来的に（ビッグ）データが

⁴ オブジェクト指向のプログラミングにおいて、インスタンス（クラスを基にした実際の値としてのデータ）を生成すること。



どのようにガバナンスを受けるかに影響を与えうるデータのための社会技術的なイマジナリー⁵を開発するための想像力に富む文脈である。

(本稿の研究・政策的意義)

- 本稿の政策的関連性は、ますますデジタル化する私たちの社会の発展の中心であると見られる人工知能（AI）を巡る現在の地政学的競争に由来する。
- 政策や学術的な議論の現在の段階では、データ・トラストやデータ主権などの多くの用語や概念が出現し、定義や使い方が不明確か矛盾している。本稿のデータガバナンスモデルの分析は、これらの概念のいくつかを明らかにし、したがって研究と政策の両方に関連するものである。

⁵ 技術がどのようにして社会的グループによって実現され、使用されるかについてのビジョンのこと。

ガバナンスの定義

- ここ 20 年でガバナンスという語が盛んに使われるようになったが、その意味は曖昧なままである。
- 本稿のガバナンスに対する理解は、例えばガバナンスを拘束力のある決定に繋がる多数の主体と過程だと捉える政治科学とリスク研究に基づいている。
- ガバナンスは概して、制度を管理する過程において、様々な役割を担う関連主体の網である。この語は国家によるいわゆる「指揮統制」との非連続性を強調し、そしてより広範な一連の主体と機関（民間部門、市民社会、そしてその他の NGO）もまた社会の管理に関与することを認めている。
- ガバナンスとは規則の作成や執行を通じて生じるだけでなく、水平レベルでの利害関係者間の（社会的）相互作用、協力そして交渉によっても発展する過程の結果である。
- データ・ポリティクス概念は、「データの遂行的な力」を強調する。データがどのように収集され、処理されるかによって、そのようなデータが関係する主体や客体を生み出す上で、力の不均衡や情報の非対称性が生じることから、それはデータを代表する能力としてだけではなく、“新たな形の力関係を創造する”力としても理解している。
- 以上の検討を踏まえ、ここでは、データガバナンスを次のように定義する：データが精査、支配、共有および利活用される方法、データから価値を生み出すためになされるさまざまな社会技術的構造(arrangements)、およびそうした価値がアクター間で再配分される方法、によって影響を受ける、またはそれらに対して影響をもつすべてのアクターの間における権力の諸関係である。

分析の諸局面

- 新興モデルの分析と説明を導くため、我々は以下の分析の次元を使用した（表1参照）。以下の次元は、データガバナンスに關与する主要な利害關係者の利益と目標を前面に出すことで、STS と CDS に關連しており、各モデルでの力關係と媒介の異なる形態を可視化するのに役立つ。

次元	定義
利害關係者	データの管理方法や創出された価値に影響を受ける、または影響を及ぼす個人、機関、組織、集団
ガバナンスの目的	データにどのようにガバナンスを及ぼすかに影響を与える主体が持つ目的
データから得られる価値	データの利用から期待される資源と、それが主体や社会全体にどのように分配されるのか
ガバナンスのメカニズム	その根底にある様々な原則を含む、具体的なガバナンスの目標を達成するために採用された様々な手段
互惠性	データへのアクセスと利用のための利害關係者間の力關係

利害關係者

- 利害關係者とは、個人、組織、グループなど、データの管理方法やそこから生み出される価値に影響を受ける、あるいは影響を及ぼす全ての主体を指す。
- 利害關係者は、どのようにデータが集められ、扱われるかについての知識とともにデータへのアクセス、制御、処理の可能性の点で大きく異なっている。彼らはまた、データに関するさまざまな価値観や関心、およびその使用に関する規範を保持している。
- 彼らには「データ主体、データ管理者（および処理者）、第三者のデータ利用者」、民間部門、公共部門、学界、科学・市民団体、活動家、社会起業家、市民が含まれる。

ガバナンスの目的

- ガバナンスの目的とは、様々な利害関係者がデータを統制する為に設定した価値ベースの目標であり、彼らにとってデータが表す意味である。
- 企業にとっての目標はデータの共有と集約を通じた、財務的な利益を最大化することであろうし、政策文書では、データ共有協定で追求すべき重要な目標の一つとして公共の利益を挙げているかもしれない。
- あるいは、データ主体によるデータ制御を強めることや恵まれない人々に声を上げる力を与えることかもしれない。

データから得られる価値

- データから得られる価値とは、これらのプロセスから異なる利益を得る可能性のある様々な利害関係者により、集約、分析、ビジネスインテリジェンス⁶を通じデータから生み出される価値の種類を指し、データが「社会的に進歩的な方法」で使用されることをモデルがどの程度予見しているか評価する。
- 何らかの形の公共的な価値が創出されているのか、それとも企業および（または）個々の利用者にとっての私的な価値だけなのか。この次元を通し、データの集約と分析を通して生み出された価値や知識が、主体間や社会全体に再分配されているかを問う。

ガバナンスのメカニズム

- ガバナンスのメカニズムとは、目標を達成し、社会技術システムの変化を導くため、異なる主体が採用する戦略と手段のことで、データがいかに制御され、いかなる価値が創造され、誰が利益を得るかの枠組みをつくるデータ集合体の要素で構成され、それは以下を含む：思考の枠組み、政策、規制、委員会、契約、利用規約、基準、アルゴリズム、インターフェイス、その他の社会技術システム。
- これらは今日のデータインフラストラクチャのガバナンスのメカニズムを形成し、一部の主体が他の主体より優れた方法で利用できる。STS に基づき、この次元は、

⁶ 経営・情報処理などの用語で、企業などの組織のデータを、収集・蓄積・分析・報告することで、経営上などの意思決定に役立てる手法のこと。

データインフラストラクチャに埋め込まれたより広範な倫理的、政治的、経済的原則をも含み、データガバナンスが行われる複雑な社会技術的な「集合体」を表す。

互惠性

- 互惠性とは、データへのアクセス、制御、利用における利害関係者間の力関係を指し、巨大テクノロジー企業が意思決定権の大部分を握っている一方的なアプローチと、より多くの利害関係者が参加している相互的なガバナンスモデルとの違いを強調しており、データを「新しい形の力関係を生み出す」とするデータ・ポリシークスの概念と関連している。

4つの新興データガバナンスモデル

図 2

モデル	重要な主体	目標	価値	メカニズム
DSP	<ul style="list-style-type: none"> 企業主体 公共団体 	<ul style="list-style-type: none"> データ共有を通じた知識の格差の是正 新たなサービスの開発 	<ul style="list-style-type: none"> 私的な利益 経済成長 	<ul style="list-style-type: none"> ”商品としてのデータ”という原則 協力関係 契約（例 反復可能な枠組み）
DC	<ul style="list-style-type: none"> 市民団体 データ主体 	<ul style="list-style-type: none"> 現在のデータ経済の力の不均衡を解消 社会的課題に取り組む 価値創出の為、社会正義とより公平な条件を促進 	<ul style="list-style-type: none"> 公共の利益 科学的研究 権限を付与されたデータ主体 	<ul style="list-style-type: none"> 協力的な動きからの原則 データ・コモンズ ボトムアップのデータトラスト GDPRに規定されたデータ転送の権利
PDT	<ul style="list-style-type: none"> 公共団体 	<ul style="list-style-type: none"> 政策決定に影響 社会的課題に取り組む イノベーション データに責任ある手法の採用 	<ul style="list-style-type: none"> 公共の利益 より効率的な公共サービスの提供 	<ul style="list-style-type: none"> ”公共インフラとしてのデータ”という原則 信頼醸成の取り組み 信頼できる仲介業者 法的枠組み
PDS	<ul style="list-style-type: none"> 企業主体 データ主体 	<ul style="list-style-type: none"> データ主体の自己決定権 現在のデータ経済の力の不均衡の解消 新たなデジタルサービスの開発 利用者のニーズを中心におく 	<ul style="list-style-type: none"> 権限を付与されたデータ主体 経済成長 私的な利益 知識 	<ul style="list-style-type: none"> ”テクノロジー主権”の原則 共同体と運動（例 My Data） 媒介型デジタルサービス（個人データスペース） GDPRに規定されたデータ転送の権利

- このセクションでは、上述した 5 つの次元に沿って特定された 4 つのデータガバナンスモデルについて説明する。
- 各モデルは最新の状態を網羅的に説明するものではなく、新興のデータガバナンスモデルを合成する上での貢献として意図されている。
- 分析には、現在の支配的なモデルとは程度の差こそあれ、異なるモデルが含まれている。故にプラットフォームが他の主体とのデータ共有に関与しつつ、データに対する完全な制御を保持し、どの利害関係者を内部に入れるか、どのようなデータにアクセスし、何ができるかを一方的に決定している事例は考慮していない。

データ共有プール (DATA SHARING POOL、以下 DSP)

- DSP では、ビッグデータの古典的なレトリックの 1 つである「データは集約すればより多くの価値を生み出す」という原則が採用されている。データを市場商品として扱い、2 人以上のデータ保有者（私的、公的の両方）が力を合わせ、データ共有契約を結ぶ。彼らはお互いのデータを分析して知識のギャップを埋め、データ主導のイノベーションを促進する。
- このモデルは水平的な関係に基づいているため、表面上は、潜在的に多くのデータ保有者間の互恵性を促進している。しかし、力の非対称性も助長する。より多くのリソースを持つ、またはより価値のあるデータセットを所有しているデータ保有者は、データへのアクセスや使用方法の条件を設定する力が大きい。
- さらに、データ主体（および一般的に市民）は、このモデルでは発言権を持たない。彼らは関係の中には含まれておらず、せいぜい関係の中で開発されたイノベーションの受け手として描かれている。

データ協同組合 (DATA COOPERATIVE、以下 DC)

- DC は草の根主導の分散型データガバナンスモデルで、市民が自発的にデータを共有し、彼らに代わってデータを管理する協同組合との信頼関係を確立する。データ主体は、データに対する民主的な制御を維持し、生み出される利益の中で公平な分け前を持ち、それらはしばしば（医療研究などの）公的な利益を目指す。
- 協同組合運動とプラットフォーム協同組合の原則から導き出されたこのモデルは、プラットフォーム資本主義とは正反対のものであり、より公正で透明性の高い非独占的な代替手段を目指している。巨大テクノロジープラットフォームは、たとえ非公式に「敵対的主体」として存在していたとしても、DC におけるデータガバナンスから完全に排除されるかもしれないし、利用者がデータを取得するデータ・プロバイダーとしてのみ含まれるかもしれない。
- DC はデータ共有への倫理的アプローチとして政策論争の中で多くの関心を集めているが、規模を拡大して決定的な利用者数に到達するのに苦労している。



パブリックデータトラスト (PUBLIC DATA TRUST、 以下 PDT)

- PDT には、より幅広い主体が関与しており、公的機関が主導権を握っている。PDT の基本的な原則は、公益性の要素を持つデータはすべて国家のインフラの一部であり、故に、それが提供する情報は、市民や社会全体の為に価値を生み出すように「社会化」されるべきであるというものである。また、公的機関は時として信頼できる仲介者を設けるかもしれない。
- PDT には、公共部門、非営利団体、企業、学界などの主体が参加している。しかし、データ主体は、データを通じて開発されたサービスの単なる受け手ではない。PDT が効果的であるためには、個人情報保護され、公共の利益のために利用されることを安心させなければならない市民との信頼関係を確立することが必要である。
- この目的を達成するためには、公的機関が市民社会からの要望に耳を傾けて信頼関係を構築する必要がある。COVID-19 接触者追跡アプリのための議論と同様に、最近の失敗した取り組み(SideWalk Toronto⁷)が、示すように、これは公共団体が(まだ)何をすべきかを学んでいるものである。
- PDT にとってのもう一つの課題は、公益性の高いデータを持っているかもしれないが、規制で義務付けられていない限り、データを共有しようとはしないであろう私企業とデータ共有の取り決めを確立することである。

個人データ主権 (PUBLIC DATA SOVEREIGNTY、 以下 PDS)

- PDS は、データ主体を、データへのアクセス、制御、共有、分析の手段を提供する新しいサービスのエコシステムの中心に置く。主権の原則に基づいたこのモデルは、データに対する個人の制御と自己決定を強調しており、監視資本主義とは対照的である。

⁷ アルファベット社が担当し、テクノロジーを用いたトロントのウォーターフロント開発(港湾地域の再開発)だが、巨大テクノロジー企業であるアルファベット社がプロジェクトを担当することにプライバシーの観点から多くの批判が殺到した。



- データ活動家の運動がこの進歩的な目標を推進しているが、同時に商業主体は新しいサービスのエコシステムを支える手段として関心を持っている。PDS では、異なる利益を持つ異なる種類の主体が、「より公正なデータ経済」の促進のために集まっている。この緊張は、いくつかの矛盾をもたらす。
- 中間サービスの完全に私的なエコシステムは、たとえ利用者がデータを制御できるようにしたとしても、現在のデータ・領域のパワーの非対称性に対処するという地位を市場に委ねることになる。デジタルデータに対する市民の視点に関する研究からは、大多数の人は、これらの仲介サービスが提供する機会を活用するスキルや関心を持っていないように見える。

各モデルについてのまとめ

- DSP は主に経済的価値の創出に焦点を当てているが、残りのモデルでは、社会変革、公益、公平性、データ主体の自己決定など、他の形態の価値が徐々に「加わって」きていることがわかる。
- PDS では個人（自己決定）と新たな商業主体（データサービス）のための価値の創造に向けられており、公共の利益はこれらの副産物である。残りのモデルは、公益を表情豊かに追求している。
- これらのモデルは、大抵、ニッチな取り組みや試験的なプロジェクトの中で見られるもので、それらが生み出す価値と長期間での各モデルの持続可能性に関する研究は限られている。故に、価値の生産と再分配は、具体的な成果の評価よりも、イマジナリーのレベルで評価できる。しかし、CDS に基づく規範的な視点を採用することで、これらのモデルがどの程度まで価値の再分配を助長しているのかを問える。
- 故に、我々は、より公正なデータ環境のために「推奨」や「促進」をされるべき単独のモデルを見つけることはできなかったが、むしろ、これら全ての組み合わせが構想されるべきである。特に、インターネット・ガバナンスの民営化や、一部のプラットフォームによる支配的モデルに反対するためには、市民社会や公的機関の創意工夫に富んだありのままの実践に目を向けることが望ましい。
- 議論する上で重要な側面は、これらのモデルがデータガバナンスをどの程度民主化しているかということである。この問いに答えるために、データ主体を巻き込んだ以下の3つのモデルに注目する。

- ① 個人データ空間のエコシステムからなる商業サービス
 - ② データを民主的に管理し、責任を共有することを可能にする協同組合（DC）
 - ③ （彼らの）データを倫理的かつ公益のために利用することを市民から委託された公的機関（PDT）
- いずれの場合も、データ主体は自分たちのデータのために信頼できる仲介者を選ぶことができる。データのガバナンスに主体を関与させることは、データの監視、ファンクションクリープ⁸、テクノクラティック⁹なガバナンスなど、データガバナンスがもたらしうる負の結果の多くに対処し、それを回避するための重要な戦略である。
 - データガバナンスモデルでは、データ主体の力が強ければ強いほど、データ保有者にはより大きな説明責任が求められ、その結果、リスクやデータの誤用が制限されることになる。
 - 反対に、DSP はデータ保有者にしかアクセスできない。このモデルでは、データ主体（一般市民や疎外されたグループ）のニーズや利益が説明されることをどのように保証するのか？この問題に対処し、優れたデータガバナンスを実現するためには、少なくとも原理的には説明責任の観点から、より多くの保証を提供する他のモデルと DSP を組み合わせることが望ましいかもしれない。
 - 本稿の知見は、データ仲介者とデータ・トラストの概念がすべてのモデルに何らかの形で含まれているため、同じ「バズワード」であっても、データガバナンスの異なる合理性に関連付けられる可能性があることを強調している。このことは、単なる道具的・技術的な側面を超えて、社会技術的な産物としてのデータインフラストラクチャを批判的に考えることがいかに重要であるかを強調している。
 - データ・トラストは、もし DC で採用されれば、現在のデータ経済の力の不均衡を低減する強力な手段となるかもしれないが、DSP では全く異なる目的を促進するかもしれない。実際、前者の場合は、市民の利益と選好を代表して行動する「ボトムアップ型のデータ・トラストであるが、後者の場合は、データの共有を容易にするための条件とメカニズムの反復可能な枠組みである。逆に、データ・トラストは、PDT のように、信頼を獲得し、公益を促進するために公共部門により、トップダウンで提供されるサービスである可能性もある。

⁸ システムや技術が本来の目的を超えて拡大すること。プライバシー保護などの観点から問題視されている。

⁹ 技術者による支配。



- 最後の検討事項として、本稿が調査したデータの実践と、それらが存在する規制の枠組みとの間の絡み合った関係が挙げられる。これらのデータガバナンスモデルは、個人データに関する GDPR や、公益性の高い商業データへのアクセスを義務付ける新しい法律など、適切な法的枠組みに支えられて初めて、さらに発展できる。

制約

- この分野で確立された用語がないことや、テーマの性質が常に変化していることから、本稿では文献の選択が複雑になっており、欧米の文献に偏っている。
- さらに、記載されているモデルは、データのガバナンスのために現在開発されている新興モデルの全てを代表するものではない。理想的なタイプではあるが、今後の研究に採用すべきツールとしてのヒューリスティック¹⁰な価値がある。
- もう一つの制約は、本稿のリサーチクエストに由来する多次元の枠組みに関連している。利害関係者、ガバナンスの目的、データから得られる価値、ガバナンスのメカニズム、互惠性といった次元にのみ焦点を本稿では当てたが、例えば、信頼と不信感、プライバシーとデータ保護、権限と公的機関、法律と規制といった観点が採用されえた。
- 最後の制約として、従事者の文献の裏に隠されたものを調査するために収集した一次データを含まないことが挙げられる。本稿は、主に都市の文脈において、これらのモデルのいくつかの特徴とその実施のための課題を検討するための手段として、実証研究を通じてこの側面に実際に取り組んでいる。実際、この研究は、データガバナンスのさまざまな顕在化の中で、社会技術的な集合体がどのように展開しているのかについて、将来の調査の基礎を築くことを目的としている。
- これらのモデルは実際にどこまで機能するのか？異なる文脈でどの程度まで再現可能なのか？主な実現要因と障害となるものは何か。この種のミクロ的分析は、どのようなメカニズムがそれぞれの特定の文脈で用いられているのか（あるいは欠けているのか）、そしてどのような戦略的実践が不可欠かを認識する手助けをする。

¹⁰ 必ず正しい答えを導けるわけではないが、ある程度のレベルで正解に近い解を得られること。

結論

- 本研究は、多くの主体が、支配的なデータガバナンスモデルに代わるものを模索していることを示しており、その過程で4つのモデルを特定した。

(本稿の意義)

- データへのアクセス、共有、制御、利用に関する社会的実践と、そこから派生したモデルは、データガバナンスとデータ主権に関する議論が起これり、AI が社会の未来のための政策の戦略的な分野として登場し、欧州で新たな機運を醸成しつつある今、非常に重要な時期に来ている。
- 故に、データガバナンスの取り決めに迅速かつタイムリーに検討し、公共の利益のためにデータの利用を奨励する方法を理解することが重要である（そして今後とも重要になる）と考えている。
- そのためには、「インフラストラクチャ」を通じて見ることで以下の質問を投げかけることを可能にするデータガバナンスに関する社会科学的な視点を採用することも有用かもしれない。
 - データの共有と利用を導く原則は何なのか？
 - データを使って何が行われ、誰がそのガバナンスにアクセスし、参加できるのか？
 - どのような価値が生まれ、どのように再分配されるのか？
- 最終的に、本稿では社会技術的なイマジナリーと「望ましい未来」について、より多くの「規範的な対話」を促したいと考えている：私たちが生きたい社会の種類と、それに応じてDXをどのように形成できるか。