

日本企業におけるデータ利活用に関する質問票調査

—上場企業と非上場企業の差異に着目して—

平井祐理 (文部科学省 科学技術・学術政策研究所 上席研究官)

立本博文 (筑波大学 ビジネスサイエンス系 教授)

生稲史彦 (中央大学 戦略経営研究科 教授)

Questionnaire Survey on Data Utilization in Japanese Companies : Focusing on the Differences Between Listed and Unlisted Companies

Yuri Hirai

Senior Research Fellow, National Institute of Science and Technology Policy, Ministry of Education, Culture, Sports, Science and Technology

Hirofumi Tatsumoto

Professor, Graduate School of Business Sciences, University of Tsukuba

Fumibiko Ikuine

Professor, Chuo Graduate School of Strategic Management, Chuo University

【要旨】 第5期科学技術基本計画が策定されて5年余りが経過した。この間、Society 5.0の実現に向けてあらゆる産業の企業がデジタル技術を活用した事業活動を行ってきた。加えて、新型コロナウイルス感染症のパンデミックによって世界的にデジタルトランスフォーメーションが加速しており、無形資産としてのデータの利活用はますます重要となっている。本稿では、独立行政法人経済産業研究所が実施した質問票調査を用いて、①データの量及びデータを利活用する組織の能力に着目してデータ利活用の成果に影響を与える要因について検討すること、②その影響が上場企業と非上場企業で異なるのかについて検討すること、を目的とした。分析の結果から、データ利活用の成果に対してデータの量の影響は限定的である一方、データを利活用する組織の能力は成果の獲得に大きく貢献していることが示唆された。また、これらの影響は上場企業と非上場企業で変わらないことがわかった。

【キーワード】 データ利活用 データ資産 ビッグデータ 人工知能 質問票調査

【Abstract】 More than five years have passed since the 5th Science and Technology Basic Plan was formulated. During this time, companies in all industries have been conducting business activities utilizing digital technologies toward the realization of Society 5.0. In addition, the COVID-19 pandemic is accelerating digital transformation worldwide, making the utilization of data as an intangible asset increasingly important. In this paper, using a questionnaire survey conducted by the Research Institute of Economy, Trade and Industry, we aimed to examine (1) the factors that influence the business results of data utilization, focusing on the amount of data and the organizational capability to utilize data, and (2) whether the influences differ between listed and unlisted companies. The results of our analysis suggest that while the impact of the amount of data on the business results of data utilization is limited, the organizational capability to utilize data contributes significantly to the business results. It is also found that these effects are the same for listed and unlisted companies.

【Keywords】 Data Utilization Data Asset Big Data Artificial Intelligence Questionnaire Survey

1. はじめに

第5期科学技術基本計画が策定されて5年余りが経過した。この間、Society 5.0の実現に向けて、あらゆる産業の企業がIoT (Internet of Things)、ビッグデータ、AI (artificial intelligence) といったデジタル技術を活用した事業活動を行ってきた。加えて、新型コロナウイルス感染症のパンデミックによって経済活動の多くがデジタル化されるようになった。世界的にDX (digital transformation) が加速する中、無形資産としてのデータの利活用はますます重要となっている。

このような中、筆者らは、独立行政法人経済産業研究所 (RIETI) の研究プロジェクト「データとAI利活用促進をグローバルに展開するための制度とマネジメントに関する研究—グローバルデータサプライチェーンの確立に向けて」¹⁾において、我が国のデータ利活用の実態を把握するために質問票調査を実施し、その研究結果をディスカッションペーパーにまとめた (渡部ほか, 2021)。本稿では、この質問票調査を用いて、①データ利活用の成果に影響を与える要因について検討すること、②その影響が上場企業と非上場企業で異なるのかについて検討すること、を目的とする。

DXの実施においてデータそれ自体は必要不可欠である。しかし、データは解析等を通じてはじめて本来の価値を引き出すことができる。そのため、企業にはデータを利活用する能力が必要となると考えられる。このような観点から、本稿ではデータ利活用の成果に影響を与える要因として、データの量及びデータを利活用する組織の能力に着目する。加えて、当該質問票調査では上場企業リストと非上場企業リストの2つの企業リストを用いて調査を実施したため、上場企業と非上場企業との差異についても検討する。一般に非上場企業と比較して上場企業では企業規模も大きくアクセスできるリソースも多いため、上場企業と非上場企業では利活用できるデータの量や組織能力が異なり、それらがデータ利活用の成果に与える影響も異なる可能性が考えられる。

本稿では、まず続く第2章において関連する先行

研究について紹介する。第3章では質問票調査の概要や分析で用いる変数について説明し、第4章では分析結果を示す。最後に第5章でまとめとして考察と結論を示す。

2. 先行研究

経営学における近年のデータ利活用に関する研究は、情報システム (IS: information systems) 分野の研究の流れを汲んでいる。1980年代に情報通信技術 (IT: information technology) が競争優位性を獲得する手段であると言われるようになると、情報システム分野ではIT投資と企業の生産性の関係について多くの研究が行われた (Gupta and George, 2016)。しかし、IT投資が生産性を向上させるという明確な研究結果は得られず、この現象は「IT生産性パラドクス (IT productivity paradox)」と呼ばれてきた。この現象の説明として、ITの価値を適切に評価するための研究データや指標が確立されていないからではないか、IT投資がビジネス価値を生むまでには一定のタイムラグが存在するためより長期的な観察が必要ではないか、企業パフォーマンスに対するIT投資の影響は間接的でありIT投資と企業パフォーマンスとの間にはそれを仲介する要因が存在するのではないか、といった様々な観点から研究が行われた (Brynjolfsson and Yang, 1996; Wamba et al., 2017)。

他方、経営学の戦略論の分野では、1990年代以降、企業が保有する経営資源 (management resources) の観点から収益性や競争力を説明する枠組みとしてRBV (resource-based view) が現れた。RBVでは、希少で、模倣困難で、代替不可能な企業特有のリソースが競争優位の源泉であると考えられる (Barney, 1991)。ここでいう経営資源とは、人と物と資金、そして、技術やノウハウ等の知識 (情報的経営資源) を指す。RBVでは、これらの経営資源と、それを使いこなす組織能力に着目し、各企業の生産性や収益性等の事業成果を説明する。

これら、IT投資と企業パフォーマンスの関係をつなぐ要因を探る研究と、RBVが接合し、RBVの組織能力を踏まえたITケイパビリティ (IT capabil-

ity) の概念が提唱された (Bharadwaj, 2000). Bharadwaj (2000) は、IT ケイパビリティを「他のリソースや能力と併せて IT リソースを動員し利用する企業の能力 (a firm's ability to mobilize and deploy IT-based resources in combination or co-present with other resources and capabilities)」と定義している。組織能力の概念を敷衍した IT ケイパビリティの概念を用いると、ハードウェアやソフトウェアといったリソースは比較的社外から調達可能である一方、それらを使いこなす組織の能力は他社が容易に模倣できるものではなく企業固有の希少なリソースであると切り分けられる。そのため、IT ケイパビリティが競争優位の源泉であるという見通しが立つ。

加えて、RBV の理論と IT 分野の現状を接合すると、IT ケイパビリティをより詳しく分類できるようになった。Grant (1991) は経営資源を、有形の資源 (tangible)、人材 (human skills)、無形の資源 (intangible) の 3 つに分類しているが、この分類は IT ケイパビリティの研究にも応用できる。Bharadwaj (2000) は、IT ケイパビリティの研究を行う上で、IT リソースを有形の IT infrastructure、人材に由来する human IT resources、無形の IT-enabled intangibles の 3 つに分類している。

さらに、この研究領域では、5 年ほど前から IT ケイパビリティを近年の文脈にあわせて発展させる研究と概念提唱が盛んになっている。例えば、ビッグデータの登場に合わせたビッグデータ解析能力 (big data analytics capability) (Akter et al., 2016; Dubey et al., 2019; Garmaki et al., 2016; Gupta and George, 2016; Mikalef et al., 2018; Wamba et al., 2017) や、データ解析能力 (data analytics competency) (Ghasemaghaei et al., 2018; Ghasemaghaei, 2019) といった概念が登場している。概念の名称や定義は研究によって異なるものの、RBV 及び IT ケイパビリティを基にしていること、有形の資源 (tangible)、人材 (human skills)、無形の資源 (intangible) という 3 つの分類を用いていることは概ね共通している。例えば、Ghasemaghaei et al. (2018) はデータ解析能力 (data analytics competency) を「正確で行動指向的なデータ解析を行うために、データ

解析のリソースを展開したり結合したりする企業の能力 (a firm's ability to deploy and combine data analytics resources for rigorous and action-oriented analyses of data)」と定義し、IT infrastructure、human IT resources、IT-enabled intangibles の 3 つに分類して調査を行っている。このように、情報システムに代表される経営資源の存在を前提にして、それを使いこなす組織としての能力に焦点を当てる研究が、現在は精力的に進められていると言える。

この研究動向を踏まえ、筆者らは日本企業を対象とした研究として、本稿で用いる質問票調査に先駆けて、RIETI の研究プロジェクト「企業において発生するデータの管理と活用に関する研究」²において日本企業 6278 社を対象として平成 29 年度「データ利活用に関するアンケート調査」³を実施した (渡部ほか, 2018)。この質問票調査では、ビッグデータ解析能力やデータ解析能力に類似した項目として「ビッグデータの利活用を行える体制が整備されているか」や「ディープラーニング等の高度なデータの処理・解析を行える体制が整備されているか」について調査した。分析の結果、特に製造業に該当しない企業においてビッグデータの体制整備や高度なデータ処理・解析能力を備えた体制整備がデータ利活用の成果 (特に間接的成果) に正の影響を与えている一方、データの総容量ははっきりとした影響が見られないということを明らかにしている (Hirai et al., 2019; 渡部・平井, 2020)。加えて、立本ほか (2018) は、この質問票調査を用いてデータ量の影響を産業別に比較している。データの総容量を入力としたときの出力 (=データ利活用の成果) に対する変換効率 (=効果) として反応係数を産業別に推定した結果、反応係数は産業によってマイナスからプラスまでばらついており、データ利活用の成果に対するデータの総容量の影響は産業ごとに違いがあることを明らかにしている (立本ほか, 2018)。また、データの量に関しては、総務省情報通信政策研究所が国内企業に対してアンケート調査を実施している (総務省, 2021)。総務省 (2021) は、企業が付加価値を生み出すための生産要素として「資本」「労働」「データ」を位置付けて生産関数モデルを推定し、活用データ容量は資本や労働と同様に付加価値

値に対してプラスの関係性を持っていることや、外部入手したデータは内部入手したデータと比較して貢献が大きいことを示している。加えて、大橋・中村（2020）は、総務省情報通信政策研究所が実施した2018年度の状況に関するアンケート調査「データの活用に関する調査」を用いて分析を行い、日本企業においてはデータ資産の保有量がマークアップの上昇にはつながっていない一方、データ資産を多く保有する企業ではTFP水準が有意に高いことを示している。

以上で紹介した既存研究を踏まえると、情報システム等を経営資源の一部と考え、それを使いこなす能力が企業の競争力に影響を及ぼすという基本的図式は、この20年ほどで定着した。しかしながら、そうした能力がどのようなものであるのか、それがどのように企業の競争力に影響を及ぼすのかについては、実証研究によって異なる結果が報告されており、より詳細な検証が必要とされていると言える。

3. 方法

3.1. 質問票調査の概要

本稿では、RIETIが実施した2020年度「データ利活用に関するアンケート調査」⁴を使用した。この質問票調査では、上場企業3685社、非上場企業2252社の計5937社を対象とした。上場企業については、第一部、第二部、マザーズ、JASDAQスタンダード、JASDAQグロース、TOKYO PRO Marketに上場する企業を株式会社東京商工リサーチのデータベースから2020年6月3日に抽出した。非上場企業については、スタートアップ情報プラットフォーム「INITIAL」⁵を用いて、従業員数が20人以上の未公開企業を2020年5月29日に抽出した。抽出した2301社から、株式会社以外の企業や倒産した企業等は対象外とし、2252社を対象とした。調査期間は2020年6月26日～2020年8月31日であった。回答方法はWeb回答と郵送回答とを併用し、657社からの有効回答を得た（有効回答率：11.1%）。なお、この調査では2019年度（2020年3月末時点）の状況について回答を依頼した。より詳細な調査概要については、渡部ほか（2021）を参照されたい。

3.2. 変数

本稿では、図表1に示す枠組みに基づいて分析を行った。以下に各変数の詳細を記す。

・被説明変数

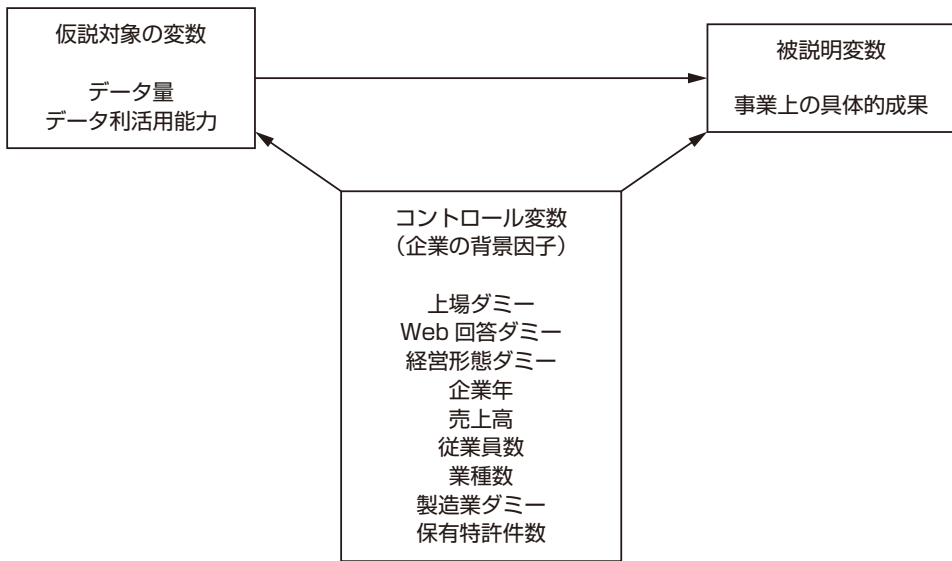
当該質問票調査では、自社におけるデータ利活用のこれまでの成果について、「1. まだ成果は得られていない」「2. 具体的成果は得られていないが、少なくとも1つの事業で間接的成果が得られている」「3. 具体的成果は得られていないが、複数の事業で間接的成果が得られている」「4. 少なくとも1つの事業で具体的成果が得られている」「5. 複数の事業で具体的成果が得られている」の5段階評価で回答してもらった。ここで、「具体的成果」とは売上やコストダウンといった財務的な成果を表しており、「間接的成果」とは事業活動に役立つノウハウやアイデアの獲得等を示している。このうち、具体的成果が得られている場合（「4. 少なくとも1つの事業で具体的成果が得られている」もしくは「5. 複数の事業で具体的成果が得られている」と回答した場合）を1、そうでない場合を0とした（「具体的成果ダミー」と表記）。

・説明変数（仮説対象の変数）

データの量を示す説明変数として、自社に利用権限があるデータのうち、利活用を行っている、または、今後利活用することを期待しているデータの総容量を示す「データ総容量」(3段階評価:「1. 1台のPCで管理できる程度」「2. 数台のサーバで管理できる程度」「3. 専用のサーバ室、サーバセンターで管理する程度」)を使用した。また、データの量についてより詳しく検討するために、「データ総容量ダミー_1台のPC」(「1. 1台のPCで管理できる程度」と回答した場合は1、そうでない場合は0)、「データ総容量ダミー_サーバ室」(「3. 専用のサーバ室、サーバセンターで管理する程度」と回答した場合は1、そうでない場合は0)という2つのダミー変数を作成した。

データを利活用する組織の能力として、本稿ではデータ利活用能力（DUC: data utilization capability）という変数を作成した。当該質問票調査では、ビッグデータやAIを用いたデータ利活用に限らず広くデータ利活用の実態について調査したため、

図表 1 分析の枠組み



Ghasemaghaei et al. (2018) を参考により包括的にデータを利活用する組織の能力を測る 5 つの質問項目を作成した (図表 2)。回答企業には、これら 5 つの質問項目について、「1. 全くそう思わない」～「5. 強くそう思う」の 5 段階評価でそれぞれ回答してもらった。これら 5 項目についての Cronbach の α は 0.880 であり、十分に高い信頼性を示す値であったことから、5 項目の 5 段階評価の合計値を DUC とした (最小値 5, 最大値 25)。

・コントロール変数

コントロール変数としては、上場ダミー (上場企業であれば 1, 非上場企業であれば 0)、Web 回答ダミー (回答方法が Web 回答であれば 1, そうで

なければ 0)、経営形態_純粋持株会社ダミー (経営形態が純粋持株会社であれば 1, そうでなければ 0)、経営形態_事業持株会社ダミー (経営形態が事業持株会社であれば 1, そうでなければ 0)、企業年 (設立年から 2020 年までの年数)、売上高 (単位: 億円)、従業員数、業種数 (36 業種のうち、自社の事業がカバーする業種の数)、製造業ダミー (自社の事業を最もよく表す業種が製造業であれば 1, そうでなければ 0)、保有特許件数 (6 段階評価: 「1. 0 件」「2. 1-9 件」「3. 10-99 件」「4. 100-999 件」「5. 1000-9999 件」「6. 10000 件以上」) を使用した。

図表 2 DUC の質問項目

Ghasemaghaei et al. (2018) との対応	質問項目
tools sophistication (IT infrastructure)	①自社では、高度なデータ分析ツールを用いている (モデリング、自然言語解析、リアルタイム分析等)
analytical skills (human IT resources)	②自社のデータ分析担当者は、高い分析スキルを有している (データ分析の深い専門性を有している、分析ツールを使いこなしている等)
domain knowledge (human IT resources)	③自社では、深いドメイン知識を有している (外部環境、自社の目標や目的、自社が成功するために鍵となる要因等)
data quality (IT-enabled intangibles)	④自社では、データ分析に用いているデータの質が高い (データの信頼性・安全性・正確性が高い、タイムリーである、自社のタスクとの関連性が深い等)
bigness of data (IT-enabled intangibles)	⑤自社では、ビッグデータを処理している (大容量のデータ、リアルタイムデータ、様々な形式のデータ)

4. 結果

4.1. ロジスティック回帰分析の結果（全サンプル）

図表 3 に使用した変数の平均，標準偏差，相関係数を，図表 4 に具体的成果ダミーを被説明変数とし

たロジスティック回帰分析の結果を示す．図表 4 において，Model 1 はコントロール変数のみのモデル，Model 2 はコントロール変数及びデータ総容量を投入したモデルである．データ総容量は $p<0.01$ で正に有意であった．Model 3 はコントロール変数及びデータ総容量ダミーを投入したモデルである（デー

図表 3 平均，標準偏差，相関係数

	平均	標準偏差	相関係数					
			1	2	3	4	5	6
1 具体的成果ダミー	0.223	0.417	1					
2 データ総容量	2.226	0.632	.182**	1				
3 データ総容量ダミー_1台のPC	0.112	0.316	-0.049	-.690**	1			
4 データ総容量ダミー_サーバ室	0.338	0.473	.209**	.876**	-.254**	1		
5 DUC	12.254	4.336	.387**	.284**	-.143**	.283**	1	
6 上場ダミー	0.566	0.496	0.004	.303**	-.146**	.308**	-0.034	1
7 Web 回答ダミー	0.796	0.403	0.049	0.056	0.030	.095*	0.063	-0.009
8 経営形態ダミー_純粋持株会社	0.076	0.265	-0.021	0.059	-0.007	0.075	.095*	.134**
9 経営形態ダミー_事業持株会社	0.373	0.484	0.034	0.065	-0.020	0.074	-0.032	.190**
10 企業年	44.245	32.417	0.022	.253**	-.144**	.242**	-0.068	.508**
11 売上高	1587.360	8721.047	0.052	.188**	-0.062	.209**	.123**	.151**
12 従業員数	2349.035	13170.081	0.078	.184**	-0.060	.206**	.149**	.151**
13 業種数	1.851	2.379	0.076	0.078	-0.055	0.067	0.045	-0.005
14 製造業ダミー	0.340	0.474	-0.041	-0.013	-0.037	-0.041	-.146**	0.000
15 保有特許件数	1.974	1.165	0.052	.252**	-.140**	.243**	.099*	.270**

	相関係数								
	7	8	9	10	11	12	13	14	15
7 Web 回答ダミー	1								
8 経営形態ダミー_純粋持株会社	0.009	1							
9 経営形態ダミー_事業持株会社	0.073	-.221**	1						
10 企業年	-.081*	-0.062	.140**	1					
11 売上高	0.058	0.039	.085*	.117**	1				
12 従業員数	0.043	.122**	0.018	.093*	.445**	1			
13 業種数	0.003	-0.015	-0.014	.083*	.170**	-0.010	1		
14 製造業ダミー	-.127**	-0.078	-0.009	.274**	0.004	-0.005	.086*	1	
15 保有特許件数	-0.068	-0.018	.106**	.370**	.238**	.306**	0.040	.516**	1

* $p<0.05$, ** $p<0.01$

図表 4 ロジスティック回帰分析の結果 (全サンプル)

	Model 1a			Model 2a			Model 3a			Model 4a			Model 5a			Model 6a		
	B	有意 確率	標準誤差	B	有意 確率	標準誤差	B	有意 確率	標準誤差	B	有意 確率	標準誤差	B	有意 確率	標準誤差	B	有意 確率	標準誤差
データ総容量																		
データ総容量ダミー_1台のPC																		
データ総容量ダミー_サーバ室																		
DUC		0.577**	(0.202)										0.093	(0.221)				
上場ダミー	-0.247	(0.282)		-0.393	(0.294)		-0.392	(0.296)		0.262**	(0.033)		0.254**	(0.035)		0.484	(0.431)	
Web回客ダミー	0.398	(0.312)		0.413	(0.323)		0.371	(0.326)		0.017	(0.316)		-0.022	(0.323)		0.359	(0.283)	
経営形態ダミー_純粋持株会社	-0.634	(0.531)		-0.609	(0.529)		-0.620	(0.529)		0.264	(0.339)		0.294	(0.344)		0.253**	(0.035)	
経営形態ダミー_事業持株会社	0.181	(0.231)		0.188	(0.236)		0.164	(0.237)		-0.681	(0.550)		-0.655	(0.548)		-0.045	(0.324)	
企業年	-0.001	(0.004)		-0.001	(0.004)		-0.001	(0.004)		0.331	(0.259)		0.306	(0.260)		-0.695	(0.551)	
売上高	-0.00004	(0.00004)		-0.00004	(0.00004)		-0.00004	(0.00004)		0.003	(0.005)		0.004	(0.005)		0.306	(0.261)	
従業員数	0.0001**	(0.00005)		0.0001**	(0.00005)		0.0001**	(0.00005)		-0.00003	(0.00004)		-0.00003	(0.00004)		0.003	(0.005)	
業種数	-0.015	(0.080)		-0.031	(0.080)		-0.023	(0.080)		0.001+	(0.00004)		0.001+	(0.00004)		-0.00003	(0.00004)	
製造業ダミー	-0.132	(0.299)		0.022	(0.309)		0.062	(0.311)		-0.109	(0.096)		-0.108	(0.096)		-0.098	(0.097)	
保有特許件数	-0.053	(0.128)		-0.119	(0.131)		-0.125	(0.132)		0.434	(0.339)		0.434	(0.341)		0.466	(0.342)	
定数	-1.420**	(0.402)		-2.553**	(0.561)		-1.527**	(0.420)		-0.216	(0.140)		-0.220	(0.141)		-0.221	(0.141)	
Cox-Snell R ² 乗	0.038			0.055			0.061			0.178			0.172			0.176		
Nagelkerke R ² 乗	0.058			0.084			0.093			0.270			0.263			0.268		
N	510			507			507			503			500			500		

* $p<0.1$, ** $p<0.05$, *** $p<0.01$

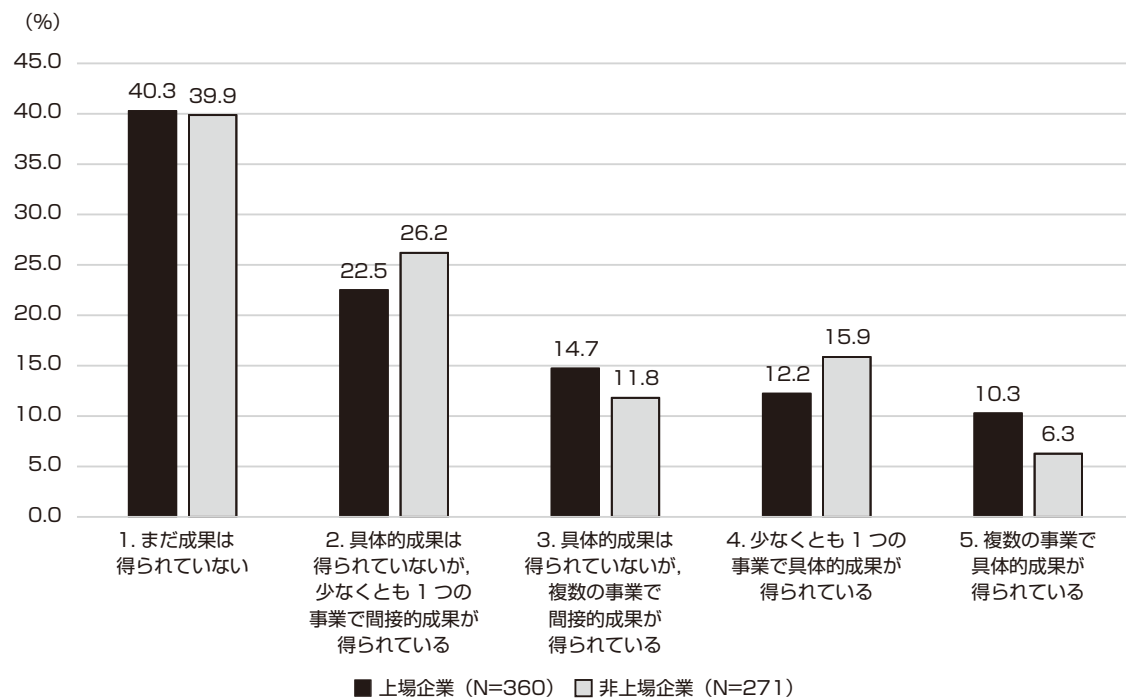
タ総容量に関して、「2. 数台のサーバで管理できる程度」がリファレンスである)。「データ総容量ダミー_1台のPC」は有意ではないが、「データ総容量ダミー_サーバ室」は $p<0.01$ で正に有意であり、データ総容量が数台のサーバで管理できる程度である場合と比較して専用のサーバ室、サーバセンターで管理する程度のデータ量であればデータ利活用の具体的成果に対して有意に効果があることが示唆された。Model 4はコントロール変数及びDUCを投入したモデルである。DUCは $p<0.01$ で正に有意であった。また、データの量に関する変数を追加したModel 2やModel 3ではコントロール変数のみのModel 1と比較して R^2 の値がさほど変わらなかった一方、DUCを追加したModel 4ではModel 1と比較して R^2 の値が大きく向上した。Model 5及びModel 6では、コントロール変数に加えてデータの量に関する変数とDUCとを同時に投入した。DUCはModel 4と同様に $p<0.01$ で正に有意であった一方、データの量に関する変数は有意ではなかった。

4.2. 上場企業、非上場企業別の集計結果

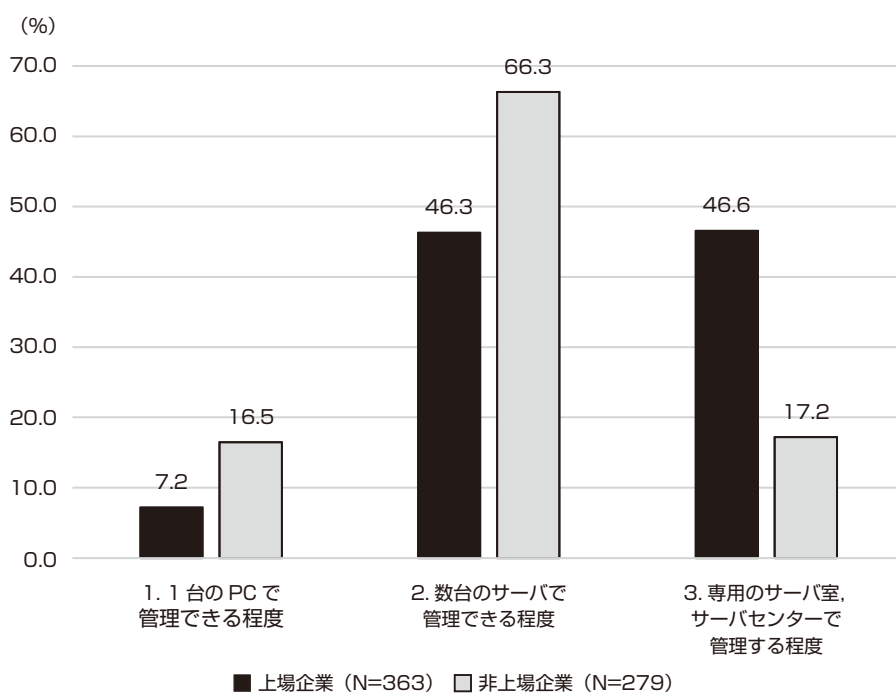
次に、主な変数について上場企業、非上場企業別に集計した結果を示す。企業年（設立年から2020年までの年数）の平均値は、上場企業で58.6年、非上場企業で25.4年であり、上場企業の方が2.3倍長かった。また、売上高の平均値は、上場企業で2698.3億円、非上場企業で21.5億円であり、非上場企業と比較して上場企業が125.4倍であった。従業員数の平均値は、上場企業で4102.8人、非上場企業で101.4人であり、非上場企業と比較して上場企業が40.5倍であった。回答企業において、企業規模を表す売上高や従業員数は上場企業と非上場企業で大きな差があった。

図表5は、自社におけるデータ利活用のこれまでの成果についてのグラフである。「1. まだ成果は得られていない」と回答した企業の割合は、上場企業で40.3%、非上場企業で39.9%であり、大きな差は見られなかった。また、具体的成果に関しては、「5. 複数の事業で具体的成果が得られている」と回答した企業の割合は上場企業の方が高いものの、具体的成果が得られている企業の割合（「4. 少なくとも1

図表 5 データ利活用のこれまでの成果



図表 6 データ総容量



つの事業で具体的成果が得られている」+「5. 複数の事業で具体的成果が得られている」は、上場企業で22.5%、非上場企業で22.1%であり、大きな差は見られなかった。

図表6は、データ総容量についてのグラフである。「3. 専用のサーバ室、サーバセンターで管理する程度」と回答した企業の割合は、上場企業で46.6%、非上場企業で17.2%であり、上場企業の方が多量のデータを扱っている企業が多いことがわかる。

図表7は、DUC（図表2の5項目の5段階評価の合計値）についてのグラフである。DUCの平均値は、上場企業で12.1、非上場企業で12.4であり、非上場企業でやや高く、大きな差は見られなかった。

4.3. 上場企業、非上場企業別のロジスティック回帰分析の結果

続いて、図表4に示した具体的成果ダミーを被説明変数としたロジスティック回帰分析を上場企業、非上場企業別に行った結果を示す。図表8は上場企業をサンプルとしたロジスティック回帰分析の結果、図表9は非上場企業をサンプルとしたロジスティック回帰分析の結果である。いずれのサンプルでも、DUCを投入しないモデルでは「データ総容量」や「データ総容量ダミー_サーバ室」は正に有意で

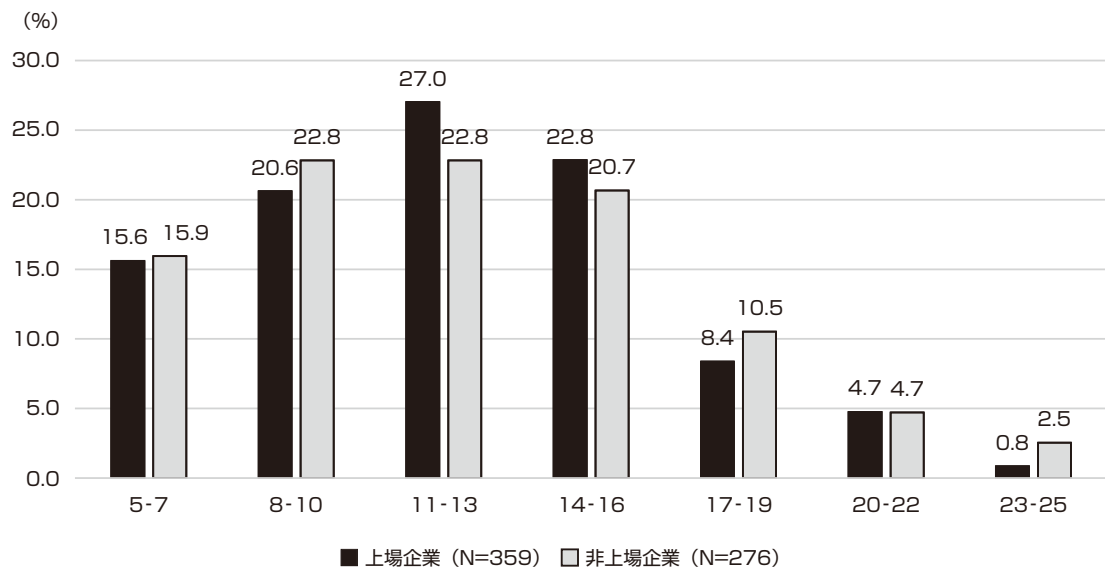
あったが（Model 2 及び Model 3）、DUCを投入したモデルでは有意ではなかった（Model 5 及び Model 6）。一方、いずれのサンプルでもDUCは正に有意であり、データの量に関する変数を投入したモデルと比較して R^2 の値が大きく向上した（Model 4～Model 6）。

5. 考察と結論

本稿では、まず、データの量及びデータを活用する組織の能力（DUC）が具体的成果の獲得に影響を与えるかについて分析を行った。分析の結果、データの量は具体的成果の獲得に対して一部正の影響が見られたものの、その影響は限定的であった。一方、DUCはロバストに正に有意であり、具体的成果の獲得に大きく貢献していることが示唆された。

データは競争優位性を生み出す源になるという意味で、新たな天然資源であると言われる（渡部ほか、2018）。この比喻になぞらえるのであれば、データの量が多いということは天然資源が多いということである。原油のような天然資源はそれ自体を多量に保有することに価値がある場合もあるが、蒸留や精製を通じてはじめてガソリンや灯油といった利用可

図表7 DUC



図表 8 ロジスティック回帰分析の結果（上場企業のみ）

	Model 1b			Model 2b			Model 3b			Model 4b			Model 5b			Model 6b		
	B	有意 確率	標準誤差	B	有意 確率	標準誤差	B	有意 確率	標準誤差	B	有意 確率	標準誤差	B	有意 確率	標準誤差	B	有意 確率	標準誤差
データ総容量																		
データ総容量ダミー_1台のPC																		
データ総容量ダミー_サーバ室																		
DUC																		
Web 回答ダミー	0.462		(0.428)	0.345		(0.436)	0.333		(0.438)	0.281**		(0.049)	0.270**		(0.050)	0.268**		(0.050)
経営形態ダミー_純粋持株会社	-0.526		(0.629)	-0.538		(0.626)	-0.576		(0.625)	-0.731		(0.659)	-0.689		(0.654)	-0.739		(0.660)
経営形態ダミー_事業持株会社	0.397		(0.316)	0.345		(0.321)	0.327		(0.322)	0.491		(0.347)	0.477		(0.350)	0.472		(0.350)
企業年	-0.001		(0.005)	-0.001		(0.005)	-0.002		(0.005)	0.001		(0.005)	0.001		(0.006)	0.001		(0.006)
売上高	-0.00004		(0.00003)	-0.00003		(0.00003)	-0.00003		(0.00003)	-0.00003		(0.00004)	-0.00003		(0.00004)	-0.00003		(0.00004)
従業員数	0.0002**		(0.00005)	0.0001**		(0.00005)	0.0001**		(0.00005)	0.0001*		(0.00005)	0.0001*		(0.00004)	0.0001*		(0.00004)
業種数	-0.044		(0.153)	-0.037		(0.150)	-0.031		(0.151)	-0.204		(0.173)	-0.188		(0.172)	-0.183		(0.172)
製造業ダミー	0.406		(0.428)	0.509		(0.442)	0.516		(0.442)	1.025*		(0.498)	1.017*		(0.501)	1.012*		(0.502)
保有特許件数	-0.281*		(0.160)	-0.328*		(0.163)	-0.330*		(0.163)	-0.463**		(0.179)	-0.463**		(0.179)	-0.460*		(0.179)
定数	-1.525*		(0.607)	-2.826**		(0.839)	-1.659**		(0.635)	-4.586**		(0.877)	-4.932**		(1.015)	-4.580**		(0.898)
Cox-Snell R2 乗	0.074			0.090			0.093			0.201			0.198			0.199		
Nagelkerke R2 乗	0.113			0.137			0.142			0.304			0.300			0.302		
N	287			285			285			282			280			280		

* $p<0.1$, ** $p<0.05$, *** $p<0.01$

図表 9 ロジスティック回帰分析の結果（非上場企業のみ）

	Model 1c			Model 2c			Model 3c			Model 4c			Model 5c			Model 6c		
	B	有意 確率	標準誤差	B	有意 確率	標準誤差	B	有意 確率	標準誤差	B	有意 確率	標準誤差	B	有意 確率	標準誤差	B	有意 確率	標準誤差
データ総容量																		
データ総容量ダミー_1台のPC																		
データ総容量ダミー_サーバ室																		
DUC																		
Web 回答ダミー	0.482		(0.472)	0.617		(0.497)	0.545		(0.503)	0.276**		(0.051)	0.275**		(0.053)	0.276**		(0.053)
経営形態ダミー_純粋持株会社	-0.946		(1.105)	-0.941		(1.115)	-0.860		(1.132)	-1.063		(1.181)	-1.053		(1.181)	-0.938		(1.176)
経営形態ダミー_事業持株会社	-0.064		(0.367)	-0.004		(0.372)	-0.026		(0.374)	0.055		(0.418)	0.060		(0.418)	0.006		(0.426)
企業年	0.010		(0.010)	0.012		(0.010)	0.011		(0.010)	0.025*		(0.012)	0.025*		(0.012)	0.024*		(0.012)
売上高	-0.008		(0.009)	-0.009		(0.010)	-0.008		(0.010)	-0.004		(0.009)	-0.004		(0.009)	-0.003		(0.009)
従業員数	0.0001		(0.001)	0.0003		(0.002)	0.0002		(0.002)	-0.0004		(0.002)	-0.0004		(0.002)	-0.0004		(0.002)
業種数	-0.009		(0.097)	-0.031		(0.098)	-0.024		(0.098)	-0.107		(0.121)	-0.108		(0.121)	-0.092		(0.123)
製造業ダミー	-1.000*		(0.477)	-0.838*		(0.490)	-0.761		(0.496)	-0.666		(0.538)	-0.653		(0.540)	-0.583		(0.549)
保有特許件数	0.570*		(0.274)	0.491*		(0.282)	0.512*		(0.283)	0.562*		(0.301)	0.552*		(0.303)	0.568*		(0.305)
定数	-2.313**		(0.716)	-3.481**		(0.978)	-2.599**		(0.752)	-6.264**		(1.152)	-6.207**		(1.282)	-6.423**		(1.177)
Cox-Snell R2 乗	0.036			0.047			0.056			0.195			0.185			0.192		
Nagelkerke R2 乗	0.055			0.072			0.085			0.296			0.283			0.294		
N	223			222			222			221			220			220		

* $p<0.1$, ** $p<0.05$, *** $p<0.01$

能な状態となる。DXにおいても同様のことが言えるのではないだろうか。本稿の分析結果が示すように、データはただ多量に保有しているからといって一様に企業の競争力に結び付くとは限らず、それを使いこなす能力を組織として整備することが肝要であると考えられる。

我々は、本稿の分析から、組織としてデータを使いこなす活用する能力について重要な示唆を得たと考えている。本稿ではさらに、サンプルを上場企業と非上場企業に分けて集計及び回帰分析を行った。これらの結果、データの量に関しては非上場企業と比較して上場企業の方が多量のデータを扱っている企業が多い傾向が見られたものの、DUCやデータ利活用の成果に関しては大きな差がないことがわかった。また、具体的成果の獲得に対してデータの量の影響は限定的であるがDUCの影響はロバストであるという傾向は、上場企業と非上場企業で変わらないということも示唆された。つまり、DUCは企業規模や事業活動の長さによって決まるとは言えず、小さい企業や新しい企業でも培うことができる能力であると推察できる。

他方、先行研究では規模の小さい企業ではデータ利活用があまり行われていないことが指摘されてきた。例えば、元橋（2016）は製造業企業にアンケート調査を行い、大企業では約2割の企業で全社的活用（全社的にデータを活用している部署・部門がある）を行っており、約6割の企業で部門活用（全社的な部門は有しないが、企業内のいずれかの部門でデータ活用を行っている）を行っている一方、このデータ利用率は企業規模とともに小さくなることを示している。また、情報通信総合研究所（2020）は、分析に活用しているデータについて大企業と中小企業別に集計を行い、大企業の方がデータの活用割合が高く、中小企業では特にIoT関連のデータの活用が進んでいないと述べている。

本稿で用いたサンプルでは、上場企業と非上場企業とで売上高や従業員数といった企業規模が大きく異なるものの、大企業と中小企業とに分けて集計や分析を行ったわけではないため、これらの先行研究の結果と単純に比較をすることはできない。また、本稿で用いた質問票調査で非上場企業の抽出を行っ

たINITIALでは、「ユニークなテクノロジーや製品、サービス、ビジネスモデルを持ち、事業成長のための投資を行い、事業成長・拡大に取り組んでいる企業」及び「これまでの世界（生活、社会、経済モデル、テクノロジーなど）を覆し、新たな世界への変革にチャレンジする企業」を未公開のスタートアップ企業として収録している⁶。そのため、資金調達を行わず自社の売上にて成長をしている未上場企業やスモールビジネス等は対象外となっている。これらを考え合わせると、本稿でサンプルとした非上場企業は一般的な非上場企業や中小企業と比較してデータ利活用を積極的に行っている企業が含まれていた可能性が高い。それでも、非上場企業では取り扱っているデータの量が比較的少ないものの、DUCやデータ利活用の成果の水準は上場企業と差がなく、成果の獲得に対するデータの量やDUCの影響も差異がないという本研究の分析結果は注目に値するだろう。なぜなら、本稿でサンプルとした非上場企業は未公開のスタートアップ企業であるため、上場企業と比較してデータの量以外の他の様々な経営リソースも限りがあると推察される。たとえそのような状況に置かれていても、自社の事業に沿ってDUCを整備しデータを活用していけば、財務的な成果に結び付けることができるということを、本研究の分析結果は指し示しているからである。さらに言えば、データの量によって企業の成果や競争優位が決まってしまうとなれば、GAFA（グーグル、アップル、フェイスブック、アマゾン）やBAT（バaidu、アリババ、テンセント）に代表される巨大ITプラットフォーム企業に、現状の日本企業は太刀打ちできない。だが、もしDUCというデータを使いこなす能力が重要で、しかもそれが自社の事業や業務に紐付いているとなれば、データ量は少なくとも、豊富な業務、事業を抱える日本企業が、それらとデータを組み合わせて競争上の優位を構築する道筋も描けるのではないだろうか。

そこで、どうすればDUCという組織としての能力が得られるのかという次の課題が浮かび上がってくる。この課題に対し、例えば、高い分析スキルを有しているデータサイエンティストを雇用したり、育成したりすれば、図表2の分析能力（analytical

skills) を高め、DUC を高められるだろう。加えて、DUC の質問項目には、データの分析能力だけではなく、外部環境、自社の目標や目的、自社が成功するために鍵となる要因等の「ドメイン知識」を有していることや、自社のタスクとの関連性が深く質の高いデータを用いていることが含まれている(図表2)。それゆえ、自社ではどのような事業領域でデータを利活用すべきかを適切に判断し、それに適したデータを取得したり処理したりすることがDUCを高めることだと考えられる。この方向でDUCを高める具体的な打ち手もいくつか考えられるが、例えば、事業部門等のビジネスを担う部門とAI部門等のデータ解析を担う部門を橋渡しできるような人材の存在が有効ではないだろうか。その人材の候補としては知財部門の担当者に期待したい。例えば顧客のデータを扱うような場合には、預かるデータの管理と取り扱い、成果物の提供方法等、事業プロセス全般において知財リスクを管理する必要がある(渡部ほか, 2019)。また、近年では他社と連携してデータ利活用を行うような場合に、事業の初期段階(事業計画や提案等の段階)から知財・法務部門が関与するケースも増えてきている(渡部ほか, 2021)。データに関する権利や制度に関してはまだ定まっていない領域があるからこそ、知財部門の担当者がデータ利活用のプロセス全体を見通すことで、ビジネスを担う部門とデータ解析を担う部門が安心してデータを利活用できるようになる。その結果、組織としてのDUCを養い、データ利活用の成果を向上させることにつながるのではないだろうか。

Society 5.0の実現に向けたデータ利活用に関する研究はまだはじまったばかりであり、さらなる研究が必要であると思われる。本稿がその良いきっかけとなれば幸いである。

【謝辞】 本稿は、独立行政法人経済産業研究所(RIETI)における研究プロジェクト「データとAI利活用促進をグローバルに展開するための制度とマネジメントに関する研究—グローバルデータサプライチェーンの確立に向けて」の成果の一部であり、当該研究プロジェクトのディスカッションペーパー(渡部ほか, 2021)の一部を発展させたものである。

本研究成果の一部はJSPS科研費(基盤B(19H01517),

基盤C(18K01788), 若手研究(18K12835))の助成を受けたものである。

注

- 1 https://www.rieti.go.jp/jp/projects/program_2020/pg-04/001.html (2021年10月13日最終アクセス)
- 2 https://www.rieti.go.jp/jp/projects/program_2016/pg-04/005.html (2021年10月16日最終アクセス)
- 3 https://www.rieti.go.jp/jp/projects/research_activity/data_utilization/ (2021年10月16日最終アクセス)
- 4 https://www.rieti.go.jp/jp/projects/research_activity/data_utilization2020/ (2021年10月13日最終アクセス)
- 5 <https://initial.inc/> (2021年10月13日最終アクセス)
- 6 <https://help.initial.inc/ja/articles/4570807-%E3%82%B9%E3%82%BF%E3%83%BC%E3%83%88%E3%82%A2%E3%83%83%E3%83%97%E3%81%A8%E3%81%AF> (2021年10月20日最終アクセス)

参考文献

- Akter, S., Wamba, S.F., Gunasekaran, A., Dubey, R. and Childe, S.J. (2016) "How to improve firm performance using big data analytics capability and business strategy alignment?", *International Journal of Production Economics*, Vol. 182, pp. 113-131.
- Barney, J. (1991) "Firm resources and sustained competitive advantage", *Journal of management*, Vol. 17, No. 1, pp. 99-120.
- Bharadwaj, A.S. (2000) "A resource-based perspective on information technology capability and firm performance: an empirical investigation", *MIS Quarterly*, Vol. 24, No. 1, pp. 169-196.
- Brynjolfsson, E. and Yang, S. (1996) "Information technology and productivity: a review of the literature", *Advances in computers*, Vol. 43, pp. 179-214.
- Dubey, R., Gunasekaran, A., Childe, S.J., Roubaud, D., Wamba, S.F., Giannakis, M. and Forooun, C. (2019) "Big data analytics and organizational culture as complements to swift trust and collaborative performance in the humanitarian supply chain", *International Journal of Production Economics*, Vol. 210, pp. 120-136.
- Garmaki, M., Boughzala, I. and Wamba, S.F. (2016) "The effect of big data analytics capability on firm performance", *In Proceedings of PACIS (Pacific Asia Conference on Information Systems) 2016*, p. 301.
- Ghasemaghaei, M., Ebrahimi, S. and Hassanein, K. (2018) "Data analytics competency for improving firm decision making performance", *The Journal of Strategic Information Systems*, Vol. 27, No. 1, pp. 101-113.
- Ghasemaghaei, M. (2019) "Does data analytics use improve firm decision making quality? The role of knowledge sharing and data analytics competency", *Decision Support Systems*, Vol. 120, pp. 14-24.
- Grant, R.M. (1991) "The resource-based theory of competitive advantage: implications for strategy formulation", *California management review*, Vol. 33, No. 3, pp. 114-135.
- Gupta, M. and George, J.F. (2016) "Toward the development of a big data analytics capability", *Information & Management*, Vol. 53, No. 8, pp. 1049-1064.
- Hirai, Y., Tatsumoto, H. and Watanabe, T. (2019) "The effect of big data and advanced analytics in Japan", *In Proceedings of ISPIM (International Society for Professional Innovation Management) Connects Ottawa*.
- Mikalef, P., Pappas, I.O., Krogstie, J. and Giannakos, M. (2018) "Big data analytics capabilities: a systematic literature review and research agenda", *Information Systems and e-Business Management*, Vol. 16, No. 3, pp. 547-578.
- Wamba, S.F., Gunasekaran, A., Akter, S., Ren, S.J.F., Dubey, R.

- and Childe, S.J. (2017) "Big data analytics and firm performance: Effects of dynamic capabilities", *Journal of Business Research*, Vol. 70, pp. 356-365.
- 大橋弘・中村豪 (2020) 「データ資産と企業のマークアップ: 定量的なアプローチ」『情報通信政策研究』第4巻第1号, pp. 33-46.
- 情報通信総合研究所 (2020) 「デジタルデータの経済的価値の計測と活用の現状に関する調査研究の請負報告書」総務省情報流通行政局情報通信政策課情報通信経済室 (https://www.soumu.go.jp/johotsu_sintokei/linkdata/r02_05_houkoku.pdf, 2021年10月20日最終アクセス).
- 総務省 (2021) 「AIネットワーク社会推進会議 AI 経済検討会報告書 2021」総務省, (https://www.soumu.go.jp/main_content/000765261.pdf, 2021年10月16日最終アクセス).
- 立本博文・平井祐理・渡部俊也 (2018) 「データ利活用に関する質問票調査を用いた産業別比較」*RIETI Discussion Paper Series* 18-J-032, 経済産業研究所, (<https://www.rieti.go.jp/jp/publications/dp/18j032.pdf>, 2021年10月16日最終アクセス).
- 元橋一之 (2016) 「日本の製造業におけるビッグデータ活用とイノベーションに関する実態」*RIETI Policy Discussion Paper Series* 16-P-012, 経済産業研究所, (<https://www.rieti.go.jp/jp/publications/pdp/16p012.pdf>, 2021年10月20日最終アクセス).
- 渡部俊也・平井祐理・阿久津匡美・日置巴美・永井徳人 (2018) 「企業において発生するデータの管理と活用に関する研究」*RIETI Discussion Paper Series* 18-J-028, 経済産業研究所, (<https://www.rieti.go.jp/jp/publications/dp/18j028.pdf>, 2021年10月16日最終アクセス).
- 渡部俊也・平井祐理・阿久津匡美・前田三奈 (2019) 「データとAIがもたらす第四次産業革命: データ利活用の推進を巡る産学官の取り組み」『日立評論』第101巻第2号, pp. 49-54.
- 渡部俊也・平井祐理 (2020) 「データの経済的価値と法的保護 実証分析とケースによる考察」『情報通信政策研究』第4巻第1号, pp. 19-31.
- 渡部俊也・平井祐理・吉岡(小林) 徹・金間大介・立本博文・古谷真帆・永沼麻奈香 (2021) 「企業において発生するデータの管理と活用—質問票調査による実態把握」*RIETI Discussion Paper Series* 21-J-017, 経済産業研究所, (<https://www.rieti.go.jp/jp/publications/dp/21j017.pdf>, 2021年10月13日最終アクセス).