

データ専門分科会 検討報告

2020年6月2日

- 1. データ専門分科会の検討枠組み**
- 2. データの価値測定手法**
 - 2-1. データの機能・役割**
 - 2-2. 測定手法の整理**
 - 2-3. 実証分析**
- 3. データの効果・価値に応じた正当な報酬のあり方**

1. データ専門分科会の検討枠組み

主査	大橋 弘	東京大学大学院公共政策大学院・経済学研究科教授【AI経済検討会構成員】
主査代理	立本 博文	筑波大学ビジネスサイエンス系教授【AI経済検討会構成員】
構成員	荒井 ひろみ	国立研究開発法人理化学研究所革新知能統合研究センター研究員
	伊藤 萬里	青山学院大学経済学部経済学科教授(第6回以降)
	生貝 直人	東洋大学経済学部総合政策学科准教授
	高口 鉄平	静岡大学大学院情報学領域准教授
	高崎 晴夫	株式会社KDDI総合研究所フューチャーデザイン1部門研究員
	田原 健吾	公益社団法人日本経済研究センター主任研究員
	平井 祐理	東京大学未来ビジョン研究センター特任助教
	渡辺 智暁	国際大学グローバル・コミュニケーション・センター主幹研究員・教授

第1回(2019年12月27日)

- ① 運営方針
- ② 今後の検討事項及び検討の進め方等
- ③ A I 経済に関する現状等
- ④ データ価値測定手法に関する先行研究（高崎構成員）
- ⑤ データの価値測定手法の方針（案）（高口構成員）
- ⑥ データの価値測定のためのアンケート調査項目（案）

実証分析のために必要となる
アンケート調査項目

第2回(2020年1月20日)

- ① 今後の検討の進め方
- ② データの機能・役割の検討
- ③ データの価値測定に係る実証分析の進め方①
- ④ データの価値測定のためのアンケート調査項目（案）

実証分析の方向性の確認

第3回(2020年1月27日)

- ① データの価値測定のためのアンケート調査項目 (案)

第4回(2020年2月4日)

- ① データの価値測定手法及びその課題の整理①
 - ② データの価値測定に係る実証分析の進め方②
- 実証分析以外の手法についての課題整理
実証分析の具体的作業の検討

メールにて御確認

- ① AI経済検討会への中間報告 (案)

第5回(2020年3月31日)

- ① データの価値測定のためのアンケート調査結果(概況) } アンケート結果の全体像を大括りに把握
- ② データの価値測定に係る実証分析作業① } ①を踏まえた作業の方向性を共有
- ③ データの効果・価値に応じた正当な報酬のあり方① } 事務局整理の論点紹介

第6回(2020年4月16日)

- ① データの価値測定のためのアンケート調査結果(確定値) } ②の基礎となるアンケート結果情報を整理
- ② データの価値測定に係る実証分析作業② } ①を踏まえた作業の進捗状況を共有し、議論
- ③ データの効果・価値に応じた正当な報酬のあり方② } 論点に関する意見交換

第7回(2020年4月27日)

- ① データの価値測定に係る実証分析作業③
- ② A I 経済検討会への検討報告(案)

第8回(メール審議)(2020年5月8日~12日)

- ① データの価値測定に係る実証分析作業④
- ② A I 経済検討会への検討報告(案)

※その他、有識者(大手電機メーカー、情報銀行事業者、学識経験者)へのヒアリングを実施

2. データの価値測定手法

2-1. データの機能・役割

- AIの社会実装が本格化する時代（AI時代）においては、データが価値を生み出す源泉となり、生産性向上に寄与する重要な役割を担うことが考えられる。
- しかしながら、データ活用によってもたらされる価値の程度やそのプロセスなど、実態は極めて分かりづらいものとなっている。そのため、データ活用の社会・経済に対するインパクトや、当該インパクトにおいて制度・技術がもたらす影響の大きさ等について、説明が困難な状況である。
- 以上を踏まえ、各産業・各国等に適用し得る汎用性・再現性を確保することを念頭に、データの価値測定手法を検討していくことが必要である。同検討に当たっては、測定対象とするデータ及びデータの効果を明確にすると共に、データの機能・役割について理解する必要がある。

（1）測定対象

測定対象	検討事項
データ	<ul style="list-style-type: none">● 本分科会におけるデータは、（A I 経済検討会と同じく）「現にデジタル化されているか、または容易にデジタル化することができる状態のもの」と定義。● データは様々な側面から分類可能であることに留意しつつ、採用する価値測定手法を踏まえて対象を検討することが必要。● データは種類によって単位が異なる（Byte、人、件、分等）ため、複数のデータを扱う際には単位に注意することが必要。
データの効果	<ul style="list-style-type: none">● データに価値を付加するためには、処理と分析が必要。その結果として、アイデアやイノベーションの向上、予測精度の向上などが期待され、更に、それらの向上を通じて社会全体（CO₂削減等）、企業（売上高、利益等）、個人（効用等）への効果が期待される。● 様々な効果が考えられるため、どの効果を計測対象とするのかを明確にする必要。● 対象範囲と結果の解釈しやすさはトレードオフの関係（個別効果のみを対象とすると、結果を解釈しやすいものの、全体を把握できない。一方、広く全体を対象とすると、結果の解釈が困難になる。）。

(2) データの機能・役割

データの機能・役割	検討事項
データが価値を生むために必要な補完的要素	<ul style="list-style-type: none"> • データを活用する主体は様々であり、それぞれの主体が同じデータを活用した場合、同一の価値が付加されるとは限らない。 • データに価値を付加するために必要な補完的要素（人材、投資、組織、ビジネス形態、市場環境等）を検討する必要。
データの価値の時間による変化	<ul style="list-style-type: none"> • 単独のデータにおいては、データが示す時点（今日のデータ、1か月前のデータ、10年前のデータ等）によってビジネスにおけるデータ活用の効果が異なる。一方、データを組み合わせて活用する場合は、古いデータが存在することによって予測の精度等が向上する可能性。 • 活用用途や分析方法によってデータの価値の時間による変化は異なる可能性があり、どのように取扱うかを検討する必要。
負の経済的価値	<ul style="list-style-type: none"> • データを保有することによって、セキュリティ対策などの管理コストや情報漏洩によるイメージ低下等マイナス面のリスクが発生する（負の経済的価値）。 • データの価値計測時にこれらの負の経済的価値をどこまで考慮するのかを検討する必要。
ICTとの関係	<ul style="list-style-type: none"> • データ分析にはICT（ハード、ソフト）が必要であるため、データの価値は広く捉えるとICTの価値に置き換えて考えることもできるという見方もある。 • データの価値とICTの価値をどこまで明確に区別するのか検討する必要。

データの機能・役割	検討事項
無形資産との関係	<ul style="list-style-type: none"> • データ（の一部）は無形資産に含まれるものとして捉えることができ、両者に共通する特性もあるため、機能・役割等に係る類似点・相違点を考慮しつつ、無形資産に関する分析手法を参考にすることが有用である可能性がある。 <p>【類似点と相違点】</p> <ul style="list-style-type: none"> • 無形資産の内、市場取引されるソフトウェアなどは資産として計上されているが、データや技術力、ノウハウ、企業文化など多くは資産計上されていない。 <p>※国連が公表する2008SNAの指針では、ソフトウェアとともにデータベースについても固定資産（知的財産生産物）に含めることを推奨している。我が国の国民経済計算（JSNA）では、ソフトウェア関連品目に属するようなデータベースについてはソフトウェアと一体のものとして固定資産に扱われている。</p> <p>【アナロジーの適用】</p> <ul style="list-style-type: none"> • 特許の企業パフォーマンスへの影響などについて先行研究があり、データの価値を計測するうえで参考とすることが有用か検討する必要。

先行研究論文等	データの機能・役割
Avi Goldfarb and Catherine Tucker (2017) "Digital Economics," <i>NBER Working Paper</i> , No.23684.	<ul style="list-style-type: none">デジタル技術によってデータの保存、計算、転送コストが削減され、その結果として、以下に掲げるコストも削減される。※データの価値よりもデジタル技術の価値に近い。<ul style="list-style-type: none">(1) 探索コスト：情報を検索するためのコスト(2) 複製コスト：品質を低下させることなく、ゼロに近いコストで誰でも複製することが可能(3) 輸送コスト：ネット経由での転送コスト。データの流通コストは0に近い(4) 追跡コスト：個人等を追跡するコスト。ネット上のアクティビティは簡単・自動で記録できるため、過去の行動に基づいた個別化が容易になる(5) 検証コスト：身元や評判の検証に関連するコスト
Yan Carriere-Swallow and Vikram Haksar (2019) "The Economics and Implications of Data : An Integrated Perspective," <i>IMF Departmental Paper</i> , No.19/16.	<ul style="list-style-type: none">データには公共政策にとって重要な3つの経済的特性がある。<ul style="list-style-type: none">(1) 非競合性 同じデータを多くの人々が利用できる。新しいアイデア等、データが広く共有されれば、より多くのユーザがデータを利用して効率を高め、革新を起こせるため、社会はデータから最も恩恵を受けることになる。しかし、実際にそうなるかどうかは、ポリシーと個人の意思決定によって左右される。 現在の政策では、民間企業が競合他社に自社が収集したデータへのアクセスを許可するインセンティブを持っていない可能性が高い。そのため、市場での競合性やデータから得られる社会的利益が制限されている可能性がある。(2) 外部性 あるエージェントによる個人データの収集、共有及び処理は、他者のプライバシーに影響を与えることによって、他者にコストを課す。つまり、データ収集者が収集したデータを好きなように扱うデータ市場では、過剰なデータ収集とプライバシーの欠如につながり、非効率が発生する可能性がある。(3) 部分的排除性 データは部分的にしか排除できないため、相互接続されたシステム上にデータを保存するということは、サイバー攻撃によるデータの損失を防ぐために継続的な投資が必要であることを意味する。

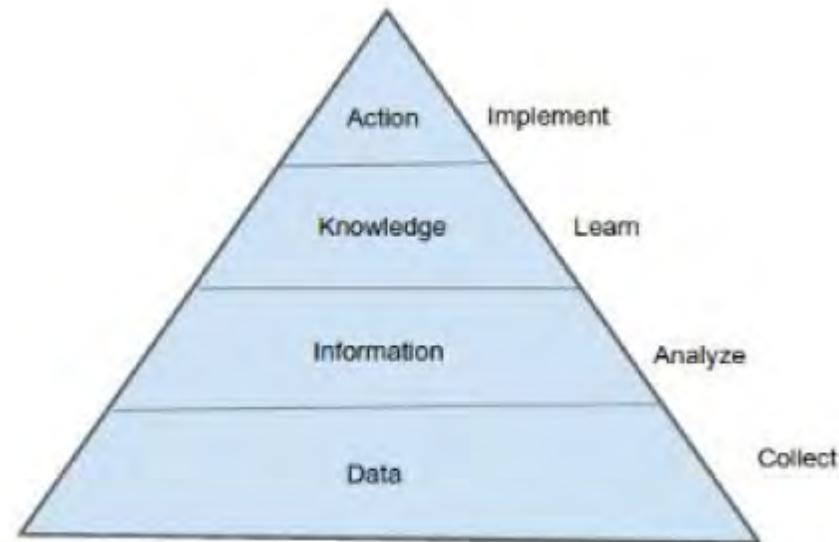
先行研究論文等	データの機能・役割
<p>Yan Carriere-Swallow and Vikram Haksar (2019) "The Economics and Implications of Data : An Integrated Perspective," <i>IMF Departmental Paper</i>, No.19/16.</p>	<ul style="list-style-type: none">● 需要面からみたデータの経済的価値（データの2つの役割）：<ul style="list-style-type: none">（1）モノやサービスの生産にインプットされ、イノベーションや効率化に貢献 財やサービスの生産における入力としての機能。データは知識の創造を可能にし、それは既存の財の継続的な生産又は新しい製品・サービスの開発に向けられる。 入力としてのデータから価値を引き出すには、コストのかかる処理と分析が必要であり、それは補完的な熟練労働者によって提供される。（2）経済主体全体で情報を創造及びシフトし、戦略的相互作用に影響 データには、消費者や企業を含む経済主体に関する情報が埋め込まれているため、このデータにアクセスすることで、これらの主体が参加する市場における情報の非対称性が変化する。データへのアクセスが買い手と売り手の間の情報の非対称性を減らすのに役立つとき、より効率的な経済取引につながる。 例えば、潜在的な消費者の興味や購買習慣などの特性に関するデータにアクセスできる売り手は、より個人々々人に向けた商品やサービスを提供することができる。同様に、潜在的な製品の特性に関するデータを持つ消費者は、より多くの情報に基づいた購買決定を行うことができる。 <p>データの不均一性： サブジェクト、タイミング、フォーマット、品質など、多くの次元で不均一であり、これらは経済的特性に影響を与える。</p> <p>データの収集コスト： データの収集量が増えるにつれて、データ収集にかかる平均コストが減少（規模の経済が生じる可能性）。</p>

先行研究論文等	データの機能・役割
<p>Yan Carriere-Swallow and Vikram Haksar (2019) "The Economics and Implications of Data : An Integrated Perspective," <i>IMF Departmental Paper</i>, No.19/16.</p>	<ul style="list-style-type: none">データの価格を決める要素： (1) データの中身 データの価格は均質ではなく、誰が、いつ、どこで収集したのか、データ構造がどのようになっているか、他のデータと統合できるかなど、多数の属性に基づいて決まる。また、時間の経過とともに大きく変化する。(2) 市場環境 顧客や競合他社との情報の非対称性がどの程度存在するか、市場における影響力の程度、市場の規模など、多くの要因によって決まる。 個人データの価値を評価することは、直接的なインセンティブを持つ場合でさえ、かなり困難である。例えば、広告主は、ターゲットを絞った広告を表示することで売上げが増加するという前提のもと、オンラインユーザに関する個人データ等に多額の資金を投じている。しかし、ターゲティング広告から得られる利益は統計的に有意であるように見えるが、売上への因果的な影響は小さく、ターゲティングに費やされる支出よりも劣る (Marotta, Veronica, Vibhanshu Abhishek, and Alessandro Acquisti. 2019. "Online Tracking and Publishers' Revenues: An Empirical Analysis".)。
<p>Charles I. Jones, Christopher Tonetti (2019) "Nonrivalry and the Economics of Data," <i>NBER Working Paper</i>, No.26260.</p>	<ul style="list-style-type: none">データの非競争性から同じデータを複数の企業で使用することが可能であり、それによって大きな社会的利益がある。ただ、企業は創造的な破壊に対する懸念のため、データを販売することをためらうかもしれない。データは消費の副産物であり、消費されるたびに新たなデータが作成され、生産性の向上に役立つ。データは、アイデアの品質と考えることができる。データは生産要素である一方、新しいアイデアは生産のための新しい機能である。アイデアとデータは情報の一種である。アイデアは経済的利益を生み出すための一連の命令であり、情報の一部であり、他のアイデアを含む場合がある。データは情報のその他の形式を示す。データ自体は良いものを作るための命令ではないが、新しいアイデアの作成を含む生産プロセスに役立つ可能性がある。アイデアは生産関数であり、データは生産要素である。

先行研究論文等	無形資産の機能・役割
<p>Corrado, C., C. Hulten, and D. Sichel (2006) "Intangible Capital and Economic Growth," <i>NBER Working Paper</i>, no.11948. Cambridge, MA: National Bureau of Economic Research.</p> <p>Corrado, C., C. Hulten, and D. Sichel (2009) "Intangible Capital and U.S. Economic Growth," <i>Review of Income and Wealth</i>, 55, pp.658-660.</p>	<ul style="list-style-type: none">• 無形資産の分類（一例）：<ul style="list-style-type: none">(1) 情報化資産 ソフトウェア、データベース(2) 革新的資産 科学・工学分野における研究開発、資源探索権、著作権・ライセンス等、他の商品開発・デザイン等(3) 経済的競争力 ブランド資産、企業特殊的人的資本、組織変革
<p>Jonathan Haskel, Stian Westlake (2018) "Productivity and secular stagnation in the intangible economy".</p>	<ul style="list-style-type: none">• 無形資産の経済的性質：<ul style="list-style-type: none">(1) 限界費用が低い 無形資産による場合は有形資産による場合よりも容易に事業規模の拡大ができる。 (例) Uberは既存の無形資産（ソフトウェア）でより多くの顧客にサービスを提供できるものの、タクシー会社は規模を拡大するためには、より多くの有形資産（車）を購入する必要がある。(2) 波及効果が大きい 無形資産は広く共有できる。他者の有形資産は利用できないが、他者の無形資産は潜在的に利用可能である。(3) サunkコスト（埋没費用）が大きい 多くの無形資産は転売することが困難である。(4) 相乗効果（シナジー効果）が大きい 無形資産は、他の無形資産と組み合わせることで大きな価値を生み出すことが可能である。 (例) iPhoneは、R&Dとデザイン、サプライチェーンにおける組織デザインとブランドが組み合わせられることにより、収益性の高い製品となった。

- 情報科学では、データ、情報、知識、行動というデータピラミッドの概念が存在する。
- データを収集し、そのデータを整理・分析することによって情報に変換され、その情報からの洞察が具体化されて知識に変わる。
- データはビット単位、情報はドキュメント、知識は人に蓄積される。
- 情報（書籍、記事、Webページ、音楽、ビデオ）と知識（労働市場、コンサルタント）に関しては十分に発展した市場と規制環境が存在する。一方、データ（未整理のビットの集合）の市場は発展していない。これは、データがコンテキスト（文脈）に大きく依存することが多く、情報に変換されるまであまり役に立たないためである。

図表1 データピラミッド



2-2. 測定手法の整理

データの価値測定手法

● データ（又はデータに関係するもの）の価値を測定するためのアプローチは3つに大別される。

アプローチ	考え方	分類	課題	論点	主な先行研究
コストベースのアプローチ	<ul style="list-style-type: none"> データの作成、管理、活用に関わるコスト（人員、時間等）に着目して算出する方法。コストをかけるだけの（コストと等しい）価値があるという仮定に基づいている。 	<ul style="list-style-type: none"> 人員コストの算出 時間コストの算出 セキュリティ対策コストの算出 	<ul style="list-style-type: none"> 品質や生産性の違いを考慮できない。 ICT機器等の価格や性能にも左右されるため、データの価値が別の要素で変化する可能性がある。 コストと価値の関係が直観的にはイメージしにくい。 	<ul style="list-style-type: none"> どの側面のコストをデータの価値として捉えるのが適切なのか 	<ul style="list-style-type: none"> Klein, Thilo (2016) "Monetizing the Value of Official Statistics: A case study". UNECE(2018)"Recommendations for Promoting, Measuring and Communicating the Value of Official Statistics".
マーケットベースのアプローチ	<ul style="list-style-type: none"> 類似製品の市場価格又はユーザの支払う意思に基づいて価値を算出する方法。 	<ul style="list-style-type: none"> 市場価格（データ売買等） 表明選好法 企業買収、損害賠償額等の事例分析 実験的取組 	<ul style="list-style-type: none"> 測定に用いるためのデータが不足している。 表明選好法の場合、調査対象ユーザの理解に左右される。 企業視点、消費者視点によって価値が異なると予想される。 	<ul style="list-style-type: none"> 個別データではなく、どのようにデータ全般を対象として計測するのか 企業視点、消費者視点のどちらでデータの価値を捉えるのが適切なのか 	<ul style="list-style-type: none"> UNECE Task Force on the Value of Official Statistics (2017) "Value of official statistics: Recommendations on promoting, measuring and communicating the value of official statistics".
インパクトベースのアプローチ	<ul style="list-style-type: none"> データを活用することによる効果（生産性、収益等）を推定することで価値を算出する方法。 	<ul style="list-style-type: none"> 計量分析（生産関数分析、回帰分析等） 実験的取組 	<ul style="list-style-type: none"> 統計的な推定が必要であり、「データ」の効果とそれ以外の要素の効果を明確に分離することが難しい。 	<ul style="list-style-type: none"> データの価値計測対象（データの定義、価値の定義）を明確化 データが価値を生むために必要な補完的要素等について検討した上で、どのようにして適切な分析モデルを構築するか 	<ul style="list-style-type: none"> Erik Brynjolfsson, Kristina McElheran (2016) "Data in Action: Data-Driven Decision Making in U.S. Manufacturing". Thomas Niebel, Fabienne Rasel, Steffen Viète (2017) "BIG Data – BIG Gains? Empirical Evidence on the Link Between Big Data Analytics and Innovation". LI, Wendy C.Y., NIREI, Makoto and YAMANA Kazufumi (2019) "Value of Data: There's No Such Thing as a Free Lunch in the Digital Economy".

- 入力データから得られる情報や知識は、通常、市場取引の対象とはならないため、様々な推計手法が提案されている。
- ただし、それぞれの手法には課題が存在する。

アプローチ	考え方	課題
マーケットベースのアプローチ	<ul style="list-style-type: none">● 市場における比較可能な製品の市場価格に基づいて評価	<ul style="list-style-type: none">● 市販されているサードパーティのデータベースを除き、市場で販売されている真に比較可能な製品が存在しない。● 「市販されているような未処理のデータ」と「整理、クリーニング、使用に適した加工が行われたデータ」は価値が異なり、市場取引価格を採用することは本来の価値とは異なる可能性。
コストベースのアプローチ	<ul style="list-style-type: none">● データから得られる情報やノウハウを生み出すコストによって評価	<ul style="list-style-type: none">● コスト範囲の設定（データサイエンティストのコストも考慮するなどして、データ資産全体に拡張する必要性）。● プラットフォームのコストをデータ取得コストとして扱う場合、プラットフォームがユーザから得るそれ以外の利益を無視する可能性。
インカム（収入）ベースのアプローチ	<ul style="list-style-type: none">● データから得られる将来のキャッシュ・フローを見積もることによって評価	<ul style="list-style-type: none">● データから期待される収入の割引現在価値（将来に受け取れる価値が、もし現在受け取れるとしたらどの程度の価値をもつかを表すもの）は、用途に応じて、コストベースのアプローチとは異なる価値を表す可能性。● 2008SNAでは、資産の評価に収入を使用する際には注意が必要（資産の耐用年数、将来キャッシュ・フロー、割引率に関する適切な仮定を決定することが困難）としている。● ほとんどの場合、プラットフォームの収益は、技術的およびビジネス上の能力、顧客との関係、ネットワーク外部性、イノベーションなど、その他の有形および無形の資産に帰属される。

(参考) デューデリジェンスにおける企業価値評価、無形資産評価の手法 20

- 企業の合併や買収（M & A）を検討する際の企業価値評価の手法は、大きく3つのアプローチに分類され、無形資産評価も、同様に3つのアプローチから成る。
- しかしながら、各アプローチについては、データの持つ価値やデータに関連するリスクを評価することの必要性、評価する際の視点についての言及・指摘はあるものの、具体的な評価手法については明示されていない。

	企業価値評価	無形資産評価
目的	企業がM&Aや組織再編等の「企業結合」に該当する取引を行うに当たって、価格決定の参考とするため	企業結合会計への情報提供
ネットアセット・アプローチ (コスト・アプローチ)	主として会社の貸借対照表上の純資産に注目して評価するアプローチ。	同等の資産を受け入れるのに要するコストをもって評価する方法
マーケット・アプローチ	上場している同業他社や類似取引事例などと比較することによって相対的に価値を評価するアプローチ。	同等の資産が市場で実際に取引される価格をもって評価する方法
インカム・アプローチ	評価対象会社から期待される利益、ないしキャッシュ・フローに基づいて価値を評価するアプローチ。	同等の資産を利用して将来における期待される収益をもって評価する方法

出典：日本公認会計士協会「企業価値評価ガイドライン」、日本公認会計士協会「無形資産の評価実務－M&A 会計における評価とPPA 業務－」

- M & A 等で考慮すべきデータの持つ価値やデータに関連するリスク

データの持つ価値	法的側面のリスク	技術的側面のリスク	その他の側面のリスク
<ul style="list-style-type: none"> ● 資産としてのデータの価値 ● 使用することによるデータの価値 ● 期待されるデータの将来価値 	<ul style="list-style-type: none"> ● データの転送又は使用を制限する法規制の有無 ● 契約上の利用制限の有無 ● データに関連する重大な責任の有無 	<ul style="list-style-type: none"> ● サイバーセキュリティリスク ● データストレージの形式 ● データの管理方法 ● データの品質 	<ul style="list-style-type: none"> ● 人材の流出リスク ● 知財リスク ● データ保護/プライバシーポリシー

- 標準的な統計分析においては、一定の観測数で推計できると、その後のデータの限界価値は急激に逡減する。(例：所得の平均は百万円レベルで把握できればよく、十万円、一万円レベルで把握する必要はない)
- 機械学習においては、取り組むタスクが複数あり、問題解決に観測対象の複雑性が求められるタスクであるほど、データの価値が増加する。また、特定のタスクの中で考えると、標準的な統計分析と同じく、一定のデータ量に達するとデータの限界価値は逡減する。

図表2 標準的な統計分析におけるデータの価値

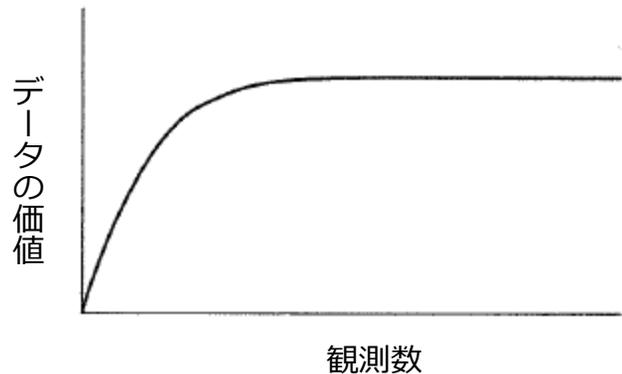


FIGURE 5.4: The value of data as a function of the number of observations in a standard statistical estimation problem. The marginal value declines rapidly. Thanks to Nicole Immorlica for providing this graph.

図表3 機械学習におけるデータの価値

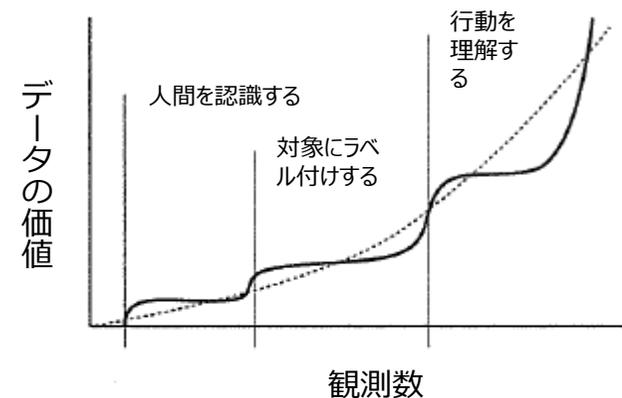
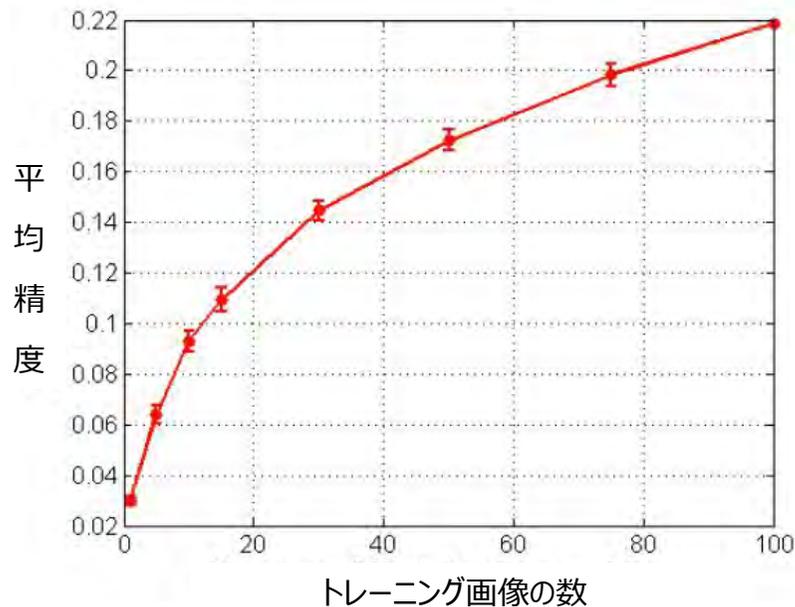


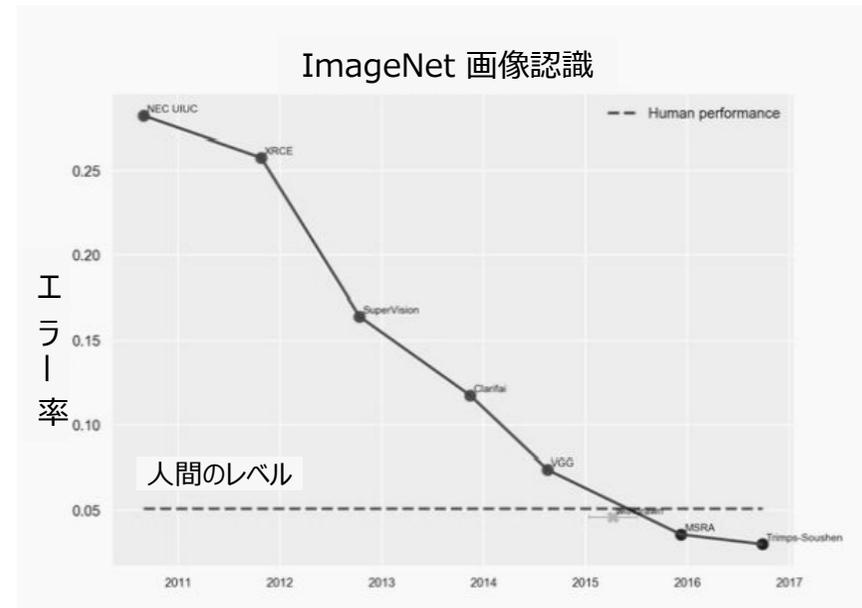
FIGURE 5.5: The value of data as a function of the number of observations in a typical ML domain, here machine vision. Each vertical line represents the sample complexity of a particular problem. Thanks to Nicole Immorlica for providing the graph.

- 一般的に、データは他の生産要素と同様に規模の収穫逓減となり、これは機械学習にも適用される。
- 学習データ量が増えるにつれて、平均精度が向上するものの、その向上率は徐々に小さくなる（図表4）。
- 近年の画像認識コンペティション（期間中において学習データ量は固定されている）の結果をみるとエラー率が大幅に低下しており、学習データ量よりもアルゴリズムの改善、ハードウェアの改善、専門知識の改善など他の要因がはるかに重要である（図表5）。

図表4 学習データ量と平均精度

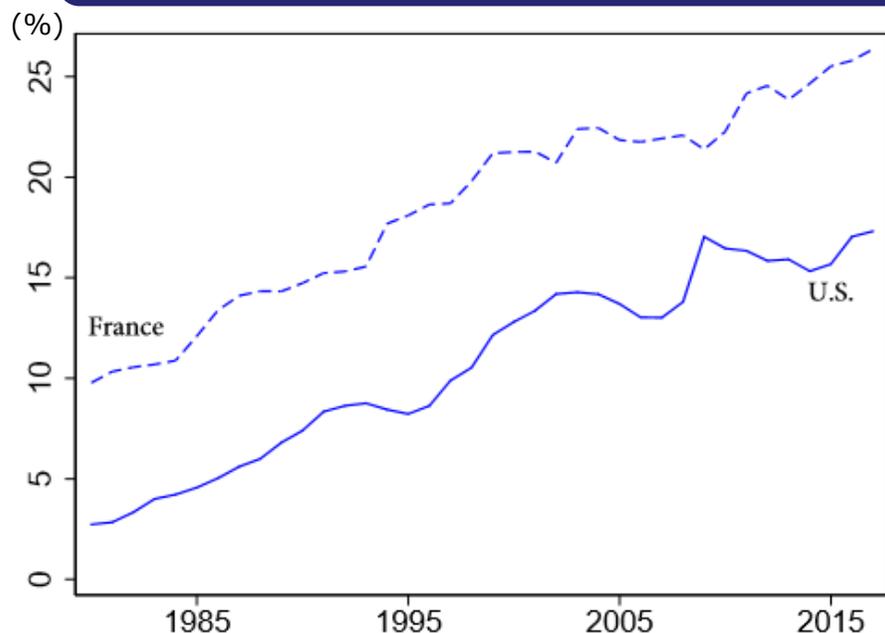


図表5 画像認識エラー率の推移



- 米国、仏国におけるソフトウェア投資の割合は増加しており、仏国の民間企業では総投資の26%を占めている（図表6）
- 無形資産は開発と維持に投資が必要なものの、ほぼゼロの限界費用で複製できるため、生産の拡大にあたりこれらの無形資産への追加費用を最小限に抑えることができる。したがって、無形資産が増加すると、生産コストが変動費から固定費にシフトする。
- 限られた「ITスーパースター」が無形資産投資を効率的に展開できる場合、限界費用を大幅に削減できるため、商品を低価格で販売することが可能になる。無形資産投資が少ない企業（イノベーター）が、「ITスーパースター」が販売する商品を高品質化したものを開発した場合、「ITスーパースター」は価格戦略によってイノベーターを弱体化し、市場でのリーダーシップを維持することができる。無形資産投資が多い企業の存在は、参加者が高品質の製品を開発することを阻害する力を持つ。

**図表6 ソフトウェア投資の増加
(総投資に占めるソフトウェア投資の割合)**



注：住宅投資、研究開発、娯楽を除く民間部門の固定投資総額に占めるソフトウェア投資の割合。米国は、米国商務省データ、仏国は、EU-KLEMSデータ（データベース投資を含む）。

図表7 無形資産の増加前後における定常状態の比較

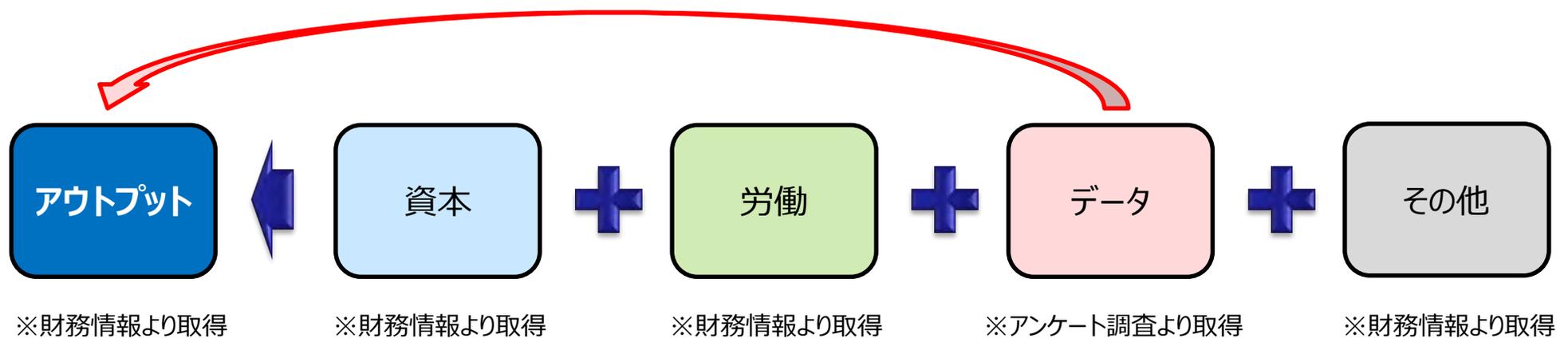
無形資産投資の増加を固定費シェアの増加（4.5 p.p.）、新規参入率の低下（3.6 p.p.）として捉えると、生産性成長率は1.7%から1.1%まで低下する。これは、一部の企業がより高い効率で無形資産投資を展開できる場合、長期的な生産性成長率が低下することを示している。

Variable	Previous St. State	New St. State	Change (model)	Change (data)
<i>Targeted</i>				
Avg. fixed cost share	10%	14.5%	4.5 p.p.	4.5 p.p.
Entry rate	10.5%	6.9%	3.6 p.p.	3.6 p.p.
<i>Untargeted</i>				
Productivity growth	1.7 %	1.1 %	-0.6 p.p.	-1.5 p.p.
Reallocation rate	33%	20%	-38%	-23%
Avg. markup	1.19	1.33	0.14 p.p.	0.11 p.p.

2-3. 実証分析

- 企業へのアンケート調査を通じてデータの保有・利用状況を把握し、企業の財務情報とデータの保有・利用状況との関係性を分析した上で、企業のアウトプットを生み出す要素として「資本」、「労働」、「データ」、「その他」を位置付け、実証分析を通じてデータの保有・利用とアウトプットとの関係性を分析した。
- データの機能・役割の整理や各産業・各国等に適用し得る汎用性・再現性の確保に留意しつつ、データの価値を計測する現状とり得る実現可能な手法として、インパクトベースのアプローチである生産関数の推定を行った。

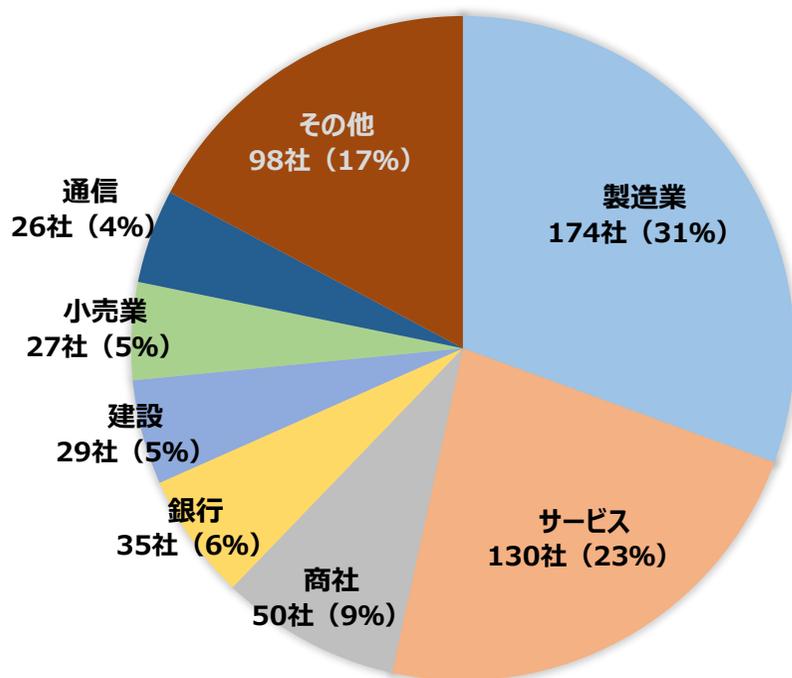
データの保有・利用とアウトプットとの関係性を分析



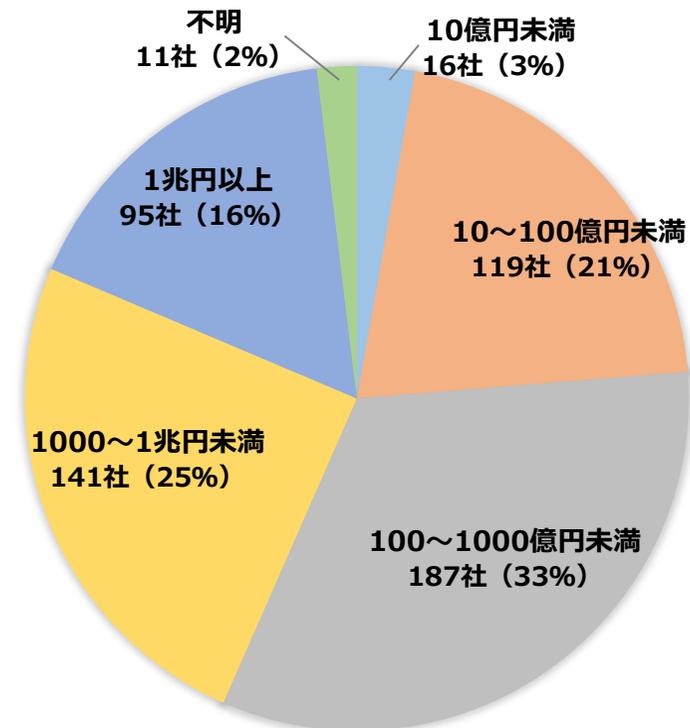
企業向けアンケート調査(郵送調査)

- ① 調査対象：全上場企業（3,819社）+非上場企業（467社）：合計4,286社
- ② 調査時期：2020年2月7日（金）～3月27日（金）
- ③ 回収数：569回答（回収率13.3%）

図表8 業種別

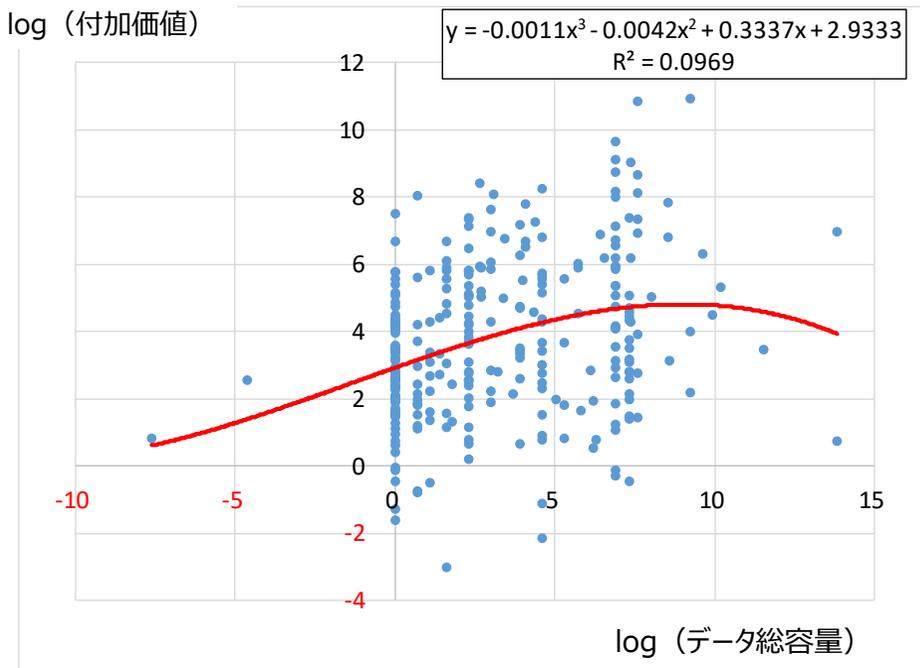


図表9 総資産別

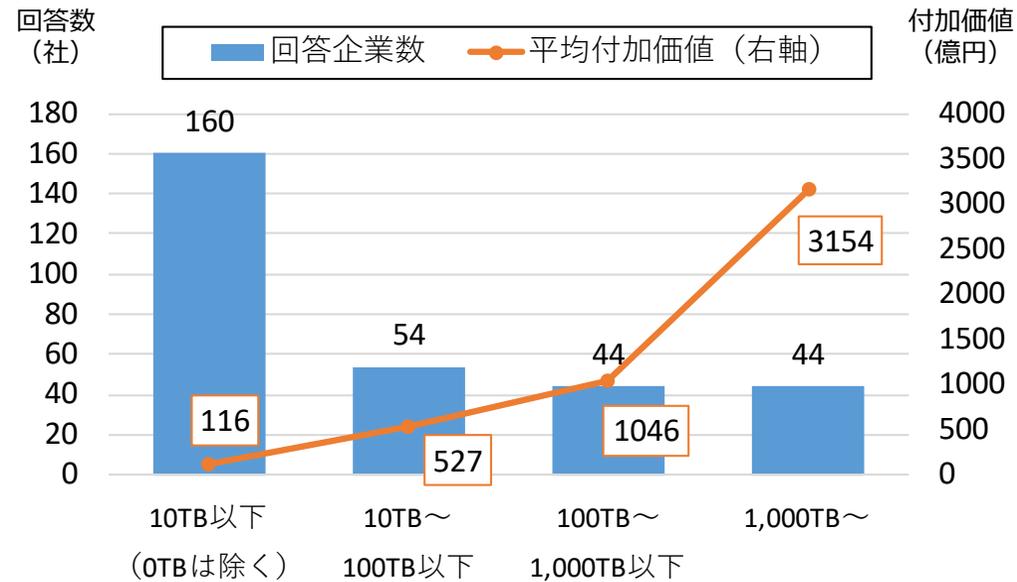


- 「付加価値 (※)」とアンケート調査で確認した企業が蓄積している「データ総容量」との関係性を分析した。(※)「売上高」-「売上原価」-「販売費及び一般管理費」+「減価償却費」
- その結果、データ総容量が大きい企業では付加価値も大きいという関係性が見られ、データが生産活動において正の影響が示唆される。

図表10 対数値をとったデータ総容量と付加価値との関係



図表11 データ総容量と付加価値との関係

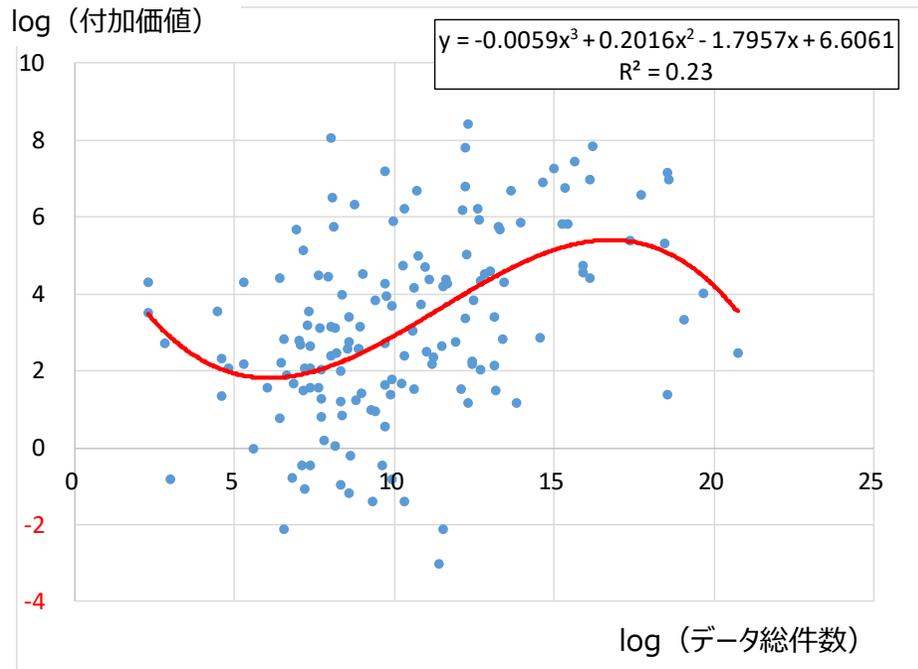


- 同様に、「付加価値」とアンケート調査で確認した企業が蓄積している「データ総件数 (※)」との関係性を分析した。

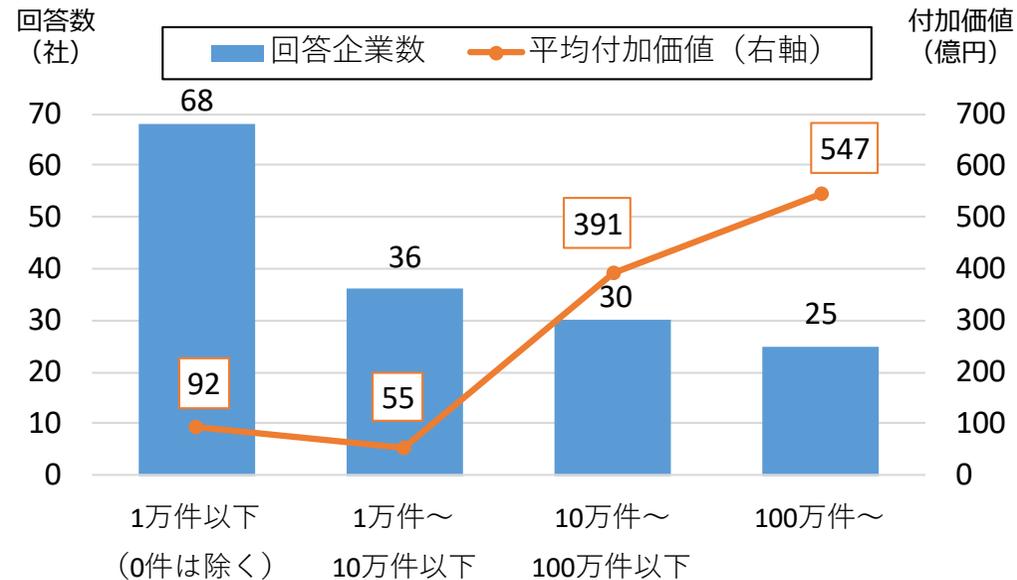
(※) データの対象数 (人・社・者・台数) の合計

- その結果、データ総容量と同様に、データ総件数が大きい企業では付加価値も大きいという関係性が見られた。

図表12 対数値をとったデータ総件数と付加価値との関係



図表13 データ総件数と付加価値との関係



- 資本、労働、データを生産要素とした下記の生産関数を推定した。

$$V = A_0 K^\alpha L^\beta Data^\gamma, \quad \log(V) = \log A_0 + \alpha \log(K) + \beta \log(L) + \gamma \log(Data) + \text{業種ダミー}$$

※製造業、非製造業に分類

変数	定義	出所
V	付加価値 (「売上高」-「売上原価」-「販売費及び一般管理費」+「減価償却費」)	日経NEEDS-FinancialQUEST
K	有形固定資産 + 無形固定資産	日経NEEDS-FinancialQUEST
L	総従業員数	日経NEEDS-FinancialQUEST
Data	データ変数 (データの保有・利用状況を変数化)	アンケート調査

データ変数	n数	修正R ²	K (資本)	L (労働)	Data (データ)
活用データ容量 (=データ総容量×分析に活用するデータの割合)	258	0.8343	0.44 ◎	0.50 ◎	0.05 ◎
活用データ件数 (=データ総件数×分析に活用するデータの割合)	135	0.8157	0.34 ◎	0.55 ◎	0.07 ◎
活用データ容量 (2015年度)	258	0.8332	0.44 ◎	0.51 ◎	0.05 ◎
活用データ件数 (2015年度)	135	0.8136	0.34 ◎	0.56 ◎	0.06 ○
外部入手データ容量 (=データ総容量×外部から入手したデータの割合)	267	0.8401	0.47 ◎	0.47 ◎	0.05 ◎
外部入手データ件数 (=データ総件数×外部から入手したデータの割合)	140	0.8193	0.35 ◎	0.51 ◎	0.07 ◎
内部保有データ容量 (=データ総容量 - 外部入手データ容量)	267	0.8379	0.47 ◎	0.48 ◎	0.05 ○
内部保有データ件数 (=データ総件数 - 外部入手データ件数)	140	0.8134	0.35 ◎	0.55 ◎	0.06 ○
データ総容量×データ活用度 (活用領域・種類・処理方法の多さ)	261	0.8334	0.44 ◎	0.49 ◎	0.05 ◎
データ総件数×データ活用度	137	0.8150	0.32 ◎	0.58 ◎	0.05 ○
データ総容量×データの多様性 (データの入手・提供の多様性)	173	0.8537	0.53 ◎	0.42 ◎	0.04 ○
データ総件数×データの多様性	85	0.8576	0.41 ◎	0.49 ◎	0.06 ◎

※2015年度と記載のない変数は2018年度の値、外部から入手したデータの割合は、データの外部入手状況を踏まえて補正を行った。

(注) ◎:有意水準1%、○:有意水準5%、△:有意水準10%

● データの価値測定に係る実証分析結果から、主に以下のことが確認された。

1. 企業の保有するデータ容量・件数と付加価値とはある程度の関係性が見られ、データが生産活動において正の影響をもたらしていると推察できる。
2. 3年前の活用データ容量・件数を用いた分析において、活用データ容量・件数は付加価値に対してプラスの関係性を示しており、時間軸を考慮してもデータ活用が付加価値の増加に影響を与えていることが推察できる。
3. 実証分析の結果から、活用データ容量・件数は、他の生産要素（資本、労働）と同様に、付加価値に対してプラスの関係性を持っていること。また、関係性の程度は、活用データ容量・件数が1%増えると付加価値が0.05%増える程度であること。これは、現状のデータ活用の取組状況を反映したものであり、活用データ容量・件数を増やすこと自体が必ずしも付加価値増加と結びつくものではないことに留意。
4. 今後、継続的に情報を収集することができれば、今回の回答データも合わせた時系列データを構築し、更なる分析が可能になると期待される。

- データに関連するソリューションやデータの活用状況と、売上高や付加価値との関係を分析した先行研究は存在するものの、データ容量やデータの質などデータそのものに焦点を当てた分析は皆無である。

先行研究	Oliver Müller, Maria Fay & Jan vom Brocke (2018) "The effect of big data and analytics on firm performance: An econometric analysis considering industry characteristics." <i>Journal of Management Information Systems</i> , Vol.35-2, pp.488-509.	Erik Brynjolfsson and Kristina McElheran (2016) "Data in Action: Data-Driven Decision Making in U.S. Manufacturing." <i>US Census Bureau Center for Economic Studies Paper</i> , No.CES-WP-16-06.
目的	ビッグデータ分析 (BDA) 資産が企業業績に及ぼす影響を計測すること	データ主導型意思決定 (DDD) が企業業績に与える影響を計測すること
分析対象	米国株式市場で上場している企業814社	米国製造業の約18,000事業所
実証分析	<ul style="list-style-type: none"> 814社のパネルデータ (2008~2014年) 被説明変数は、売上高 説明変数は、労働 (従業員数)、資本、資材、BDA 資産、コントロール変数 (業種、年、IT資産) BDA変数は、ビッグデータ分析に関連する資産 (データベース技術、データマイニングソリューション、データ視覚化ツール) の有無を表す0-1変数 分析モデルは、コブ・ダグラス型の生産関数 	<ul style="list-style-type: none"> 約18,000事業所のパネルデータ (2005年、2010年) 被説明変数は、付加価値 説明変数は、資本、労働、消費エネルギー、IT資本ストック、DDD、マネジメント指標、従業員の学歴 DDD変数は、アンケート調査で把握したものであり、企業が経営上の意思決定を行う際のデータの活用状況等を表す変数 (0~10の値をとる) 分析モデルは、コブ・ダグラス型の生産関数
結果	BDA資産は企業の売上高を約4%増加させる。	DDD指標が1増加すると事業所の付加価値を1%増加させる。
課題	BDA資産がその有無を表す0-1変数として扱われており、分析ではデータ容量の大小やデータ活用の違いなどが検証できていない。	DDD指標はデータの活用状況を表したものであり、データ容量の大小やデータの質などによる違いは検証できていない。

(参考) アンケート調査項目と調査票設問

アンケート調査項目	調査票設問
(Q1)データの分析手法 [種類別×領域別]	貴社では、どのようなデータをどのように処理したものを各領域で活用していますか？当てはまるものすべてをお答えください。
(Q2)分析に活用するデータの量（期間） [種類別]	貴社では、データ分析を行う際、おおよそどれくらい前のデータまでを利用しますか？
(Q3)分析の頻度 [領域別]	貴社では、データ分析をどのくらいの頻度（間隔）で行いますか？
(Q4)データの品質に関連した取り組み	貴社ではデータを活用する際に、データの品質に関連したどのような取り組みを行っていますか？
(Q5)分析の体制、データ分析の専門部署が立ち上がった時期	貴社がデータ分析を行う際、どのような体制（部署・担当者）で分析（外部企業等が提供しているデータ解析ツール・サービスの利用も含みます）を行いますか？ データ分析を行う専門部署がある場合は、どのくらい前に当該部署が立ち上がったのかについてもお答えください。
(Q6)分析の人員数と学歴 ※現在、3年前	貴社で、データ分析を主な業務とする従業員（正社員だけではなく、データ分析に寄与するパートタイマー・アルバイト・契約社員・派遣社員等を含みます）はおおよそ何人くらいいますか？ また、その内、大学院卒（修士または博士の学位を有する）の学歴をもつ従業員の割合はどの程度ですか？
(Q7)分析の手法（AIの活用状況）	貴社では、データ活用において、どのようなAI技術を用いていますか？
(Q8)データ活用の各プロセスにおける効果への影響	データを活用することによる効果（企業活動に対する貢献）についてお伺いします。 貴社においては、データ活用に関わる各プロセスが効果にどの程度貢献していると思いますか。
(Q9)データの多様性（データの入手元、提供状況） [種類別]	貴社で蓄積（保有）されているデータについて、データの入手元はどちらですか？また、保有しているデータを必要に応じて加工した上で、他社等に提供または共有していますか？
(Q10)外部から購入したデータの割合	貴社で蓄積（保有）しているデータの内、外部から購入したデータ（無料で公開されているデータやアライアンス等による無償の共同利用のデータは含みません）はおおむねどの程度を占めていますか？データの容量ではなく、件数ベースでお答えください。

(参考) アンケート調査項目と調査票設問

アンケート調査項目	調査票設問
(Q11)データの蓄積量（容量割合、件数）[種類別]、データの総容量 ※現在	貴社で蓄積（保有）されているデータについて、2018年度末時点におけるおおよその構成割合（容量ベース）を合計が100%になる形でお答えください。 また、蓄積（保有）データの対象数（人・社・者・台数）を、有効数字2桁の概数（例：12,000、450,000）でお答えください。対象数が不明な項目については、お答えいただかなくても構いません。併せて、蓄積（保有）されている具体的なデータの種類について可能な範囲でお答え下さい。 上記でお答えいただいたデータについて、2018年度末時点における容量の総計を有効数字2桁の概数（例：12,000、450,000）でお答えください。正確な値を把握されていない場合は、想定される範囲の中間値（例：1,000～2,000TB程度と思われる場合は1,500TB）をお答えください。
(Q12)データの蓄積量（自社以外から購入・入手したデータの件数割合）[種類別] ※現在	貴社で蓄積（保有）されているデータについて、2018年度末時点で自社以外から購入・入手したデータ（無料で公開されているデータやアライアンス等による無償の共同利用のデータも含みます）は蓄積されているデータのどれくらいの割合を占めていますか？データの容量ではなく、件数ベースでお答えください。
(Q13)データの蓄積量（件数）[種類別] ※3年前からの変化	貴社で蓄積（保有）されているデータについて、2015年度末時点と比べて2018年度末時点のデータはどの程度変化（増減）しましたか？データの容量ではなく、件数ベースでお答えください。
(Q14)データの蓄積量（非構造化データの件数割合）[種類別]	貴社で蓄積（保有）されているデータについて、非構造化データ（業務日誌やSNSの書き込み、音声/画像/映像など規則（構造）が明確に定義されていないデータ）はどれくらいの割合を占めていますか？データの容量ではなく、件数ベースでお答えください。
(Q15)分析に活用するデータの件数割合 ※現在、3年前	貴社で蓄積（保有）している全データの内、1年間（2018年度）に実際にデータ活用に用いたデータの割合はおおむねどの程度ですか？データの容量ではなく、件数ベースでお答えください。また、3年前（2015年度）時点の状況と合わせてお答えください。
(Q16)新製品・新サービスの投入数	貴社では2018年度に新製品・新サービスをどれくらい市場に投入しましたか？
(Q17)競合数 ※現在、3年前からの変化	貴社の代表的な製品・サービスを投入している市場において、2018年度末時点で、貴社と競合する企業は国内外合わせて何社程度ありましたか？ また2015年度末時点から2018年度末時点の変化についても合わせてお答えください。

3. データの効果・価値に応じた正当な報酬のあり方

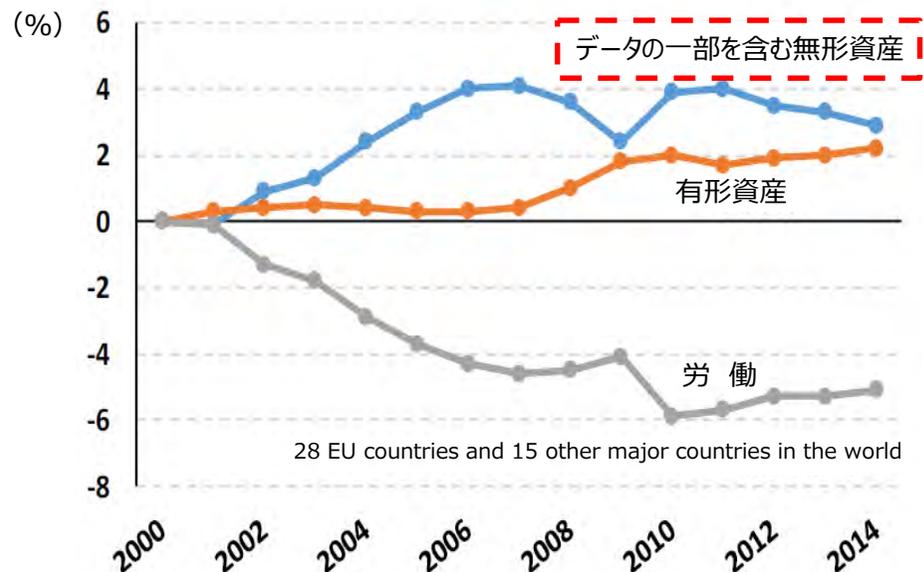
問題設定

- 分配状況は、データの一部を含む無形資産へシフトしている（図表14）。データが生み出す富がデータを収集・活用する企業に分配されている可能性があるとするれば、今後、経済活動において重要な地位を占めることが見込まれるデータの効果・価値に応じた正当な報酬のあり方を検証すべきではないか。

検討事項

- 生産要素としてのデータを資本又は労働として捉えるアプローチについての論点を整理する
- 「データ生成したユーザに報酬を支払う」というアプローチについての論点を整理する

図表14 無形資産・有形資産・労働の分配率（累積変化幅）



出典：Chen, Los & Timmer (2018) Factor Incomes in Global Value Chains: The Role of Intangibles, NBER Working Paper 25242.

- データは一般的に資本として扱われることが多いが、データを労働として扱い、データを生成する個人に対して金銭的な報酬を与えるべきという意見も存在。
 - **「資本としてのデータ」(Data as Capital)** は、起業家精神とイノベーション促進の観点から、データが生み出す価値の貢献者としてデータの活用者（AI企業・プラットフォーマー・データサイエンティスト等）を重視する考え方。
 - **「労働としてのデータ」(Data as Labor)** は、データを生成する個人の貢献を通常の労働と等しく認める観点から、データが生み出す価値の貢献者としてデータの生成者を重視する考え方。
- ※データが生み出す価値の貢献者としてデータの生成者を重視するアプローチについては、過去 E U で議論された「データ製作者権」(Data producer's right) (スライド37～39参照) や、プラットフォーム協同組合主義 (Platform Cooperativism) (スライド40参照) も存在。
- それぞれインセンティブや労働の未来、その反面として想定される懸念が異なる。

図表15 「資本としてのデータ」及び「労働としてのデータ」の主な特徴

	資本としてのデータ	労働としてのデータ
データが生み出す価値の貢献者	データの活用者を重視	データの生成者を重視
インセンティブ	起業家精神	通常の（労働と等しい）貢献
労働の未来	ユニバーサル・ベーシック・インカム	データ労働
想定される懸念	<ul style="list-style-type: none"> ・AIは、関連データを生成する個人の積極的な参加に依存するため、生産性の向上に対するAIの貢献を妨げる可能性 ・個人が生成したデータが特定の企業に集中したまま、市場のダイナミズムが停滞する可能性 	<ul style="list-style-type: none"> ・アンチコモنزの悲劇（共有されるべき財産が細分化されて私有され、社会にとって有用な資源の活用が妨げられること）が発生する可能性 ・金銭的対価を与えることにより、活動に対する本来のモチベーションを下げる可能性

※出典論文、他の参考情報、分科会での議論を参考に事務局作成

- EUにおいては一般データ保護規則 (**General Data Protection Regulation : GDPR**) が2018年5月に全面施行され、**パーソナルデータ保護に関するEU域内での統一的なルール**を制定。その中に**データポータビリティ権** (自らのパーソナルデータを、機械可読性のある形式で取り戻す権利であり、技術的に可能な場合には、自らのパーソナルデータを、ある管理者から別の管理者に直接的に移行させる権利も認められる。) も含まれている。
- 一方、**IoTデバイス・コネクテッドカー等のマシンを介して収集される非パーソナルデータ**について、欧州委員会において将来におけるデータアクセスのためのフレームワーク (A future EU framework for data access) の一項目として、**データ製作者権 (Data producer's right) の提唱**が行われたが、ステークホルダーに対するパブリックコメントが行われた後、**政策オプションから外れた**。

※ なお、上記フレームワークの中で言及された企業間のデータ共有という視点は、欧州データ戦略(A European strategy for data) (2020年2月)におけるデータ法(Data Act) (2021年制定予定)に係る構想の中で継承(産業用IoTなどにおける)共同生成データの使用权に関する問題対処のサポートや、データ共有に関する法的責任の明確化等)されている。

将来におけるデータアクセスのためのフレームワーク (A future EU framework for data access)

- ・企業のデータ共有の誘引付けに関するガイダンス (Guidance on incentivising businesses to share data)
- ・信頼できるデータの識別と交換のための技術ソリューション開発促進 (API等) (Fostering the development of technical solutions for reliable identification and exchange of data)
- ・デフォルト契約ルール (不公正契約条項指令の見直しを含む) (Default contract rules)
- ・公益及び科学的目的のためのアクセス (B2G含む) (Access for public interest and scientific purposes)
- ・**データ製作者権 (Data producer's right)**
- ・アクセスへの対価 (FRAND (Fair, Reasonable And Non-Discriminatory) 等) (Access against remuneration)

- データ製作者権に関し、フランス法学者から、当該データの財産権化を求める論文が出された一方、ドイツ法学者からはデータ製作者権の創設は時期尚早との反論がなされていた。

フランス法学者の主張

- ・個人情報保護に関する議論では2つの立場（市場と既存ビジネスモデルを守る立場と、デジタルプラットフォームの透明性を持たせる法律を提案しユーザーを守る立場）が対立。
- ・個人情報の財産権化が、これら2つの見方を補完調整する。財産権化はデータ利用のための個別価格を導入する。

出典: Lucas Leger, Pierre Bentata "Aux data, citoyens ! Pour une patrimonialité des données personnelles," (市民よデータのために立ち上がれ！個人情報の財産権化のために) Generation Libre Publishing 2019

ドイツ法学者の主張

- ・第1に現在のEU法枠組みがデータに対し財産権を付与しているか、著作権やデータベース保護のような確立されたIPRのデータ取扱いを調査。第2にデータ制作者権を財産所有権とみなせるか論述。第3に、データ製作者権を財産権の導入に関する法的及び経済的正当性に照らし評価。
- ・結論としては、データ製作者権が生産的かは懐疑的であり、安易に導入すべきではないとしている。

出典: Ivan Stepanov Introducing a property right over data in the EU: the data producer's right – an evaluation (International Review of Law, Computers & Technology, 1-22.) 2019
(https://www.researchgate.net/profile/Ivan_Stepanov5/publication/334169019_Introducing_a_property_right_over_data_in_the_EU_the_data_producer's_right_-_an_evaluation/links/5d493f42a6fdcc370a7fbb02/Introducing-a-property-right-over-data-in-the-EU-the-data-producers-right-an-evaluation.pdf)

- データ製作者権に関するステークホルダーに対するパブリックコメントにおいては、多くのステークホルダーより、データに関する新たな所有権を設けることについて懸念が表明され、むしろ企業間共有データへのアクセスを確保することに対する必要性が主張された。

パブリックコメント結果の概要

・・・会議やワークショップにおいて、多くのステークホルダーより、B2Bデータ共有における重要な問題が、データに何らかの「オーナーシップの資格」があるかどうかではなく、アクセスがどのように構成されているかであるというコメントがあった。欧州政治戦略センターからの論文や学界から受け取った意見は、この見解を強く支持していた。両者（欧州政治戦略センターや学界）は、データの所有権の創設に向けて取り組むことと、データアクセスをさらに開放することに取り組むこととの間に、方針選択が必要であると主張した。センサーを装備したマシン、ツール又はデバイスからのデータの使用を許可する権利の考え方は、OEM（相手先ブランド名製造者）又はセンサーを装備したマシン、ツール若しくはデバイスのユーザーに独占的に与えられた場合、懐疑的にみなされる。ステークホルダーは、その法的地位を強化することによって、データの取引可能性を促進するという目標を達成する可能性は低いと考えている。この潜在的な方法は、むしろ、事実上の所有者によるデータに対するアクセス制御を強化し、実際の適用に法的な不確実性をもたらす、追加の法的取引コストを生み出すことになるだろう。一方、OEMやセンサーを備えたマシン、ツール又はデバイスのユーザーの間で共有されるデータ使用をライセンスする権利の考え方は、比較的好意的に見られていた。特に中小企業の代表は、このようなソリューションを支持している。・・・

- **プラットフォーム協同組合主義 (Platform Cooperativism)** とは、デジタル情報財を取扱うインターネット上のプラットフォームを組織運営者、財の提供者、(組織によっては) プラットフォームの利用者が協同組合を所有し、意思決定における投票権を有する仕組み。
- 協同組合を所有する組合員の異質性が低い場合 (特に組合員間で事業貢献上の大きな差が生まれにくい場合)、事業内容が単純な場合、外部環境変化が乏しいために頻繁な資金調達が必要でない場合に協同組合によるガバナンスが適性を持つが、プラットフォーム協同組合においては外部環境変化が激しい点が当てはまらない。
- データのフォーカスした取組としては、**データ協同組合 (Data Cooperatives)** があり、**パーソナルデータの保有者自身による主導管理に焦点**を当てている。

データ協同組合 (Data Cooperatives)

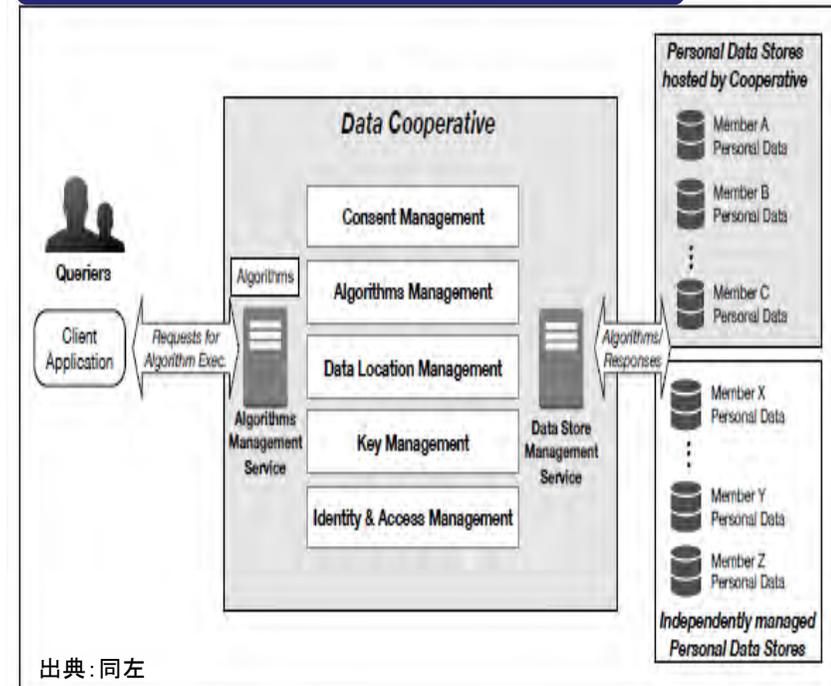
・データ協同組合は、グループ又はコミュニティの組合員の便益のために、当該組合員の有するパーソナルデータの自発的な協同共有管理を行い、データを通じて得られる共通の洞察 (現在の経済、健康、社会状況の理解) を共有することがデータ提供・管理の動機となる。

<主な特徴>

- 1 : 組合員個人は自身のパーソナルデータを加除したり、アクセスの一時停止をすることができる一方、組合員全員のパーソナルデータのコピーを収集できる。データ協同組合は預かったパーソナルデータを保護し、オプションで組合員の利益のために当該データの整理・要約を行う。
- 2 : データ協同組合は組合員に対して法的義務 (データの使用・アクセス・プライバシー等に対するポリシー等) を負う。
- 3 : データ協同組合の目的はデータの収益化ではなく、組合員のニーズをよく理解し、組合員間で洞察を共有するために、継続的な分析を行うこと。

出典: Thomas Hardjono and Alex Pentland "Data Cooperatives (2019): Towards a Foundation for Decentralized Personal Data Management" MIT Connection Science Massachusetts Institute of Technology, Cambridge, MA, USA

図表16 データ協同組合のエコシステム



出典: 同左

- データ生成したユーザに報酬を支払うことについては「労働としてのデータ」の立場では肯定的である一方、「資本としてのデータ」の立場では賛否両論ある。
- それぞれの立場において想定される課題が存在する。

図表17 「データ生成したユーザに報酬を支払う」というアプローチについての論点

データを生成したユーザに対する報酬の支払	資本としてのデータ		労働としてのデータ
	肯定的	否定的	肯定的
報酬を支払うべき／支払うべきではない理由	<ul style="list-style-type: none"> ・A I の進歩に伴い、データの重要性が高まるため、データを提供するインセンティブが必要 ・データが生み出す価値の大部分を私的利益として獲得している企業の存在 	<ul style="list-style-type: none"> ・無料のデータセットや機械学習アルゴリズムを無料で入手できる ・データは分析して初めて価値を生むものであり、データの分析者が報酬を受け取るべき ・金銭的対価を与えることにより、活動に対する本来のモチベーションを下げる可能性 	<ul style="list-style-type: none"> ・労働としてのデータは収入を補完する重要な機会となり、格差の拡大に苦しむ市民に社会に貢献しているという意識をもたらす
報酬支払のプロセス	<ul style="list-style-type: none"> ・市場メカニズム（マイクロペイメント・情報信託機能（情報銀行）） ・規制（競争政策／税制／社会保障等） 	-	<ul style="list-style-type: none"> ・（限界）貢献を測定し、個人のユーザが生み出す価値を遡及し追跡するための適切なテクノロジーシステムを構築
想定される課題	<ul style="list-style-type: none"> ・市場メカニズム：ビジネスモデルの確立 ・規制：導入の客観的正当性（定量的指標等） 	<ul style="list-style-type: none"> ・企業・データ分析者とそれ以外の個人との富の偏在（労働分配率の更なる低下）を生む可能性 	<ul style="list-style-type: none"> ・（限界）貢献の測定やテクノロジーの実現性 ・貢献の測定が監視社会化や格差・差別の助長をもたらす可能性

※スライド42～44の出典論文、他の参考情報、分科会での議論を参考に事務局作成

(参考) 「資本としてのデータ」に関連し、【「データ生成したユーザに報酬を支払う」というアプローチに肯定的な考え方】

- 「資本としてのデータ」に関連し、【「データ生成したユーザに報酬を支払う」というアプローチに肯定的な考え方】

データ生成したユーザに報酬を支払うべき理由

- ・Jaron Lanier (Microsoftの研究者) は、**人々が提供するデータに対して金銭的な報酬を支払うべきだ**と主張している。
- ・人々を観察することで、ロボットが自然な会話をできるようになったりするため、その貴重なデータを使用するための対価を支払うべきという考え方である。

出典：Jamie Susskind (2018) 「Future Politics: Living Together in a World Transformed by Tech」 Oxford Univ Pr.

- ・個人データの使用に対する懸念の高まりにより、EU一般データ保護規則 (GDPR) など、行動とポリシーの変更が促された。
- ・Jaron Lanierは、データをロックアップするのではなく、データを提供するように奨励する必要があると述べている。
- ・人工知能技術が進歩するにつれてデータがさらに重要になるため、データが貴重であることを認め、人々にお金を支払い、それを改善するようにするべきだと主張している。

出典：「Tech pioneer Jaron Lanier says companies should pay for data: 'Let's get out of the manipulation business」
<https://www.cnn.com/2018/06/21/tech-pioneer-jaron-lanier-says-companies-should-pay-for-data.html>

- ・現在、既存の大企業はデータが生み出す価値の大部分を私的利益として獲得している。

出典：Bennett Institute for Public Policy (2020) 「The Value of Data - Summary report 2020」.

報酬支払のプロセス

- ・(Jaron Lanier (Microsoftの研究者) は) 「マイクロペイメント」のシステムを提案し、これが構築できれば人々は、自分が生涯にわたって行った「数万の小さな貢献」に対して「ロイヤリティ」を獲得できるという。

出典：Jamie Susskind (2018) 「Future Politics: Living Together in a World Transformed by Tech」 Oxford Univ Pr.

- ・競争政策によって、データ主導型の市場を他の供給者に開放することを推奨する。

出典：Bennett Institute for Public Policy (2020) 「The Value of Data - Summary report 2020」.

(参考)「資本としてのデータ」に関連し、【「データ生成したユーザに報酬を支払う」というアプローチに否定的な考え方】

- 「資本としてのデータ」に関連し、【「データ生成したユーザに報酬を支払う」というアプローチに否定的な考え方】

データ生成したユーザに報酬を支払うべきではない理由

- ・「データは新しい石油」と言われており、両方とも有用なものにするために洗練される必要がある。
- ・しかし、重要な違いがある。石油は私的な財であり、競合財である。つまり、ある人が石油を消費すると、他の人が消費できる量が少なくなる。
- ・一方、データは非競合財である。ある人がデータを使用しても、他の人の使用が減ったりすることはない。
- ・したがって、データの「所有権」に焦点を当てるのではなく、データアクセスについて考えるべきである。

出典：Hal Varian (2018) "Artificial Intelligence, Economics, and Industrial Organization," *NBER Working Paper*, No.24839.

- ・金銭的対価は活動に対する本来のモチベーションを下げる（例：娯楽としてのネットサーフィンに対価を与えると、支払を求める労働に転化し、娯楽としての効用が下がる）。

出典：Eric A. Posner and E. Glen Weyl (2018) "Radical Markets".

データの効果・価値に応じた報酬を与えるべき主体

- ・政府や企業（GoogleやKaggle）から無料のデータセットが入手できるようになっており、機械学習アルゴリズムもTensorFlowおよびCaffeから無料で入手できる。
- ・ただ、多くの企業はその可能性を十分に活用できておらず、問題はリソースの不足ではなく、スキルの不足である。優秀な人材は、競争上の優位性を確保するための重要である。
- ・データがあるものの、それを分析する人がいない会社は、データを利用するのに不十分な立場にある（＝データは分析してはじめて価値を生むものであり、データの生成者ではなく、データの分析者にこそ報酬を支払うべき）。

出典：「Google's chief economist thinks the world needs more data scientists」
<https://qz.com/1296930/hal-varian-googles-chief-economist-thinks-the-world-needs-more-data-scientists/>

(参考) 「労働としてのデータ」に関連し、【「データ生成したユーザに報酬を支払う」というアプローチに肯定的な考え方】

- 「労働としてのデータ」に関連し、【「データ生成したユーザに報酬を支払う」というアプローチに肯定的な考え方】

データ生成したユーザに報酬を支払うべき理由

- ・労働としてのデータは収入を補完する重要な機会となり、格差の拡大に苦しむ市民に社会に貢献しているという意識をもたらす。
- ・世界が多種多様な機会を提供するようになって、個人がいまよりも幅広いニッチ分野に特化できるようになって欲しい。

報酬支払のプロセス

・第一に、データの価値を定量的に把握する必要がある。**個人の貢献の価値を評価する第一歩は、こうした（限界）貢献を測定すること。**

・第二に、個人のユーザが生み出す価値を遡及し追跡するための適切なテクノロジーシステムを構築し、どのユーザがどのデータ貢献に責任を負っているのかを測定することを試みるべき。特に、特定の個人の貢献が不釣り合いに大きい場合や、そうした個人が金銭的なインセンティブを受けなければ特別な貢献につながる独自のデータを供給しなそうなきはなおさらである。

- ・データがあるプロセスを通じて何かの価値を生み出しているとすると、そのプロセスはブラックボックスでありつつも何らかの関数（生産関数）で表現できる。
- ・**生産関数の推定**によって、データを一定単位増やした時の限界的な貢献を測定することができるため、それがデータの適正な価値とみなすことができる。

報酬の程度

・ある種のデータクラスはコモディティ化し、全体的な質の基準を満たすことを条件に「平均価格」を支払うようにするべきだ。だが、たとえデータ労働が大勢の人にとって重要な所得源になったとしても、その果実が均等に分配される保証はない。

・機械学習（ML）に極めて高い価値がある独自の文化的な知識や能力を持っている人もいるだろうし、提供するデータがありきたりすぎて限界価値がそれほどない人もいる。

・幅広いMLプロセスに少しずつ貢献するデータ労働者もいれば、一つの分野には大きく貢献するが、それ以外の分野にはほとんど、あるいはまったく貢献しないという人もいるだろう。