



RIETI Discussion Paper Series 21-J-017

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

渡部 俊也
経済産業研究所

平井 祐理
文部科学省

吉岡（小林） 徹
一橋大学

金間 大介
金沢大学

立本 博文
筑波大学

古谷 真帆
一般社団法人日本知的財産協会

永沼 麻奈香
株式会社日立製作所



Research Institute of Economy, Trade & Industry, IAA

独立行政法人経済産業研究所

<https://www.rieti.go.jp/jp/>

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¹

渡部 俊也（経済産業研究所）

平井 祐理（文部科学省） 吉岡（小林） 徹（一橋大学）

金間 大介（金沢大学） 立本 博文（筑波大学）

古谷 真帆（一般社団法人日本知的財産協会） 永沼 麻奈香（株式会社日立製作所）

要 旨

第 4 次産業革命の進展に伴って、あらゆる産業の企業にとってデータの効果的な利活用は不可避の課題となっている。また、データ利活用は自社のみで行うだけでなく、他組織と連携して行うケースも増えてきている。そこで本研究では、データ利活用の実態を把握し、それにより成果を得るために重要な要因は何かについて明らかにすることを目的として日本企業を対象とした質問票調査を実施した。日本企業がどの程度データ利活用を行っているのか、データ利活用の際にどのように他組織と連携しているのか等について分析を行うとともに、過去に実施した質問票調査も用いて 3 年前との比較を行った。あわせて今後重要となるクロスボーダーのデータ取引についても実態について検討し、クロスボーダー契約のあり方についても考察を加えた。

キーワード：データ利活用、ビッグデータ、人工知能、組織能力、データ契約

JEL classification: O34

RIETI ディスカッション・ペーパーは、専門論文の形式でまとめられた研究成果を公開し、活発な議論を喚起することを目的としています。論文に述べられている見解は執筆者個人の責任で発表するものであり、所属する組織及び（独）経済産業研究所としての見解を示すものではありません。

¹本稿は、独立行政法人経済産業研究所（RIETI）におけるプロジェクト「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」の成果の一部である。本稿の分析に当たっては、RIETI が実施した平成 29 年度「データ利活用に関するアンケート調査」及び 2020 年度「データ利活用に関するアンケート調査」を利用した。また、本稿の原案に対して、本プロジェクトのプロジェクトメンバー、オブザーバー、経済産業省経済産業政策局知的財産政策室、経済産業省商務情報政策局情報経済課、ならびに RIETI ディスカッション・ペーパー検討会の方々から多くの有益なコメントを頂いた。ここに記して、感謝の意を表したい。

目次

■概要編

1. はじめに
2. 最近の国内制度整備と政策の状況
3. インダストリアルデータの越境問題
4. インダストリーデータについての国内外における最新動向と本報告書の構成

■分析編

5. 質問票調査
 5. 1. 調査概要
 5. 2. 集計
 5. 3. 回帰分析①：2017 年との比較
 5. 4. 回帰分析②：Data Analytics Competency についての分析
 5. 5. 回帰分析③：データ利活用の戦略的取組みの進捗プロセスと成果の関係の分析
 5. 6. 回帰分析④：データ取得・提供プロジェクトにおける法務体制の影響の分析
 5. 7. 産業別 Data Analytics Competency についての分析

■調査編

6. 調査概要
7. データの越境フローに関する視点と規制
 7. 1. データフリーフロー（DFF）とデータローカライゼーション
 7. 2. 諸外国のデータ越境フロー規制
8. 諸外国における企業のデータ利活用と越境移転にかかるモデル契約
 8. 1. 外国における企業のデータ利活用に関する調査
 8. 2. 外国におけるデータ移転のモデル契約

■Appendix 2020 年度「データ利活用に関するアンケート調査」調査票

1. はじめに

新型コロナウイルス感染症パンデミックは否が応でも社会と産業の急速なデジタル化を促進している。デジタル化が進んだ経済社会において、価値を生み出す源泉として最も重要視されているのがデータである。データから生み出される価値は、富を生み出す経済的価値や、多様な個人に合わせたサービスを提供するための基盤となることで生み出される社会的価値、そして国家にとっては安全保障にかかわる情報としての価値をも有する。それゆえにデータの利活用は、意図しない形で悪用されるリスクも伴う。パーソナルデータの不用意な取り扱いに関してはプライバシーにかかわる問題や人権問題などにも及ぶが、インダストリーデータについても、そのデータが企業の生産活動に基づくものであれば、企業の保有する技術ノウハウが一举に流出してしまうリスクも伴う。パーソナルデータもインダストリーデータも等しく、利活用による大きな価値とそれに伴うリスクの両面をどのように社会経済が管理するのかといった問題が、デジタル化に伴う大きな課題となっている。

RIETI のプロジェクトとして 2016 年 8 月 1 日～2018 年 9 月 30 日に実施した「企業において発生するデータの管理と活用に関する実証研究」においては、インダストリーデータについて国内企業のデータ利活用パフォーマンスに対する詳細な質問票調査を実施した。その結果はデータの利活用の成果に対しては、アクセス可能なデータ量が大きいことが重要なのではなく、データを取り扱う組織の能力や、他者と連携してデータを利用するための契約習熟度などが重要であることを明らかにするとともに、データ利活用契約の考え方について具体的事例を示して整理した。

今回の「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」においては、前回と同様、日本の企業におけるデータ利活用実態について、前回との変化を把握するとともに、前回対象としていなかったデータのクロスボーダー取引の実態や、平成 30 年の不正競争防止法改正において導入された限定提供データの活用実態に関しても調査を行い、分析を行ったものである。

あわせてクロスボーダーデータ取引に及ぼす各国の関連データ保護制度及び企業のデータ利活用の実情やデータ越境移転にかかるモデル契約についての実情について文献調査を行った。

2. 最近の国内制度整備と政策の状況

日本政府によるデータに関する施策としては、改正不正競争防止法（2019 年 7 月 1 日施行）において悪質性の高いデータの不正取引を不競法に基づく「不正競争行為」と位置づけ、差止請求権、損害賠償請求）が認められている。この限定提供データは、業として特定の者に提供する情報として電磁的方法（電子的方法、磁気的方法その他の知覚 によっては認識することができない方法をいう。）により相当量蓄積され、及び管理されている技術上又は営業上の情報（秘密として管理されているものを除く）であって「特定の者に提供する情報」でなければならない。そしてデータ提供契約に基づいて提供しようとするデータがこの制度によって保護される条件としては、データ提供契約の中で第三者提供を禁止することが求められる。その場合のデータは暗号化をしたり、提供データが蔵置されたサーバへのアクセスを ID とパスワードで管理したりするなどの手段をとる必要がある。

本制度は 2016 年から 2017 年にかけて知的財産戦略本部および経産省、特許庁において当時の欧州の動向として Sui generis right（特別の権利）の見直しや、データを生成している主体にデータの使用权（データ生成権）を付与する等の議論も踏まえた議論がなされたものであり、2017 年 3 月に示された「知的財産戦略本部 検証・評価・企画委員会 新たな情報財検討委員会 報告書（抄）」においては、データの保護については利用を拒否することができる排他的な権利として物権的な権利を設定することについて望ましいとは言えない。としつつ、価値あるデータの保有者及び利用者が、安心してデータを提供しかつ利用できる公正な競争秩序を確保するため、新たな不正競争行為の対象となる行為や保護対象となるデータについて、先端ビジネスや事業に及ぼす影響に留意しつつ、産業の実態を踏まえ、具体的に検討を進めることが適当と結論付けている。

この報告と相前後して 2017 年 4 月に第四次産業革命を視野に入れた知財システムの在り方に関する検討会においても、排他的な知的財産権を付与することは、データの流通を阻害するのではな

いかといった意見が強く、むしろ価値あるデータの不正取得については、不正競争防止法における行為規制アプローチを考慮すべきという結論が示唆されたものである¹。

そしてその後 2017 年 12 月より、産業構造審議会「営業秘密の保護・活用に関する小委員会」及び「不正競争防止小委員会」において審議を行い、2018 年 1 月 19 日に中間報告を策定・公表。第 196 回国会にて、「不正競争防止法等の一部を改正する法律」が成立し、2018 年 5 月 30 日に公布されている。

またこのような新たな保護制度の創設と並行して、データの契約による保護の在り方については 2018 年 6 月に発表された「AI・データの利用に関する契約ガイドライン」において詳しく検討されている。経緯としては先立つ 2017 年 5 月に「データの利用権限に関する契約ガイドライン Ver1.0」が策定しているものの、さらに詳しい検討が必要との事業者・事業者団体等からの意見も踏まえ、AI・データ契約ガイドライン検討会が設置されて、データの利用に関する契約類型の整理・深掘りやユースケースについての検討や、AI 開発・利用に関する権利関係・責任関係等の考え方で射程に入れた議論を行い抜本的な改訂が行われたものである²。

本研究においてはこのような新たな制度の導入がデータ利活用実態にどのような影響を及ぼしているかについても検討する。

3. インダストリアルデータの越境問題

機械学習等で優れた推論エンジンを得るためには、データの収集量が重要であり、工作機械や建築機械、自動運転など世界中でデータ収集を行って学習させることができることが望ましい。その点産業界からはデータの自由な流通が求められる。

一方インダストリアルデータについて、新興国を中心にデータローカライゼーション（data localization）規制が拡大しつつある。具体的にはデータが格納されるサーバー設備の国内設置を求めるものであり、個人データに関する越境データ移転規制と異なり、ある国において特定の事業活動を営む場合に、当該事業活動に必要なサーバーやデータ自体の国内設置・保存を求める規制である。また、越境個人データ移転規制では原則として本人の同意があれば海外への移転が可能であるが、データローカライゼーション規制では、対象データが個人データに限られないため、本人の同意による移転は行い得ず、データの越境移転にあたっては、当該国政府の許可等が必要となる。

特に中国においては 2017 年 6 月に中国で施行された、サイバーセキュリティ法（中華人民共和国网络安全法）11 第 37 条において、何らかのネットワークを所有・運営する「ネットワーク運営者」のうち、特に国の安全や国民経済と民生、公共の利益に与える影響が大きいと指定される「重要情報インフラストラクチャーの運営者」について、それが保有する個人情報及び「重要データ」を中国国外に移転するにあたっては、国が定める基準に従い「安全評価」を行わなければならないことを定めていることが注目される。

データの利活用においては、クロスボーダーのデータ取引や収集は不可欠となるが、実際にデータのクロスボーダー取引がどの程度行われており、これらのデータ越境規制がどのような影響を与えているのかについての情報はなかったことから、本研究ではこのようなデータのクロスボーダー取引についても調査対象として検討した。

4. インダストリーデータについての国内外における最新動向と本報告書の構成

EU は最近産業データ等を対象とする政策に取り組んでおり、2018 年には非個人データの EU 域内自由流通枠組み規則³の制定、2020 年の 2 月 19 日に、“A European Strategy for Data”を発表するなどしてインダストリーデータについて新たな制度導入の動きがある。この提言では、最終目

¹ <https://www.meti.go.jp/report/whitepaper/data/pdf/20170419001-1.pdf>

² その後限定提供データ制度などの総説も加味した AI・データの利用に関する契約ガイドライン 1.1 版が 2019 年 12 月に公表されている。

<https://www.meti.go.jp/press/2019/12/20191209001/20191209001.html>

³ the Regulation on the free flow of non-personal data (FFD)

的としてヨーロッパの単一のデータスペースを作成し個人データと非個人データについて、企業が簡単にアクセスでき安全性の高いシステムの構築を目指すとしており、クロスセクターのデータ共有が進まないことやデータの公益性の観点から必要なデータにアクセスできない課題を解決するための「信頼に基づくデータガバナンス」を提言してデータガバナンス法を適用するとしている。また 2019 年 10 月 29 日ドイツで開催されたデジタルサミットで、欧州クラウド／データインフラ構想である GAIA-X プロジェクトが発表された。使いやすく、競争力のある、安全で信頼できるデータインフラを整備するとして、GAIA-X を進める非営利組織 Gaia-X Foundation を設立し Deutsche Telekom、Orange、SAP、Bosch、Siemens などの欧州地元企業を中心に、300 以上の企業や組織が参画しており、インダストリーデータ中心のデータ利活用の大規模な枠組みが実装されつつある。

これらの欧州の動向は、先行してパーソナルデータ中心にデータ利活用を進めた GAFA のようなプラットフォームをインダストリーデータでは EU のイニシアティブを回復させたいという意図があることは明確である。このレポートについては、欧州内外からの反響があり、米国と欧州にまたがる団体などからも反応が示されている。

一方、我が国は昨年の G20 大阪サミットにおいて議長国として「信頼ある自由なデータ流通」(DFFT)のコンセプトを提案しておりこれが我が国のデータガバナンスの基本的ポリシーとなっている。その後の EU の動きを受けて、DFFT の考え方に沿った国際的なルール作りを加速する方針が知的財産戦略計画において示されていることに加え日本経団連からは、過度の規制や域外適用を懸念する指摘を含む「欧州委員会“**A European strategy for data**”への意見」⁴が、またや自由民主党からは「わが国が G20 で提唱した信頼性のある自由なデータ流通 (DFFT) の推進方法を具体化し、産業データの利活用に関する国際ルールの形成を主導し、わが国の国益を追求するためにも、2021 年に法整備を目指す欧州に先駆けて、わが国として実効性あるルール整備を行うことが必須である」とする「リアルデータの利活用推進について」が自民党新国際秩序創造戦略本部へ報告されている(令和 2 年 12 月 4 日)⁵。

このように、EU の動向が核となってインダストリーデータ政策についての議論が行われているといつてよい。

先述したように、日本においては当時の欧州の動向も踏まえ、データの保護について様々な議論を経た結果、行為規制としての限定提供データの保護という一つの結論にたどり着いたといえるが、最近の欧州の動向から、改めてインダストリーデータのガバナンスの具体的なあり方について様々な議論がなされるに至っている。

その一つは中小企業の活動などに伴うデータの流通が円滑に促進されるためには、データが生み出される企業に対して、個人情報の保護制度に類する保護制度が必要なのではないかとする意見である。欧州でも 2017 年に欧州委員会が「欧州データ経済の構築 (Building a European Data Economy)」に係るスタッフ作業文書の中で、機械的に生成される生データの保護措置として「データプロデューサーの権利 (data producer's right)」を物権的な権利として創設することを議題として提案した。そののちの学会や経済界の強い批判があり、現段階でこのような権利の導入が進められるという状況にはなっていない⁶。一方欧州委員会は、「物権的な権利ではなく防御的な権利 (a set of purely defensive rights) として考えることも可能である」との記載もあり、これについても議論がなされており一定の評価が得られている。これは後述する日本で導入された限定提供データに近い制度であり、欧州でも日本と同様の行為規制としての対応がなされる可能性があるものと思われる。

いずれにしても、現段階でデータの保護がどうあるべきかがますます重要な政策的焦点となっていることを踏まえ、エビデンスとしてのデータ利活用実態とそれを阻む問題について、前回の調査に引き続いて本研究で把握し、データ利活用のパフォーマンスを向上させる要因と、その阻害要因等の分析、さらにはクロスボーダー取引に関して影響を与えると思われる各国データ関連制度につ

⁴ <https://www.keidanren.or.jp/policy/2020/050.html>

⁵ https://jimin.jp-east-2.storage.api.nifcloud.com/pdf/news/policy/201021_3.pdf

⁶ <https://system.jpaa.or.jp/patent/viewPdf/3556>

いての調査を行ったものである。

本報告書は、大きく分けて質問票調査の結果解析を中心に記載した分析編（５章）と、クロスボーダーデータ取引に影響すると思われる各国データ関連制度について文献調査を行った調査編（６，７，８章）の２つのパートからなる。また巻末に実施した質問票を付した。

■分析編

5. 質問票調査

5. 1. 調査概要

今回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」⁷では、2020 年度「データ利活用に関するアンケート調査」⁸を実施した。このアンケート調査では、①上場企業 3685 社、②未上場企業 2252 社の合計 5937 社を対象とした。①上場企業の企業リストに関しては、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 に上場する企業について、株式会社東京商工リサーチのデータベースより 2020 年 6 月 3 日時点に切り出したものを使用した。②未上場企業は、スタートアップ情報プラットフォーム「INITIAL」⁹を用いて、2020 年 5 月 29 日に検索式を「タイプ：未公開企業、かつ、従業員数：20 人以上、かつ、調査状況¹⁰：未評価 and 調査継続中」として抽出した。抽出した 2301 社のうち、株式会社以外の企業（22 社）及び倒産等の企業（27 社）は除外し、最終的に 2252 社を調査対象とした。回答方法は Web 回答もしくは郵送回答とし、回収期間は 2020 年 6 月 26 日～2020 年 8 月 31 日であった。657 社からの有効回答を得て、有効回答率は 11.1%であった。回収状況の詳細を図表 1、図表 2 に示す。また、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。調査票は 2 部構成であり、第 I 部（問 1）では従業員数、業種等の企業概要について質問し、第 II 部ではデータ利活用について質問をした。第 II 部では、データ利活用の全般的な状況について（問 2）、自社で保持しているデータを自社内で利活用する場合について（問 3）、データ利活用を行うことを主として他組織が保持しているデータを受領し提携契約を結んだプロジェクト（データ取得プロジェクト）について（問 4）、データ利活用を行うことを主として自社で保持しているデータを他組織に提供し提携契約を結んだプロジェクト（データ提供プロジェクト）について（問 5）、のそれぞれについて質問をした（図表 3）。なお、経営形態が純粋持株会社である企業に関しては、子会社等の傘下の主要なグループ会社を 1 社選び、その企業について回答を依頼した。

また、本稿では、3 年前との比較を行うために、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」¹¹において 2017 年 9 月～2017 年 11 月に実施した平成 29 年度「データ利活用に関するアンケート調査」¹²も利用した。この調査に関する詳細は渡部ら（2018）¹³を参照されたいが、この調査では①日本における平成 24 年の特許出願件数上位 5000 社から自治体や大学等を除いた企業、②東証一部上場企業、③ビジネス SNS サイト「WANTEDLY」¹⁴から抽出

⁷ https://www.rieti.go.jp/jp/projects/program_2020/pg-04/001.html （2021 年 1 月 28 日最終アクセス）

⁸ https://www.rieti.go.jp/jp/projects/research_activity/data_utilization2020/ （2021 年 1 月 28 日最終アクセス）

⁹ <https://initial.inc/> （2021 年 1 月 28 日最終アクセス）

¹⁰ 調査状況の定義については INITIAL のホームページ

（<https://help.initial.inc/ja/articles/4570991-%E8%AA%BF%E6%9F%BB%E7%8A%B6%E6%B3%81>）を参照されたい（2021 年 1 月 28 日最終アクセス）。

¹¹ https://www.rieti.go.jp/jp/projects/program_2016/pg-04/005.html （2021 年 1 月 28 日最終アクセス）

¹² https://www.rieti.go.jp/jp/projects/research_activity/data_utilization/ （2021 年 1 月 28 日最終アクセス）

¹³ 渡部俊也，平井祐理，阿久津匡美，日置巴美，永井徳人，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（2018）。

<https://www.rieti.go.jp/jp/publications/dp/18j028.pdf> （2021 年 1 月 28 日最終アクセス）

¹⁴ <https://www.wantedly.com/> （2021 年 1 月 28 日最終アクセス）

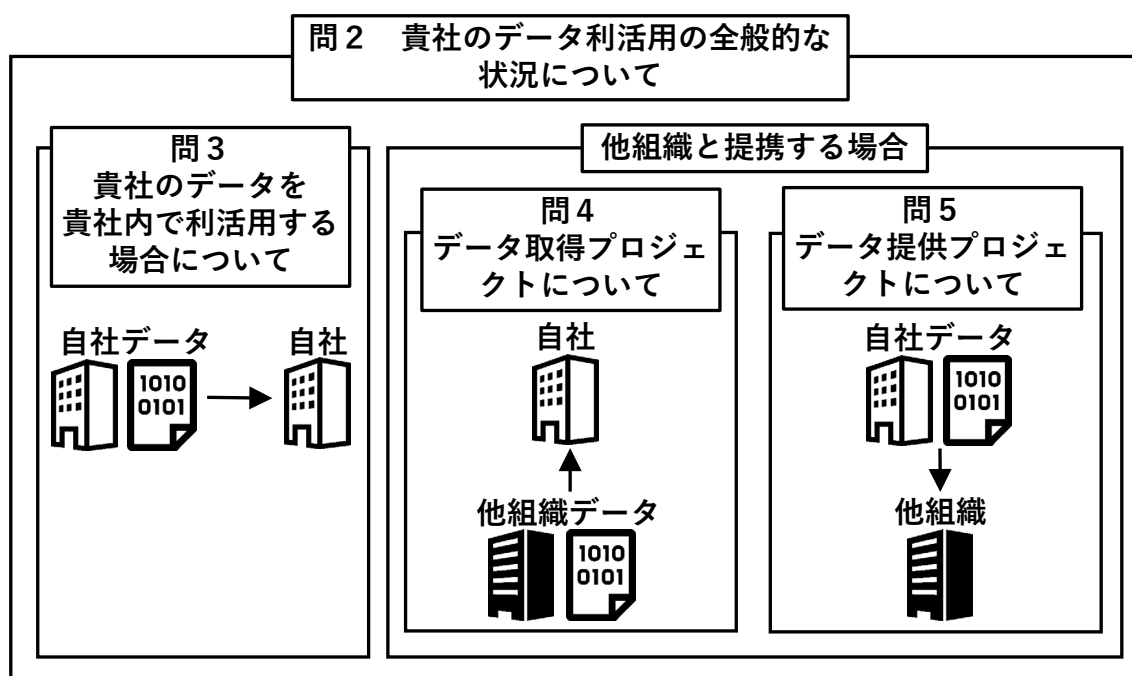
した企業、を調査対象としており、2020 年度のアンケート調査とは調査対象が異なることは注意が必要である。なお、平成 29 年度のアンケート調査では 2016 年度（2017 年 3 月末時点）の状況について回答を依頼した。

	上場	未上場	計
有効回答(社)	372	285	657
調査票送付(社)	3685	2252	5937
有効回答率(%)	10.1	12.7	11.1

図表 1 回収状況の詳細 1

	上場	未上場	計
Web回答(社)	295	228	523
郵送回答(社)	77	57	134
計(社)	372	285	657

図表 2 回収状況の詳細 2



図表 3 調査票第Ⅱ部の構成

5. 2. 集計

ここでは、2020 年度「データ利活用に関するアンケート調査」（以下、「2020 年」と表記）の主な集計結果について記載する。3 年前との比較を行うため、一部に平成 29 年度「データ利活用に関するアンケート調査」（以下、「2017 年」と表記）の結果も掲載する¹⁵が、特に断りのない場合は 2020 年の結果である。

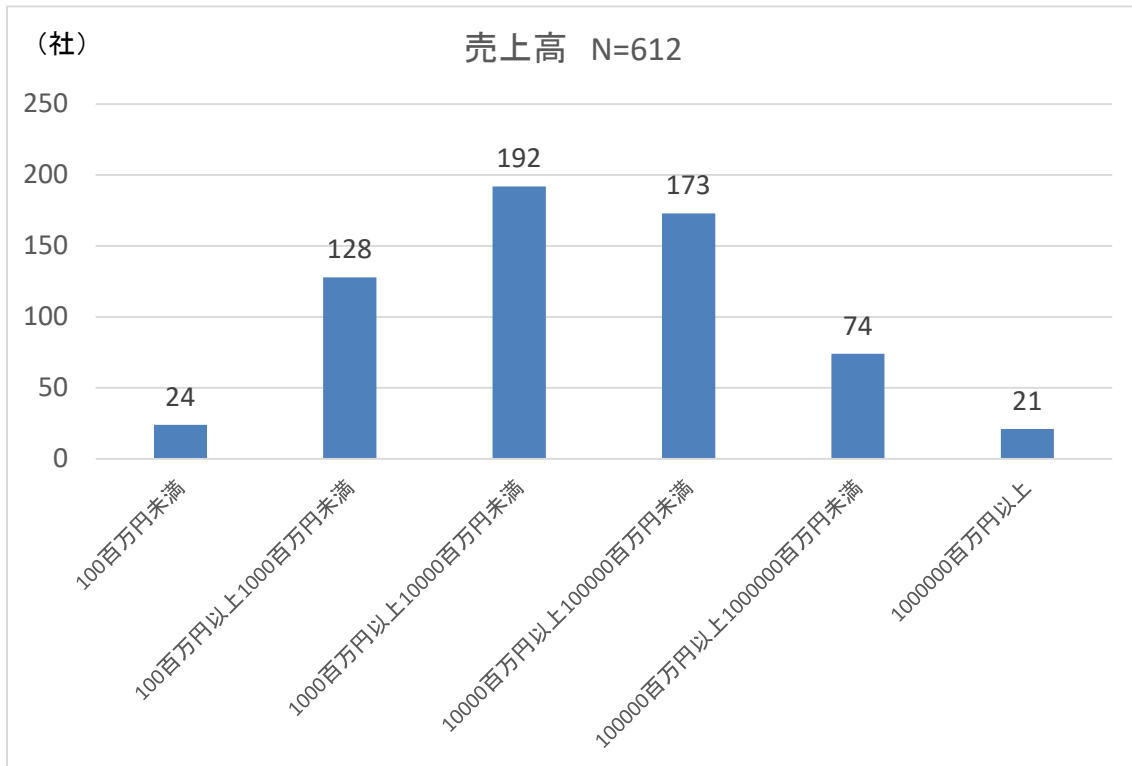
5. 2. 1. 回答企業の概要について（問 1）

図表 4 は、回答企業の売上高についてのグラフである。売上高が 10 億円以上 100 億円未満の企業が最も多いが、1 億円未満や 1 兆円以上の企業も含まれている。図表 5 は、従業員数（非正社員も含む）のグラフである。10 人以上 100 人未満の企業が最も多いが、10 人未満や 1 万人以上の企業も含まれる。また、図表 6 は、回答企業の業種（自社の事業を最もよく表す業種についての単一回答）である。「32 情報／システム／ソフト」が最も多く、次いで「28 卸売／小売」、「35 その他サービス」となっている。

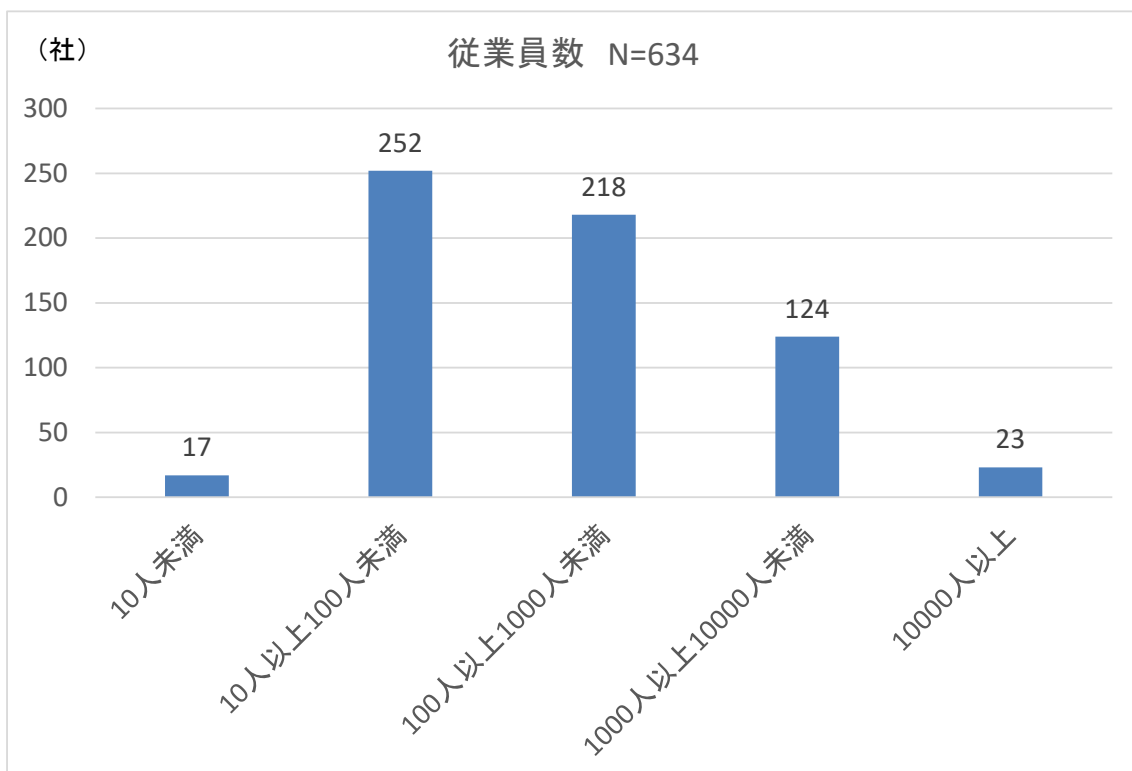
¹⁵ 2017 年の集計結果は、下記からの再掲（もしくは再編の上、再掲）である。

渡部俊也，平井祐理，阿久津匡美，日置巴美，永井徳人，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（2018）。

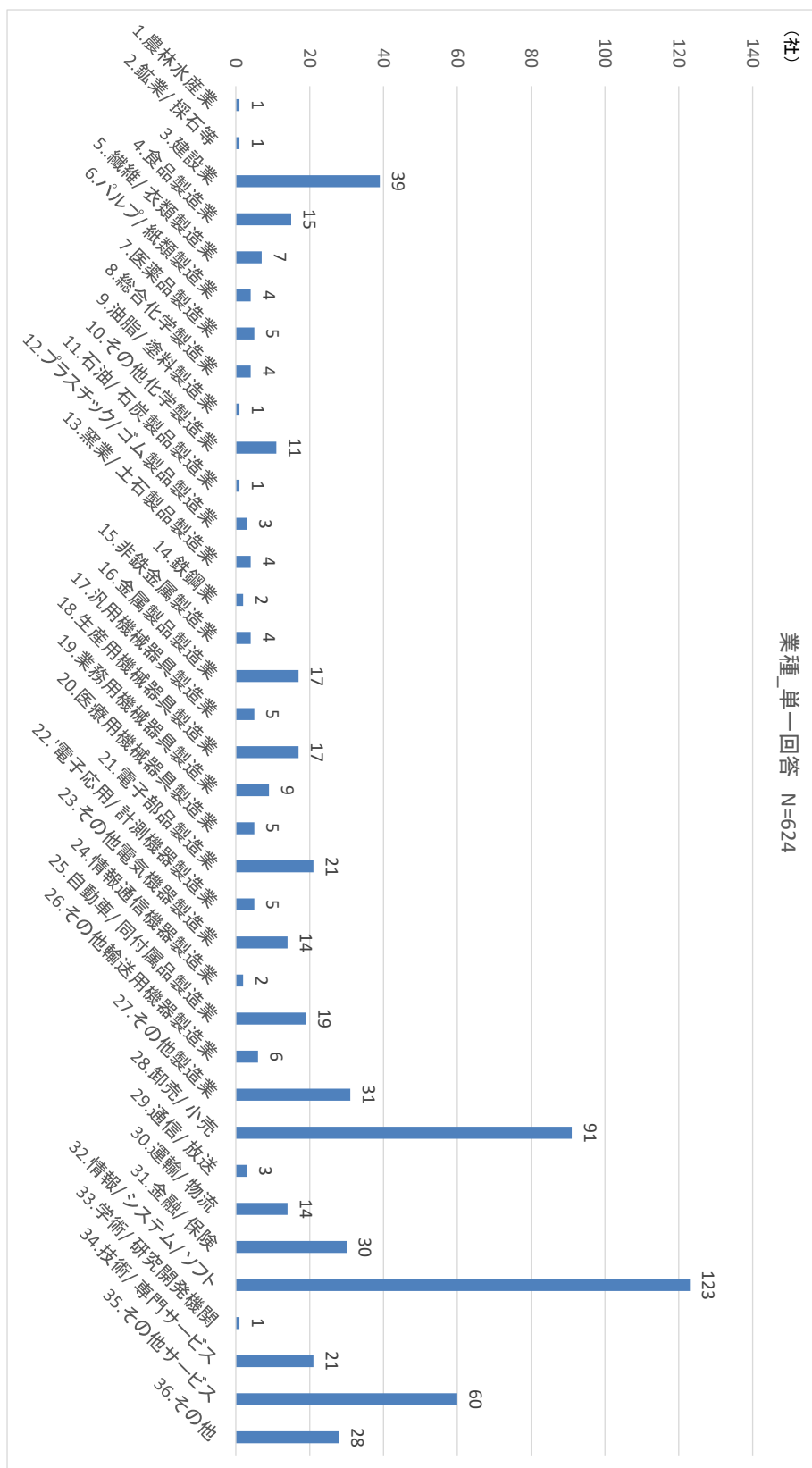
<https://www.rieti.go.jp/jp/publications/dp/18j028.pdf>（2021 年 1 月 28 日最終アクセス）



図表 4 売上高 (2020 年)



図表 5 従業員数 (2020 年)



図表 6 業種（単一回答、2020 年）

5. 2. 2. データ利活用の全般的な状況について（問2）

図表 7、図表 8 は、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についてのグラフである。「数台のサーバで管理できる程度」と回答した企業の割合は 2017 年では 51.1%、2020 年では 55.0%とやや増加している。しかし、「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 2017 年では 33.3%、2020 年では 33.8%であり、ほとんど増加しておらず、多量のデータを扱う企業が大きく増加したわけではないと思われる。

図表 9、図表 10 は、自社に利用権限があるデータの総容量のうち、実際に利活用を行っているデータの容量についてのグラフである。「20%未満」と回答した企業の割合は 2017 年と比較して 2020 年ではやや減少したものの、大まかな傾向としては 3 年間であまり変化がない。

図表 11、図表 12 は、全社的なデータ利活用を推進する専門部門や担当者についてのグラフである。「担当を設けていない」と回答した企業の割合は、2017 年と比較して 2020 年ではやや減少している。

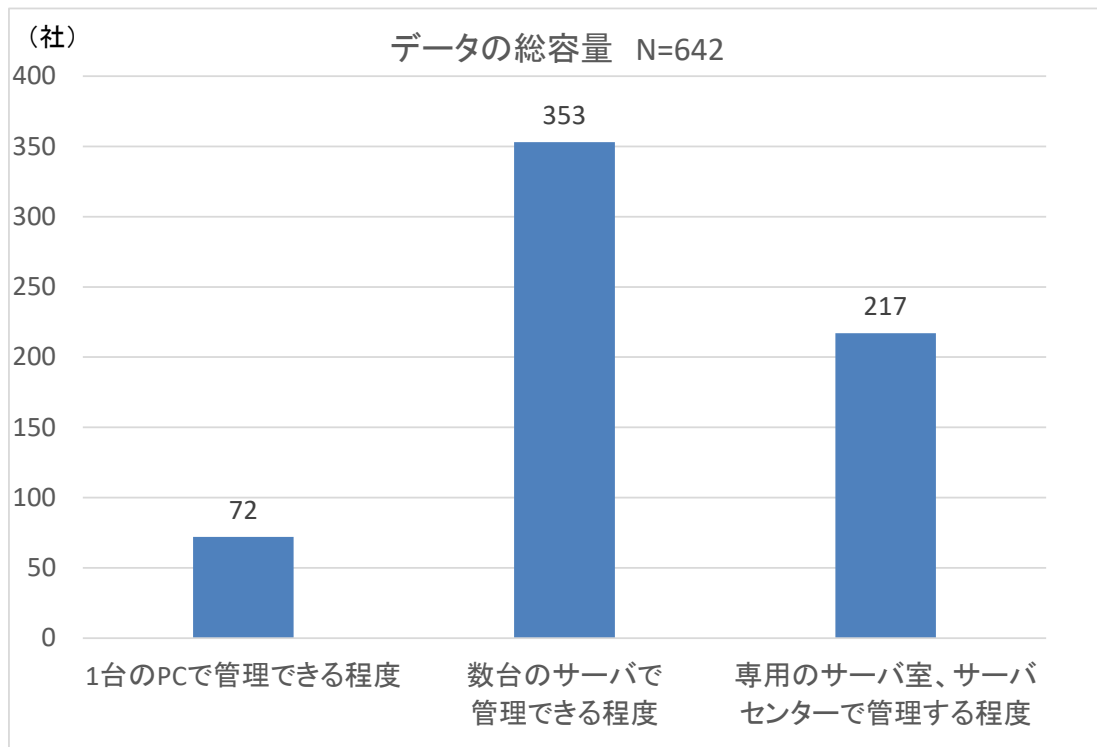
図表 13 は、データサイエンティストの人数（2017 年には同様の質問なし）に関するグラフである。67.5%の企業が「0 人」と回答している。

図表 14、図表 15 は、「ビッグデータの利活用を行える体制が整備されている」かどうかについての 5 段階評価に関するグラフである。「強くそう思う」「そう思う」と回答した企業の割合は 2017 年では 12.3%、2020 年では 12.1%でありほとんど変化がない。

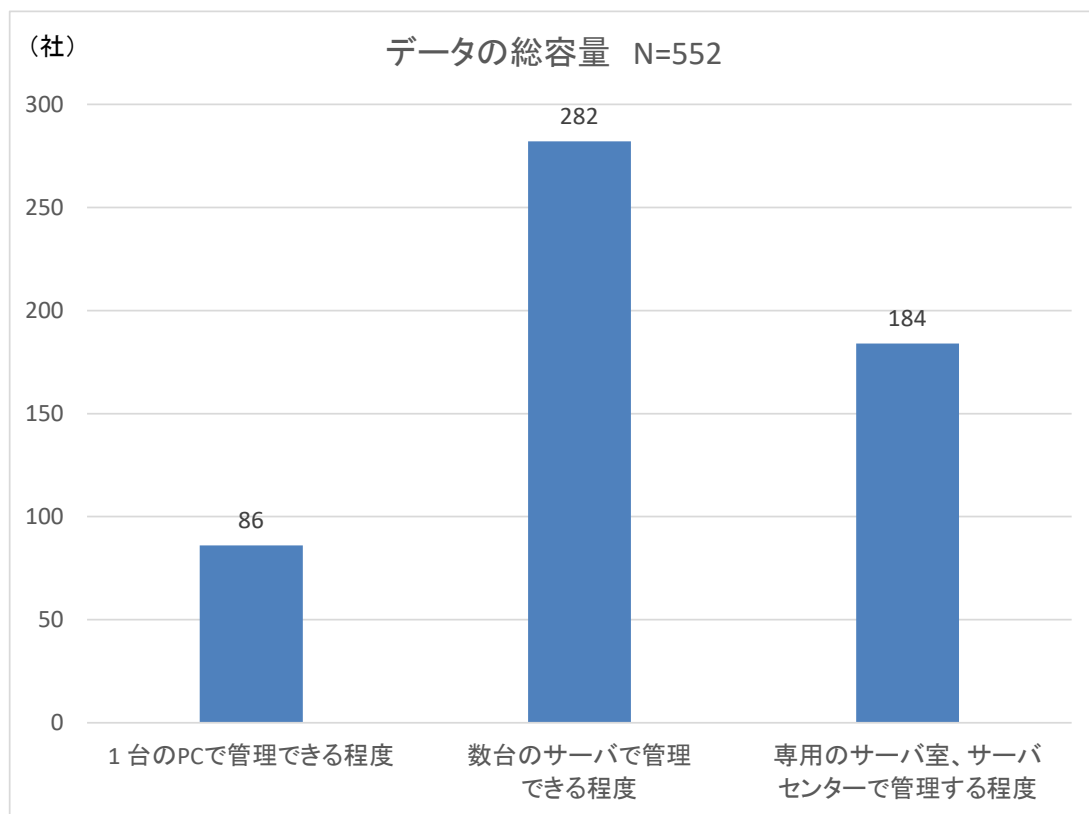
図表 16、図表 17 は、「ディープラーニング等の高度なデータの処理・解析を行える体制が整備されている」かどうかについての 5 段階評価に関するグラフである。「強くそう思う」「そう思う」と回答した企業の割合は 2017 年では 8.5%、2020 年では 10.9%であり、やや増加している。

図表 18、図表 19 は、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。具体的成果とは売上やコストダウンといったフィナンシャルな成果であり、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表している。具体的成果が得られている企業の割合は 2017 年と比較して 2020 年では減少している一方、「まだ成果は得られていない」と回答した企業の割合は 2017 年と比較して 2020 年では増加している。2020 年の調査では 2017 年の調査と調査対象企業が異なるため単純に比較できないが、3 年間でデータ利活用による成果が得られた企業が大きく増加したわけではないと推察される。

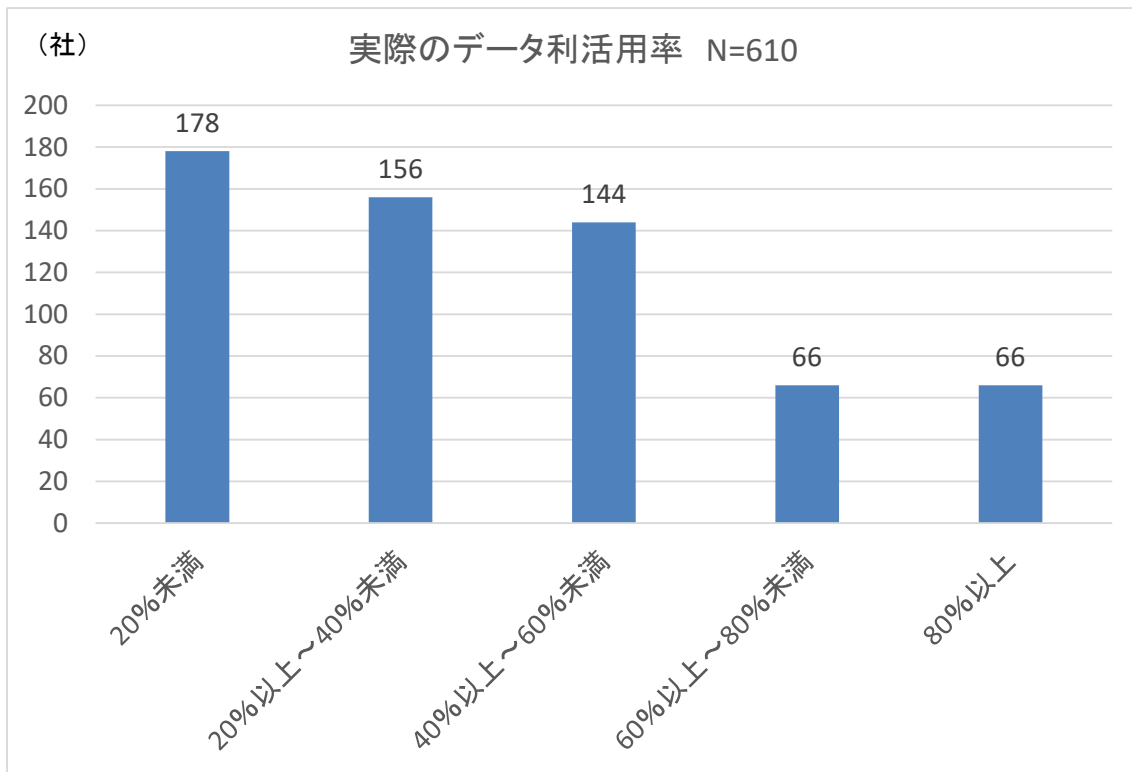
図表 20、図表 21 は、データ利活用の今後の方針についてのグラフである。2017 年と比較して 2020 年では「拡大する見通し」と回答した企業が増加しており、データ利活用の裾野が広がったことがうかがえる。



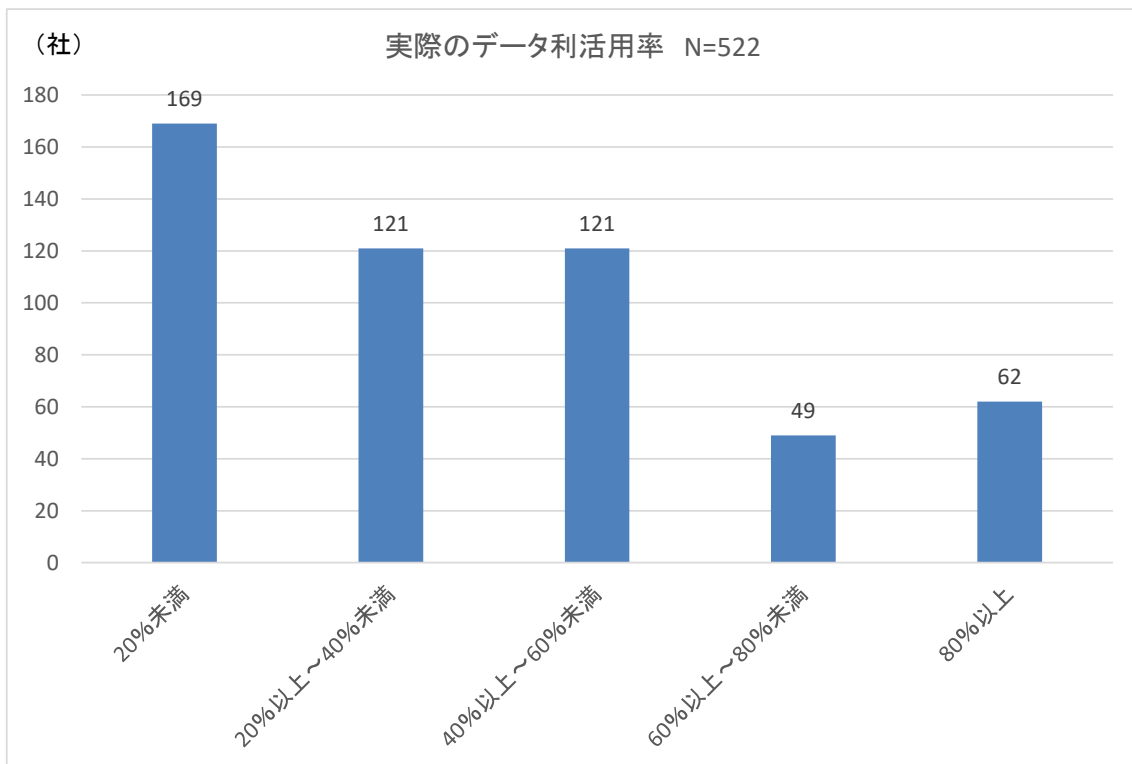
図表 7 データの総容量 (2020 年)



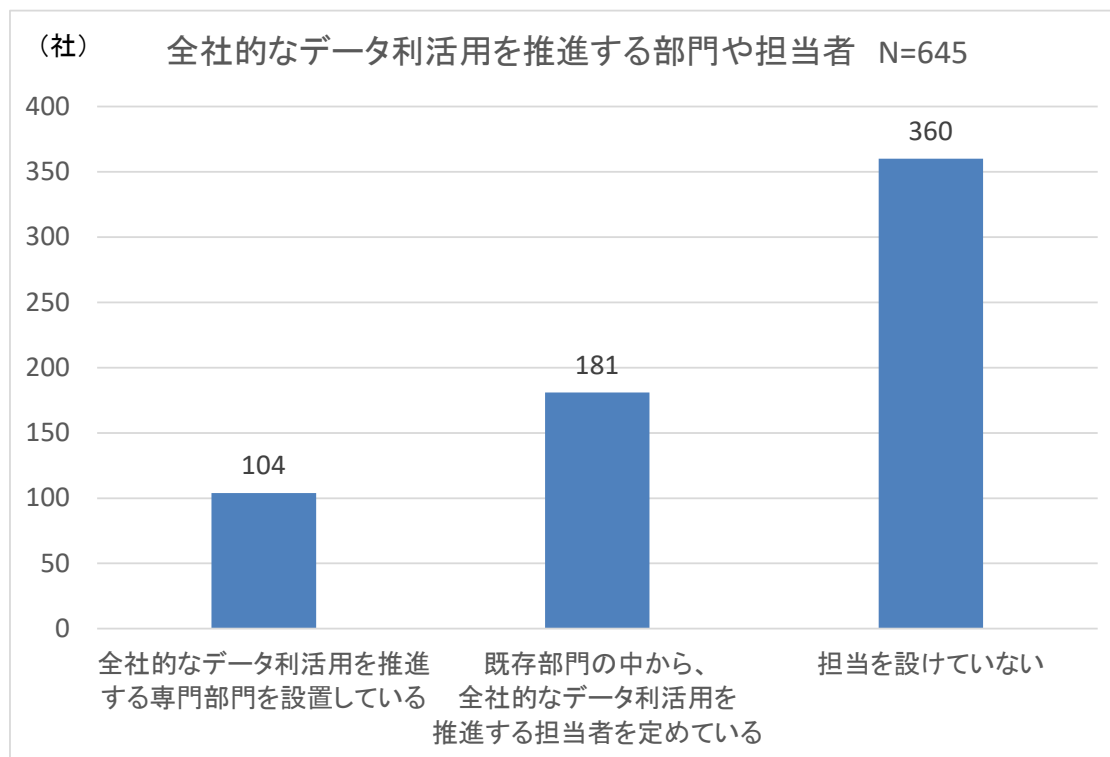
図表 8 データの総容量 (2017 年)



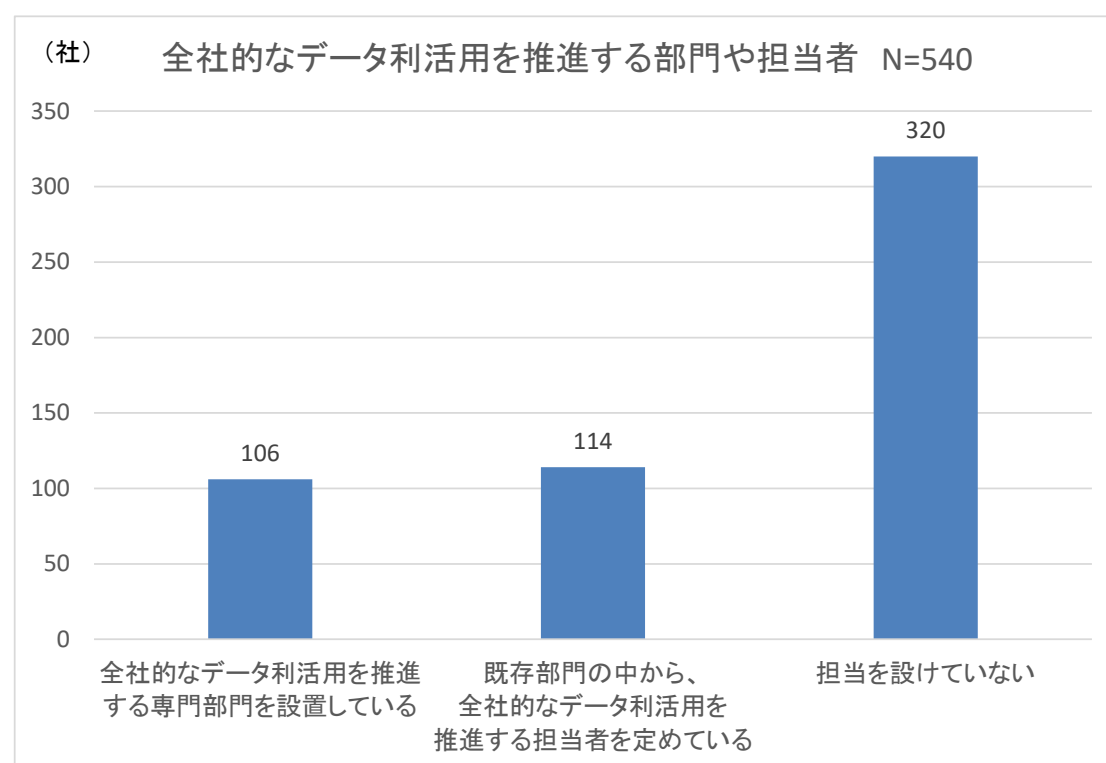
図表 9 実際のデータ利活用率（2020 年）



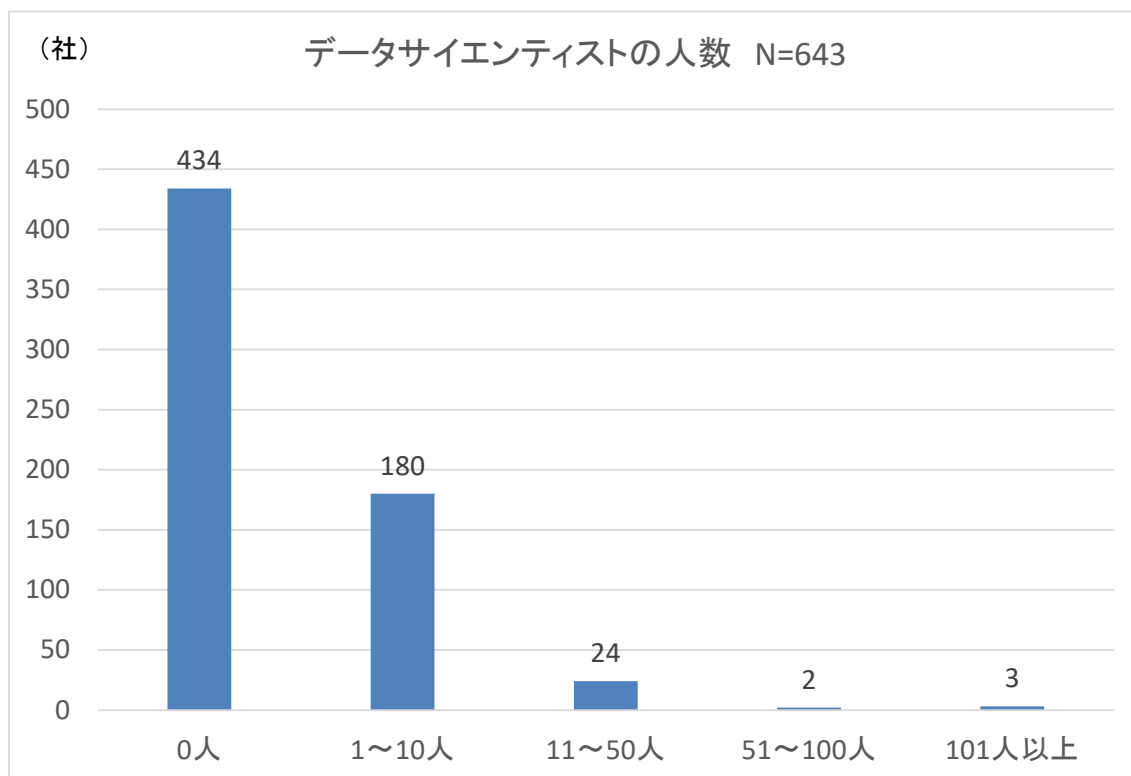
図表 10 実際のデータ利活用率（2017 年）



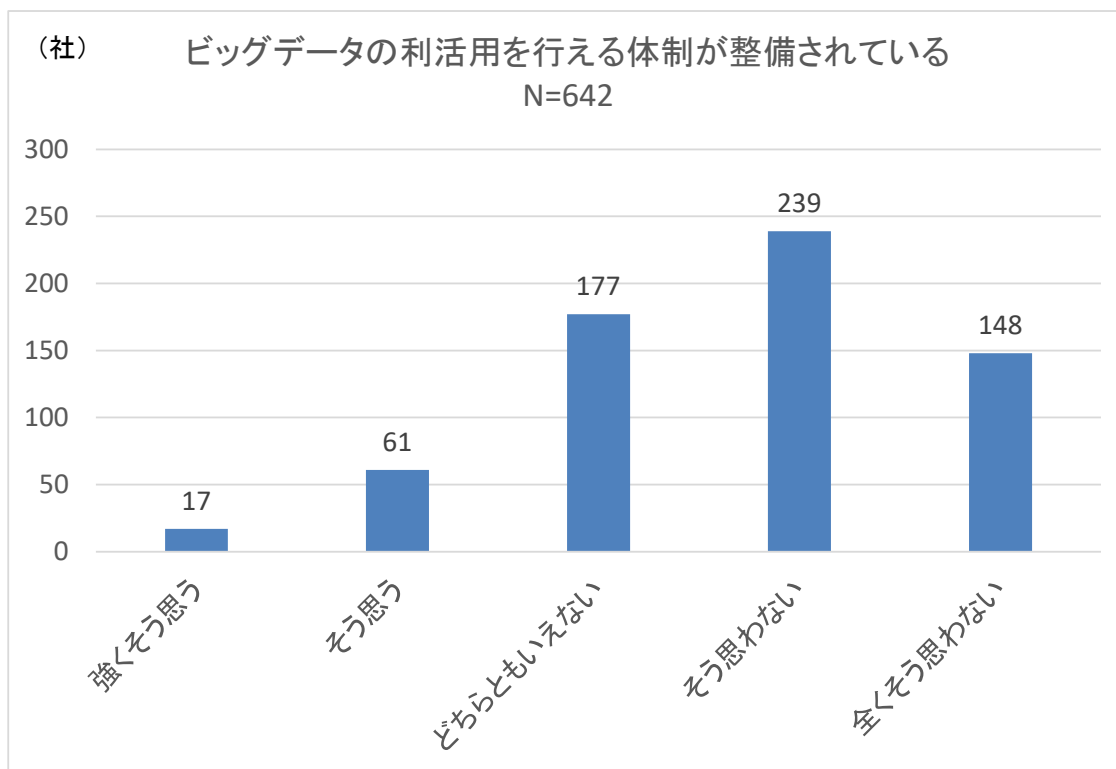
図表 11 全社的なデータ利活用を推進する部門や担当者（2020 年）



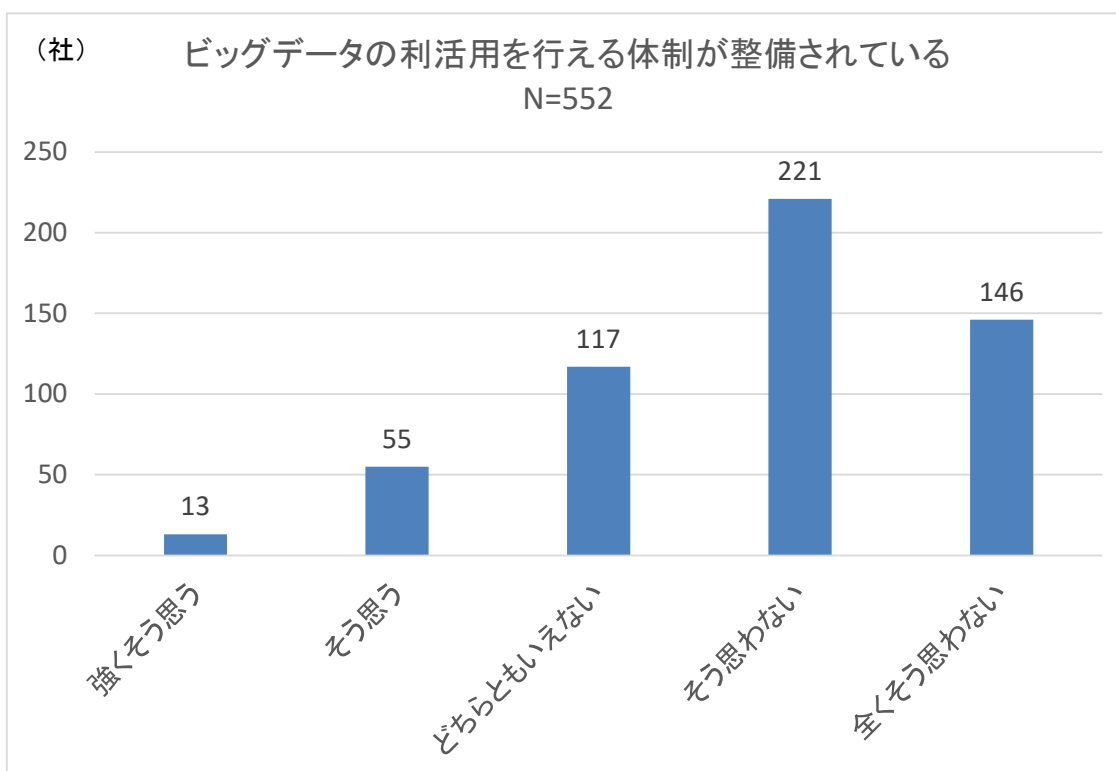
図表 12 全社的なデータ利活用を推進する部門や担当者（2017 年）



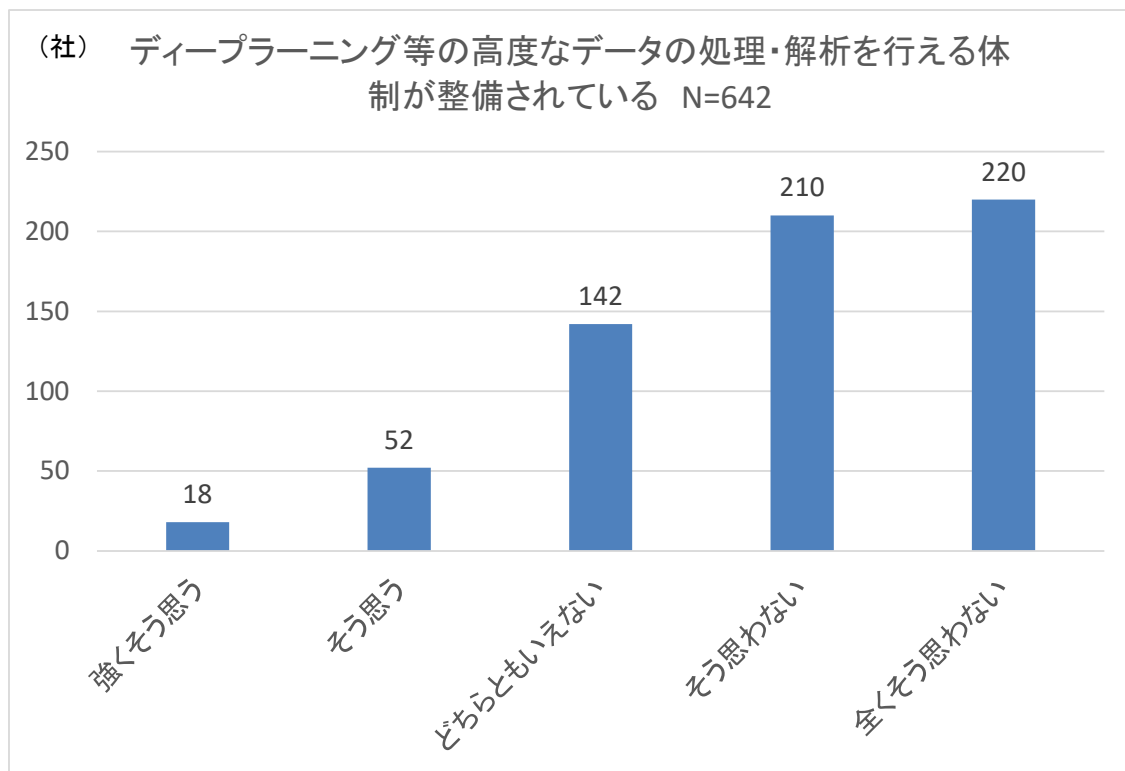
図表 13 データサイエンティストの人数 (2020 年)



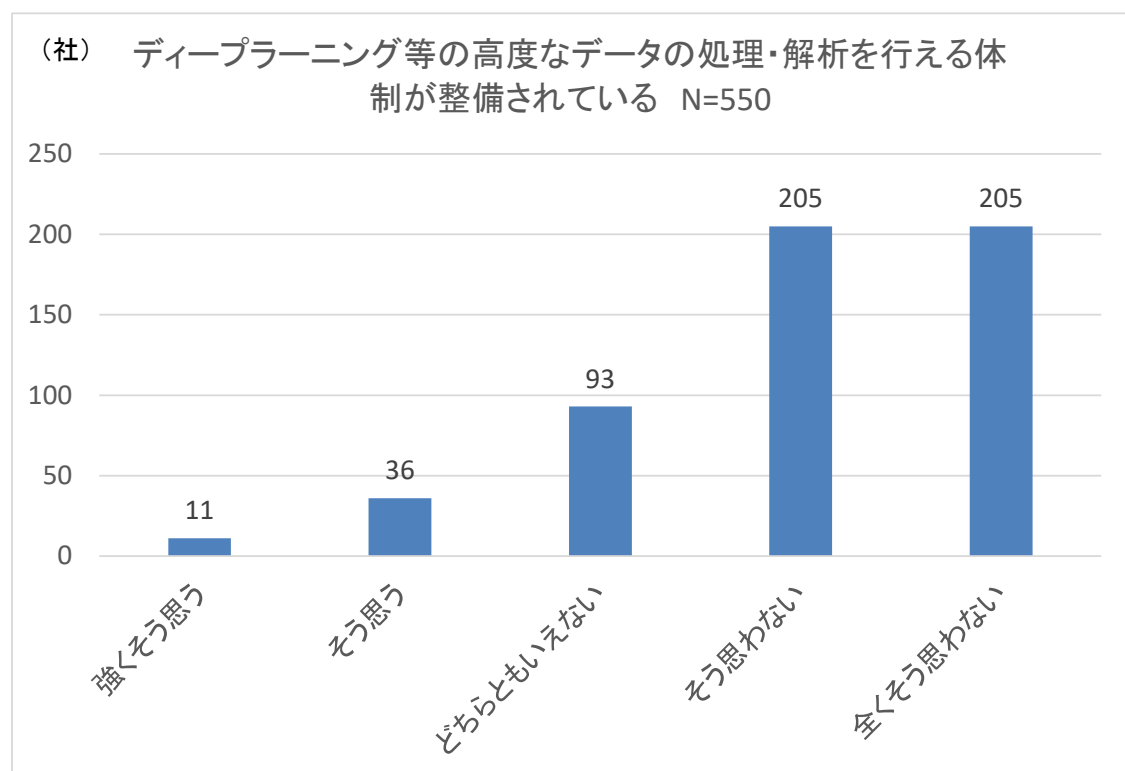
図表 14 ビッグデータの体制整備 (2020 年)



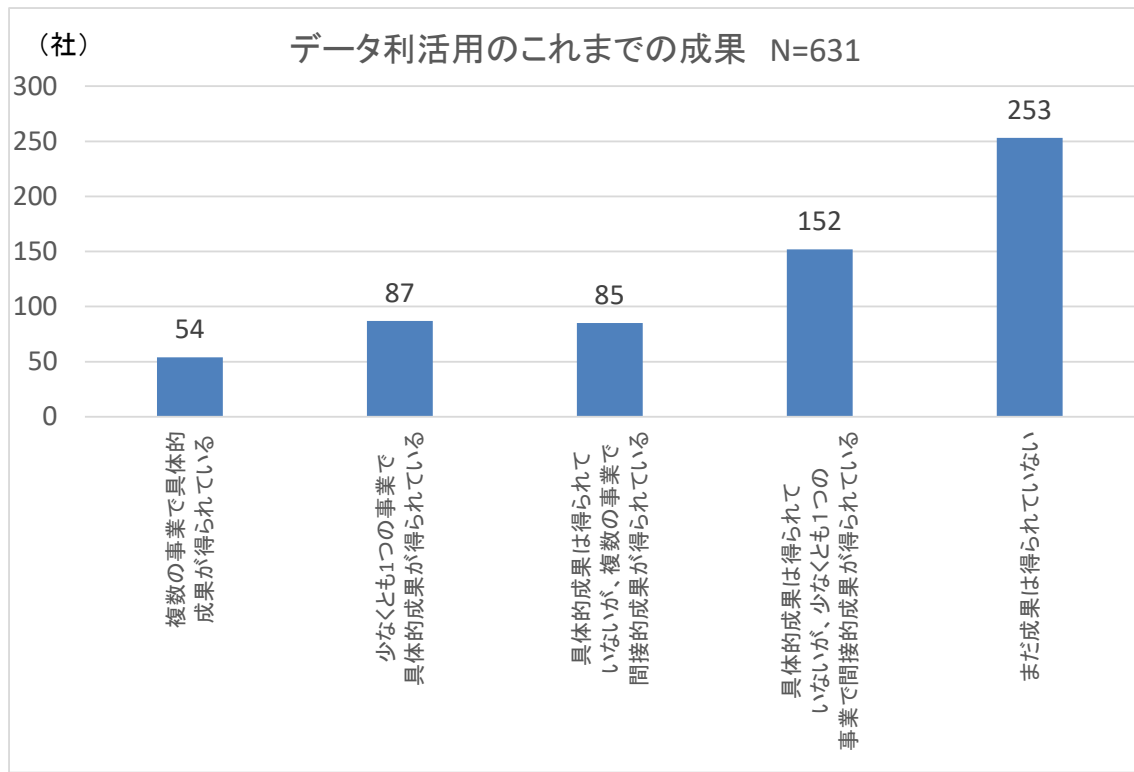
図表 15 ビッグデータの体制整備 (2017 年)



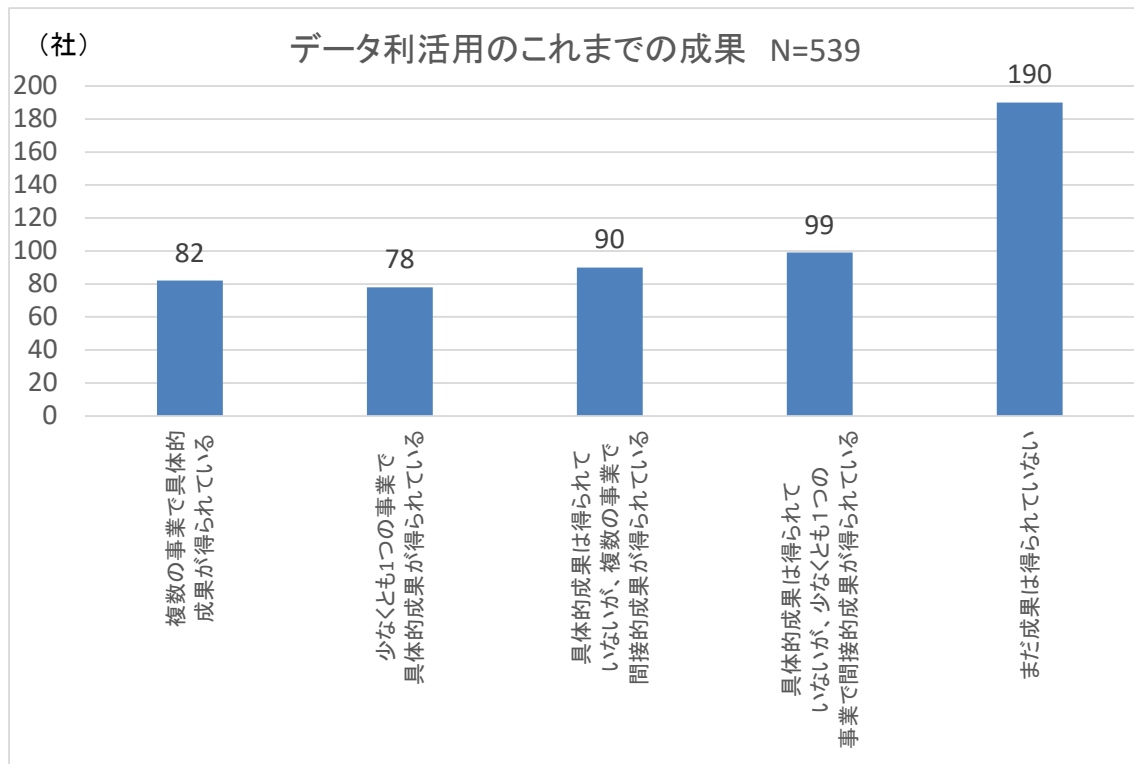
図表 16 高度なデータの処理・解析の体制整備 (2020 年)



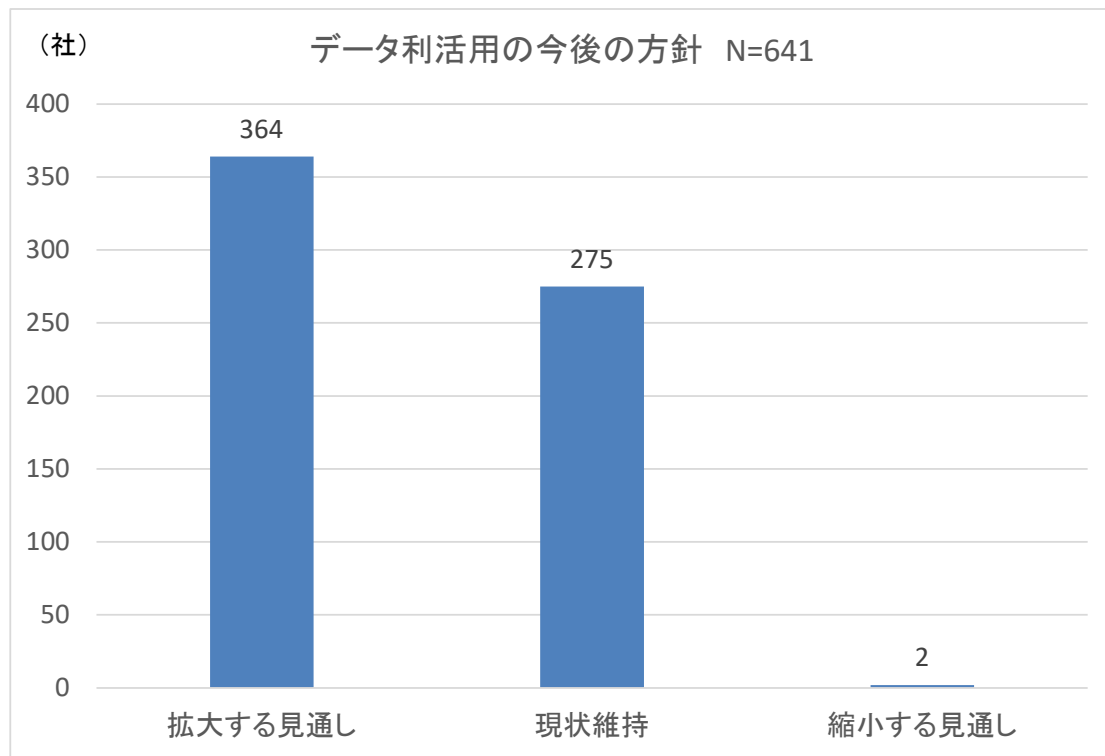
図表 17 高度なデータの処理・解析の体制整備 (2017 年)



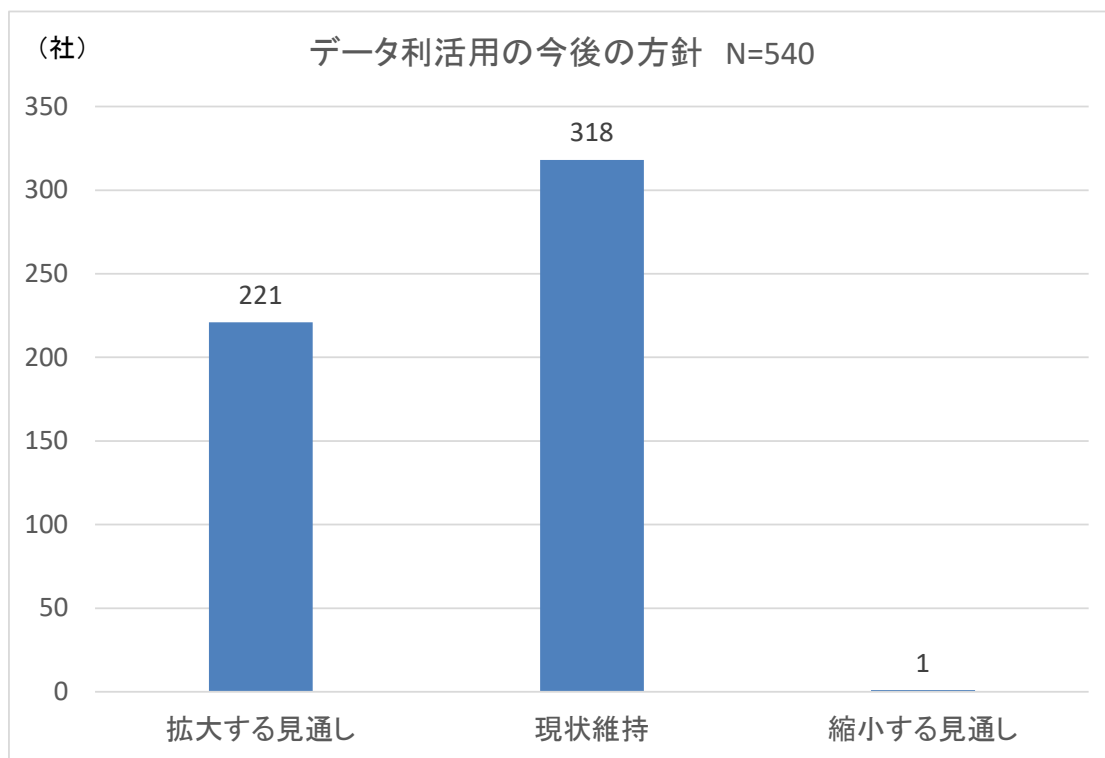
図表 18 データ利活用のこれまでの成果 (2020 年)



図表 19 データ利活用のこれまでの成果 (2017 年)



図表 20 データ利活用の今後の方針 (2020 年)

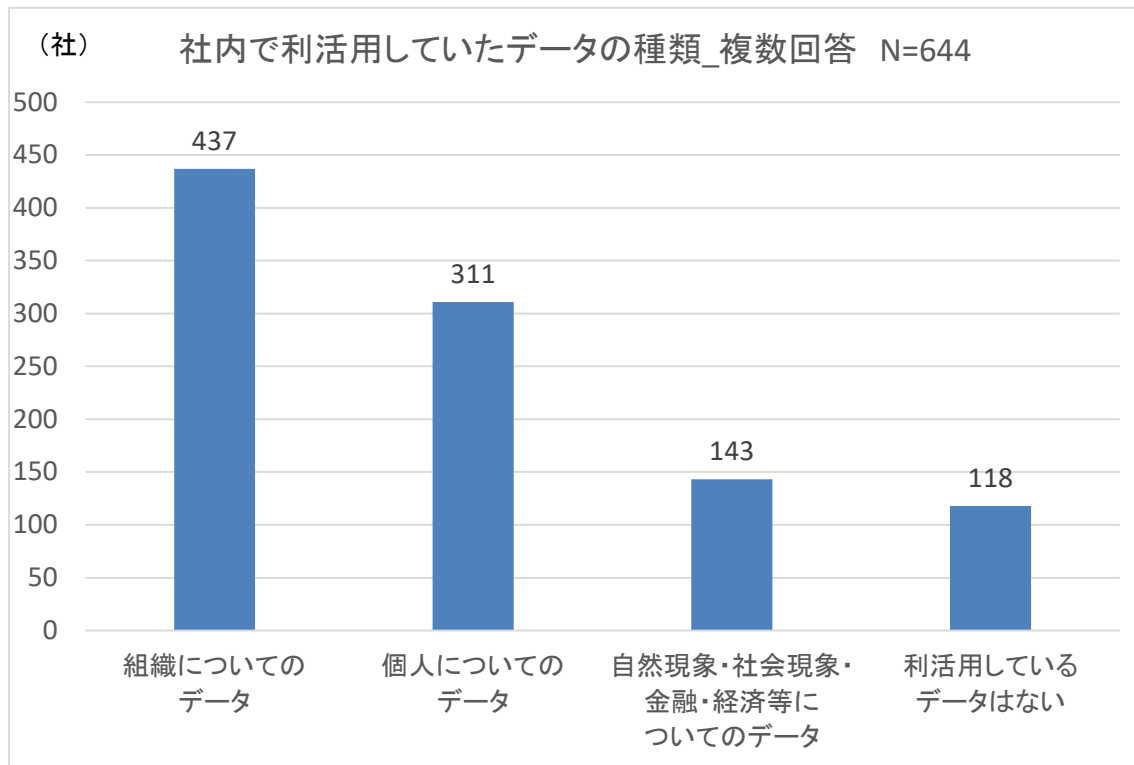


図表 21 データ利活用の今後の方針 (2017 年)

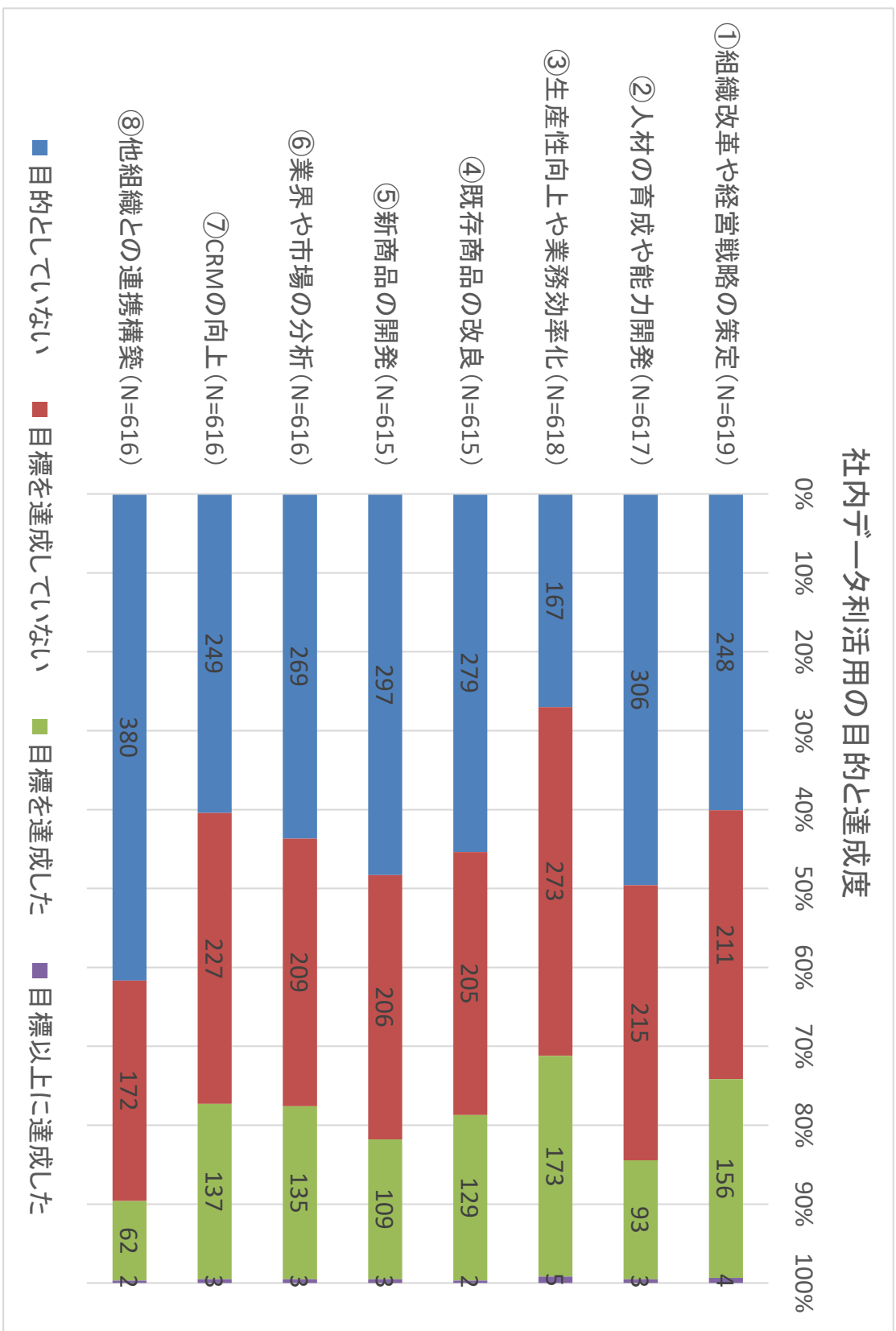
5. 2. 3. 自社のデータを自社内で利活用する場合について（問3）

図表 22 は、自社で保持しているデータのうち、直近 1 年間で社内で利活用していたデータの種類に関するグラフ（複数回答）である。「組織についてのデータ」が最も多く、次いで「個人についてのデータ」となっている。また、「利活用しているデータはない」と回答した企業は 18.3% である。

図表 23 は、直近 1 年間における社内でのデータ利活用の目的と（目的としていた場合の）目標達成度についてのグラフである。「③生産性向上や業務効率化」を目的としていた企業の割合が最も高く、また目標を達成した企業の割合も高くなっている。



図表 22 社内で利活用していたデータの種類（複数回答、2020 年）



図表 23 社内データ利活用の目的と達成度（2020 年）

5. 2. 4. データ取得プロジェクト（問4）及びデータ提供プロジェクト（問5）について

図表 24、図表 25 は、それぞれこれまでのデータ取得プロジェクト、データ提供プロジェクトの有無についてのグラフである。データ取得プロジェクトが「ある」と回答した企業は 15.1%、データ提供プロジェクトが「ある」と回答した企業は 11.8%であり、いずれも 2 割未満とそれほど多くはない。

図表 26、図表 27 は、直近 1 年間の各プロジェクトにおいて提携契約を結んだ相手組織についてのグラフ（複数回答、それぞれ直近 1 年間で提携契約を結んだ各プロジェクトが 1 件以上あった企業のみ回答）である。日本の相手組織としては、データ取得プロジェクト、データ提供プロジェクトのいずれも「大学、研究機関、公的機関等、その他の組織」が最も多い。また、いずれのプロジェクトでも海外の相手組織は少ないが、データ取得プロジェクトではアジア（中国（香港含む）、その他アジア）の「顧客またはクライアント」がやや多くなっている。

図表 28、図表 29 は、直近 1 年間の各プロジェクトにおいて限定提供データとして扱ったデータについてのグラフ（それぞれ直近 1 年間で提携契約を結んだ各プロジェクトが 1 件以上あった企業のみ回答）である。「全くない」と回答した企業の割合は、データ取得プロジェクトでは 20.3%、データ提供プロジェクトでは 9.3%であり、限定提供データとして扱っていた企業は比較的多い。

図表 30、図表 31 は、直近 1 年間の各プロジェクトの目的と（目的としていた場合の）目標達成度についてのグラフ（それぞれ直近 1 年間で提携契約を結んだ各プロジェクトが 1 件以上あった企業のみ回答）である。いずれのプロジェクトでも「⑤新商品の開発」を目的としていた企業の割合が最も高く、また目標を達成した企業の割合も高くなっている。

図表 32、図表 33 は、それぞれ各プロジェクトにおいてプロジェクトを行うかどうかを判断する上でリスクとして特に懸念することについてのグラフ（複数回答、それぞれこれまでに各プロジェクトがあった企業のみ回答）である。データ取得プロジェクトでは「受領するデータが個人情報を含んでいる」が最も多く、次いで「受領するデータが重要な営業秘密を含んでいる」「受領するデータの希少価値が高い」となっている。データ提供プロジェクトでは「提供するデータが重要な営業秘密を含んでいる」が最も多く、次いで「提供するデータが個人情報を含んでいる」「提供するデータの希少価値が高い」となっている。いずれのプロジェクトにおいても、データそのものの内容に関して懸念する企業が多いことがうかがえる。

図表 34、図表 35 は、各プロジェクトにおける契約書のひな型についてのグラフ（それぞれこれまでに各プロジェクトがあった企業のみ回答）である。いずれのプロジェクトでも「契約書のひな型はない」と回答した企業が最も多いが、「すでに契約書のひな型があり、それを使いこなしている」と回答した企業も次いで多い。

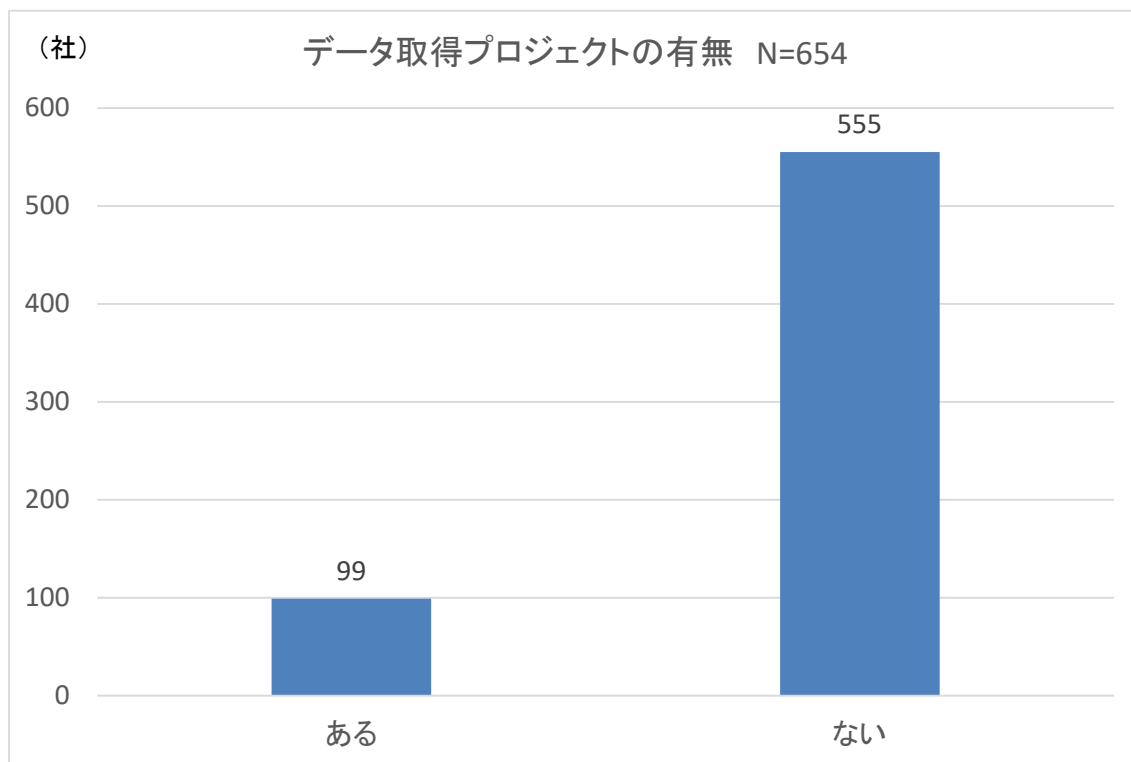
図表 36、図表 37 は、各プロジェクトを行う場合、データ利活用を行うことを主としない他組織との提携プロジェクトの契約と比較して、その契約書はかなりカスタマイズする必要があり、相当な法律業務が必要かどうかについてのグラフ（それぞれこれまでに各プロジェクトがあった企業のみ回答）である。いずれのプロジェクトでも「全くそう思わない」「そう思わない」と回答した企業の割合よりも「強くそう思う」「そう思う」と回答した企業の割合の方が高いが、その割合はデータ提供プロジェクトの方がやや高くなっている。

図表 38、図表 39 は、海外とデータのやり取りをして各プロジェクトを行う場合、国内でデータのやり取りをする場合と比較して、その契約書はかなりカスタマイズする必要があり、相当な法律業務が必要かどうかについてのグラフ（それぞれこれまでに各プロジェクトがあった企業のみ回答）である。いずれのプロジェクトでも国内でデータをやり取りする場合より海外とやり取り場合の方が契約書のカスタマイズや相当の法的業務が必要であることがうかがえる。

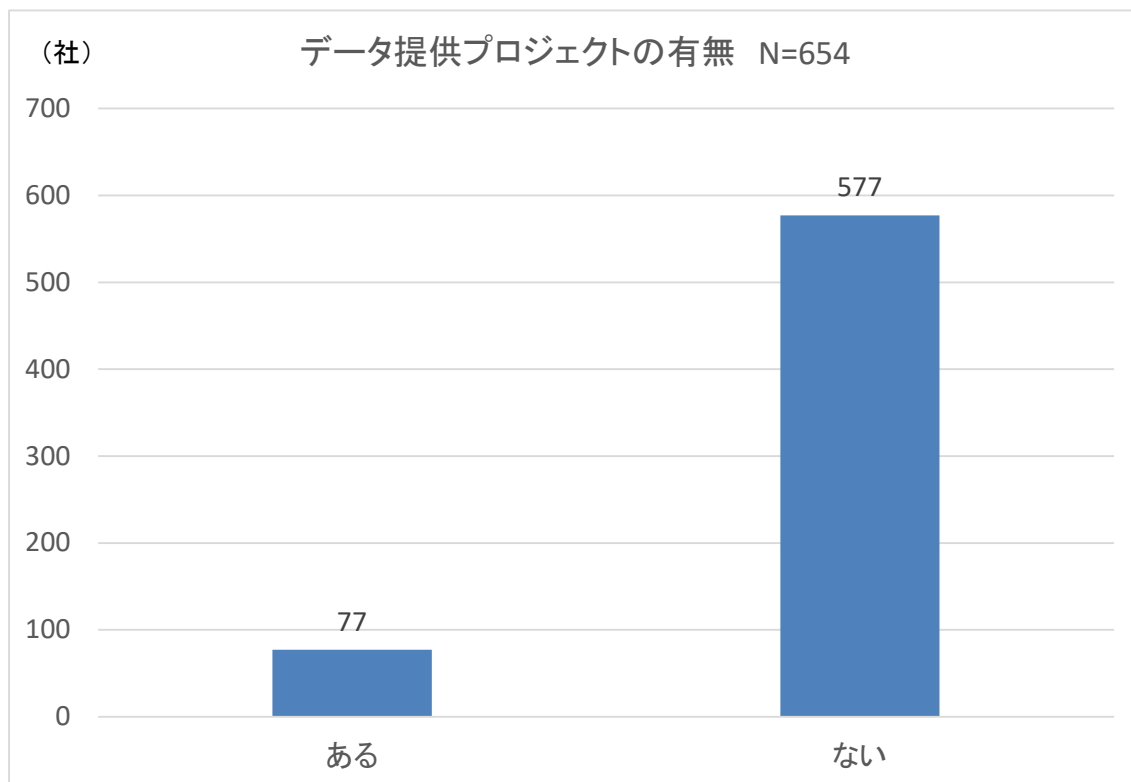
図表 40、図表 41 は、各プロジェクトを行う場合にデータの来歴管理やそれに関連する法的責任について検討する部門（担当者）についてのグラフ（複数回答、それぞれこれまでに各プロジェクトがあった企業のみ回答）である。いずれのプロジェクトでも「知財・法務部門（担当者）」が最も多く、次いで「各プロジェクトを担う事業部門（担当者）」「情報システム部門等のデータを処理する部門（担当者）」となっている。

図表 42、図表 43 は、各プロジェクトを行う場合に知財・法務部門（担当者）が関与する段階に

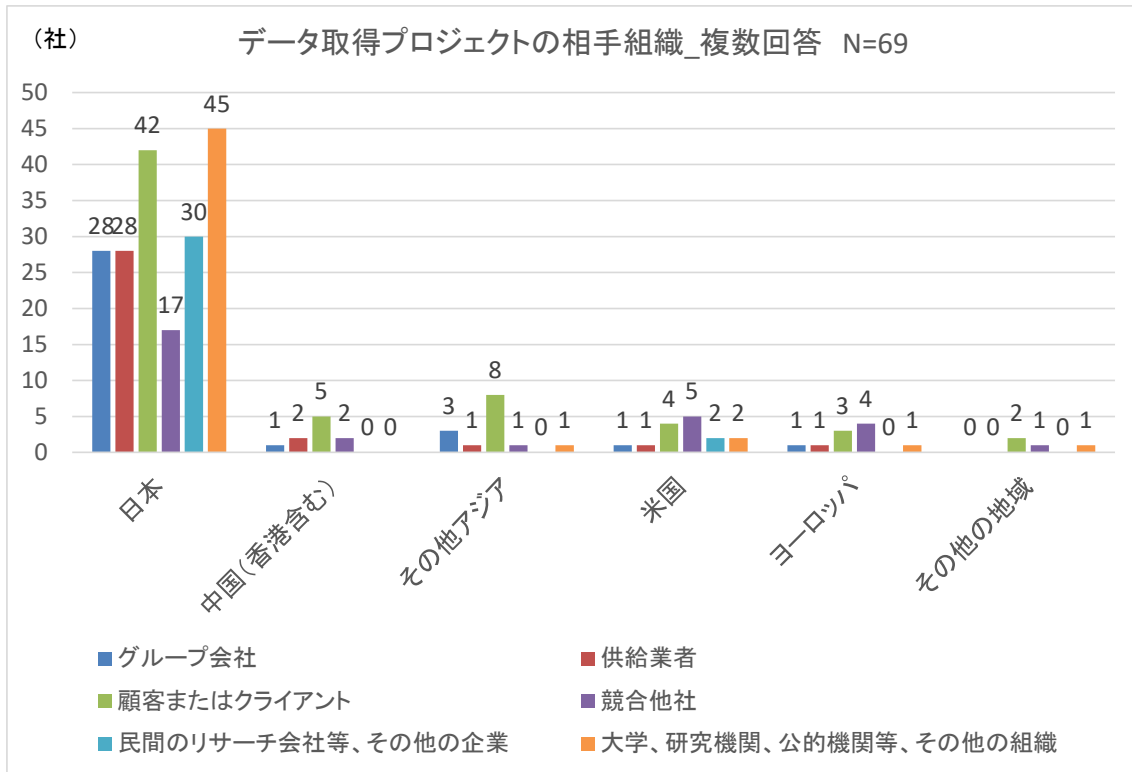
についてのグラフ（それぞれこれまでに各プロジェクトがあった企業のみ回答）である。いずれのプロジェクトでも「知的財産に関連する課題が生じたり契約交渉が必要になったりする段階から関与していることが多い」が最も多い。



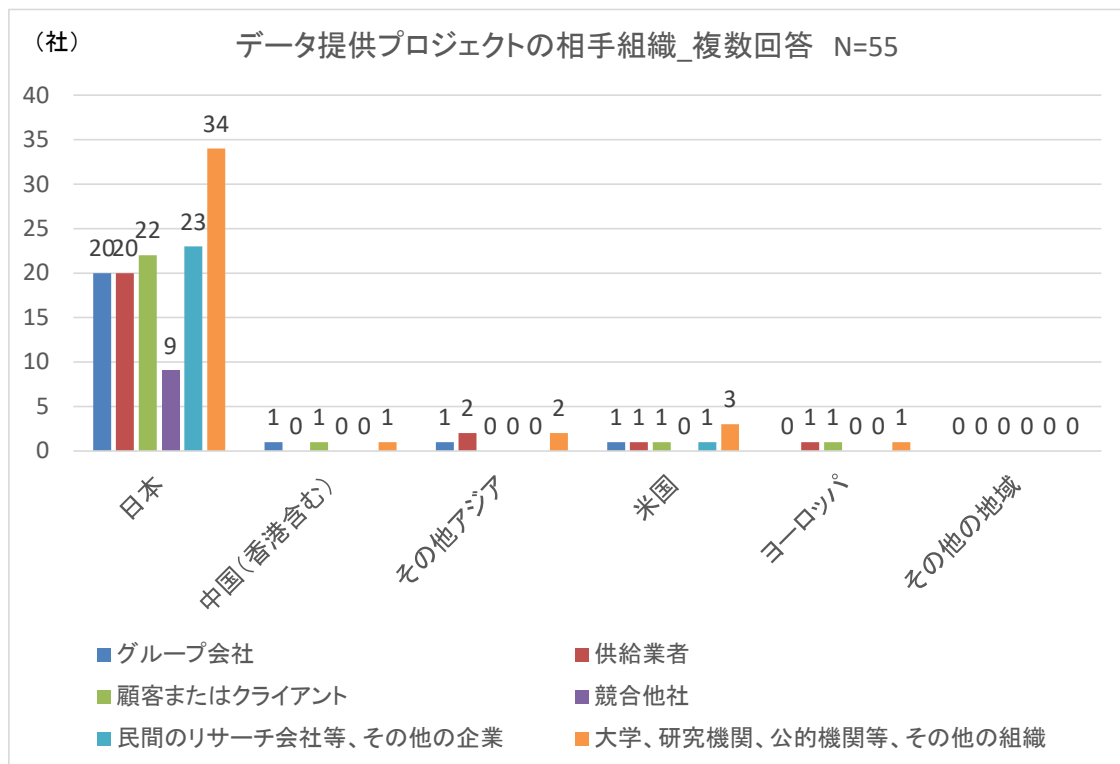
図表 24 データ取得プロジェクトの有無（2020 年）



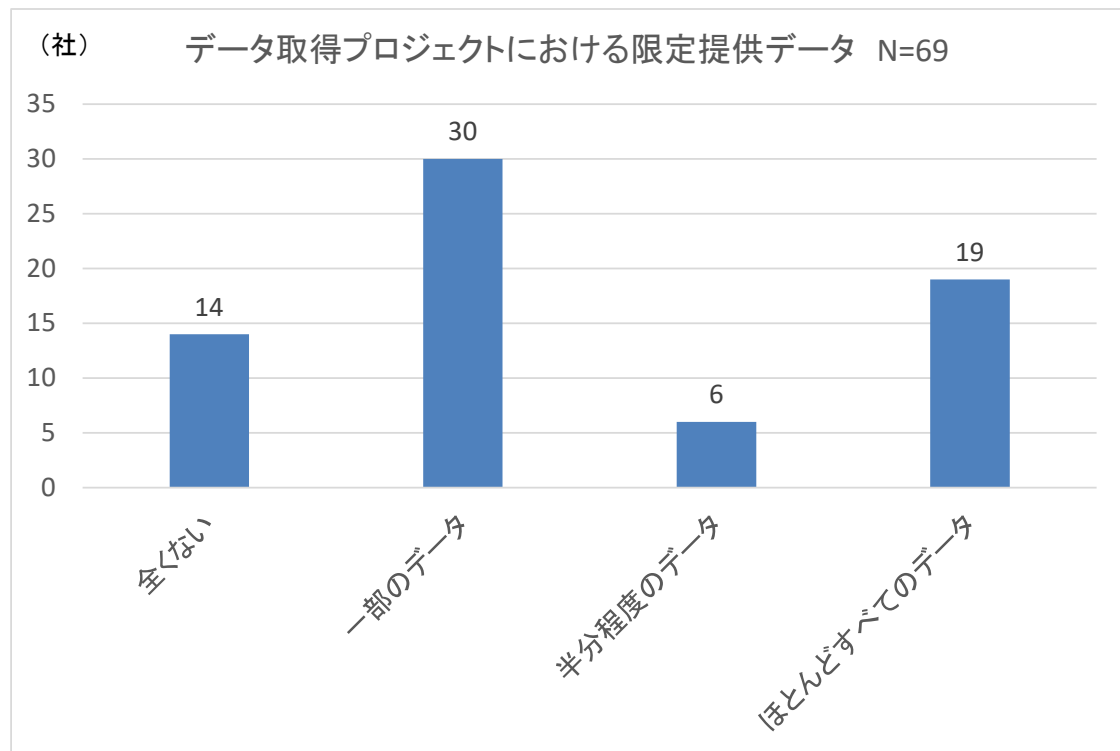
図表 25 データ提供プロジェクトの有無（2020 年）



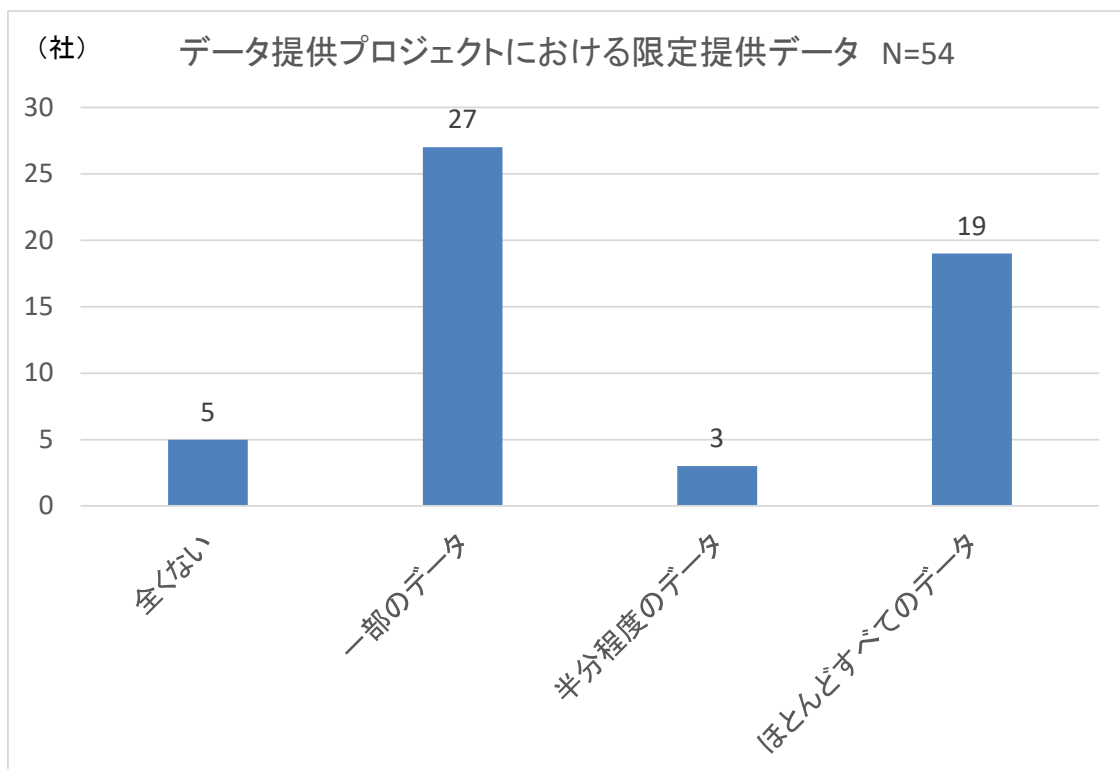
図表 26 データ取得プロジェクトの相手組織（複数回答、2020 年）



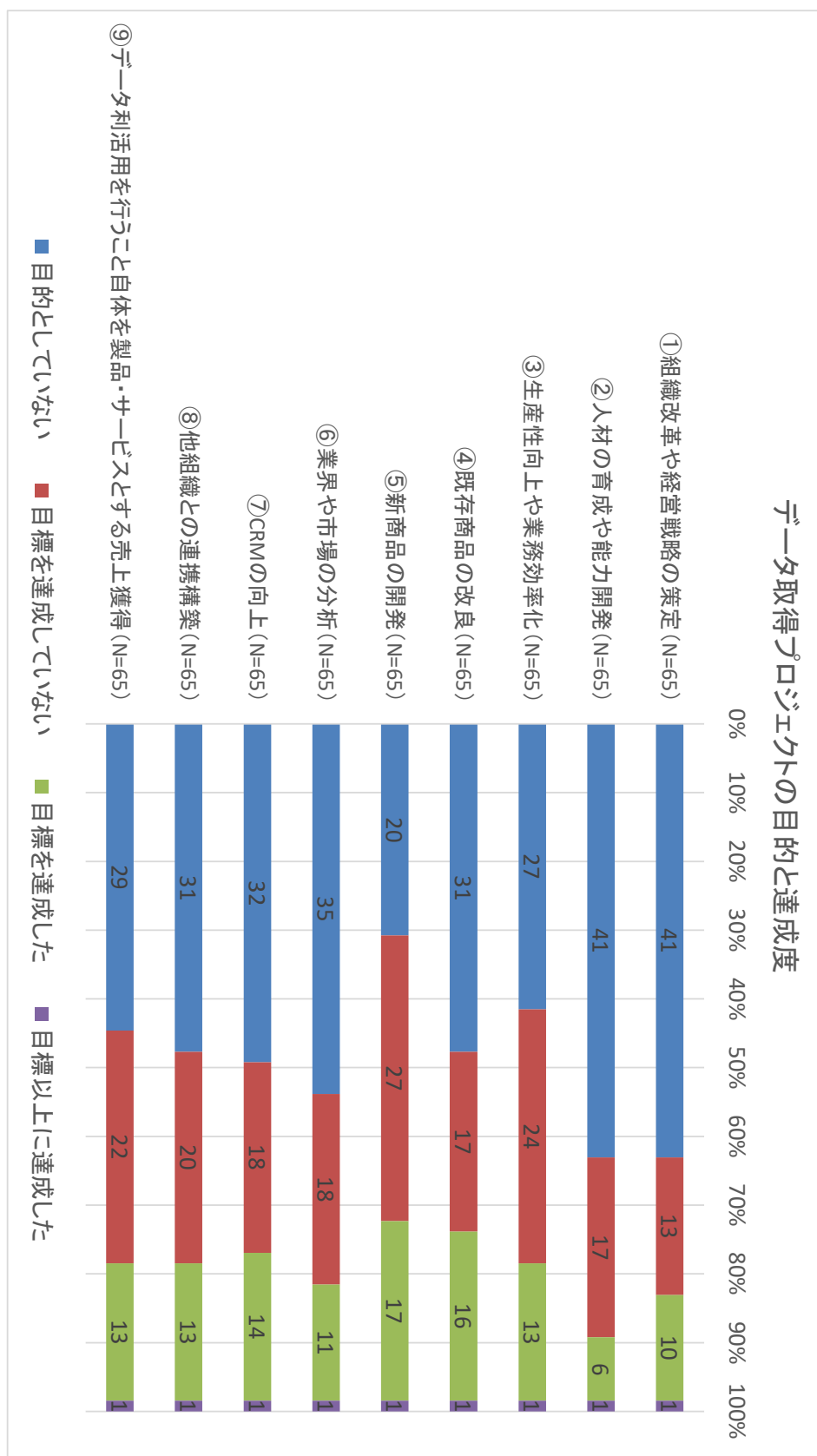
図表 27 データ提供プロジェクトの相手組織（複数回答、2020 年）



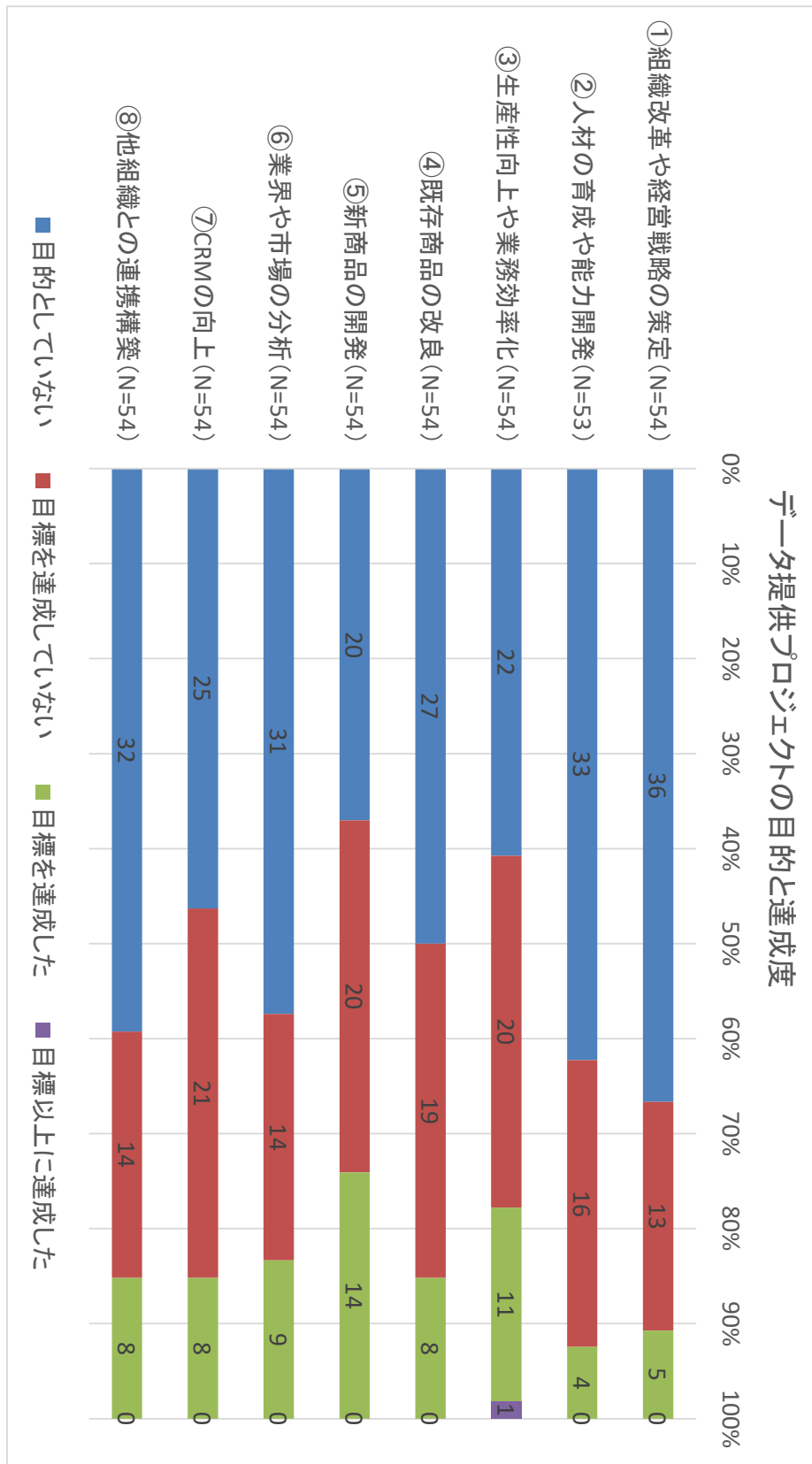
図表 28 データ取得プロジェクトにおける限定提供データ (2020 年)



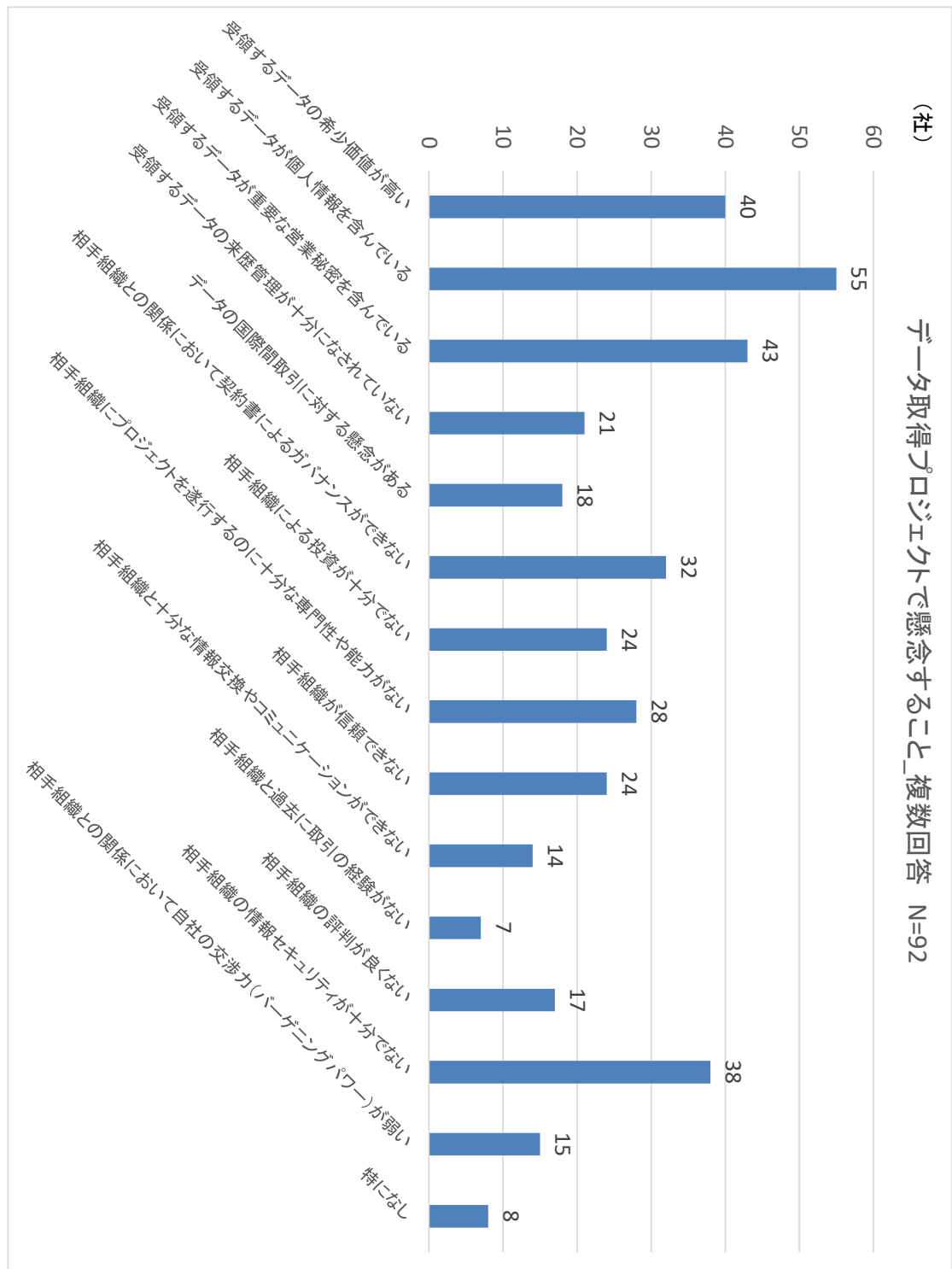
図表 29 データ提供プロジェクトにおける限定提供データ (2020 年)



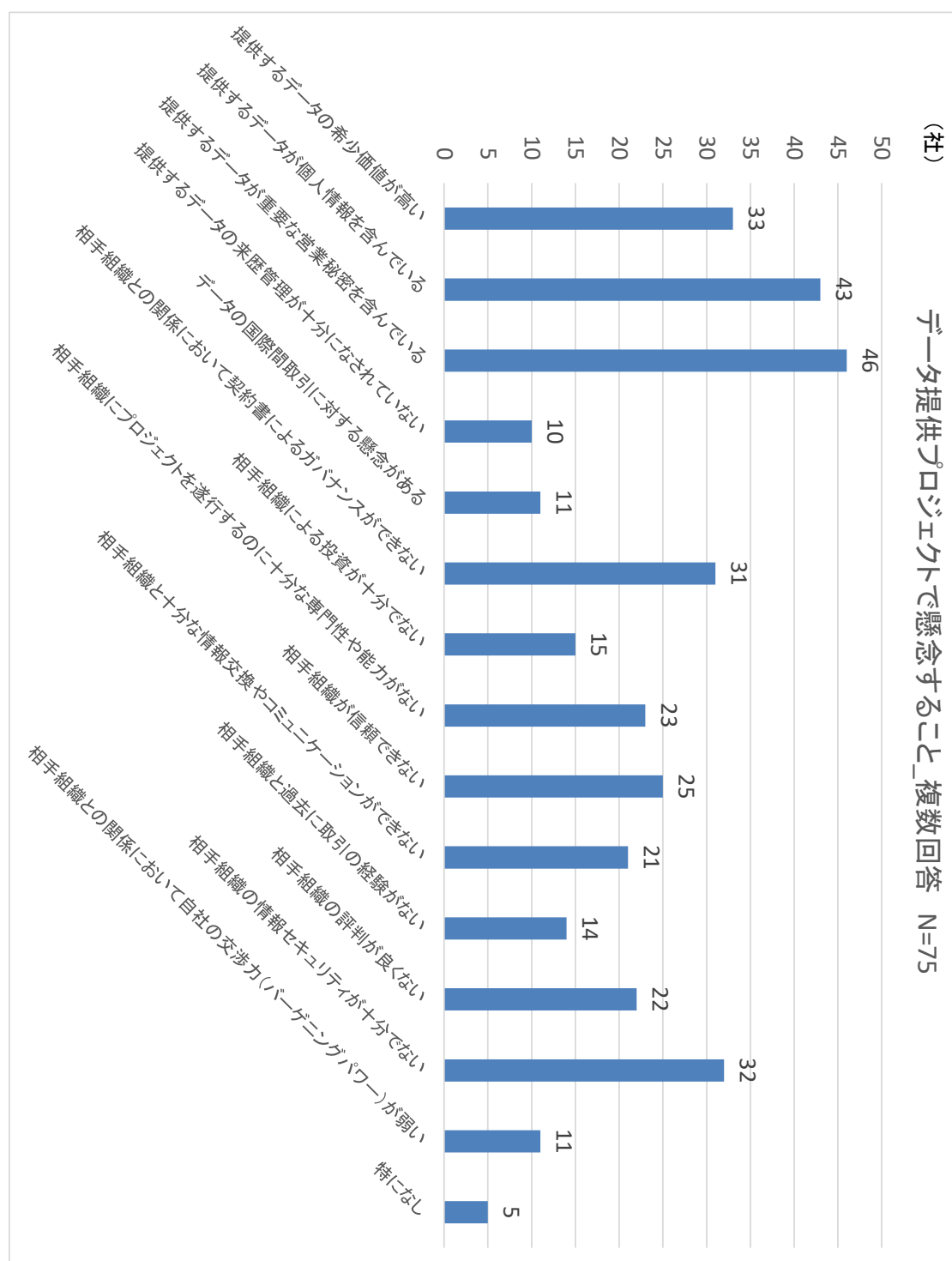
図表 30 データ取得プロジェクトの目的と達成度 (2020 年)



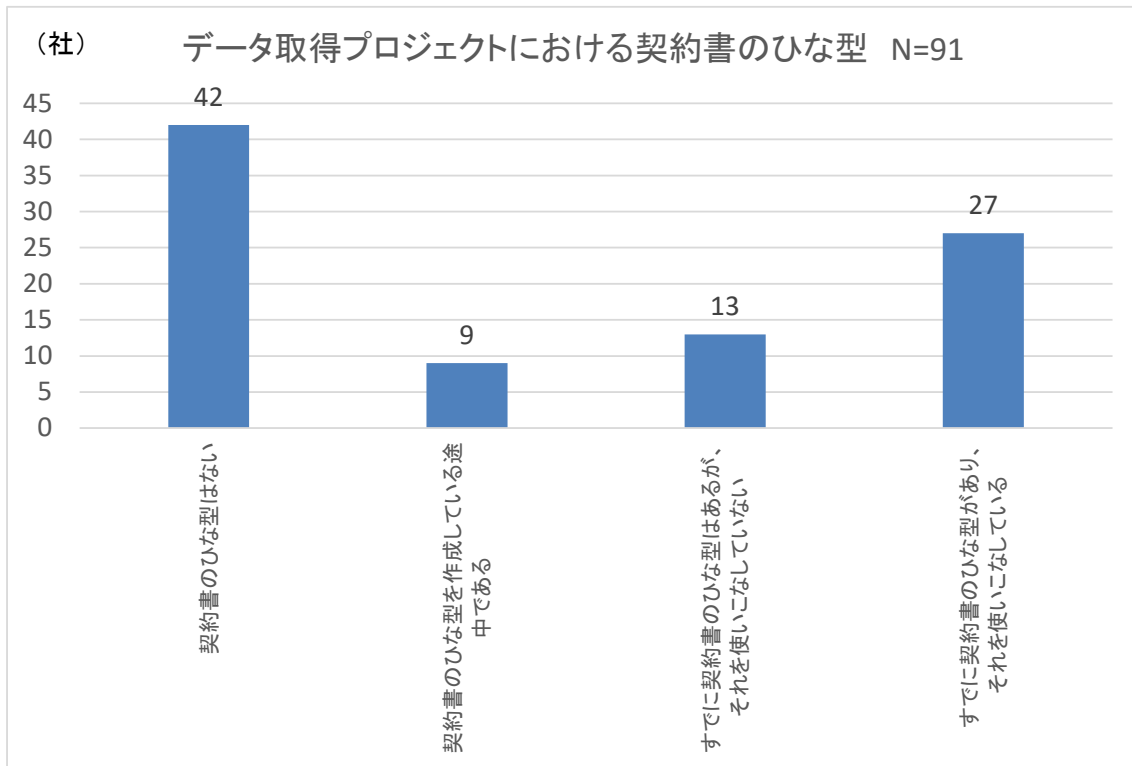
図表 31 データ提供プロジェクトの目的と達成度（2020 年）



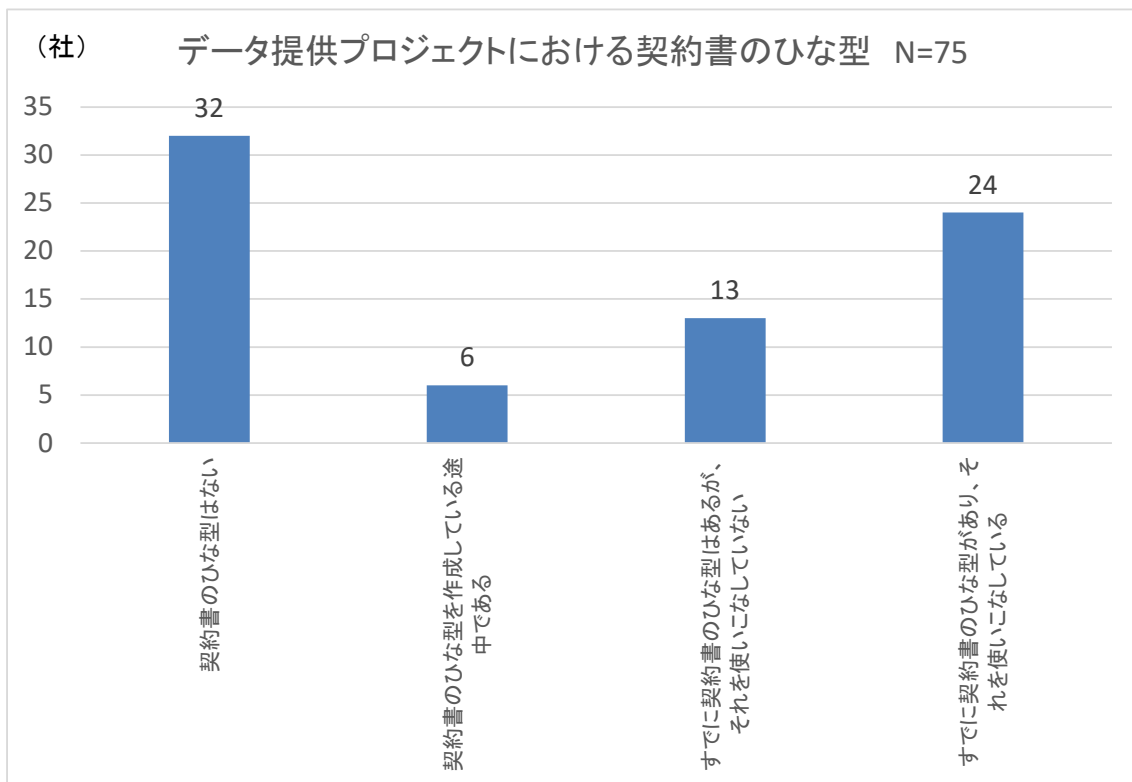
図表 32 データ取得プロジェクトで懸念すること（複数回答、2020 年）



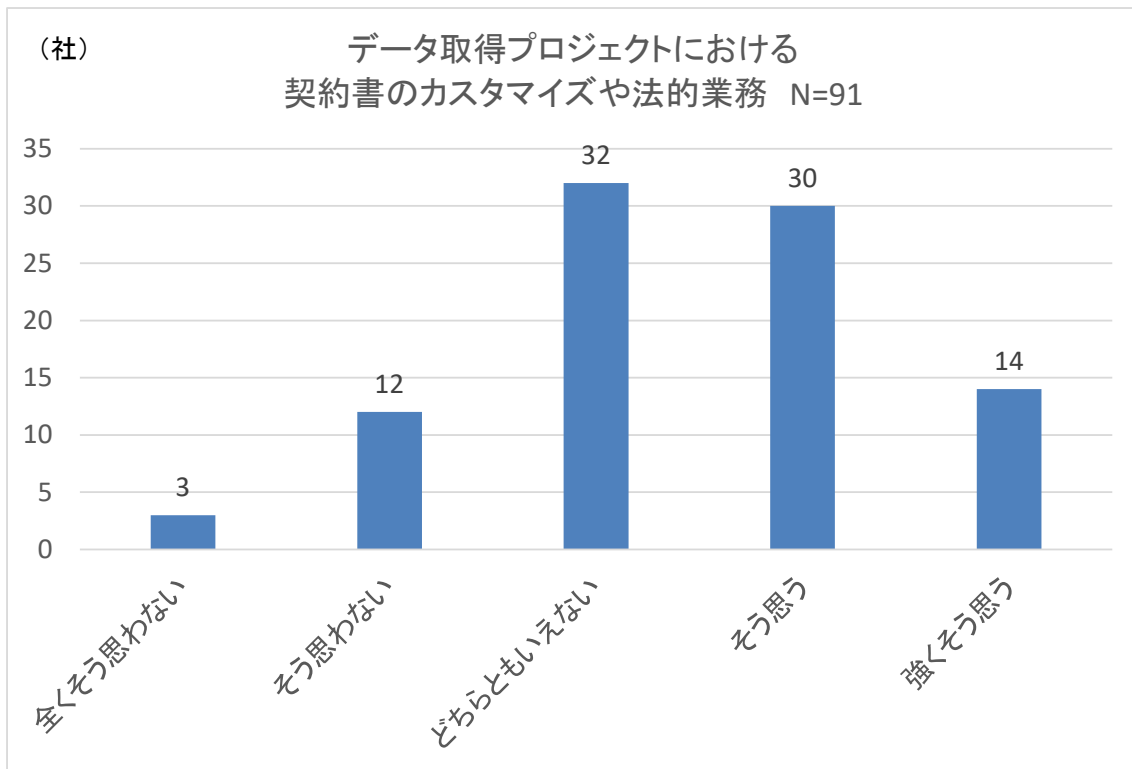
図表 33 データ提供プロジェクトで懸念すること（複数回答、2020 年）



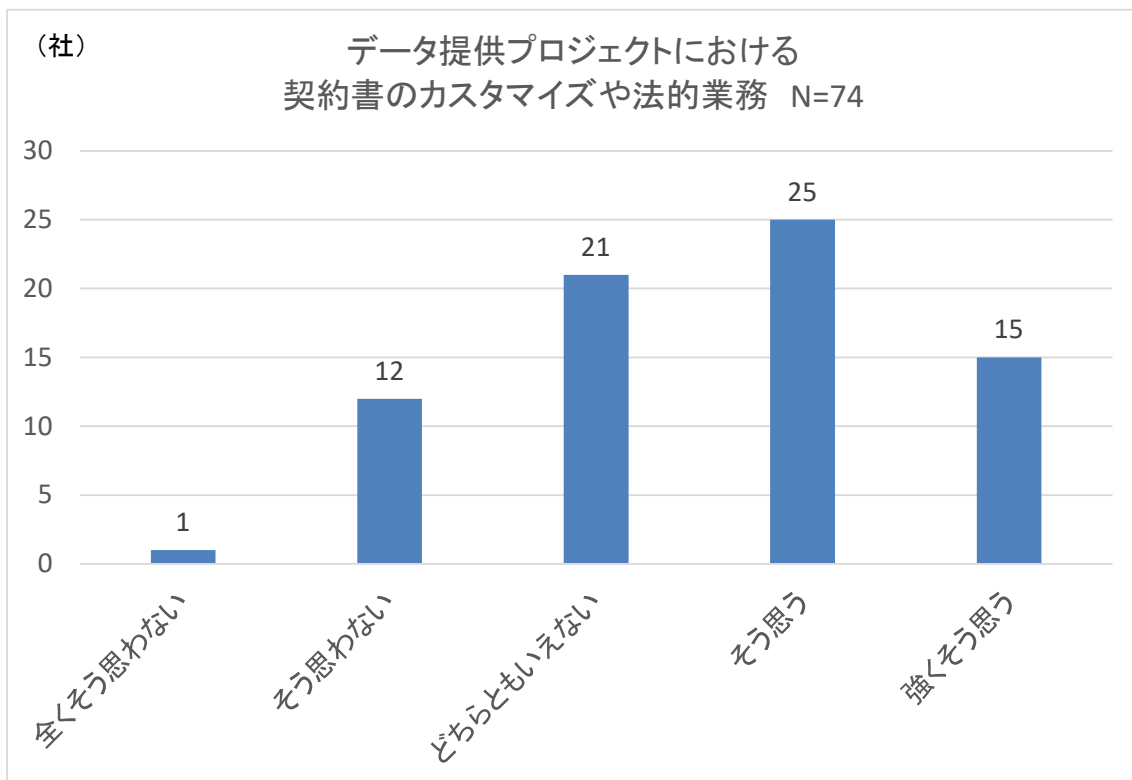
図表 34 データ取得プロジェクトにおける契約書のひな型 (2020 年)



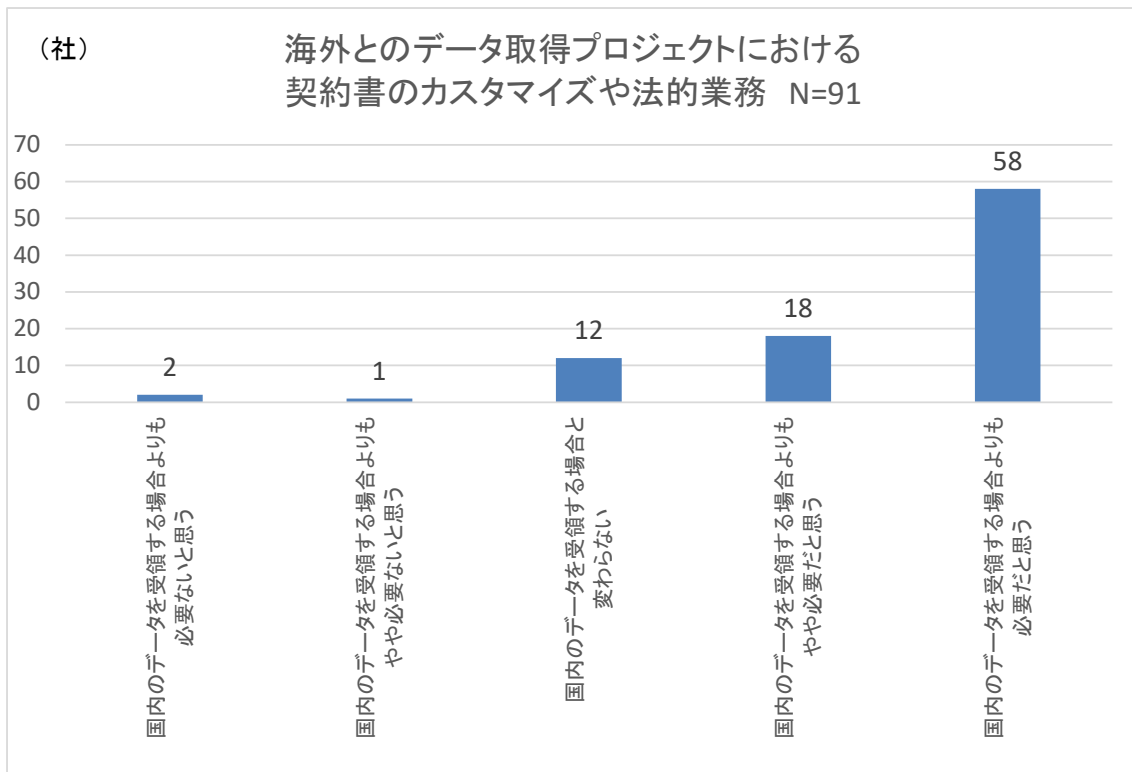
図表 35 データ提供プロジェクトにおける契約書のひな型 (2020 年)



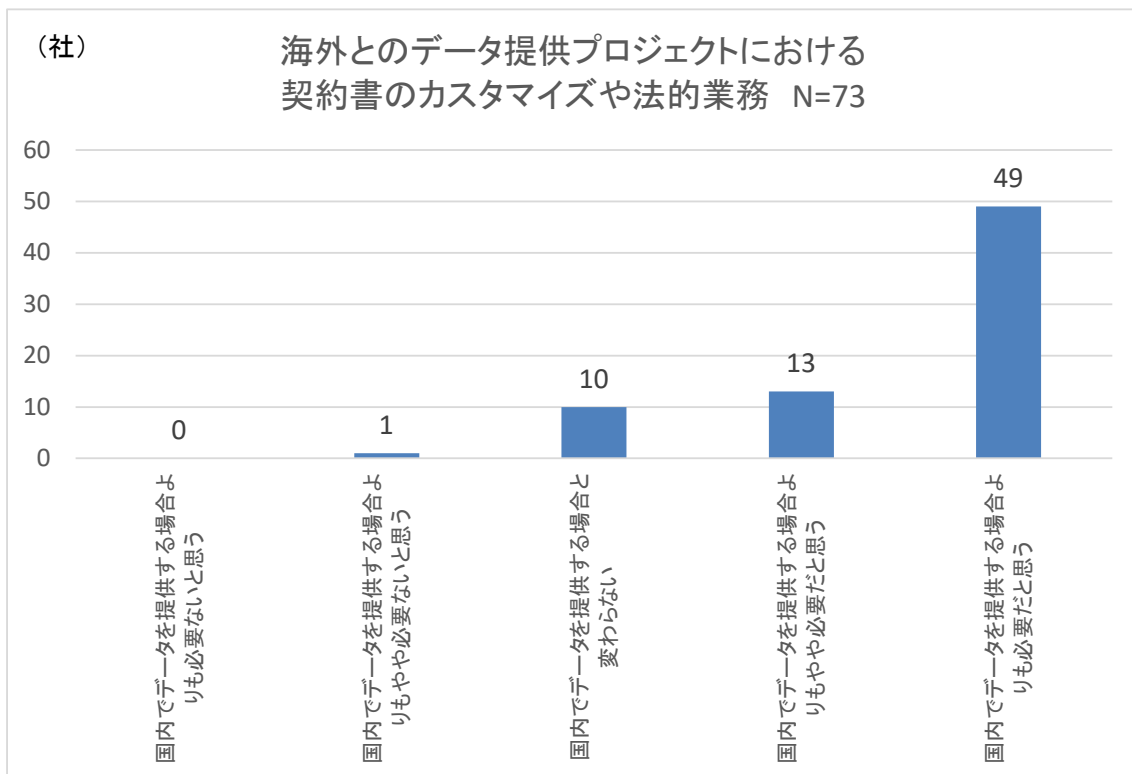
図表 36 データ取得プロジェクトにおける契約書や法的業務 (2020 年)



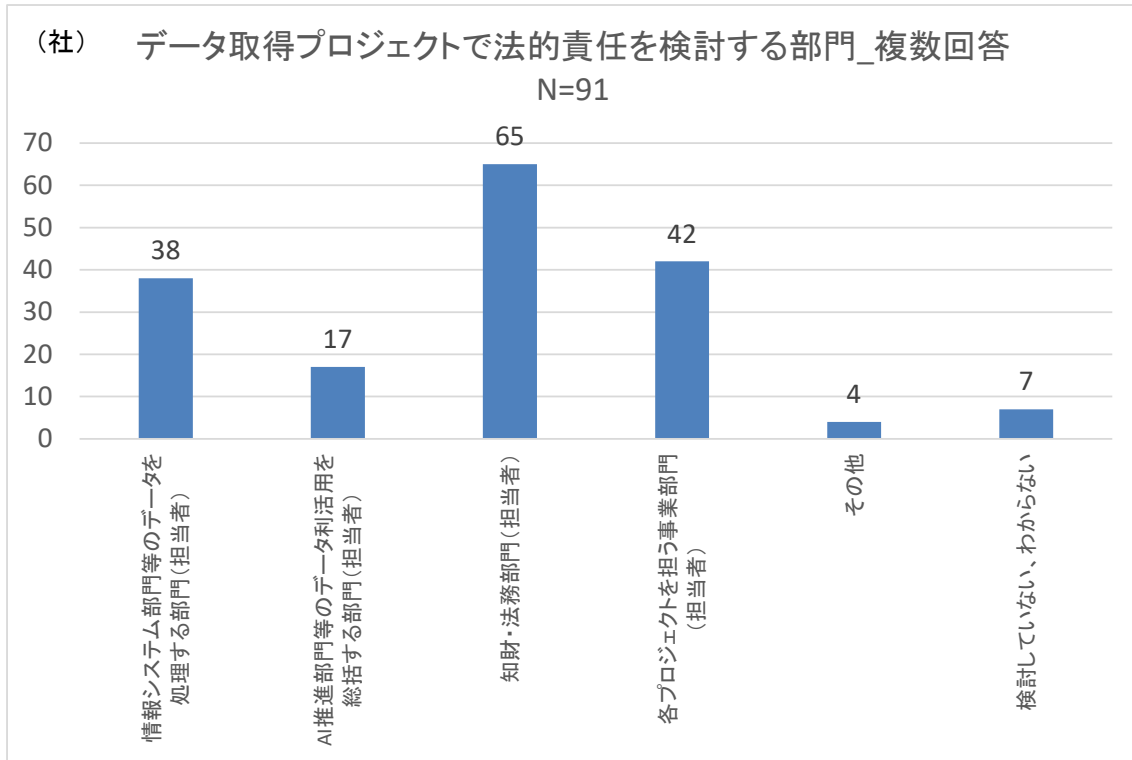
図表 37 データ提供プロジェクトにおける契約書や法的業務 (2020 年)



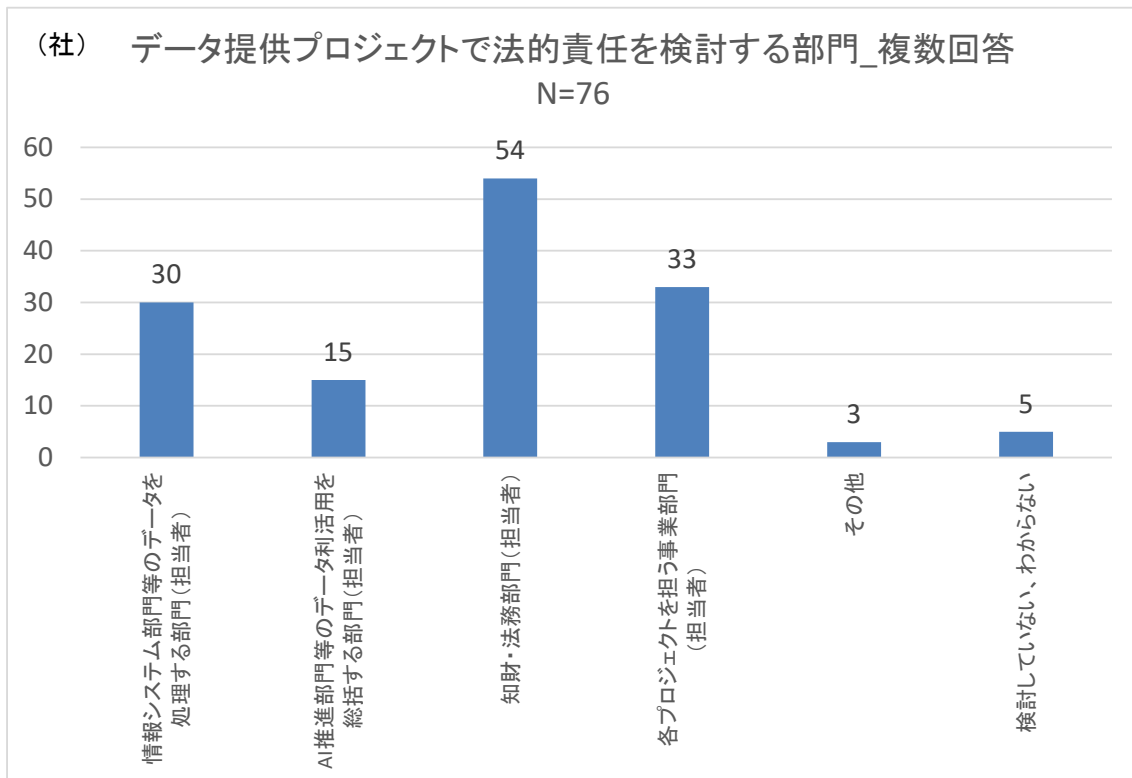
図表 38 海外とのデータ取得プロジェクトにおける契約書や法的業務 (2020 年)



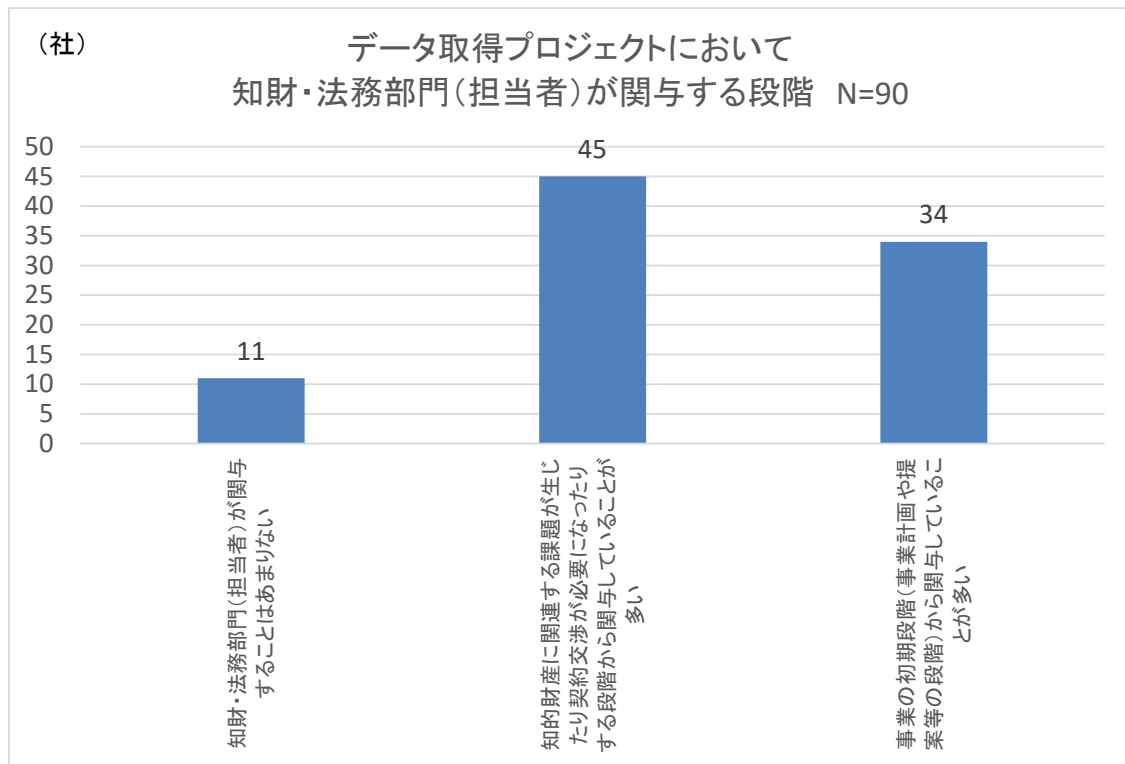
図表 39 海外とのデータ提供プロジェクトにおける契約書や法的業務 (2020 年)



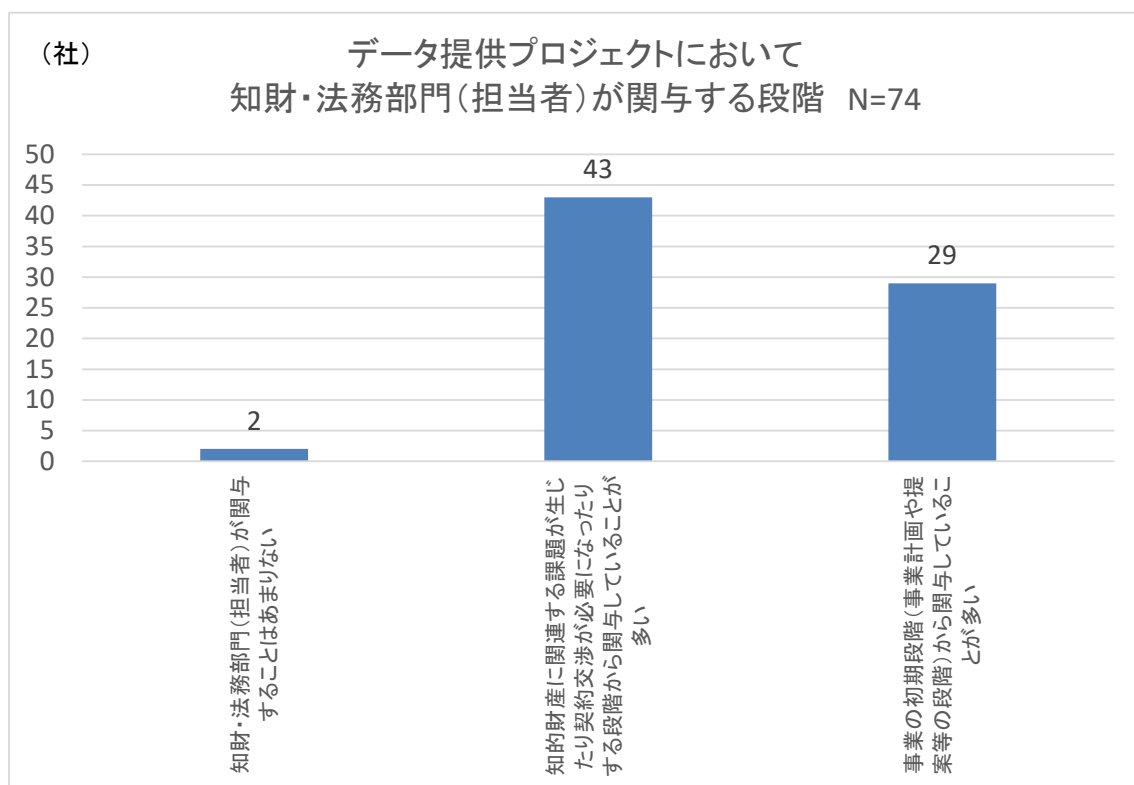
図表 40 データ取得プロジェクトで法的責任を検討する部門（複数回答、2020 年）



図表 41 データ提供プロジェクトで法的責任を検討する部門（複数回答、2020 年）



図表 42 データ取得プロジェクトで知財・法務が関与する段階 (2020 年)



図表 43 データ提供プロジェクトで知財・法務が関与する段階 (2020 年)

5. 2. 5. 小括

2017 年との比較について、データの総容量やビッグデータの体制整備、高度なデータの処理・解析の体制整備、データ利活用のこれまでの効果等、全体的なデータ利活用の状況としてはこの 3 年間で大きく進展している傾向は見られなかった。一方で、データ利活用の今後の方針については拡大する見通しと回答した企業の割合は増加しており、データ利活用の裾野が広がった可能性がうかがえた。

また、社内でデータ利活用を行う場合と他組織と提携してデータ利活用を行う場合では、その目的がやや異なる傾向が見られた。社内でのデータ利活用では生産性向上や業務効率化を目的としていた企業の割合が最も高かったのに対し、データ取得プロジェクトやデータ提供プロジェクトでは新商品の開発を目的としていた企業の割合が最も高かった。

加えて、これまでにデータ取得プロジェクトやデータ提供プロジェクトがあると回答した企業はいずれも 2 割未満であり、他組織と提携してデータ利活用を行う企業はまだそれほど多くはないことがうかがえた。中でも、海外の組織と提携するケースはごく一部の企業に限られるということがわかった。また、これまで企業がどの程度限定提供データの制度を活用しているかは明らかではなかったが、データ取得プロジェクト、データ提供プロジェクトにおいて限定提供データとして扱っていた企業は比較的多いことがうかがえた。

5. 3. 回帰分析①：2017 年との比較

ここでは、近年特に注目されているビッグデータの活用とディープラーニング等の高度なデータの処理・解析に着目し、その事業成果への影響について分析した結果を示す。分析にあたっては2017年の調査と2020年の調査を用いて同様の分析を行い、それらの影響がこの3年間で異なるのかどうかについて検証した。

5. 3. 1. 変数

被説明変数としては、「具体的間接的成果有無」及び「具体的成果有無」を使用した。「具体的間接的成果有無」は、図表 18、図表 19 の質問において「まだ成果は得られていない」と回答した場合は 0、そうでない場合は 1 とした。「具体的成果有無」は、図表 18、図表 19 の質問において「複数の事業で具体的成果が得られている」「少なくとも 1 つの事業で具体的成果が得られている」と回答した場合は 1、そうでない場合は 0 とした。

説明変数としては、「ビッグデータ体制整備」及び「高度なデータ処理・解析体制整備」を使用した。「ビッグデータ体制整備」は図表 14、図表 15 で示した質問についての 5 段階評価（1「全くそう思わない」～5「強くそう思う」）、「高度なデータ処理・解析体制整備」は図表 16、図表 17 で示した質問についての 5 段階評価（1「全くそう思わない」～5「強くそう思う」）とした。

制御変数としては、2020 年の調査、2017 年の調査で同様の質問をしたものとして、「従業員数」、「製造業ダミー」（自社の事業がカバーする業種についての複数回答で、図表 6 の「4.食品製造業」～「27.その他製造業」のうち 1 つ以上に該当する場合は 1、そうでない場合は 0）、「サービス業ダミー」（自社の事業がカバーする業種についての複数回答で、図表 6 の「1.農林水産業」～「3.建設業」及び「28.卸売／小売」～「36.その他」のうち 1 つ以上に該当する場合は 1、そうでない場合は 0）、「グループ企業数」、「海外進出国数」（自社及び自社のグループ企業の拠点が所在する海外の国数）、「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」）、「データ総容量」（図表 7、図表 8 で示した質問についての 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」）、「実際のデータ利活用率」（図表 9、図表 10 で示した質問についての 5 段階評価：1「20%未満」、2「20%以上～40%未満」、3「40%以上～60%未満」、4「60%以上～80%未満」、5「80%以上」）、「担当者数」（全社的なデータ利活用を推進する担当者数）を使用した。

5. 3. 2. 分析結果

図表 44、図表 45 は、平均、標準偏差、相関係数を示しており、図表 46、図表 47 は、ロジスティック回帰分析の結果を示している。2020 年でも 2017 年でも、被説明変数が「具体的間接的成果有無」の場合、「ビッグデータ体制整備」も「高度なデータ処理・解析体制整備」も制御変数に加えてこれら単独で投入した場合には $p < 0.01$ で正に有意であるが（モデル 2 及びモデル 3）、両方投入したモデル 4 では「ビッグデータ体制整備」のみが正で有意であった。被説明変数が「具体的成果有無」の場合は、「ビッグデータ体制整備」は 2020 年でも 2017 年でも正で有意であったが（モデル 6 及びモデル 8）、「高度なデータ処理・解析体制整備」は 2020 年のモデル 7 を除いて有意ではなかった。

5. 3. 3. 小括

2017 年と 2020 年の分析結果を比較すると、「高度なデータ処理・解析体制整備」が「具体的成果有無」に与える影響について若干の相違はあったものの、ビッグデータの体制整備や高度なデータの処理・解析の体制整備が事業成果に与える影響は大きく変わらなかった。ビッグデータの利活用を行える体制が整備できていることや高度なデータの処理・解析を行える体制が整備できている

ことはいずれも事業成果に正の影響を与えるが、高度なデータの処理・解析の体制整備の影響は限定的であった。

加えて、制御変数として投入した「データ総容量」に関しては、正に有意であったモデルもあるが有意でないモデルもあり、ロバストな結果ではなかった。

			相関													
	平均値	標準偏差	具体的成果有無	具体的間接的成果有無	従業員数	製造業タミー	サービス業タミー	グループ企業数	海外進出国数	保有特許件数	データ総容量	実際のデータ利用率	担当者数	ビッグデータ体制整備	高度なデータ処理・解析体制整備	
具体的成果有無	0.223	0.417	1													
具体的間接的成果有無	0.599	0.490	.439**	1												
従業員数	2349.035	13170.081	0.078	.094*	1											
製造業タミー	0.396	0.489	-0.036	0.018	-0.017	1										
サービス業タミー	0.739	0.440	0.024	-0.034	-0.005	-.735**	1									
グループ企業数	17.645	70.465	.118**	.106**	.309**	.099*	-0.043	1								
海外進出国数	3.411	9.290	.117**	.110**	.277**	.184**	-.133**	.670**	1							
保有特許件数	1.974	1.165	0.052	0.073	.306**	.500**	-.473**	.415**	.522**	1						
データ総容量	2.226	0.632	.182**	.221**	.184**	-0.004	-0.030	.192**	.301**	.252**	1					
実際のデータ利用率	2.485	1.305	.178**	.309**	-0.015	0.077	-.096*	-0.057	-0.020	0.066	.219**	1				
担当者数	2.802	7.638	.221**	.177**	.232**	0.052	-0.001	.368**	.560**	.230**	.272**	.110**	1			
ビッグデータ体制整備	2.315	1.014	.307**	.335**	.136**	-.124**	.098*	.138**	.144**	0.047	.256**	.245**	.236**	1		
高度なデータ処理・解析体制整備	2.125	1.061	.283**	.266**	.136**	-.090*	0.068	.154**	.153**	0.061	.210**	.179**	.230**	.738**	1	

**p<0.01, *p<0.05

**p<0.01, *p<0.05

図表 44 平均、標準偏差、相関（2020 年）

			相関												
	平均値	標準偏差	具体的成果有無	具体的間接的成果有無	従業員数	製造業タミー	サービス業タミー	グループ企業数	海外進出国数	保有特許件数	データ総容量	実際のデータ利活用率	担当者数	ビッグデータ体制整備	高度なデータ処理・解析体制整備
具体的成果有無	0.297	0.457	1												
具体的間接的成果有無	0.647	0.478	.479**	1											
従業員数	1576.742	4616.803	.180**	.156**	1										
製造業タミー	0.645	0.479	0.009	-0.002	0.015	1									
サービス業タミー	0.487	0.500	0.043	0.010	-0.012	-.762**	1								
グループ企業数	26.284	103.318	.152**	.141**	.791**	0.018	0.035	1							
海外進出国数	5.187	15.612	.137**	.155**	.435**	.098*	-0.082	.583**	1						
保有特許件数	2.854	1.188	0.067	0.083	.352**	.395**	-.361**	.241**	.363**	1					
データ総容量	2.178	0.677	.228**	.291**	.284**	-0.040	0.006	.214**	.266**	.272**	1				
実際のデータ利活用率	2.452	1.341	.239**	.268**	0.019	-0.039	0.028	0.043	-0.063	0.017	.269**	1			
担当者数	3.048	10.861	.153**	.152**	.176**	-0.004	0.048	.096*	.183**	.209**	.254**	.175**	1		
ビッグデータ体制整備	2.217	1.021	.307**	.369**	.241**	-.101*	.106*	.201**	.213**	0.055	.397**	.214**	.205**	1	
高度なデータ処理・解析体制整備	1.987	0.993	.251**	.282**	.203**	-.151**	.152**	.180**	.186**	0.073	.345**	.210**	.187**	.773**	1

*p<0.01, **p<0.05

図表 45 平均、標準偏差、相関（2017 年）

被説明変数：具体的間接的成果有無												
	モデル1			モデル2			モデル3			モデル4		
	B	標準 誤差	有意 確率	B	標準 誤差	有意 確率	B	標準 誤差	有意 確率	B	標準 誤差	有意 確率
定数	-1.650	0.543	**	-2.610	0.597	**	-2.274	0.578	**	-2.607	0.598	**
従業員数	0.00004	0.00004		0.00003	0.00004		0.00003	0.00004		0.00003	0.00004	
製造業ダミー	-0.006	0.301		0.174	0.313		0.084	0.309		0.174	0.313	
サービス業ダミー	-0.171	0.326		-0.213	0.337		-0.182	0.332		-0.213	0.337	
グループ企業数	0.001	0.005		0.001	0.005		0.001	0.005		0.001	0.005	
海外進出国数	-0.019	0.026		-0.017	0.026		-0.021	0.026		-0.017	0.026	
保有特許件数	-0.062	0.120		-0.070	0.122		-0.067	0.121		-0.070	0.122	
データ総容量	0.425	0.175	*	0.330	0.180	+	0.384	0.177	*	0.330	0.180	+
実際のデータ利活用率	0.485	0.082	**	0.432	0.085	**	0.466	0.084	**	0.432	0.085	**
担当者数	0.045	0.026	+	0.028	0.023		0.031	0.024		0.028	0.024	
ビッグデータ体制整備				0.600	0.116	**				0.609	0.164	**
高度なデータ処理・解析体制整備							0.384	0.103	**	-0.011	0.148	
Cox-Snell R2 乗	0.130			0.180			0.157			0.180		
Nagelkerke R2 乗	0.175			0.242			0.210			0.242		
N	511			510			510			510		

**p<0.01, *p<0.05, +p<0.1

被説明変数：具体的成果有無												
	モデル5			モデル6			モデル7			モデル8		
	B	標準 誤差	有意 確率	B	標準 誤差	有意 確率	B	標準 誤差	有意 確率	B	標準 誤差	有意 確率
定数	-2.790	0.664	**	-3.544	0.700	**	-3.431	0.699	**	-3.633	0.707	**
従業員数	0.0001	0.00003		0.00004	0.00003		0.00004	0.00003		0.00004	0.00003	
製造業ダミー	-0.095	0.368		0.031	0.381		-0.031	0.374		0.032	0.380	
サービス業ダミー	-0.225	0.399		-0.305	0.409		-0.251	0.409		-0.287	0.410	
グループ企業数	-0.003	0.005		-0.003	0.005		-0.003	0.005		-0.003	0.005	
海外進出国数	-0.003	0.026		-0.001	0.027		-0.002	0.026		-0.002	0.027	
保有特許件数	-0.103	0.136		-0.092	0.140		-0.093	0.139		-0.091	0.141	
データ総容量	0.448	0.209	*	0.294	0.217		0.353	0.212	+	0.297	0.217	
実際のデータ利活用率	0.258	0.086	**	0.175	0.091	+	0.201	0.089	*	0.173	0.091	+
担当者数	0.054	0.021	**	0.043	0.020	*	0.044	0.021	*	0.041	0.020	*
ビッグデータ体制整備				0.552	0.124	**				0.392	0.170	*
高度なデータ処理・解析体制整備							0.454	0.109	**	0.208	0.152	
Cox-Snell R2 乗	0.079			0.116			0.110			0.119		
Nagelkerke R2 乗	0.122			0.179			0.170			0.184		
N	511			510			510			510		

**p<0.01, *p<0.05, +p<0.1

図表 46 ロジスティック回帰分析結果（2020 年）

被説明変数：具体的間接的成果有無												
	モデル1			モデル2			モデル3			モデル4		
	B	標準 誤差	有意 確率	B	標準 誤差	有意 確率	B	標準 誤差	有意 確率	B	標準 誤差	有意 確率
定数	-2.048	0.713	**	-2.803	0.763	**	-2.514	0.744	**	-2.790	0.767	**
従業員数	0.0001	0.0001		-0.00001	0.0001		0.0001	0.0001		-0.00001	0.0001	
製造業ダミー	0.533	0.432		0.400	0.446		0.510	0.437		0.397	0.446	
サービス業ダミー	0.590	0.413		0.426	0.428		0.495	0.417		0.428	0.428	
グループ企業数	0.006	0.008		0.006	0.008		0.006	0.008		0.006	0.008	
海外進出国数	0.044	0.043		0.029	0.044		0.029	0.044		0.030	0.044	
保有特許件数	-0.126	0.147		-0.027	0.152		-0.086	0.149		-0.028	0.152	
データ総容量	0.461	0.225	*	0.235	0.239		0.338	0.233		0.238	0.239	
実際のデータ利活用率	0.417	0.109	**	0.376	0.110	**	0.398	0.111	**	0.376	0.110	**
担当者数	0.084	0.049	+	0.075	0.044	+	0.075	0.046	+	0.075	0.044	+
ビッグデータ体制整備				0.636	0.160	**				0.661	0.216	**
高度なデータ処理・解析体制整備							0.407	0.157	**	-0.036	0.213	
Cox-Snell R2 乗	0.163			0.208			0.183			0.208		
Nagelkerke R2 乗	0.221			0.282			0.248			0.282		
N	328			326			326			326		

**p<0.01, *p<0.05, +p<0.1

被説明変数：具体的成果有無												
	モデル5			モデル6			モデル7			モデル8		
	B	標準 誤差	有意 確率	B	標準 誤差	有意 確率	B	標準 誤差	有意 確率	B	標準 誤差	有意 確率
定数	-4.597	0.829	**	-4.967	0.846	**	-4.798	0.843	**	-4.950	0.847	**
従業員数	-0.0001	0.0001		-0.0001	0.0001		-0.0001	0.0001		-0.0001	0.0001	
製造業ダミー	0.447	0.425		0.435	0.428		0.489	0.427		0.417	0.431	
サービス業ダミー	0.865	0.414	*	0.806	0.415	+	0.835	0.415	*	0.812	0.416	+
グループ企業数	0.004	0.004		0.004	0.004		0.004	0.004		0.004	0.004	
海外進出国数	0.001	0.017		-0.001	0.018		0.001	0.017		-0.0005	0.018	
保有特許件数	0.052	0.144		0.073	0.147		0.039	0.145		0.079	0.148	
データ総容量	0.965	0.251	**	0.782	0.263	**	0.898	0.259	**	0.784	0.263	**
実際のデータ利活用率	0.219	0.105	*	0.189	0.107	+	0.195	0.106	+	0.190	0.107	+
担当者数	0.006	0.011		0.005	0.011		0.005	0.011		0.005	0.011	
ビッグデータ体制整備				0.371	0.146	*				0.430	0.205	*
高度なデータ処理・解析体制整備							0.215	0.144		-0.085	0.206	
Cox-Snell R2 乗	0.114			0.132			0.121			0.133		
Nagelkerke R2 乗	0.168			0.194			0.177			0.195		
N	328			326			326			326		

**p<0.01, *p<0.05, +p<0.1

図表 47 ロジスティック回帰分析結果（2017 年）

5. 4. 回帰分析② : Data Analytics Competency についての分析

前節では、ビッグデータの体制整備や高度なデータの処理・解析の体制整備が事業成果に与える影響について 3 年前との比較を行った。2020 年の調査では、これらの影響についてさらに分析するために、data analytics competency という概念をもとにした質問項目を設けた。ここでは、この data analytics competency を説明変数として分析した結果を示す。

5. 4. 1. Data Analytics Competency

経営学における IT (information technology) 分野の研究では、resource-based view を IT の文脈に当てはめた IT capability という概念を用いた研究が行われてきた。IT capability は、IT によるリソース (IT-based resources) を IT infrastructure、human IT resources、IT-enabled intangibles の 3 つとした上で、「他のリソースや能力と併せて IT によるリソースを動員して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)」と定義される¹⁶。このように組織の能力が注目されてきた背景には、データそのもののようなリソースは社外から調達することが比較的容易である一方、IT capability のような組織の能力は企業特有のものであり他社による模倣や代替が困難であるため競争優位性の源泉になり得るという考えがある。近年では、さらにこの IT capability を第 4 次産業革命の流れの中で発展させた data analytics competency (DAC) という概念や、よりビッグデータに焦点を当てた big data analytics capability (BDAC) という概念が登場している。

2020 年の調査では、ビッグデータの利活用だけでなく広くデータ利活用について調査を行ったため、DAC の概念を用いた。質問項目の作成にあたっては Ghasemaghaei et al. (2018)¹⁷を参考にした。この研究においても IT capability と同様に IT infrastructure、human IT resources、IT-enabled intangibles という 3 分類を採用しているが、human IT resources を analytical skills と domain knowledge に、IT-enabled intangibles を data quality と bigness of data にさらに細分化している。これを基に図表 48 に示す 5 つの質問項目を作成した。なお、Ghasemaghaei et al. (2018) は DAC を "a firm's ability to deploy and combine data analytics resources for rigorous and action-oriented analyses of data" と定義している。

¹⁶ Bharadwaj, A. S. (2000). A resource-based perspective on information technology capability and firm performance: an empirical investigation. *MIS Quarterly*, 24(1), 169-196.

¹⁷ Ghasemaghaei, M., Ebrahimi, S., & Hassanein, K. (2018). Data analytics competency for improving firm decision making performance. *The Journal of Strategic Information Systems*, 27(1), 101-113.

変数名	質問項目
Tools Sophistication	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	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	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	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	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）

図表 48 DAC に関する質問項目

5. 4. 2. 変数

被説明変数は、5. 3 節と同じ「具体的間接的成果有無」及び「具体的成果有無」を使用した。

説明変数は、図表 48 に示した各質問項目についての 5 段階評価（1「全くそう思わない」～5「強くそう思う」）を使用した。加えて、個々の影響だけでなく DAC の総合的な影響についても分析するため、これら 5 つの質問項目の 5 段階評価について主因子法（回転なし）により因子を抽出したところ、1 つの因子が抽出された（図表 49）。この因子を「DAC 因子」と名付け、これも説明変数とした。

制御変数は、5. 3 節で使用した変数に加えて、「上場ダミー」（上場企業であれば 1、そうでなければ 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）、「純粋持株会社ダミー」（純粋持株会社であれば 1、そうでなければ 0）、「事業持株会社ダミー」（事業持株会社であれば 1、そうでなければ 0）、「存続年数」（設立年から 2020 年までの年数）、「資本（出資）金額」（資本金額または出資金額（単位：百万円））、「データサイエンティスト数」（図表 13 で示した質問についての 5 段階評価：1「0 人」、2「1～10 人」、3「11～50 人」、4「51～100 人」、5「101 人以上」）とした。

	第1因子
	DAC因子
Tool Sophistication	0.811
Analytical Skills	0.858
Domain Knowledge	0.773
Data Quality	0.677
Bigness of Data	0.736
因子寄与	2.990
寄与率（％）	59.798
累積寄与率（％）	59.798

主因子法による

図表 49 因子行列及び因子寄与、寄与率、累積寄与率（2020 年）

5. 4. 3. 分析結果

図表 50 は、平均、標準偏差、相関係数を示しており、図表 51 は、被説明変数を「具体的間接的成果有無」とした場合のロジスティック回帰分析の結果、図表 52 は、被説明変数を「具体的成果有無」とした場合のロジスティック回帰分析の結果を示している。

DAC の 5 項目に関して、制御変数に加えてそれぞれを単独で投入した場合には、いずれの項目も $p < 0.01$ で正に有意であった（モデル 2～モデル 6）。DAC の 5 項目をすべて投入したモデル 7 では、「Data Quality」のみが $p < 0.01$ で正に有意であった。また、「DAC 因子」も $p < 0.01$ で正に有意であったが（モデル 8）、Cox-Snell R^2 乗や Nagelkerke R^2 乗の値は「Data Quality」が投入されているモデル 5 やモデル 7 の方が大きかった。なお、これらの結果は被説明変数がいずれの場合も同じであった。

5. 4. 4. 小括

本節の分析結果から、DAC はデータ利活用の成果に対して正の影響があることが示唆された。中でも、「Data Quality」の影響が大きかった。加えて、制御変数として投入した「データ総容量」に関しては、正に有意であったモデルもあるが、DAC を含むモデルでは有意でない場合もあり、ロバストな結果ではなかった。

以上の結果から、データ利活用の成果を得るためには DAC を高めることが重要であると考えられる。特にデータの量というよりは質を高めることが肝要である。図表 9 でみたように、データがあったとしてもそれを実際に十分利活用できていない企業も存在する。そのため、本当に使えるデータとしてデータの質を高めていくことが成果につながると考えられる。

	平均値	標準偏差	相関										
			具体的成果有無	具体的間接的成果有無	従業員数	製造業ダミー	サービス業ダミー	グループ企業数	海外進出国数	保有特許件数	データ総容量	実際のデータ利活用率	担当者数
具体的成果有無	0.223	0.417	1										
具体的間接的成果有無	0.599	0.490	.439**	1									
従業員数	2349.035	13170.081	0.078	.094*	1								
製造業ダミー	0.396	0.489	-0.036	0.018	-0.017	1							
サービス業ダミー	0.739	0.440	0.024	-0.034	-0.005	-.735**	1						
グループ企業数	17.645	70.465	.118**	.106**	.309**	.099*	-0.043	1					
海外進出国数	3.411	9.290	.117**	.110**	.277**	.184**	-.133**	.670**	1				
保有特許件数	1.974	1.165	0.052	0.073	.306**	.500**	-.473**	.415**	.522**	1			
データ総容量	2.226	0.632	.182**	.221**	.184**	-0.004	-0.030	.192**	.301**	.252**	1		
実際のデータ利活用率	2.485	1.305	.178**	.309**	-0.015	0.077	-.096*	-0.057	-0.020	0.066	.219**	1	
担当者数	2.802	7.638	.221**	.177**	.232**	0.052	-0.001	.368**	.560**	.230**	.272**	.110**	1
上場ダミー	0.566	0.496	0.004	-0.004	.151**	0.006	0.020	.194**	.268**	.270**	.303**	-0.045	.177**
Web回答ダミー	0.796	0.403	0.049	0.004	0.043	-.113**	.140**	0.054	0.038	-0.068	0.056	0.015	0.008
純粋持株会社ダミー	0.076	0.265	-0.021	0.047	.122**	-0.047	0.073	0.003	0.013	-0.018	0.059	-0.028	.106**
事業持株会社ダミー	0.373	0.484	0.034	0.021	0.018	0.008	0.020	.089*	.124**	.106**	0.065	-0.049	0.037
存続年数	44.245	32.417	0.022	0.033	.093*	.257**	-.269**	.190**	.268**	.370**	.253**	0.047	.200**
資本（出資）金額	13509.331	79116.877	.110**	.084*	.195**	-0.030	0.026	.272**	.385**	.174**	.167**	-0.051	.311**
データサイエンティスト数	1.383	0.621	.274**	.249**	.219**	-0.010	0.019	.222**	.296**	.201**	.235**	.246**	.444**
Tool Sophistication	2.191	1.058	.296**	.294**	.160**	-.121**	0.072	.161**	.170**	.129**	.274**	.194**	.255**
Analytical Skills	2.423	1.081	.334**	.360**	.133**	-.084*	0.053	.133**	.132**	0.074	.238**	.275**	.276**
Domain Knowledge	2.669	1.022	.293**	.358**	.083*	-0.074	0.036	.082*	.100*	.090*	.186**	.238**	.194**
Data Quality	2.688	1.018	.390**	.409**	0.063	-.082*	0.045	0.038	0.062	0.029	.212**	.255**	.194**
Bigness of Data	2.266	1.087	.289**	.309**	.169**	-.120**	.094*	.096*	.154**	.085*	.262**	.214**	.252**
DAC因子	0.000	0.944	.378**	.412**	.152**	-.114**	0.070	.134**	.154**	.102*	.282**	.288**	.291**

	相関												
	上場ダミー	Web回答ダミー	純粋持株会社ダミー	事業持株会社ダミー	存続年数	資本（出資）金額	データサイエンティスト	Tool Sophistication	Analytical Skills	Domain Knowledge	Data Quality	Bigness of Data	DAC因子
上場ダミー	1												
Web回答ダミー	-0.009	1											
純粋持株会社ダミー	.134**	0.009	1										
事業持株会社ダミー	.190**	0.073	-.221**	1									
存続年数	.508**	-.081*	-0.062	.140**	1								
資本（出資）金額	.144**	0.032	.134**	0.011	.077*	1							
データサイエンティスト数	0.055	0.034	0.071	0.061	.094*	.276**	1						
Tool Sophistication	0.020	0.072	.134**	-0.0005	-0.057	.148**	.471**	1					
Analytical Skills	-0.071	0.046	.085*	-0.035	-0.069	.150**	.531**	.717**	1				
Domain Knowledge	-.080*	.133**	0.044	-0.002	-.084*	.114**	.422**	.577**	.700**	1			
Data Quality	-0.058	-0.025	0.029	-0.071	-0.050	0.060	.327**	.490**	.560**	.588**	1		
Bigness of Data	0.032	0.045	.098*	-0.029	-0.020	.146**	.450**	.675**	.586**	.521**	.514**	1	
DAC因子	-0.038	0.067	.099*	-0.028	-0.070	.156**	.551**	.859**	.909**	.819**	.717**	.780**	1

**p<0.01, *p<0.05

図表 50 平均、標準偏差、相関（2020 年）

被説明変数：具体的間接的成果有無																								
	モデル1			モデル2			モデル3			モデル4			モデル5			モデル6			モデル7			モデル8		
	B	標準誤差	有意確率	B	標準誤差	有意確率	B	標準誤差	有意確率	B	標準誤差	有意確率	B	標準誤差	有意確率	B	標準誤差	有意確率	B	標準誤差	有意確率	B	標準誤差	有意確率
定数	-2.282	0.636	**	-2.555	0.653	**	-2.775	0.656	**	-3.000	0.673	**	-3.655	0.704	**	-2.701	0.655	**	-3.757	0.727	**	-0.947	0.694	
従業員数	0.00002	0.00005		0.00001	0.00004		0.00002	0.00004		0.00004	0.00005		0.00001	0.00004		0.00001	0.00004		0.00002	0.00004		0.00002	0.00004	
製造業ダミー	-0.084	0.311		0.082	0.318		0.013	0.317		-0.026	0.319		0.034	0.325		0.024	0.319		0.072	0.331		0.102	0.325	
サービス業ダミー	-0.163	0.341		-0.141	0.347		-0.163	0.348		-0.162	0.353		-0.219	0.360		-0.234	0.350		-0.267	0.365		-0.192	0.358	
グループ企業数	0.001	0.005		0.001	0.005		0.0005	0.005		0.0002	0.005		0.0001	0.005		0.001	0.005		-0.0001	0.005		0.0003	0.005	
海外進出国数	-0.021	0.026		-0.017	0.027		-0.015	0.026		-0.016	0.026		-0.011	0.027		-0.018	0.027		-0.007	0.027		-0.009	0.027	
保有特許件数	-0.014	0.125		-0.072	0.127		-0.035	0.127		-0.068	0.128		-0.067	0.132		-0.057	0.127		-0.101	0.134		-0.103	0.131	
データ総容量	0.436	0.181*		0.323	0.185+		0.332	0.186+		0.388	0.186*		0.283	0.192		0.361	0.185+		0.234	0.197		0.254	0.192	
実際のデータ利用率	0.469	0.086**		0.460	0.088**		0.448	0.089**		0.437	0.089**		0.454	0.092**		0.457	0.088**		0.440	0.093**		0.435	0.091**	
担当者数	0.023	0.026		0.019	0.025		0.012	0.025		0.016	0.025		0.017	0.025		0.017	0.025		0.015	0.025		0.013	0.024	
上場ダミー	-0.347	0.250		-0.325	0.254		-0.260	0.258		-0.279	0.257		-0.322	0.264		-0.325	0.254		-0.293	0.268		-0.267	0.262	
Web回答ダミー	0.067	0.270		-0.004	0.276		0.097	0.280		-0.118	0.282		0.167	0.284		0.086	0.278		0.128	0.294		0.030	0.286	
純粋持株会社ダミー	0.407	0.422		0.313	0.430		0.347	0.430		0.369	0.424		0.495	0.425		0.317	0.423		0.427	0.429		0.292	0.432	
事業持株会社ダミー	0.234	0.216		0.221	0.219		0.234	0.221		0.260	0.222		0.331	0.230		0.259	0.221		0.330	0.233		0.251	0.226	
存続年数	0.0004	0.004		0.002	0.004		0.002	0.004		0.003	0.004		0.002	0.004		0.002	0.004		0.003	0.004		0.004	0.004	
資本（出資）金額	0.000002	0.000004		0.000001	0.000003		0.000001	0.000003		0.000001	0.000003		0.000003	0.000004		0.000002	0.000004		0.000002	0.000004		0.000001	0.000003	
データサイエナリスト数	0.469	0.217*		0.230	0.228		0.089	0.233		0.155	0.226		0.127	0.228		0.194	0.227		-0.063	0.241		-0.126	0.240	
Tool Sophistication				0.421	0.119**														-0.005	0.164				
Analytical Skills							0.498	0.117**					0.543	0.115**					0.070	0.168				
Domain Knowledge																			0.159	0.152				
Data Quality																			0.627	0.139**				
Bigness of Data																0.464	0.112**		0.198	0.137				
DAC因子																						0.870	0.150**	
Cox-Snell R2 乗	0.150			0.173			0.186			0.191			0.234			0.182			0.242			0.215		
Nagelkerke R2 乗	0.201			0.232			0.250			0.256			0.314			0.245			0.325			0.289		
N	503			500			500			501			501			502			496			496		

*p<0.01, *p<0.05, +p<0.1

図表 51 ロジスティック回帰分析結果（2020 年）

被説明変数：具体的成果有無																								
	モデル1			モデル2			モデル3			モデル4			モデル5			モデル6			モデル7			モデル8		
	B	標準誤差	有意確率	B	標準誤差	有意確率	B	標準誤差	有意確率	B	標準誤差	有意確率	B	標準誤差	有意確率	B	標準誤差	有意確率	B	標準誤差	有意確率	B	標準誤差	有意確率
定数	-3.686	0.765**		-4.009	0.796**		-4.456	0.823**		-4.665	0.833**		-5.839	0.917**		-4.120	0.793**		-5.891	0.939**		-2.429	0.831**	
従業員数	0.00004	0.00004		0.00004	0.00004		0.00005	0.00004		0.00005	0.00004		0.00003	0.00004		0.00003	0.00004		0.00003	0.00004		0.00004	0.00004	
製造業ダミー	-0.138	0.378		0.046	0.383		-0.048	0.388		-0.049	0.385		-0.094	0.407		-0.043	0.384		-0.015	0.417		0.115	0.394	
サービス業ダミー	-0.223	0.419		-0.145	0.426		-0.188	0.433		-0.129	0.430		-0.282	0.451		-0.269	0.425		-0.336	0.462		-0.169	0.440	
グループ企業数	-0.003	0.005		-0.002	0.005		-0.003	0.005		-0.003	0.005		-0.003	0.005		-0.002	0.005		-0.003	0.005		-0.003	0.005	
海外進出国数	-0.009	0.028		-0.009	0.029		-0.003	0.029		-0.005	0.029		-0.004	0.031		-0.006	0.028		-0.002	0.032		-0.003	0.030	
保有特許件数	-0.073	0.143		-0.107	0.146		-0.081	0.145		-0.101	0.145		-0.037	0.152		-0.095	0.145		-0.068	0.154		-0.123	0.149	
データ総容量	0.448	0.220*		0.237	0.228		0.281	0.230		0.324	0.224		0.194	0.238		0.306	0.224		0.100	0.243		0.119	0.234	
実際のデータ利用率	0.218	0.091*		0.196	0.093*		0.154	0.095		0.169	0.094+		0.116	0.098		0.181	0.094+		0.092	0.101		0.140	0.097	
担当者数	0.025	0.022		0.026	0.022		0.017	0.022		0.024	0.022		0.025	0.023		0.024	0.022		0.020	0.023		0.020	0.022	
上場ダミー	-0.013	0.309		0.087	0.317		0.190	0.323		0.142	0.322		0.105	0.335		0.035	0.318		0.220	0.341		0.245	0.330	
Web回答ダミー	0.253	0.354		0.214	0.361		0.287	0.366		0.143	0.365		0.375	0.372		0.304	0.358		0.385	0.384		0.217	0.369	
純粋持株会社ダミー	-0.817	0.586		-0.930	0.588		-0.894	0.585		-0.878	0.595		-0.513	0.595		-0.800	0.583		-0.556	0.595		-0.889	0.587	
事業持株会社ダミー	0.153	0.254		0.175	0.260		0.199	0.264		0.145	0.261		0.327	0.277		0.185	0.260		0.369	0.283		0.230	0.270	
存続年数	-0.003	0.005		-0.0001	0.005		-0.0001	0.005		0.0004	0.005		0.001	0.005		-0.001	0.005		0.002	0.005		0.002	0.005	
資本（出資）金額	0.000001	0.000002		0.000001	0.000002		0.000001	0.000002		0.000001	0.000002		0.000001	0.000003		0.000001	0.000002		0.000001	0.000002		0.000001	0.000002	
データサイエンティスト数	0.654	0.220**		0.387	0.235+		0.185	0.244		0.334	0.235		0.328	0.239		0.357	0.235		0.169	0.261		0.052	0.254	
Tool Sophistication				0.451	0.126**														-0.049	0.177				
Analytical Skills							0.648	0.135**											0.316	0.193				
Domain Knowledge										0.585	0.140**								-0.071	0.198				
Data Quality													1.014	0.154**					0.899	0.178**				
Bigness of Data																0.478	0.120**		0.143	0.150				
DAC因子																					0.976	0.173**		
Cox-Snell R2 乗	0.104			0.127			0.146			0.137			0.195			0.132			0.204			0.167		
Nagelkerke R2 乗	0.161			0.197			0.226			0.212			0.302			0.204			0.316			0.259		
N	503			500			500			501			501			502			496			496		

**p<0.01, *p<0.05, +p<0.1

図表 52 ロジスティック回帰分析結果（2020 年）

5. 5. 回帰分析③：データ利活用の戦略的取組みの進捗プロセスと成果の関係の分析

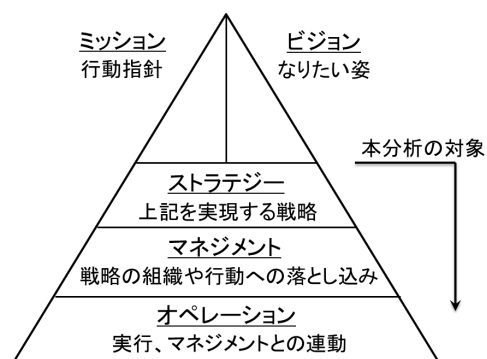
5. 5. 1. 本分析の視座

本節における分析の焦点は、企業におけるデータ利活用に係る戦略の策定と、その後のデータ利活用の浸透や成果との関係にある。本来、多くの企業では、経営に関する上位概念としてビジョン（なりたい姿や目指す方向性等）やミッション（会社全体の行動指針等）を掲げ、それらを実現していく方策として戦略（経営戦略や事業戦略、知財戦略等）を策定することが多い（図表 53）。このうち、本分析では、データ利活用に関する戦略およびそれより下位層における組織的変化を興味の対象として、アンケート結果からそれらの関係性を追跡する。

一般に、戦略が策定され、それに合わせ必要に応じた組織の改編や整備、知識の獲得、人材育成や担当者の能力構築がなされていく。近年では、これらと並行して実際のオペレーション作業の準備が進むことも多い。大企業であれば、本社や事業部の経営層から現場へと進捗が進むイメージとなる。他方、中小企業やベンチャー企業では、上位から下位へという伝達経路を辿ることなく、同時に各階層の取組みが進むかもしれない。いずれにしても、それらの取組みに対する進捗の結果として、具体的な事業目標の達成につながる。当然、事業目標は多様であり、戦略やマネジメントも一義的ではないため、N×Nの構造となる。

本研究の「データ利活用に関するアンケート調査」では、戦略やマネジメント、オペレーションに係る進捗状況を5段階リッカートスケールで回答してもらっている。また、それらの成果にあたるデータも複数の形で取得している。よって、戦略やマネジメント、オペレーションの進捗の関係、ならびにそれらと成果との関係を回帰することが可能である。さらに、企業規模についても、売上高や研究開発費、上場企業か否かといったデータを組み合わせることで、差異を確認することができる。

以上を本分析の視座とし、2つの関連した推定モデルを構築した上で、重回帰分析を用いて戦略浸透のプロセスおよび成果や目標の達成度について推定を試みる。



図表 53 組織におけるビジョンやミッション、戦略、マネジメント、オペレーションのイメージ

5. 5. 2. 推定モデルと仮説生成

以上の問題意識の下、本分析では次の2つの推定モデルを構築した。

推定モデル（1）

本研究の「データ利活用に関するアンケート調査」では、2020年3月末時点での回答企業におけるデータ利活用の取組みがどの程度進捗しているかを問う設問を設定している（問 2-5）。この設問は、19項目に及ぶデータ利活用に関する取組みについて、「強くそう思う」、「そう思う」、「どちらともいえない」、「そう思わない」、「全くそう思わない」の5段階で回答を得ている。また、19の項目は「戦略・方針について」（3項目）、「実施体制・ツールについて」（6項目）、「スキル・知識・人材について」（4項目）、「データについて」（6項目）の4つのグループに分

けられている。

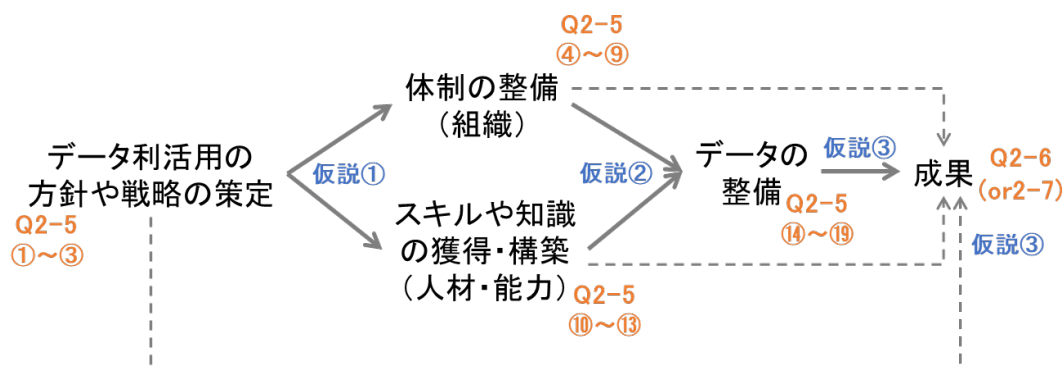
本分析では、このグループ間の関係に強い焦点を当て、モデルを構築した。具体的には「データ利活用の方針や戦略の策定」がなされた上で、「体制の整備」や「スキルや知識の獲得・構築」が進み、さらに「データの整備」が行われるというプロセスを想定した。最終的にそれらの進捗の結果として、事業活動に何らかの成果がもたらされる。

以上の流れから、以下の3つの仮説が生成される。

仮説①：データ利活用に関する戦略が策定されることで、データ利活用の体制の整備やスキル・知識の獲得が進む

仮説②：データ利活用の体制の整備やスキル・知識の獲得が進むことで、データの整備が進む

仮説③：データの整備が進むことで、データ利活用の成果が得られる



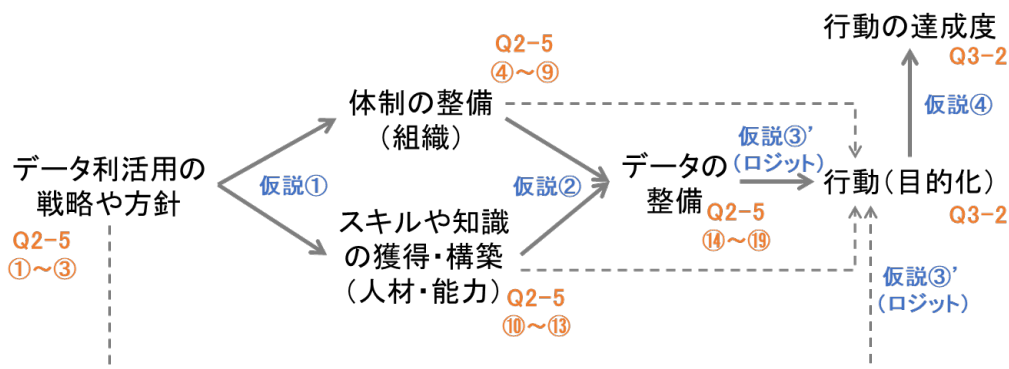
図表 54 本分析の推定モデル (1)

推定モデル (2)

2つ目のモデルは、途中まではモデル (1) と同様である。異なるのは仮説③'以降にある。モデル (1) において成果としていた箇所を、事業における行動 (目的化) へ置換したのがモデル (2) になる。データの利活用はすべての事業に活かされるわけではない。よって、どのような目的のためにデータ利活用を進めたのかを確認する必要がある (仮説③')。また、その目的が本当に達成できたかどうかを分析する (仮説④)。

仮説③'：多様な事業の目的に合わせて、データ利活用に関する戦略の策定、データ利活用の体制の整備やスキル・知識の獲得、データの整備が進められる

仮説④：データ利活用に関する戦略の策定等が、事業目的の達成につながる



図表 55 本分析の推定モデル (2)

5. 5. 3. 使用した変数

推定モデル（1）における従属変数・独立変数

本節では、上記の通り、問 2-5 の 4 つのグループのすべての回答結果を仮説①および仮説②の従属変数と独立変数に設定した。「データ利活用の戦略や方針」は独立変数のみ、「体制の整備」と「スキルや知識の獲得・構築」は、仮説①では従属変数、仮説②では独立変数という位置づけになる。また、「データの整備」は仮説②の従属変数となる。

仮説③については、問 2-5 の 4 グループを独立変数に、問 2-6 のデータ利活用のこれまでの成果に関する回答結果を従属変数に設定した。同設問は回答企業におけるデータ利活用の成果について、「まだ成果は得られていない」という段階から「複数の事業で具体的成果（売上やコストダウン等）が得られている」という段階までの 5 段階から回答を得ている。

推定モデル（2）における従属変数・独立変数

モデル（2）の仮説②までは、モデル（1）と同じ構造である。異なるのは仮説③'と仮説④である。

仮説③'の従属変数に用いたのは問 3-2 の回答結果である。問 3 の一連の設問は、回答企業自身のデータを同社内で利活用するケースを想定した設計となっている。そのうち、問 3-2 は、2019 年度において回答企業で保持しているデータの利活用の目的、ならびにその達成状況について問うたものである。「組織改革や経営戦略策定」や「人材育成・能力開発」など、8 つの項目に対し、それぞれ「目的としていない」、（目的とした場合）「目標を達成していない」、「目標を達成した」、「目標以上に達成した」の 4 段階で回答を得ている。

モデル（2）の仮説③'では、このうち“目的としたか否か”を従属変数として二項ロジスティック回帰を適用した。さらに仮説④では、“目的としたうち、どの程度目標を達成したか”を従属変数に、順序ロジットを設定し分析した。8 つの項目が用意されているため、それぞれについて推定した。なお、独立変数はモデル（1）と同様である。

制御変数

従属変数に影響を与えうる要因として、「売上高」、「研究開発費（総額）」、業種（製造業ダミー）、上場しているか否か（上場ダミー）、「国内保有特許件数」、全社的なデータ利活用担当者の有無（担当者ダミー）を用いた。また、問 4-1 では、他機関が保持しているデータを受領し、提携契約を結んだプロジェクトを実施した経験の有無を問うている。本分析では、この経験を持つ企業ほど、外発的にデータ利活用の戦略や人材育成等が進むと見て、制御変数に加えた。

5. 5. 4. 推定式

本分析の基本的な推定式は、シンプルな以下の構造をしている。

$$performance_i = \alpha + \beta \times data_i + \sum_{k=1} \gamma_k X_i + \varepsilon_i \quad (1)$$

添え字の i はデータセットに含まれる回答企業を表す。 α 、 β 、 γ は推定するパラメータ、 ε は $N(0, \delta_\varepsilon^2)$ に従う誤差項を示している。従属変数である *performance* は、すでに説明した通りである。このうち、仮説'については二項ロジスティック回帰、仮説については順序ロジット分析を用いてパラメータを推定した。分析方法は最小二乗回帰（OLS）を用いた。

5. 5. 5. 結果と仮説検証

ここからは推定の結果と仮説検証に入る。

モデル（1）の仮説①の推定結果を図表 56 に示す。上が体制の整備を従属変数にした結果で、下がスキル・知識・人材の獲得を従属変数にした結果である。有意水準は表の下に付した通りで、参考までに 4 段階で示してある。

一見して、データ利活用の戦略の策定は、体制の整備やスキル・知識・人材の獲得に有意に働いていることがわかる。ただし、社内全体で連携するような戦略よりは、社外の組織と連携する戦略を打ち出す方が、有意水準が高く、かつ係数も大きい。

制御変数の中では、上場ダミーと製造業ダミーの一部でマイナスの有意が見られる。すなわち、上場企業よりも非上場企業の方が、また製造業よりも非製造業の方が、データ利活用の体制の整備や人材育成が進みやすいことを表している。

従属変数:体制の整備	係数	標準誤差	係数	標準誤差
売上高	0.000	0.000	0.000	0.000
研究開発費	0.000	0.000	0.000	0.000
上場ダミー	-0.169 *	0.069	-0.062	0.056
製造業ダミー	-0.257 **	0.081	-0.219 **	0.065
保有特許件数	0.018	0.036	0.003	0.029
担当者ダミー	0.461 ***	0.066	0.121 *	0.057
データ取得Pjの有無	0.679 ***	0.094	0.404 ***	0.078
戦略①利活用を積極推進			0.149 **	0.048
戦略②社内全体で連携して利活用			0.124 +	0.059
戦略③社外の組織と連携して利活用			0.251 ***	0.033
Observations	487			
Adjusted R ²	0.209		0.232	
+p<.1, *p<.05, **p<.01, ***p<.001				

従属変数:スキル・知識・人材の獲得	係数	標準誤差	係数	標準誤差
売上高	0.000	0.000	0.000	0.000
研究開発費	0.000	0.000	0.000	0.000
上場ダミー	-0.311 **	0.080	-0.201 **	0.069
製造業ダミー	-0.218 *	0.093	-0.189 *	0.081
保有特許件数	0.081 *	0.041	0.064 +	0.036
担当者ダミー	0.586 ***	0.075	0.253 ***	0.070
データ取得Pjの有無	0.710 ***	0.108	0.441 ***	0.096
戦略①利活用を積極推進			0.232 ***	0.060
戦略②社内全体で連携して利活用			0.063	0.061
戦略③社外の組織と連携して利活用			0.214 ***	0.041
Observations	487			
Adjusted R ²	0.201		0.221	
+p<.1, *p<.05, **p<.01, ***p<.001				

図表 56 仮説①の推定結果

次にモデル（1）の仮説②の検証結果を示す。有意水準から明らかなように、体制の整備、スキル・知識・人材の獲得が進めば進むほど、社内データの整備がなされる傾向にあることがわかる。特に、体制の整備の影響が大きく出ている。上場ダミーや製造業ダミーでは、仮説①と類似した傾向が表れている。

従属変数: データの整備	係数	標準誤差	係数	標準誤差	係数	標準誤差	係数	標準誤差
売上高	0.000	0.000	0.000 +	0.000	0.000	0.000	0.000 *	0.000
研究開発費	0.000	0.000	0.000 +	0.000	0.000	0.000	0.000 *	0.000
上場ダミー	-0.187 *	0.075	-0.062	0.054	0.000	0.058	-0.010	0.052
製造業ダミー	-0.326 ***	0.088	-0.133 *	0.064	-0.195 **	0.067	-0.130 *	0.061
保有特許件数	0.069 +	0.039	0.055 +	0.028	0.019	0.030	0.035	0.027
担当者ダミー	0.428 ***	0.071	0.084	0.054	0.074	0.058	0.019	0.052
データ取得PJの有無	0.504 ***	0.101	0.002	0.077	0.075	0.081	-0.051	0.073
体制の整備			0.751 ***	0.036			0.516 ***	0.048
スキル・知識・人材の獲得					0.607 ***	0.034	0.296 ***	0.042
Observations	487							
Adjusted R ²	0.245		0.237		0.222		0/272	

図表 57 仮説②の推定結果

次にモデル (1) の仮説③の検証結果を示す。データ利活用の何らかの取組みが実際の成果につながるかどうかという意味では、ここが大きなポイントなる。推定の結果からは、データ利活用の戦略の策定、体制の整備、スキル・知識・人材の獲得、データの整備がそれぞれ進むことで、データ利活用の成果は上がる傾向が見られた。

興味深いのは最右の結果で、4 つのグループをすべて推定モデルに投入した場合、データの整備のみが統計的に有意な結果となった。その係数が大きさからも、やはり企業においては戦略策定からデータの整備まで、いかに効率的・効果的につなげていくかで、成果を大きく左右する可能性が高い。現状では、このプロセスは道半ばという企業も多いことから、今後も継続した取組みが重要になってくるだろう。

従属変数: データ利活用の成果	係数	標準誤差	係数	標準誤差	係数	標準誤差	係数	標準誤差	係数	標準誤差	係数	標準誤差
売上高	0.000 **	0.000	0.000 **	0.000	0.000 *	0.000	0.000 **	0.000	0.000 ***	0.000	0.000 **	0.000
研究開発費	0.000 **	0.000	0.000 **	0.000	0.000 **	0.000	0.000 **	0.000	0.000 ***	0.000	0.000 ***	0.000
上場ダミー	-0.132	0.126	-0.001	0.119	-0.021	0.118	0.043	0.119	0.012	0.113	0.055	0.113
製造業ダミー	-0.038	0.147	0.033	0.137	0.168	0.138	0.125	0.137	0.255 +	0.132	0.235 +	0.133
保有特許件数	0.068	0.064	0.029	0.060	0.025	0.060	-0.008	0.060	-0.012	0.057	-0.020	0.057
担当者ダミー	0.597 ***	0.118	0.180	0.121	0.250 *	0.117	0.215 +	0.118	0.211 +	0.110	0.110	0.116
データ利活用の戦略の導入			0.563 ***	0.067							0.142 +	0.088
体制の整備					0.643 ***	0.075					0.001	0.119
スキル・知識・人材の獲得							0.568 ***	0.066			0.155 +	0.095
データの整備									0.764 ***	0.068	0.551 ***	0.109
Observations	459											
Adjusted R ²	0.066		0.191		0.195		0.197		0.267		0.274	

+p<.1, *p<.05, **p<.01, ***p<.001

図表 58 仮説③の推定結果

ここからはモデル (2) に移り、仮説③の検証結果を検討する。先に述べた通り、従属変数とした問 3-2 は 8 項目あるため、それぞれの推定結果を並べた。統計的な有意の関係は、一見して複雑な状況にある。わかりやすいのは、「データ利活用の戦略」と「組織改革や経営戦略策定」（上段左）や、「スキル・知識・人材の獲得」と「人材育成・能力開発」（上段左から 2 番目）などだろう。独立変数と従属変数が対応したところに統計的有意が付いており、この結果からも推定の信頼性を確認することができる。ただし、いずれも有意水準は高くなく、不安定な印象を拭えない。選択する推定モデルによって結果は異なることも容易に想像される。もともと、データ利活用の戦略等によって事業の目的化が進む、という仮説自体にやや無理があるため、妥当な結果とも言えるかもしれない。

二項ロジスティック	組織改革や経営戦略策定 (目的化)		人材育成・能力開発(目的 化)		生産性向上・効率化(目的 化)		既存商品の改良(目的化)	
	係数	標準誤差	係数	標準誤差	係数	標準誤差	係数	標準誤差
売上高	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
研究開発費	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
上場ダミー	0.460 *	0.219	0.273	0.213	0.251	0.247	0.057	0.218
製造業ダミー	0.518 *	0.258	-0.009	0.249	0.592 *	0.290	0.455 +	0.258
保有特許件数	0.032	0.115	0.063	0.110	0.059	0.140	0.157	0.113
担当者ダミー	-0.227	0.224	0.086	0.214	-0.296	0.248	-0.017	0.222
データ利活用の戦略の導入	0.361 *	0.169	0.131	0.162	0.390 *	0.183	0.283 +	0.168
体制の整備	-0.069	0.230	0.076	0.222	0.155	0.257	0.086	0.230
スキル・知識・人材の獲得	0.420 *	0.186	0.394 *	0.179	-0.213	0.212	0.176	0.185
データの整備	0.065	0.206	0.042	0.201	0.426 +	0.232	0.389 +	0.209
Observations	467							
Cox-Snell R ²	0.099		0.102		0.118		0.129	

+p<.1, *p<.05, **p<.01, ***p<.001

二項ロジスティック	新商品の開発(目的化)		業界や市場の分析(目的 化)		CRM(目的化)		他組織との連携構築(目的 化)	
	係数	標準誤差	係数	標準誤差	係数	標準誤差	係数	標準誤差
売上高	0.000	0.000	0.000	0.000	0.000	0.000	0.000 +	0.000
研究開発費	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
上場ダミー	-0.086	0.217	0.136	0.215	0.063	0.222	0.069	0.230
製造業ダミー	0.469 +	0.256	0.433 +	0.254	0.348	0.260	0.379	0.265
保有特許件数	0.153	0.114	0.138	0.111	-0.085	0.120	0.071	0.118
担当者ダミー	0.025	0.220	-0.081	0.220	0.072	0.223	0.122	0.224
データ利活用の戦略の導入	0.243	0.167	0.432 **	0.167	0.353 *	0.169	0.433 +	0.176
体制の整備	0.208	0.228	0.290	0.228	0.000	0.233	0.257	0.234
スキル・知識・人材の獲得	0.256	0.182	-0.048	0.183	0.016	0.187	-0.064	0.186
データの整備	0.155	0.207	0.039	0.206	0.502 *	0.211	0.239	0.215
Observations	467							
Cox-Snell R ²	0.095		0.082		0.116		0.117	

+p<.1, *p<.05, **p<.01, ***p<.001

図表 59 仮説③'の推定結果

以下が最後の仮説検証になる。仮説④の結果である。こちらも先と同様、やや複雑な印象を受けるが、より明確な結果が出ている箇所もある。

ここで主張したい最大の論点は、上段左から3番目、4番目と、下段左から3番目の結果である。これらの従属変数はいずれも既存事業あるいは既存顧客を対象とした事業目標である。それらの目標の達成に対して、データ利活用の戦略の導入や、データの整備が強く寄与していることがわかる。特にデータの整備、具体的には質の高いデータセットの構築やビックデータの処理などの取組みが、既存事業・既存顧客への貢献を強める可能性がある。

一方、それと対比して、下段左の「新商品の開発」は係数も比較的小さく、また統計的有意も見られない。2019年度の時点では、データ利活用に対する戦略策定以下の企業努力は、少なくとも新商品開発には寄与していないと示唆される。

順序ロジット	組織改革や経営戦略策定 (達成度)		人材育成・能力開発(達成 度)		生産性向上・効率化(達成 度)		既存商品の改良(達成度)	
	係数	標準誤差	係数	標準誤差	係数	標準誤差	係数	標準誤差
売上高	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
研究開発費	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
上場ダミー	0.320 +	0.199	0.258	0.204	0.460 *	0.199	0.019	0.202
製造業ダミー	0.242	0.233	-0.103	0.241	0.413 +	0.234	0.464 +	0.238
保有特許件数	0.072	0.099	0.065	0.103	0.109	0.100	0.105	0.101
担当者ダミー	-0.146	0.200	0.199	0.204	-0.133	0.201	0.102	0.202
データ利活用の戦略の導入	0.415 **	0.153	0.124	0.156	0.392 **	0.153	0.207	0.155
体制の整備	-0.118	0.206	0.123	0.211	0.005	0.207	0.056	0.209
スキル・知識・人材の獲得	0.405 *	0.165	0.394 *	0.168	-0.101	0.165	0.031	0.165
データの整備	0.198	0.187	0.170	0.192	0.678 ***	0.190	0.673 ***	0.193
Observations	456							
Cox-Snell R ²	0.136		0.110		0.154		0.155	
+p<.1, *p<.05, **p<.01, ***p<.001								

順序ロジット	新商品の開発(達成度)		業界や市場の分析(達成度)		CRM(達成度)		他組織との連携構築(達成度)	
	係数	標準誤差	係数	標準誤差	係数	標準誤差	係数	標準誤差
売上高	0.000 +	0.000	0.000 +	0.000	0.000 +	0.000	0.000	0.000
研究開発費	0.000 +	0.000	0.000 *	0.000	0.000	0.000	0.000	0.000
上場ダミー	-0.138	0.204	0.306	0.201	0.115	0.201	0.165	0.229
製造業ダミー	0.335	0.238	0.422 +	0.235	0.163	0.237	0.276	0.268
保有特許件数	0.143	0.101	0.109	0.100	-0.030	0.101	0.105	0.116
担当者ダミー	0.120	0.204	-0.003	0.201	0.147	0.201	0.266	0.222
データ利活用の戦略の導入	0.187	0.158	0.295 *	0.154	0.364 *	0.154	0.455 **	0.177
体制の整備	0.153	0.210	0.300	0.207	-0.031	0.209	0.284	0.231
スキル・知識・人材の獲得	0.257	0.167	0.040	0.165	0.027	0.165	-0.063	0.182
データの整備	0.329 +	0.192	0.243	0.188	0.659 **	0.191	0.419 *	0.213
Observations	456							
Cox-Snell R ²	0.143		0.132		0.167		0.166	
+p<.1, *p<.05, **p<.01, ***p<.001								

図表 60 仮説④の推定結果

5. 5. 6. 本分析の課題

本分析はまだ初期段階にあり、多くの課題が存在する。その中でも特に大きな課題は次の3点である。

第一に、仮説のさらに前の段階とも言える、戦略策定の駆動力が何なのかを本研究では捉えることができていない。当然ながら、何らかの理由がなければ、戦略の策定には至らない。図表 53 に、ビジョンやミッションから、戦略、マネジメント、オペレーションまでの簡単な概念図を示したが、本研究が対象とするデータ利活用に係る戦略の検討においては、こういった要因や必要性が生じ、図表 53 でいうより下位のフェーズへ落とし込まれたのかを無視している。その要因や必要性は企業によって異なるはずで、これを仮に仮説とすると、その理由如何によって、それ以降のプロセスの進行も大きく異なってくることが予想される。本研究で実施したアンケート調査からは、仮説に該当する変数を得ることは困難であるため、すぐには着手できない課題となるが、絶えず念頭に置いておく必要がある。

第二に、本分析は重点的な先行研究を行っていない。したがって、生成した仮説はあくまでもパイロット的な位置づけに留まる。戦略の策定から実際のマネジメント、オペレーションに落とし込むプロセスは極めて複雑であり、各フェーズに対する先行研究をしっかりと把握することで、より精緻な問題意識の設定と仮説生成が可能となる。本研究からしっかりとした学術貢献を積み上げるために、早急に実施すべき課題と言える。

第三に、推定結果の解釈の深化と精緻化である。これは本研究全体の課題であるが、推定結果の解釈を補完する質的な調査を、まだほとんど実施できていない。そのため、統計的に分析した結果と実際の現場の認識との乖離があるのか、またあるとしたらどのような議論が不足しているのかに

についても、把握できていない。本研究の結果を学術のみならず、社会的な貢献にまでつなげるためには、インタビューやワークショップ、セミナーといった活動を継続的に取り入れる必要がある。

5. 6. 回帰分析④：データ取得・提供プロジェクトにおける法務体制の影響の分析

5. 6. 1. 問題意識

組織間でのデータ取得・提供プロジェクトは、データ分析やデータを起点とした組織間連携を長く行ってきた組織を除くと、必ずしも幅広い企業で行われているものではない。本調査において、データ取得プロジェクトを行っていると回答した企業は 15.1%、データ提供プロジェクトを行っていると回答した企業は 11.8%にとどまっていることが確認されたことはその証拠の一つである。

このような状況においては、データ取得・提供の取引関係のマネジメントについての知見は必ずしも蓄積されていないことが推測される。例えば、広くデータの取得・提供を行う際に、様々な不確実性が障害になっていることが予想される。ここでの不確実性とは、たとえば、データの質についての保証や、データに対する法的な制限の保証、あるいは、得られる中間生成物や成果物などである。

このような不確実性は契約により緩和できるものの、契約書上の取り決めには、現実の個々の取引・組織間関係に当てはめると多くの不完全な部分が残されていることが一般的であり、そのために、当事者は契約上の責任を回避する動機を持ちやすい¹⁸。そこで、個々の関係にカスタマイズした契約の作成過程（契約のドラフティング）を通じて双方の期待と責任範囲を明確化することが重要になる。これによって当事者それぞれの期待が可視化され、信頼構築につながる¹⁹。そして、このような信頼関係を構築していることが、取引関係の成功の確率を高めることとなる²⁰。

しかし、単に契約を作成すればよいというわけではない。的確な期待の可視化、リスクの予測、そして、責任の分配を行う、契約ドラフト能力が必要である。とりわけ、適切なマイルストーン管理を契約に落とし込む能力が取引関係の成功を左右する²¹。このときに有効であると考えられるものが、契約雛形の整備と、専門的な知識を保有しうる担当者の存在である。

第一に、契約雛形は、的確に作り込んだものであれば、双方の期待やリスクのカatalogとして機能しうると考えられる。一方で、これに固執したり、雛形に柔軟性がなければ、かえって相手方との信頼関係を削ぎかねない。とくに相手方との関係は、時間に応じて変わるものであり、one fit all となるような雛形は考えにくい²²。

第二に、知財・法務部門の関与は、法務の専門家として適切な契約のドラフトに寄与する可能性が考えられる。一方で、データ取得・提供契約について必ずしも知見が培われているものではないため、寄与が限定的である可能性もある。また、プロジェクトの窓口となる部門や相手方とのコミュニケーション上の障壁になる可能性もある。

このような契約に関する法務体制の影響に加えて、近年の研究では組織としての法制度・契約の機敏な使いこなし能力（legal astuteness）の重要性を指摘するものがある²³。データ取得・提供の文脈で考えると、平成 30 年の不正競争防止法改正で導入された限定提供データ制度の使いこなしが、企業のデータ取得・提供プロジェクトの成功に寄与している可能性が考えられる。

¹⁸ Williamson, O. E. (1999). *The Economic Institutions of Capitalism: Firms, Markets, Relational Contracting*. Free Press.

¹⁹ Lumineau, F., & Malhotra, D. (2011). Shadow of the contract: How contract structure shapes interfirm dispute resolution. *Strategic Management Journal*, 32(5), 532-555.

²⁰ 実証研究として、Poppo, L., & Zenger, T. (2002). Do formal contracts and relational governance function as substitutes or complements?. *Strategic Management Journal*, 23(8), 707-725.

²¹ 実証研究として、Shen, L., Su, C., Zheng, X., & Zhuang, G. (2019). Contract design capability as a trust enabler in the pre-formation phase of interfirm relationships. *Journal of Business Research*, 95, 103-115.

²² Lui, S. S. (2009). The roles of competence trust, formal contract, and time horizon in interorganizational learning. *Organization Studies*, 30(4), 333-353.

²³ Bagley, C. E. (2008). Winning legally: The value of legal astuteness. *Academy of Management Review*, 33(2), 378-390.

5. 6. 2. 研究上の問い

このように、広義の法務体制がデータ取得・提供プロジェクトの成功に影響することが想定される。ただし、データ取得・提供プロジェクトの中の質の違いは無視できない。組織改革や人材育成、あるいは組織間連携のためのプロジェクトはその性質上、必ずしも多数存在しないと推測される。他方、既存事業の改善、市場の分析や新規事業開発に関するものは、繰り返し実行される可能性が相対的に高く、契約に関する知見も蓄積が相対的に見れば容易であると考えられる。

そこで、既存事業、新規事業に関わるプロジェクトのみを対象として、以下の4つの要因がプロジェクトの成功に寄与しているかを分析することとする。

- 雛形の整備の程度
- 契約書のカスタマイズの程度
- 知財・法務部門の関与の程度
- 法制度の機敏な使いこなしとしての限定提供データ制度の活用

このうち、限定提供データ制度については、企業の事業内容によっては全く活用の余地が無いことが考えられる。そこで、主としてその活用の割合が中程度である場合と高程度である場合の差に注目をする。

なお、法務体制との関係で考えると、ここでの成功には、プロジェクトの目標達成という観点からの成功と、法務的な観点が入ることによるリスクの低減という観点からの成功の2つが含まれる。これは法務機能に戦略的な側面と防衛的な側面があることに由来する²⁴。本分析ではその両者に着目をする。

5. 6. 3. 分析手法

データ取得・提供プロジェクトを行っているという回答した企業（123社）の回答から、提供、取得プロジェクトそれぞれについての回答を独立した回答として分析の対象とした。これは、本分析の仮説からは取得・提供を区別する意義が乏しいためである。つまり、分析の単位は企業×プロジェクト種別であり、取得・提供の双方を行っている企業については取得・提供のそれぞれで分析対象となっている。このような二重のカウントの対象となったものは49社ある。

このうち、取得・提供において既存事業の改善、市場の分析や新規事業開発に関する目的を含むプロジェクトを実施している企業に絞った。具体的には以下の5種の目的である。

- 生産性向上や業務効率化（生産量の最適化、生産方法の改良、業務サポート等）
- 既存事業の改良（既存の商品・サービスの改良や高品質化、低価格化、メンテナンス精度の向上等）
- 新規事業開発（新商品の開発、新サービスの提案等）
- 業界や市場の分析（市場動向の分析、競合他社の調査等）
- CRM（customer relationship management：顧客関係管理）の向上（顧客満足度の向上、顧客接点の充実、営業支援等）

結果、76社の109社×プロジェクト種別のデータを分析対象とし、欠損値を除外した102社×プロジェクト種別が最終的な分析対象となった。

5. 6. 4. 使用した変数

被説明変数

被説明変数の第一は、プロジェクトのパフォーマンスである。ここではプロジェクトの目標を達成したか否かによってパフォーマンスを計測した。ただし、個々の企業では上記5種の全てを目的

²⁴ 経済産業省(2019)国際競争力強化に向けた日本企業の法務機能の在り方研究会報告書。

としてデータ取得・提供を行っているわけではない。そこで、目的としていないものを除外したうえで、目的を達成したかしていないかで区分し、最大 5 つの目的のうち半数以上で目的を達していたか否かを問題とした。半数以上で目的を達成した場合は 1 をとり、そうでない場合は 0 を取る変数『目的達成過半数ダミー』を作成した。これにより、企業としての広義の法務体制によって、平均的に円滑なプロジェクト遂行が達成できているかを検証する。

被説明変数の第二は、組織としてプロジェクトに取り組む判断の際の不安材料である。ここでは相手方組織との信頼関係に関わる次の 3 つの「受領するデータの来歴管理がきちんとなされていない」「相手組織が信頼できない（契約が遵守されない可能性がある、十分なコミットメントを期待できない等）」「相手組織と十分な情報交換ができない」不安材料とした。これらそれぞれが不安材料であると回答した場合に 1 をとるダミー変数、『データ追跡可能性リスクダミー』『相手方信頼リスクダミー』『コミュニケーションリスクダミー』を作成した。これにより、企業としての広義の法務体制により、平均的にリスク認識が緩和されているかを検証する。

これらはいずれもダミー変数である。計量経済学的な推計では、その分布はロジスティック曲線に沿うと仮定して、ロジスティック回帰により行った。

説明変数

説明変数の第一は、雛形の整備状況である。『1. すでに契約書のひな型があり、それを使いこなしている』『2. すでに契約書のひな型はあるが、それを使いこなしていない』『3. 契約書のひな型を作成している途中である』『4. 契約書のひな型はない』の 4 段階それぞれをダミー変数として用いた。

説明変数の第二は、契約のカスタマイズ割合である。データ利活用を行うことを主としない提携プロジェクトの契約と比較し、契約書はかなりカスタマイズする必要がある、相当な法律業務が必要かについて、5 段階尺度でたずねた結果を用いた。

説明変数の第三は、知財・法務部門の関与である。『1. 事業の初期段階（事業計画や提案等の段階）から関与していることが多い』『2. 知的財産に関連する課題が生じたり契約交渉が必要になったりする段階から関与していることが多い』『3. 知財・法務部門（担当者）が関与することはあまりない』の 3 段階それぞれをダミー変数として用いた。ただし、事業の初期段階からとの回答は極端に少ない。

説明変数の第四は、直近 1 年間（2019 年度）のデータ取得・提供プロジェクトにおける限定提供データ割合である。『1. 全くない』『2. 一部のデータ』『3. 半分程度のデータ』『4. ほとんどすべてのデータ』の順序尺度指標であるが、分析の便宜のため、これを連続値として取り扱った。

制御変数

説明変数に影響を与えうる要因として、『過去の取得・提供プロジェクト数規模』『従業員数』、業種（『製造業ダミー』『非製造業ダミー』）、『グループ企業数』、『保有特許件数』をそれぞれ用いた。ただし、『グループ企業数』、『保有特許件数』については統計的に有意な影響が見られなかったため、ほとんどの分析では除外をした。これに加えて、分析対象の単位（企業×プロジェクト種別）が提供プロジェクトである場合に 1 を取るダミー変数『提供プロジェクトダミー』も用いた。

記述統計

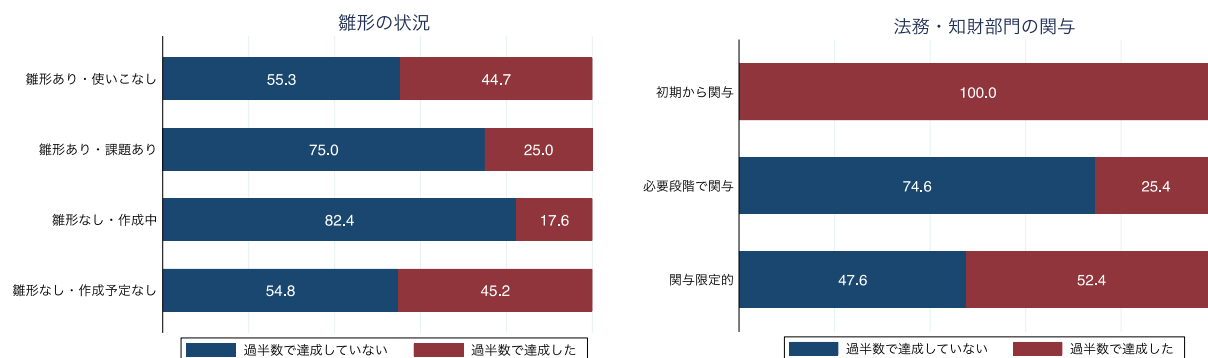
以上の変数の記述統計を図表 61 に示す。

変数	平均	標準偏差	最小	最大
目的達成過半数ダミー	0.32	0.47	0	1
データ追跡可能性リスクダミー	0.18	0.39	0	1
相手方信頼リスクダミー	0.31	0.47	0	1
コミュニケーションリスクダミー	0.19	0.40	0	1
従業員数（対数値）	6.53	2.37	2.40	12.21
製造業ダミー	0.39	0.49	0	1
非製造業ダミー	0.54	0.50	0	1
グループ企業数	23.25	34.08	0	144
保有特許件数	2.61	1.49	1	5
提供プロジェクトダミー	0.45	0.50	0	1
雛形				
雛形あり・使いこなしている	0.44	0.50	0	1
雛形あり・使いこなしていない	0.11	0.32	0	1
雛形なし・作成途中	0.16	0.37	0	1
雛形なし	0.29	0.46	0	1
契約のカスタマイズ割合（5段階）	3.44	1.07	1	5
知財・法務関与				
関与はあまりない	0.39	0.49	0	1
関与は必要時のみ	0.58	0.50	0	1
関与は初期から	0.04	0.19	0	1
限定提供データ割合（4段階）	2.49	1.07	1	4
過去の取得・提供プロジェクト数規模				
1～5 件	0.62	0.49	0	1
6～10 件	0.17	0.37	0	1
11～50 件	0.12	0.33	0	1
51～100 件	0.02	0.13	0	1
101 件以上	0.07	0.26	0	1

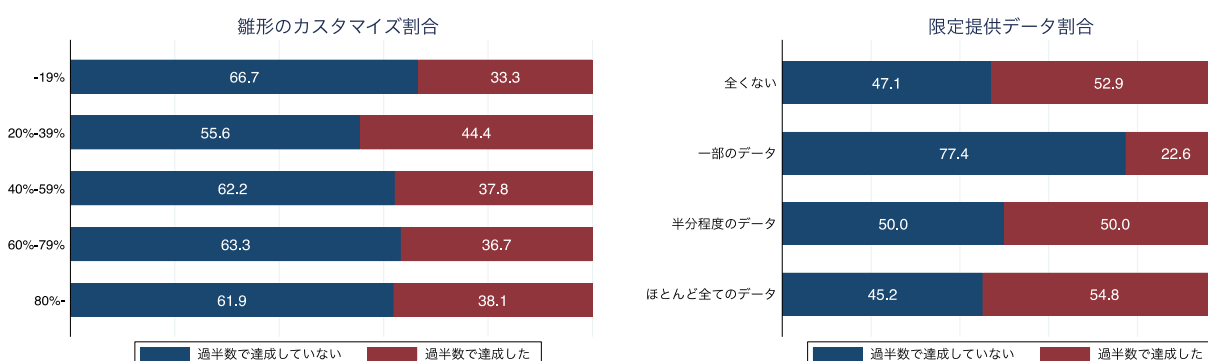
図表 61 記述統計

5. 6. 5. クロス集計結果（目標達成のみ）

目標達成の関係については、まずクロス集計により整理を行った。その結果を以下に示す。



図表 62 雛形の状況、知財・法務部門の関与と目標達成の関係



図表 63 契約のカスタマイズ割合、限定提供データ割合と目標達成の関係

5. 6. 6. 計量経済学的分析結果

ついで計量経済学的な分析により因果関係の推定を行った。

まず、目標達成との相関をロジスティック回帰により推計した結果を図表 64 に示す。Model 1 は他の節と同じ制御変数を用いた結果、Model 2、3 は、そのうち係数が統計的に有意ではなく、また、この分析の文脈では直接の関係が乏しいと考えられる一部の制御変数を除外した結果である。欠落変数の影響では後者の 2 つのモデルのほうが観測数が多い。

雛形の整備状況については、雛形があり、かつ、使いこなしている企業に比べ、雛形があるにもかかわらず使いこなしていない企業では、目標達成の確率が低い傾向があることがわかった（有意水準 10%）。ただし、頑強な結果ではない点に注意が必要である。加えて、雛形がなく、現在作成中である企業でも目標達成の確率が低い傾向があった（有意水準 5%）。他方、雛形がなく、かつ、作成予定がない企業では、雛形があり、かつ、使いこなしている企業との間で、目標達成の確率に統計的に有意な差が認められなかった。

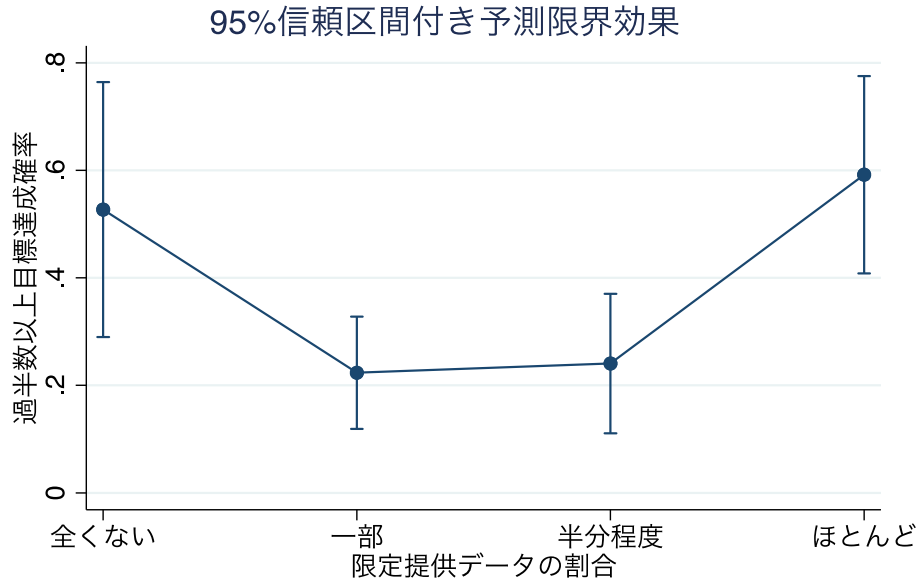
契約のカスタマイズ割合については、統計的に有意な相関が認められなかった。

知財・法務の関与については、そもそも初期から関与する事例（4 事例）全てで目標達成をしているため、これらについては分析から除外されている。その上で、知財・法務の関与があまりない企業に比べると、必要時に関与する企業の方が目標達成の確率が低い傾向があることがわかった（有意水準 1%）。ただし、この解釈については注意が必要である。注意点は考察において述べる。

限定提供データ割合については、2 乗項を用いたモデル（Model 3）で統計的に有意な下に凸（U 字型）の相関があることがわかった（有意水準 5%）。これは限定提供データの割合が少ない場合と多い場合に目標達成の確率が高い傾向があることを示している。限界効果を算定すると、限定提供データの利用が全くない場合とほぼ全てである場合に目標達成の確率が高かった（図表 65）。

	Model 1	Model 2	Model 3
従業員数（対数値）	0.198 (0.193)	-0.0123 (0.128)	-0.0490 (0.128)
製造業ダミー	1.437* (0.818)	0.581 (0.601)	1.056 (0.674)
非製造業ダミー	-0.660 (0.663)	-0.362 (0.603)	-0.610 (0.652)
グループ企業数	-0.0112 (0.0167)		
保有特許件数	-0.285 (0.327)		
提供プロジェクトダミー	-0.108 (0.541)	0.129 (0.509)	0.246 (0.535)
雛形（ベース＝使いこなしている） 雛形あり・使いこなしていない	-1.175 (1.138)	-1.878* (1.010)	-1.896* (1.036)
雛形なし・作成途中	-2.022** (0.932)	-2.101** (0.874)	-2.184** (0.904)
雛形なし	-0.306 (0.620)	-0.525 (0.607)	-0.171 (0.655)
契約のカスタマイズ割合（5段階）	0.103 (0.271)	0.0302 (0.254)	-0.186 (0.282)
知財・法務関与（ベース＝あまりない） 関与は必要時のみ	-1.848*** (0.595)	-1.772*** (0.561)	-1.622*** (0.584)
初期から	※1	※1	※1
限定提供データ割合（4段階）	0.290 (0.275)	0.398 (0.265)	-4.632** (1.827)
同2乗項			0.951*** (0.344)
（定数項）	-0.892 (1.769)	-0.448 (1.585)	5.840** (2.825)
過去の取得・提供プロジェクト数規模	あり	あり	あり
観測数	92	98	98
Pseudo R2	0.227	0.202	0.269

*** p<0.01, ** p<0.05, * p<0.1（括弧内は標準誤差）※1は該当する場合、目標過半数以上達成
 図表 64 過半数の目標達成との相関（ロジスティック回帰）



図表 65 限定提供データの割合の限界効果

次にリスクの認識との相関をロジスティック回帰により推計した結果を表に示す。

雛形の整備状況については、雛形を使いこなしている企業に比べ、雛形がありながら使いこなしていない企業では、データの追跡可能性について不安材料とを感じる傾向があることがわかった（有意水準 5%）。しかし、雛形がない場合には顕著な差は見られなかった。その他の 2 種の不安材料との間では統計的に有意な相関は見られなかった。

知財・法務の関与については、初期から関与している全ての企業で、データの追跡可能性、相手方とのコミュニケーションについては不安材料でないとの回答となっていたため、それらの企業は分析から除外されている。その上で、知財・法務の関与があまりない企業に比べると、必要時に関与する企業の方がコミュニケーションについての不安を感じない傾向があることがわかった（有意水準 5%）。

契約のカスタマイズ割合、限定提供データ割合については、いずれの不安材料との間にも統計的に有意な相関は見られなかった。

	データの 追跡可能性	相手方への 信頼	コミュニケー ション
従業員数（対数値）	0.568** (0.246)	0.224* (0.125)	0.710*** (0.230)
製造業ダミー	3.002** (1.232)	1.001* (0.605)	2.302*** (0.859)
非製造業ダミー	1.062 (1.140)	0.571 (0.605)	-0.307 (0.815)
提供プロジェクトダミー	-0.982 (0.864)	-0.0363 (0.507)	1.482* (0.762)
雛形（ベース＝使いこなしている）			
雛形あり・使いこなしていない	2.965** (1.173)	0.611 (0.764)	0.878 (1.044)
雛形なし・作成途中	0.539 (1.132)	-1.416 (0.936)	0.543 (1.247)
雛形なし	-0.962 (1.316)	-0.00214 (0.621)	-1.497 (0.999)
契約のカスタマイズ割合（5段階）	0.749 (0.464)	0.419 (0.256)	0.240 (0.335)
知財・法務関与（ベース＝あまりない）			
関与は必要時のみ	-0.469 (0.946)	-0.302 (0.591)	-2.283** (0.964)
初期から	※2	-0.403 (1.395)	※2
限定提供データ割合（4段階）	0.328 (0.384)	0.0128 (0.246)	0.451 (0.355)
（定数項）	-11.63*** (3.599)	-4.320*** (1.653)	-8.406*** (2.739)
過去の取得・提供プロジェクト数規模	あり	あり	あり
観測数	98	102	82
Pseudo R2	0.500	0.190	0.379

*** p<0.01, ** p<0.05, * p<0.1（括弧内は標準誤差）※2は該当する場合、リスクの発生なし
図表 66 リスクの認識との相関（ロジスティック回帰）

5. 6. 7. 考察

雛形の整備状況については、雛形があるものの使いこなしていない場合と、雛形を整備しようとしている場合に目標の達成がしにくい傾向があり、雛形があるものの使いこなしていない場合にデータの追跡可能性について不安を持つ傾向があった。これらの結果は、雛形を使いこなすことでデータ取得・提供に固有の不安材料を緩和させ、プロジェクトを円滑なものにしている可能性を示唆する。

一方、相手方への信頼や相手方とのコミュニケーションなどデータ取得・提供固有ではなく汎用的な不安材料との関連は見られなかった。これらについては、相手方の選択に強く影響を受けるものであり契約では緩和し難い、契約の条項よりは連携での相対の関係が重要であると受け止められており契約では不安は緩和し難いなど複数の解釈が考えられるものの、要因の特定は現段階では難

しい。

また、雛形がなく、かつ、作成を予定していない場合に、雛形を使いこなしている場合と統計的に有意な差が見られなかった点についても、複数の解釈がありうる。第一には、雛形に頼らず常にかスタマイズした契約をする方針であるため柔軟な関係性のマネジメントができている可能性が考えられる。第二には、そもそも不確実性の低いデータ取得・提供関係しかないため、雛形の用意の必要性も、成果獲得までの障害も少ない可能性が考えられる。これらについても特定の解釈に絞り込むことはできず、解明は今後の課題である。

知財・法務の関与については、解釈に注意が必要である。目標達成との関係では、知的財産に関連する課題が生じたり契約交渉が必要になったりする段階から関与していることが多いことと負の相関が見られた。これは知財・法務の必要時のみの関与が目標達成を妨げていることの現れであると解釈することは可能である。例えば、必要時のみの関与では相手方やデータ取得・提供を主体的に行う事業部門との信頼関係ができず、むしろ適切でない契約を推し進めることになり、プロジェクトに支障をきたす、という道筋がありうる。しかし、このような負の因果のみでは解釈として不十分である。逆の因果、すなわち、目標達成が困難なプロジェクトを行う企業で、相対的に知財・法務が関与する傾向があることも含んでいると解釈するべきであろう。

一方で、コミュニケーションについては少なくとも知財・法務の関与が正の寄与をしており、これについては知財・法務の関与によってコミュニケーション上のボトルネックが解消される傾向があることを示唆している。例えば、定期的なモニタリングの仕組みなどを知財・法務は提言する傾向があるのかもしれない。

限定提供データ割合については、目標達成との関係が U 字の関係にあった。これは、限定提供データの割合が全くない企業の中に、個人情報を含むデータを提供・取得しておらず、目標達成が相対的な容易な案件が多い企業が含まれていることの影響である可能性が考えられるが、詳細な検討は今後の研究に委ねられる。少なくともほとんどすべてが限定提供データである企業については制度を機敏に使いこなしたことによる便益を受けている可能性がある。

5. 6. 8. 分析の限界

他方で、本分析は以下の限界が残された初期的なものにとどまることは注記せざるを得ない。

第一に、データ取得・提供プロジェクトの性質の違いは十分に統制できていない。アンケートで把握した相手方の属性、所在国など大まかな傾向を制御変数として用いた分析を併せて行ったものの（本ペーパーでは未報告）、統計的に有意な相関が見られず、本分析では除外した。より精緻に行うには、プロジェクトごとの目標の達成度等を把握するべきであろう。

第二に、データ取得・提供プロジェクトにおける法的責任の所在部門の多様性を加味できていない。アンケートではデータ活用部門、データ処理部門、知財・法務部門にそれぞれ分かれており、これらに応じて知財・法務機能との関わり方が変わってくることが想定されたが、データの限界からこれらを織り込むことができなかった。

第三に、データ数の制約から統計的な頑強性は十分に担保されていない。これはデータ取得・提供プロジェクトを実施する企業そのものが少ないことが要因である。

第四に、先行研究²⁵で指摘されている契約の内容、とくに責任の分配には立ち入っていない。これは企業単位での傾向を把握したためであり、プロジェクト単位での分析を行う場合には考慮されるべき要素である。

5. 6. 9. 小括

²⁵ Ryall, M. D., & Sampson, R. C. (2009). Formal contracts in the presence of relational enforcement mechanisms: Evidence from technology development projects. *Management Science*, 55(6), 906-925.

上記のとおり、本分析の結果は様々な限界を伴う初期的なものにとどまるものの、次の傾向が示唆された。これらの頑強性の検証は今後の研究の対象となるべきものとする。

- 雛形の使いこなしは、プロジェクトの目標達成、および、データの信頼性確保を促す効果を持っている可能性がある。経済産業省「AI・データの利用に関する契約ガイドライン」の一層の普及とともに、これを使いこなすための知財・法務人材の能力向上の機会、知識の共有の機会が望ましい可能性がある。
- 限定提供データ制度の使いこなしは、プロジェクトの目標達成を促す効果を有している可能性がある。同制度の活用についてのベストプラクティスの普及がデータの利活用推進に有効である可能性がある。

5. 7. 産業別 Data Analytics Competency についての分析

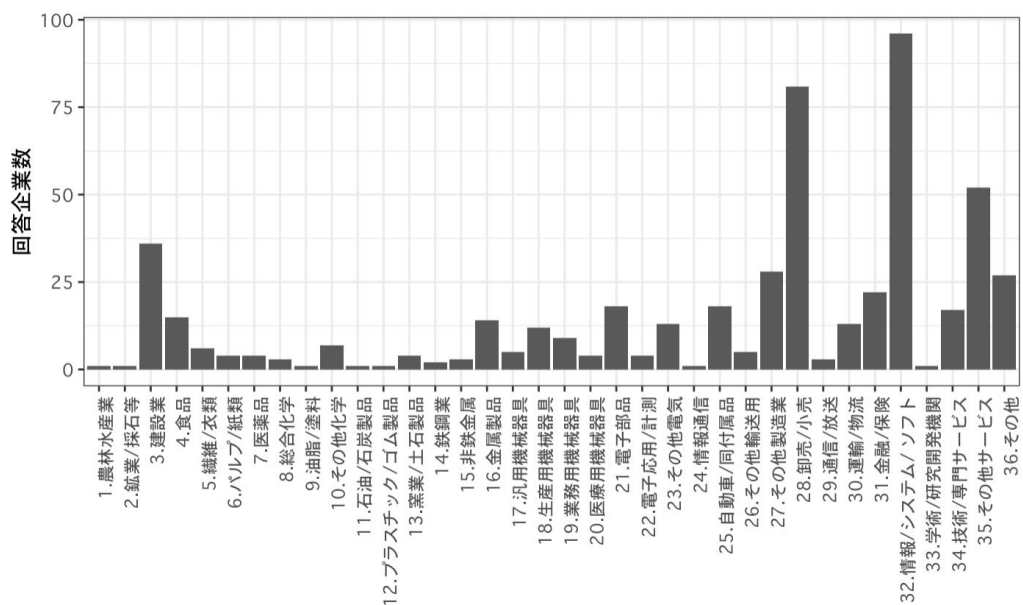
5. 7. 1. 本分析の視座

企業がデータを利活用する上でその産業がどのような状況にあるのか、というのは能力構築プロセスに大きく影響すると考えるのが自然である。例えば、対象の産業は、すでにデータを利用するビジネスモデルが当然のように確立している産業であるとか、データ資源を活用するための人材が十分に多く存在しているなどは、その産業に所属する企業のデータ利活用能力の構築に影響を与えると考えるのが素直である。このような考えに基づき、以降の分析では、所属産業毎に分析し、Data Analytics Competency(DAC)が事業成果に与える影響がどのように異なるのかを比較する。

など、本項の DAC 指標は、先述の「5. 4. 回帰分析②：Data Analytics Competency についての分析」で用いた DAC 指標を使用した。DAC は 5 つの設問の合成変数であり、1 から 25 の値を取る。値が大きいほど、企業は高い DAC を保持していると考えられる。

5. 7. 2. 産業毎の DAC の平均値の比較

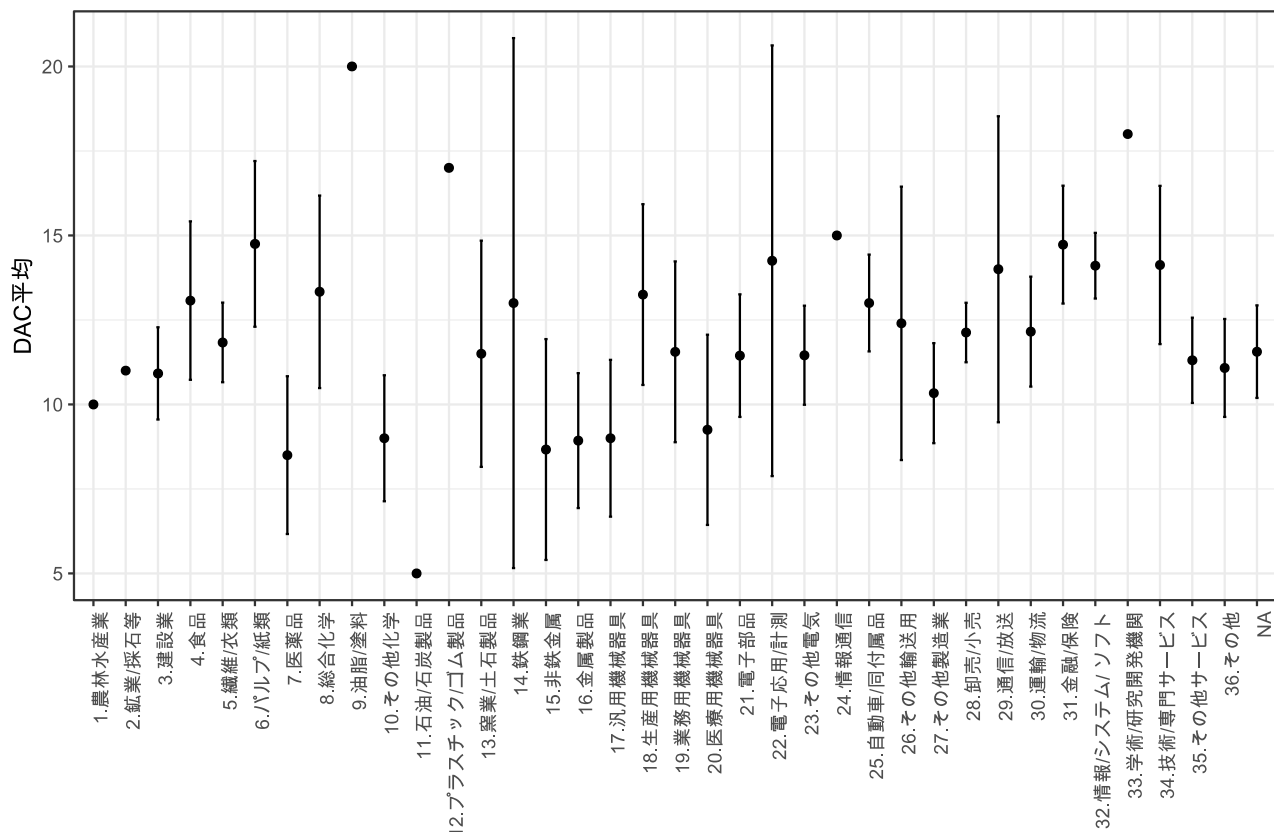
アンケート回答者の所属産業について、無回答であった 28 件を除き、有効回答であった 532 件を集計すると図表 67 となった。



図表 67 回答者の所属産業の集計

回答者の業種属性をみると、「32.情報/システム/ソフト」や「28.卸売/小売」が多いことがわかる。逆に、「1.農林水産業」や「2.鉱業/採石等」はすくないことがわかる。

産業毎に DAC に違いがあるのかを確かめるために、産業別の DAC の平均値と 95%信頼区間をしめす図表を作成した。なお、回答数が少ないため、95%信頼区間が算出できない業種については、平均値のみを表示した（図表 68）。



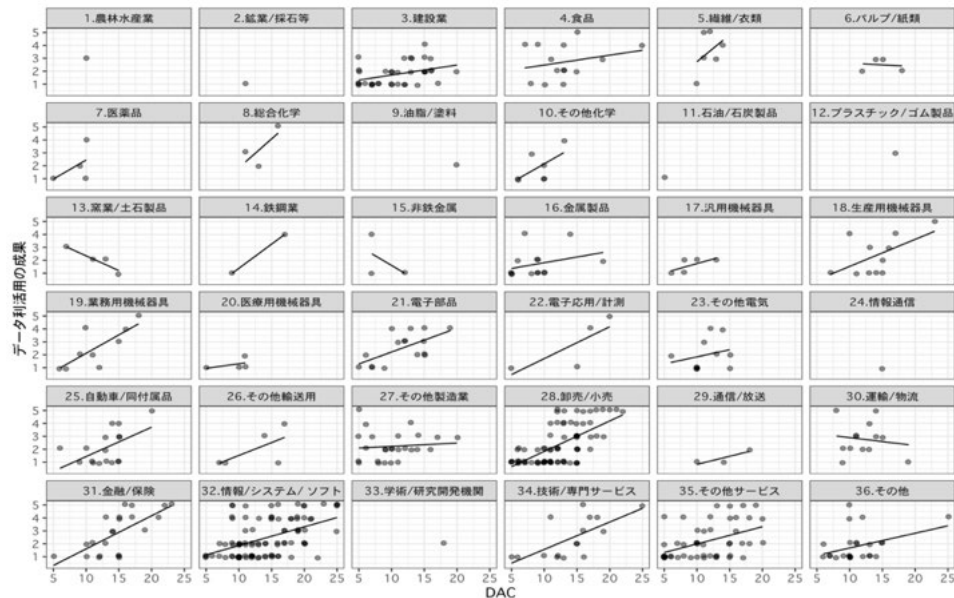
エラーバーは95%信頼区間を示す

図表 68 産業ごとの DAC の平均値と 95%信頼区間

図表 68 でみると、産業毎に DAC の平均値はばらついている。「7 医薬品」「15 非鉄金属」「16 金属製品」「17 汎用機械器具」「20 医療用機械器具」は DAC がやや低い。一方、「6 パルプ/紙類」「31 金融/保険」「32 情報/システム/ソフト」「34 技術/専門サービス」などは DAC がやや高い。

5. 7. 3. 産業毎の DAC とデータ利活用成果の関係

続いて、産業毎に DAC のデータ利活用成果に対する貢献がどのように違うのかを把握するため、まず、DAC と、データ利活用の成果(問 2-6)についての散布図を作成した(図表 69)。

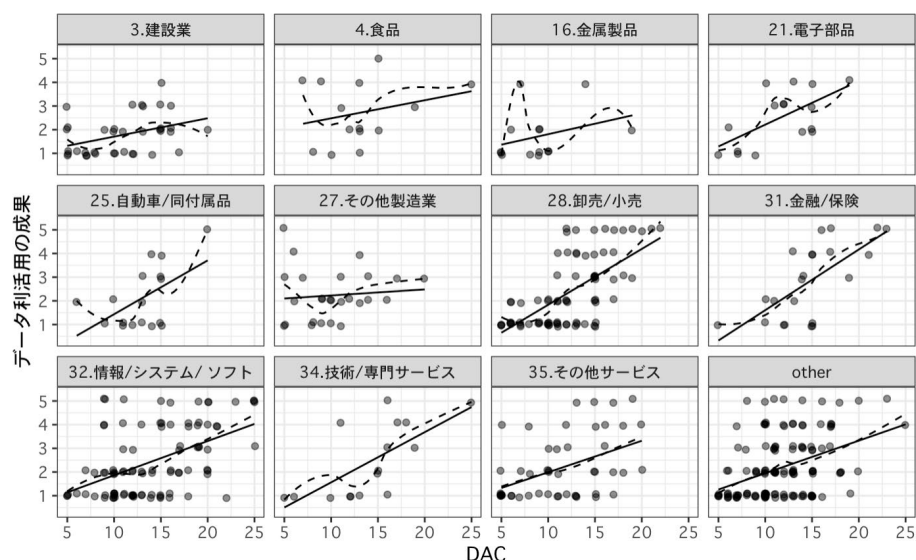


図表 69 業種ごとの DAC とデータ利活用成果の関係

DAC とデータ利活用の成果の関係を見やすくするため、回帰直線を付記した。ただし、所属する企業が少ない産業は回帰直線のパラメータを計算することができないため、回帰直線を省略した。また、DAC もデータ利活用成果も離散データであるため、散布図にすると、同じ座標に複数の回答がプロットされてしまい、回答件数の分布がわかりにくくなってしまう。そこで見やすさを優先し、プロットに対して小さいジッター（ゆらぎ）を加えて、散布図を作成した。

業種毎に回帰直線を確認すると、多くの産業で DAC とデータ利活用の成果はプラスの関係であった。すなわち、DAC が大きくなれば、データ利活用の成果が向上する傾向が見られた。ただし、DAC のデータ利活用成果に対する効果の大きさは、業種ごとに異なることも確認できる。また、産業によっては、回答者が少ないことに起因するばらつきによって、回帰直線が不自然に大きな傾きを描いているものもあった。

そこで、業種ごとに集計した際に、回答者が少ないことに起因するばらつきを抑制するため、回答者企業が 10 件以上の業種のみを選択し、分析した。これに加えて、回答者が 10 件未満の業種についてはその他(other)として、1 つのカテゴリーに集約して、図表 70 を作成した。



図表 70 業種ごとの DAC とデータ利活用成果の関係

DAC の効果の線形性を確認するため、回帰直線（実線）とともに、Loess 平滑化曲線（破線）も付記した。Loess 平滑化曲線は、移動平均線の一種であり、あるデータの周辺データによって重み付けした移動平均線で、DAC とデータ利活用成果の複雑な関係を示すことができる。回帰直線は線形の関係のみを表現するが、移動平均線はより自由度が高く、曲線の関係を表現できる。

もしも回帰直線と平滑化曲線との間に大きな乖離がある場合は、DAC とデータ利活用の関係は、回帰直線が表すような線形な関係ではなく、より複雑な曲線の関係であるといえる。DAC のような組織能力を表す指標の場合、成果との関係が複雑な関係になることはよくあることである。逆に両者に大きな乖離がなければ、単純な線形の関係で両者の関係を代表しても良いことになる。

このことを念頭に図を確かめると、「16.金属製品製造業」などの回答企業が少ない業種で、両者にやや乖離がみられるものの、他の業種では両者の乖離は大きくないことがわかる。全体にみれば乖離の大きかった業種はすくない。このことから、回帰直線（すなわち線形モデル）を用いて DAC とその成果の関係を詳細に検討する。

産業毎に、(1)式に示すモデル式を用いて回帰モデルを推定して、DAC のデータ利活用成果に対する効果を推定した。(1)式では、統制変数として上場有無、資本金、売上高、従業員数が説明変数に加えられている。各変数の説明は図表 71 の通りである。

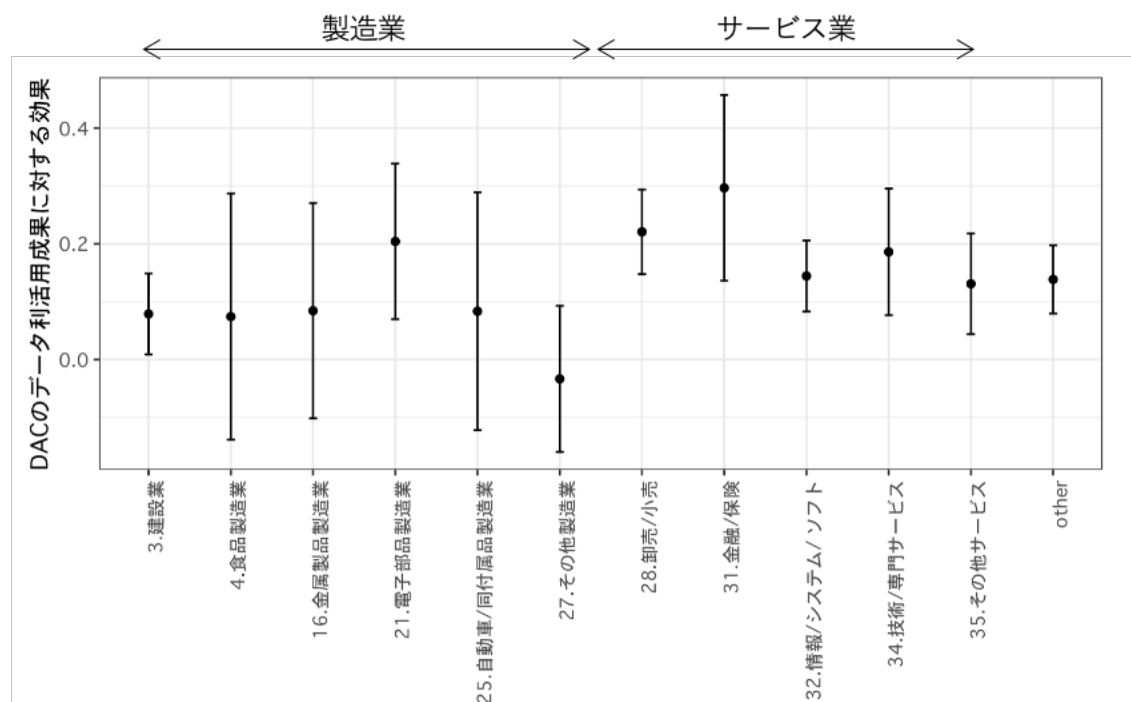
$$\text{データ利活用成果} = \beta_0 + \beta_1 \text{DAC} + \beta_2 \text{上場有無} + \beta_3 \text{資本金} + \beta_4 \text{売上高} + \beta_5 \text{従業員数} \dots(1)$$

被説明変数	
データ利活用成果	設問項目Q2-6の値。「1=まだ成果は得られていない, 2=具体的な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な成果が得られている」。1～5の値を取る。値が大きいほどデータ利活用の成果が大きい。
説明変数	
DAC	Data Analytics Capability。図表6に示す5つの設問項目（各設問は1～5の値）の合計点。1～25の値をとる。大きいほどデータ利活用能力が大きい
上場有無	上場企業=1, 非上場企業=0の値をとるダミー変数
資本金	資本金額（単位：百万円）
売上高	売上高金額(単位：百万円)
従業員数	従業員数（単位：人）

図表 71 各変数の説明

その他(other)を含め、12 個の業種毎にデータを分割し、それぞれ回帰モデルを推定した。12 個の各回帰モデルの決定係数の平均は 0.364 であり、その範囲は 0.112 から 0.695 までであった。データに対する回帰モデルの適合は、決定係数の平均が 0.364 であることを勘案すると、産業毎にばらつきがあるものの、回帰モデルはデータにやや適合していると思われる。(1)式で推定したパラメータのうち、DAC のデータ利活用への効果を示すものが B1 である。

DAC のデータ利活用成果への効果を産業毎に比較した。(1)式で推定したパラメータのうち、DAC のデータ利活用への効果を示すものが B1 である。DAC の効果と 95%信頼区間を産業毎に表示したものが図表 72 である。



図表 72 業種ごとの DAC とデータ利活用成果の関係

「4.食品製造業」や「25.自動車/同付属品製造業」などは、95%信頼区間が 0 を含んでいるため、その効果が統計的に有意ではないと解釈することができる。一方、「28.卸売/小売」や「31.金融/保険」や「32.情報/システム/ソフト」などは、DAC の効果が統計的に有意であるといえる。

左側に配置されている「3.建設業」から「27.その他製造業」までは、製造業に分別される業界である。右側に配置されている「28.卸売/小売」から「35.その他サービス」までは、サービス業に区分される業界である。この製造業とサービス業との対比でみると、製造業に区分される業界は、DAC の効果が統計的に有意であるといえない業界が多く含まれる。今回のデータからは「3.建設業」と「21.電子部品製造業」は、DAC の効果が統計的に有意であったが、それ以外の製造業は統計的に有意でなかった。

これとは対照的に、サービス業に含まれる業界は、DAC の効果はすべて統計的に有意であった。つまり、サービス業では、DAC を向上させることでデータ利活用成果の成果が大きくなる、というような道筋がみえている、と考えられる。このような業界による DAC の効果の違いは、業界ごと分析を行った結果、初めてわかったことである。

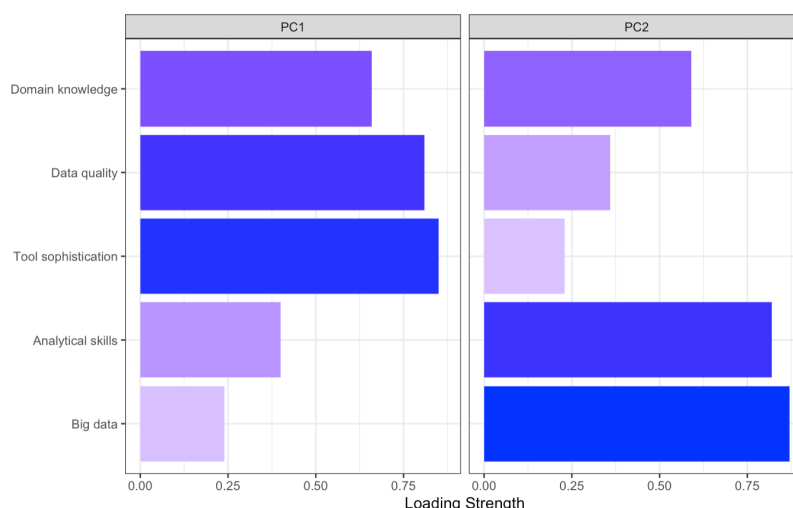
産業ごとの DAC のデータ利活用成果に対する効果の違いが、何によって引き起こされているのかについては、今後さらに精査が必要だと思われる。1つの可能性としては、製造業が主に扱っているデータと、サービス業が主に扱っているデータの種類の違いが考えられる。

製造業は主に産業由来のデータ、たとえば、製造設備のセンサーデータやサプライチェーンの在庫データなどを扱っていると考えられる。これらのデータは資源化が困難なのかもしれないし、データを資源化したとしても収益化が難しい、という状況があるのかもしれない。一方、サービス業

が扱うデータは個人データが多くふくまれる。こういった個人データに関しては、ネット EC などを中心にビジネス上の利用の仕方が既にあり、データ資源化の技術への投資が進んでいるとされている。近年、話題になっている GAF A の台頭もこのような個人データの利用が背景にある。これらを考え合わせると、一部企業で作られたビジネスモデルや技術成果がこういったサービス業の業界に波及しているのかもしれない。

5. 7. 4. DAC 能力と産業間の関係

DAC は 5 つの能力スコアの合成指標である。DAC と 5 つの能力の関係を可視化する目的で次元縮約をおこなった。次元縮約の手法としては主成分分析を用い、5 つの指標の合成変数である DAC を 2 つの主成分に分解した（図表 73）。

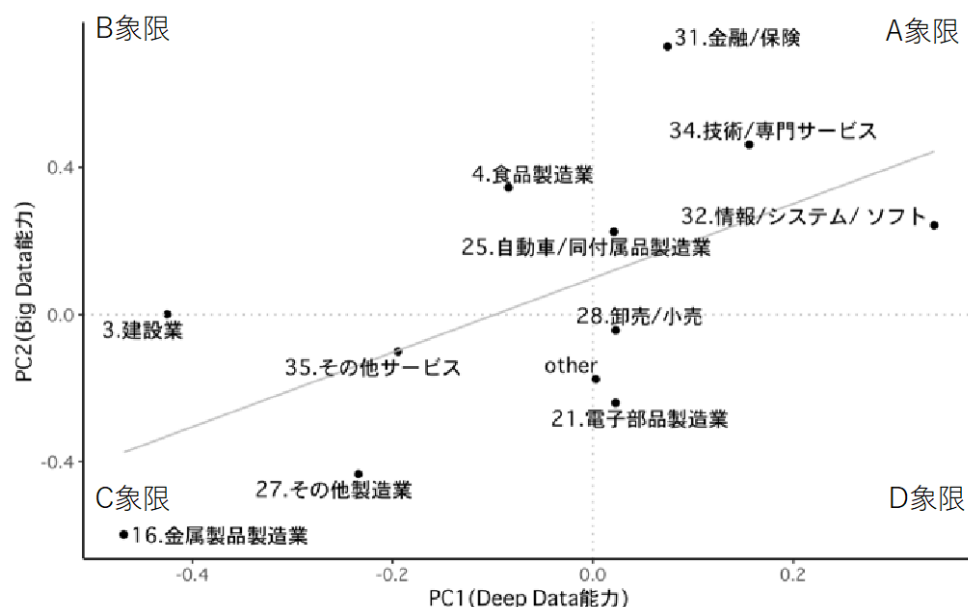


図表 73 DAC の主成分分析の結果

第 1 主成分（PC1）は主にデータの品質（Data quality）やツールの洗練（Tool sophistication）の負荷量が高い主成分である。第 2 主成分（PC2）はデータ分析スキル（Analytical Skills）やビッグデータ蓄積（Big data）で大きい主成分である。ドメイン知識（Domain knowledge）は第 1 主成分、第 2 主成分ともに同程度の負荷量であった。なお、第 1 主成分と第 2 主成分の分散寄与率はそれぞれ 41% と 39% であり、合計すると全体の 80% の分散を説明していた。

第 1 主成分はデータの品質やツール洗練などで代表される成分であり、いわゆる Deep Data を扱う能力であると考えられる。一方、第 2 主成分は、ビックデータの蓄積や分析スキルであり、大規模データである Big Data を扱う能力であると考えられる。

このようにして得られた主成分次元をつかって、各企業の第 1 主成分と第 2 主成分の主成分得点を算出した。企業毎の主成分得点をつかって、所属産業毎に平均を算出し、産業を単位としてプロットしたものが図表 74 である。



図表 74 DAC の能力地図における各産業の関係

第1主成分（Deep Data 能力）と第2主成分（Big Data 能力）で俯瞰したときの、各業界の平均値の相関関係図である。なお、第1主成分軸、第2主成分軸ともに0は、すべての産業の平均値となっている点に留意が必要である。さらに、全体の傾向を把握しやすくするために産業毎のプロットに対する回帰直線を実線で付記した。

この図を確認すると、A 象限は第1主成分得点（Deep Data 能力）と第2主成分得点（Big Data 能力）で高い得点を得ている産業が位置している。この象限には、「32.情報/システム/ソフト」や「34.技術/専門サービス」といった業界が位置取りしている。「31.金融/保険」は Deep Data 能力の得点も低くないが、特に Big Data 能力の得点が高い。

「4.食品製造業」は B 象限に属し、第2主成分得点（Big Data 能力）の観点からは、「32.情報/システム/ソフト」と同じような得点であるが、第1主成分（Deep Data 能力）の観点からは劣後しているように見える。「28.卸売/小売」は C 象限に属しているが、第1主成分、第2主成分ともに平均的な位置にあると言える。「3.建設業」は、第2主成分（Big Data 的能力）は平均的であるものの、第1主成分（Deep Data 能力）では他の業界よりも劣後していると思われる。

C 象限は、第1主成分（Deep Data 能力）も第2主成分（Big Data 的能力）も平均からは劣後するエリアである。「27.その他製造業」「16.金属製品製造業」は C 象限に属している。

図表 13 のように各産業を位置づけることによって、データ利活用能力の観点から、どのような産業がどのような産業と親しい関係と見做せるかを把握することができる。たとえば「31.金融/保険」「32.情報/システム/ソフト」や「34.技術/専門サービス」といった産業は近い関係にある。これらの産業は、業態の面からは似ているとは言えないが、データ利活用能力の構築の面からはお互いに学び合えるものがあるのかもしれない。また、産業政策の観点からは、とくに、データ利活用能力が劣後している産業を把握できるとともに、どのような意味でデータ利活用能力が不足しているのか（Deep Data 能力なのか、Big Data 能力なのか）というより踏み込んだ理解が可能になる。これらの理解をもとに、似たような産業の成功例を参考にしながら、より実効性のある産業政策の立案が可能になるとと思われる。

5. 7. 5. 分析の限界

本分析では DAC の概念を用いながら、産業毎に DAC の平均水準の違いがあるのかや、DAC のデータ利活用成果への貢献に違いがあるのか、といった点を検討した。産業毎に回答企業のばらつきが大きいため、すべての産業について分析することはできていない。また、回答企業の多少を補

正するというような処理も行っていない。そのため、本研究の推定結果はバイアスを含んでおり、あくまで今回得られたサンプルをもとに得られた結論ということになる。本研究の結果を産業政策に活かすときには、これらのバイアスについて考慮しながら慎重に事象の理解をすることが必要となる。

ただし、既存研究では、DAC の様々な性質や、DAC の事業成果への貢献について、業界ごとに比較したものは少ない。これは DAC が一般的な企業能力であり、まわりの産業環境の影響を受けないと暗黙のうちに仮定されてきたためであると思われる。しかしながら、企業の能力構築プロセスを考えれば、産業毎に企業が獲得する DAC が異なると考えるのが自然であると考えられる。本分析はそのよい出発点を与えていると思われる。

将来の分析では、とくに、産業毎の回答企業数のばらつきの影響を補正した分析が必要であると思われる。

5. 7. 6. 小括

本分析は産業毎に、DAC のデータ利活用成果への効果を比較した。その結果、以下の3つのことが明らかになった。

- ① 産業毎に DAC の水準が異なり、DAC が高い業界（「31 金融/保険」「32 情報/システム/ソフト」など）がある一方、DAC の低い業界（「7 医薬品」「15 非鉄金属」「16 金属製品」「17 汎用機械器具」「20 医療用機械器具」など）が存在する。
- ② DAC のデータ利活用成果への効果は、多くの業界でプラスであると推定された。そのプラスの効果は、サービス業では統計的有意であり、強くロバストである。一方、製造業では統計的有意ではなく、そのプラスの効果は、ロバストではない。
- ③ DAC を2つの能力（Deep Data 的能力と Big Data 的能力）にわけて、どのような産業がどのような産業と親しい関係にあるのかを把握した。今回の推計に基づけば、たとえば「31.金融/保険」「32.情報/システム/ソフト」や「34.技術/専門サービス」といった産業は近い関係にある。

これらの知見はいまだ初歩的な分析に基づくものであり、より深く精査していく必要がある。また、分析の限界で指摘したような改善が必要になる。

産業毎に DAC の違いを報告した分析は少なく、本分析がこの分野の良い出発点になれば幸いである。

■調査編

6. 調査概要

IoT や AI 等の技術革新によってデータが爆発的に増加するに伴い、事業者間の垣根を超えたデータ連携により、新たな付加価値の創出や社会課題の解決が期待されており、他方で、データや AI 技術を巡っては、契約実務の蓄積が乏しいこと、あるいは当事者間の認識・理解のギャップがあること等により、契約の締結が進まないという課題があった。

そこで、経済産業省では、①データの利用等に関する契約、及び②AI 技術を利用するソフトウェアの開発・利用に関する契約の主な課題や論点、契約条項例、条項作成時の考慮要素等を整理した「AI・データの利用に関する契約ガイドライン」（以下、「AI データ契約ガイドライン」）を策定した²⁶（2018 年 6 月発行）。AI データ契約ガイドラインは、国内においては、データ利活用や AI 技術開発に関する契約作成の手引きとして、広く参照されている。

他方で、日本は他の先進諸国に比べてデータ資源に乏しいため、活用できるデータ量に限界がある。したがって、国内のデータ利活用のみならず、国際的な連携によって国外のデータを補充していくことが、今後の技術開発上重要になってくるものと考えられる。しかしながら、今回の質問票調査の結果を踏まえても国際的なデータ取引を実施している企業はまだ少ない²⁷。一方で、国際的なデータの授受については、契約のカスタマイズや法的業務が「国内のデータを授受する場合よりも必要だと思う」と回答した企業が大多数であり、よりハードルが高くなっている状況がうかがえる²⁸。そこで、国際的な連携によるデータ利活用を促進するための施策の検討が期待される。

本プロジェクトでは、国際的なデータ利活用促進のため、AI データ契約ガイドラインの国外での適用の可能性を把握する目的で、2018 年から 2019 年にかけて、中国の大学、病院、企業と意見交換を実施した²⁹。その結果、中国においては、医療等の特定の分野を中心にデータの開放が政策的に進んでいること、一方で、データの取扱いについては、日本と同様に課題があり、特に組織対組織の契約において、日本の AI データ契約ガイドラインのようなガイドが有用であることが確認された。

日本国内でのデータの利用等に関する契約（以下、「データ契約」）を問題とする場合、当該データに適用される可能性のある個人情報保護や産業データの保護は同一であるが、国境を越えたデータ移転の場合、移転先で適用される法規制は移転元と同一のものとは直ちにいけない。そして、データ契約において、準拠法を日本と決めた場合でも、移転先の外国で強行法規とされるデータ保護規制の適用を免れるわけではないこと、データのパッケージが個人情報保護と非個人情報・産業

²⁶ 経済産業省ウェブサイトより

<https://www.meti.go.jp/press/2019/12/20191209001/20191209001.html>

²⁷ 質問票調査 問 4-6 問 5-6

²⁸ 質問票調査 問 4-13、問 5-13

²⁹ 2018 年 9 月には中南財經政法大学知財センター（中国湖北省武漢市）との意見交換を実施し、データ利活用を推進する施策としてのガイドラインに関心を持っていただいた。2019 年 9 月には鄭州大学（中国河南省鄭州市）工学学院教授、鄭州大学第一附属病院科研所所長等と意見交換を実施。中国においても組織対組織の契約において、利用権限の設定などに課題があり、日本の AI データ契約ガイドラインが参考になるとのこと。合わせて中国における医療データのオープン化状況を紹介いただいた。2019 年 10 月には、中国のインターネット関連企業の訪日の際に意見交換を実施。データの利用権限のほか、契約相手の信用度の問題についても議論。

²⁹ なお、コロナウィルス感染症の拡大により、予定していた海外調査が行えず、代わりに文献調査を実施した。

データの両方を含む場合にはより強い移転規制である個人情報保護規制を踏まえたデータ移転とならざるをえないこと、同一のデータであっても他のデータとの関連性によって性質が異なり得ること等の事情を考慮すると、データ契約で規律できない事項を確定し、データ契約で合意できる範囲を明らかにすることが、データ契約の内容を確定するにあたって不可欠の作業となるといえる。

そうすると、データの移転元または移転先となる国外法制の法制度について十分な検討が必要となるが、現時点で、各国で統一されたデータ、情報等の用語が用いられるものではなく、法令、条約や協定、決議の対象となるデータの範囲は、それぞれの文書の対象につき、用語法や概念、規制目的の際を踏まえて個別に確定していくほかない。

以下では、データの越境フローに関する視点と規制（７．）としてデータの越境移転の沿革や現在の状況を概説した上（７．１．）で、各国の関連データ保護制度（７．２．）及び企業のデータ利活用の実情にかかる調査の紹介（８．１．）やデータ越境移転にかかるモデル契約（８．２．）を概観する。

７．データの越境フローに関する視点と規制

７．１．データフリーフロー（DFF）とデータローカライゼーション

７．１．１．データフリーフローと国家間競争

データ³⁰は、現代社会の生産活動における不可欠な資源であり、資本となっている。2016年に採択された「G20 デジタル・エコノミー発展と協力イニシアティブ」³²は、現在のデジタル化社会においてデジタル化された情報や知識が生産の主たる要因であると認識し、G20におけるデジタル経済に関する政策議論が始まった。2019年6月のG20大阪会議の「G20 貿易・デジタル経済担当大臣の声明」は、日本のイニシアティブで“データの信頼された自由な流れ”を提唱した³³。国境を越えたデータの流れに関する国際ルールの設定が、グローバルなデジタル経済における核心的な問題となっていることは避けがたい趨勢である。デジタル化が進むことによる経済的恩恵を享受するためには、国際的な協力や各国のルールの調和を通じたオープン化という方向性は概ね各国共通

³⁰ データは、国際標準化機構（ISO）や国際電気標準会議（IEC）の「情報技術用語集」では、「representation of information in a formalized manner suitable for communication, interpretation or processing[<https://std.iec.ch/terms/terms.nsf/0/EC272F0FE87BE91CC12577BB002712C7?OpenDocument>]（通信、解釈、処理に適した形式化された情報の表現である）」とされ、データは情報の表現であると規定されている。そのコローラリーとして、情報の概念が先に存在し、一般的に、データは、ある種の符号化によって構成され、特定の装置または装置によって読み取ることができるものと理解される。データキャリアとしては、サーバ、デスクトップ、ラップトップ、メモリなどが挙げられるが、データが存在するためには特定のキャリアに接続されていることが前提として議論されている。

³¹ サイバー、デジタル、インターネット、データについては場面に応じて多様な用いられ方をする。それぞれに統一された定義があるとはみられない。例えば、米国では IoT という用語も使われているが、Cyber Physical System (CPS)という言葉で参照されていることも多い。定義は諸説あるが、IoT はあらゆるモノがインターネットにつながる世界のことであり、現実社会で、モノが集めたデータは、インターネットを通じてクラウドサーバーに送られる。サーバー側でデータ処理を行い、その結果に応じたインターネットを通じて現実社会のモノに指示を出すというようにインターネットを介した情報活用概念を指す。一方、CPS はセンサーデータなどリアルな情報を集めてサイバー空間で分析・解析し、機械や人・社会に反映させる概念であるとされる。

³² 000185874.pdf (mofa.go.jp)

³³ g20_2019_japan_digital_statement_jp.pdf (g20-digital.go.jp)

認識と言えそうだ。

データ資源がフローで国家間を移動する場合、デジタル経済における利益の拡大を求め、不可避免地にデータ資源獲得の競争が発生する。国境を越えたデータフローは、70年代初頭から欧米の間では競争の焦点となっており、90年代までは欧州がルールメイキングを主導し、2000年以降は米国が主導して米国版のデータの越境フロー規制を構築してきているが、この10年間で世界的なデータ越境フロー規制の新たな波が到来している。これまで50年ほどのデータ越境フロー規制において、最初の40年ほどは欧米諸国が個人のプライバシー保護に主眼を置いて規制の検討を実施していたのに対し、後半の10年は、新興国が自国の経済、産業の発展の観点からルール作りに加わる、またスノーデン事件の発生等で国家安全保障としての観点も加わり、問題が多様化した。この様にデータの越境フローに対するルール・規制は、自国民のプライバシー保護という権利保護という観点以外、国家安全保障、国家経済発展、国家競争力に関する重要な決定にあたるため、国富の奪い合いの問題としてはもちろん、グローバルのパワーゲームの様相を呈している³⁴。

7. 1. 2. データの越境フロー規制の具体的態様

データの越境フローに対する規制の特徴として、その具体的態様はデータ移転メカニズムに対応した形³⁵で行われることがある。データがインターネットを通じて移転する場合、インターネットプロトコル（「IP」）パケットが移転の単位であり、IPパケットが特定の領域に入るには、特定の物理的な装置が域内に存在する必要がある³⁶。クラウドサーバーであっても、国家は、物理的な装置を規制することでサイバー空間を規制することが可能となる。別のレベルでは、国家が階層型ネットワークを作成してから、その階層の最上位ルーター（ゲートウェイホスト）を制御することによってサブネットへのゲートウェイは地域でインターネットを調節することができる。その場合、ゲートウェイホストのコンポーネントは政府が管理することになる。ソフトウェア規制の点も、ルーターレベルおよびエンドユーザーレベルでの規制が考えられる。ルーターレベルでは、インターネット規制は通常ファイアウォール等によってデータの移動が制限される。また、国境を越えたデータフローの問題は、情報のトランスナショナルな流れとインターネットのグローバルな接続に関

³⁴ データを富として扱う国の代表は中国とインドである。中国国務院が2015年に発表した「ビッグデータ発展促進行動計画」では、「データは基本的な戦略的国家資源である」としている。2019年11月、中国共産党第19回中央委員会第4回全体会議では、「労働、資本、土地、知識、技術、管理、データなどの生産要素が貢献度に応じて分配に参加する仕組みを改善する」ことを提案した。2019年6月、Carnegie India シニアフェローである Vijay Gokhale は記者会見を開き、ドナルド・トランプ米大統領が G20 会合でデータローカライズを批判したことを受け "データは国富の新しい形（新しい形の富）であり、発展途上国はそれに注意を払う必要がある"、と述べている。参照：

“Data ‘new form of wealth’、take it into account of developing nations’ needs” *The Economic Times*, June 28, 2019, <https://economictimes.indiatimes.com/tech/internet/data-new-form-of-wealth-needs-to-take-into-account-developing-nations-needs-india/articleshow/69988888.cms>.

³⁵ データの提供、流通、利用については、三菱総合研究所「安心・安全なデータ流通・利活用に関する調査研究」（h29_02_houkoku.pdf (soumu.go.jp)）では、多様なメカニズムが提示されている。

³⁶ サイバー空間の捉え方について、レッシングのような法学者は、コモンズを他人の許可なしに誰でも利用できる資源と定義し、データの自由でオープンなアクセスを促進する。他方で、誰もがそのような資源を生産して保護するインセンティブを持たず、財産権保護のないデータの市場は、実質的な価値のないデータや法的に公開が義務付けられているデータへの自由なアクセスだけを許すことになる。土地、水などの公共資源について、所有権がないということにはなっておらず、財産権管理システムを用いて最適な資源配分を実現していることとパラレルに考えることができるかについて、データの性質に照らして検討する必要がある。

近時の各国の政策についてみると、情報それ自体の記録方式に着目して「データ」とするか、情報の表現方式に着目して「デジタル」とするか、ネットワークに着目して「インターネット」とするかで異なり、セキュリティの問題の場合はもっぱら「サイバー」が持ちいられる。しかし、論者や国によって統一的な用語法が定まっているとはいえない状態にある。

係しており、その規制の対象は、規範的な波及効果を生み出す国際的なインターネットの中核的な要素であり、本質的には国際的な相互作用と競争の問題である国内と国際の交差点にある極めて特殊な国際的なインターネットガバナンスの問題として捉え直す見方もある³⁷。

狭義でのデータローカライゼーションとは、データの保管と処理をデータの発信国内のデータセンターやサーバに配置することを要求することを意味する。国内のデータセンターやサーバに配置することを要求することを意味する。実際には、越境データフローの制限は、深刻度の異なる様々な種類の制限を網羅しており、より包括的なものとなっている。

7. 1. 3. データの越境フロー規制をめぐる各国の諸状況

各国の情報技術の強弱という国家権力構造の不均一性、富であるデータが強国に流れがちであるという状況において、その国の ICT 技術の発展、産業規模、セキュリティニーズという主要な各要因の影響の大きさは異なり、その国がデータの越境フローに対していかなる戦略を採用することが最適かは、各国ごとに異なる。データにはその国の国民や国家に関する情報が含まれており、莫大なデータを保有する国家は、その規模により、強力な情報能力を持っているため、データの自由流通のもと富を生み出す仕組みを有するものの、国外のコントローラにデータが流れてしまうというセキュリティリスクがある。データの越境フロー規制の目的は、①個人データの保護、②国家安全保障、③国内法執行、④産業界の保護等が理由とされる。

例えば、世界のインターネット帯域幅のシェアに関する統計と予測によると、2006 年には北米が全世界の帯域幅の 97% を占め、2018 年には 85% を占め、2024 年までにも 80% を占めると予測されており、大陸間の世界的な帯域幅容量に関しては、北米と西ヨーロッパが他の大陸との間で相対的に絶対的なリードを取っていると見込まれる。そうすると、現在及び将来にわたって、データの越境フロー促進について、優位なポジションにある国々は、データのフリーフローを促進した方が自国の利益になり、他方で、立場の弱い国は、自国の利益のためには自国内でのデータの越境フロー規制を強める傾向となるといえる。

データのフリーフローを奨励している国とは対照的に、データローカライゼーションを含めた越境フロー規制を選択している国は、現実の国益から抽象的な概念的価値観、社会福祉全体から個人の市民権まで、より複雑な目標を持っていることが多いとみられる。具体的には、国家安全保障、人権保護、産業保護、技術開発、国際貿易、外国投資、サイバー主権などである。急速に発展している新興国は、サイバーリスクが相対的に高く、国内 IT 産業の競争力が相対的に弱い状況にあり、開発とセキュリティのセーフガードを優先することが最も重要な目標である。この優先的目標を達成するため、データローカライゼーションを推進する過程で、多国籍企業、外国政府、国際機関の強い反対にもかかわらず、データローカライゼーション政策を堅持してきた中国、ロシア、インドのように、外部からの抵抗にもかかわらず、各国政府がその政策を遂行することが自国の利益になると考えることになる。これらの状況を踏まえ、新興国・発展途上国は、国家的な戦略的データローカライゼーションの主張をする状況が生まれている。中国、ロシア、インドネシア、ナイジェリア、ベトナムなど、データローカライゼーションに関する厳しい規則を導入している³⁸。

世界の主要 60 カ国の分布図を作成し、国境を越えたデータフローに関するその国の位置と、その国の ICT の発展度やサイバーセキュリティの脅威との関係を調査した報告では³⁹、自国の技術

³⁷ インターネットの基本的な側面は情報の交換であり、サイバースペースのガバナンスは究極的にはデータのガバナンスである。国境を越えたデータフローの規制は、国際貿易を指向しているが、本質的にはインターネットガバナンスの話題であり、一国のインターネットガバナンスの基本原則と論理に沿ったものである。

³⁸ Arindrajit Basu, Elonnai Hickok, and Aditya Singh Chawla, “The Localisation Gambit”, Centre for Internet & Society Report 49 頁を参照。

³⁹ 劉金和、崔宝国「データローカライゼーションとデータ防御主義の合理性と動向（数据本地化和数据防御主义的合理性与趋势）」『国際展望』2020 年 6 期の図 1 「サイバーセキュリティ脅威度と ICT 発展指数によるデータ越境モビリティのスタンスの分布」

開発レベルが高く、産業が十分に発達しており、安全保障上の脅威や緊急性が低い場合には、国境を越えたデータ流出を促進する戦略を採用することになる。また、国内産業の規模が小さすぎたり、ICT（情報通信技術）の開発レベルが低すぎると、海外のインターネット技術やサービスへの依存度が高くなりがちで、国として、データローカライゼーションを主張する余裕がないということもあり得る。他方で、一定の ICT 産業規模を有し、かつ急速に発展している国では、産業や技術の発展、セキュリティの観点から、データローカライゼーションの選択を行い、自国のデータをコントロールし続けることで自国の安全を確保するというデータ防衛主義を構築することを志向するとされる。ここでいう、データ防衛主義は、自己防衛の要素はあるものの、強力な外部敵対者に直面した場合の防衛手段であり、貿易を制限することを目的としたものではない。

7. 1. 4. データローカライゼーションと国家間協調

データローカライゼーションについて、理論的な説明も試みられており、データとナショナリズムをナラティブに組み合わせたデータナショナリズム⁴⁰、貿易保護の新たな形態としてのデジタル保護主義及び欧米の覇権主義への抵抗としての反データ・コロニアリズム⁴¹など、異なる視点からの理論的説明を提案しながら、データローカライゼーションという複雑な現象の説明的枠組の構築が試みられている。経済のグローバル化の時代には、ディフェンシブ主義は貿易保護主義と批判されることが多く、国際的な反発を招き、制裁の対象にさえなっている。現在のデジタル貿易交渉において、先進国は国境を越えたデータフローの制限やローカライゼーション要件を重要なデジタル貿易障壁として挙げており、一方で途上国は「データナショナリズム」等をその正当化の理由として挙げていとされる。

データローカライゼーションの問題点も顕在化しつつある。個人データの保護にする問題点としては、データの国境を越えることを拒否することは、データを保護するための最高の設備を備えた人々の手にデータが届かないようにすることを意味することが多い。例えば、データのセキュリティにとって、データがローカルに保存されているかどうかよりも、グローバルなクラウドプロバイダーの技術的・財務的能力が重要であると指摘している。データセンターには、データを物理的に国内に保持するための信頼性の高いインフラが必要だが、多くの発展途上国ではまだ十分に確保できていない可能性がある。また、仮に確保できたとしても、発展途上国の小規模な国の市場状況によっては、Amazon Web Services、Microsoft Azure、Google Cloud などの大手クラウドプロバイダーがインドでこれまで積極的に行ってきたように、現地でのデータセンター構築に投資するインセンティブが十分に得られない可能性がある。例えば、マイクロソフトとアマゾンが南アフリカにアフリカ初のデータセンターを開設しているが、多くの発展途上国の金融サービス・プロバイダーは、データを管理・保護するための最も堅牢なクラウド・コンピューティング・ソリューションを見つけるために、外部からの協力が必要とされることもある。また、地域の経済競争力の向上に関する問題点としては、国内企業をデジタル競争相手から守ろうとデータの流れを制限すると、他国が報復するきっかけとなり、データを利用してグローバルにビジネスを展開しようとする他の地域企業に損害を与える可能性があるとされる。データローカライゼーション政策がもたらす貿易障壁の一例として、金融規制当局が導入した新しい IT システムのローカライゼーション規制のために、PayPal がトルコでのライセンスを失い、事業停止を余儀なくされたこと等が指摘される。

このような状況下において、EU は、データの越境フロー規制について、GDPR が及ばない非個

⁴⁰ 毛维准、刘一燊「データナショナリズム：推進する論理と政策影響（数据民族主义：驱动逻辑与政策影响）」『国際展望』2020年第3期。

⁴¹ The Costs of Connection: How Data Colonizes Human Life and Appropriates it for Capitalism (Stanford University Press, August 2019). <https://colonizedbydata.com/>
Nick Couldry : Data Colonialism: Rethinking Big Data's Relation to the Contemporary Subject
https://www.researchgate.net/publication/326694097_Data_Colonialism_Rethinking_Big_Data's_Relation_to_the_Contemporary_Subject

個人情報について、域内における非個人情報の自由流通を促進するための規則を 2018 年 11 月 28 日に公布している⁴²。適用対象については、規則 2 条は、①域内に居住し又は事業所を有する利用者に対しサービスとして提供されるもしくは②域内に居住し又は事業所を有する個人・法人が自らの必要性のために行う、GDPR が規定する個人データ以外の電子データの取扱いであり、EU 法の範囲外の活動には適用されない。なお、1 つのデータ集合に個人データと非個人データの両方が含まれるときには、非個人データの部分に適用され、両者が不可分に関連している場合には、GDPR の適用を妨げない、と規定している。非個人情報の自由流通に関して、4 条で、データローカライゼーション規制は、公共の安全の理由に基づき正当化されるものを除き、禁止される、また、EU 各国は、これに違反する法律、規則又は行政規定による現行のデータローカライゼーション規制を、既存の EU 法に基づくデータローカライゼーション規制を除いて、2021 年 5 月 30 日までに廃止しなければならないと規定している。

今後の動向は予断し得ないが、国境を越えたデータフローの核となるデータローカライゼーションは、短期的な現象ではなく、グローバル化の深化に伴い、ますます深刻な課題となっているとみられる⁴³。科学技術の国際競争が激化し、国家間の信頼関係が希薄になっている中で、データのローカライゼーションの傾向はさらに強まることが予想され、デジタル時代の課題に対応し、各国が将来の発展のための戦略的なスペースを残すためには、多様で柔軟性のあるアレンジメントがより求められる。この点について、各国の規制を踏まえた協調の必要性はデータに限ったことではなく、これまでも様々な分野や技術に共通しており、国際標準化機構（IOS: International Organization for Standardization）によって開発された何千もの標準規格等の成果は、今後の方向性のひとつを示唆するものである。また、データ保護へのアプローチが異なる国が、国境を越えたデータ転送を可能にする一連の原則に合意することを可能にするフレームワークの好例がいくつか紹介されている。データに関するこのような協力関係はまだ初期段階にあるとされるが、今後さらに進展していく可能性がある⁴⁴。

7. 1. 5. 対国家の枠組みを超えたデータ移転規制の進展の可能性

米中の技術覇権争いの激化でビジネスや安全保障でのデータの重要性が増す中、欧州からは「デジタル主権」が提唱されている⁴⁵。具体的な概念の外延に対しても様々な議論があるが、新しいデジタル主権は、「デジタル分野で何が起きているかを決定する国家の権利」であり、デジタル分野での国内や地政学上のナショナル・インタレストを自主的かつ独立して決定する国民国家の権利を意味するものと一般的に言われている。自主的な内外の情報政策を導入する権利や自己の情報資源

⁴² Regulation (EU) 2018/1807 of the European Parliament and of the Council of 14 November 2018 on a framework for the free flow of non-personal data in the European Union, OJ L303, 2018.11.28, pp.59-68. <<https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:32018R1807>>

⁴³ 中国において、経済学者ハーバート・A・サイモンの「限定合理性（Bounded Rationality）の理論」を参照して、データの局在化は望ましくないというのは合理的であるが、実用的な意味合いを持つ戦略的な選択であり、関係する国は依然として、一見好ましくない選択をし続ける、政策立案者が課題をパーツパーツで分解する傾向があることは示唆されており、それは、政策立案者は、遠大で不可逆的な決定をするのではなく、増分的または限界的な選択をすることができるように、実利的、経験主義的なアプローチをするとされる [David Braybrooke and Charles E. Lindblom. (New York: The Free Press of Glencoe, 1963.)] こととの研究も発表されている（劉金和、崔宝国「データローカライゼーション：一限定合理性によるデータ防衛主義」）。

⁴⁴ エストニアとルクセンブルクが署名した「データ大使館」や、バーレーンが提案したデータ外交放棄に基づく国際的なデータセンター建設のための法案など、世界各地で国境を越えたデータの流れに関する新しい試み等が実施されている。

⁴⁵

https://ecfr.eu/publication/europe_digital_sovereignty_rulemaker_superpower_age_us_china_rivalry/

を処分したり、国家情報空間のインフラを形成したりする権利も含まれる。さらに、国家の電子・情報の安全性を保障する権利も持つと考えられる。

ここで振り返っておきたいのは、サイバー空間における主権概念の検討は決して新しい議論ではないということだ。1997年、米国のネットワーク法学者ティモシー・ウーが、「サイバー主権？—インターネットと国際システム」⁴⁶においてサイバー主権の概念を発表して以来、この概念は世界的に変遷してきた。2013年に、国連は国連憲章における主権平等の原則はサイバー空間にも適用される立場を示し、2015年、中国の習近平主席は第二回世界インターネットフォーラムにおいて、サイバー主権⁴⁷の尊重をグローバルインターネットガバナンスの重要な価値観⁴⁸であると発言している。この様な中国の動きの後、欧州議会より、2020年7月に「ヨーロッパのためのデジタル主権 (Digital sovereignty for Europe)」⁴⁹が公表されている。デジタルまたはサイバー主権については、従前はスノーデン事件等も含め、国家機密の漏洩等の主権確保のために議論されていたが、近時は、米国や中国のIT企業への対抗のために主張され、同様に、データコモンズ（共通データ空間）概念も戦略的な意味で用いられている。

EUは、個人情報情報の流通につき、GDPRを制定し、非個人情報につき、域内の自由流通のためのフレームワークに関する規則を制定しているが、対域外において、主に産業データの分野でのルールセッターとしての地位を獲得するため、欧州クラウドの創設、データインフラ構想「GAIA-X」などのイニシアティブを包括する概念として、上記のデジタル主権という概念を提唱している。EUは、戦略的な分野（産業・製造、環境保護、輸送、医療、金融、エネルギーなど）に関して、域内で様々な組織からのデータを集約し、共通で互換性のある汎欧州データ空間、つまりデータプールを作成することで、ビッグデータ分析やAIアプリケーションを促進することを目的とする枠組みの構築を検討している。米国及び中国の大企業の手でデータが過度に集中することを避けつつ、人工知能などの新製品やサービスの開発に必要なデータをEU企業（特に中小企業）に提供することを目的としたコモンズを提案している。EUのモデルでは、データを提供している組織は、他の提供者のデータにアクセスできるだけでなく、データプールのアナリティクス、予測メンテナンス、ライセンス料などのサービスにアクセスすることが企画されている。

中国は、サイバー主権の国際的なコンセンサスを進めようとしており、2020年11月の世界インターネットフォーラムでは、「サイバー主権：理論と実践（2.0版）」を公表している⁵⁰。その他の国でもデータローカライゼーションを正当化する概念として依拠される可能性も高い。国家主権とデジタル／サイバー主権を重ねた研究も増えてきている⁵¹。欧州域内でも、ウェストファリア体制

⁴⁶ Tim Wu, *Cyberspace Sovereignty? – The Internet and the International System*, 10 Harv. J. L. & Tech. 647 (1997).

⁴⁷ 中国で、サイバー主権の概念が初めて打ち出されたのは、2010年の「中国インターネット事情（白書）」においてである。

⁴⁸ 本精神は、2016年制定のサイバーセキュリティ法にも引き継がれ、サイバー主権の保護が同法の立法趣旨とされた。その時点で、サイバー主権が何なのかについては明らかにされていなかった。

⁴⁹ EPRS_BRI(2020)651992_EN.pdf

⁵⁰

[http://www.wuzhenwic.org/download/SovereigntyInCyberspaceTheoryandPractice\(Version2.0\).pdf](http://www.wuzhenwic.org/download/SovereigntyInCyberspaceTheoryandPractice(Version2.0).pdf)

⁵¹ インターネットがサイバースペースへと発展していく中で、インターネットガバナンスの問題は、ボトムの技術プロトコルの管理から、トップのコンテンツ・アプリケーションの開発、経済社会行動の規制へと徐々に上昇していき、インターネットガバナンスの属性は、技術的ガバナンスから総合的な社会的ガバナンスへと徐々に進化していき、国家主権がサイバースペースに確立されていく。将来的には、サイバー主権の意味も、政治的な論理から具体的な商業的な論理へと変化していくことになるだろう。一般的に、サイバースペースにおける主権の問題に関する各国の主張と実践は、現時点での自国の核心的利益を反映している。したがって、ルール優勢の国に有利なグローバルなサイバースペースガバナンスシステムにおいて、明確な非中立の特徴を示している。サイバー主権、データ主権等の主張は、排他的絶対権によって外部からの侵入を免れると同時に、領土の境界線の限界に自らを限定するという、ある意味で防衛的な性質を持っている。これは、サイバー空間や国境を越えたデータの流れにおける防御的リアリズムの典型的なパソロジックであり、サイバー主権とデータ主権は、このデータ防御性を実現するためのロジックとなる。限られた合理的なデータ防御主義の論理の下では、相対的に競争上不利な立場にある国家は、防御的な国際的なインターネットガバナンス政策を採用する可能性が高

以降の主権国家における「主権概念」との連続性を前提に議論が構築されつつある⁵²。

EU によるデジタル主権の提唱は一過性で終わる可能性もあるが、主権概念を用いつつ、その対象は民間企業でテック企業を想定していること、域内の加盟国から主権を主張される危険をも踏まえて地域共同体としての主権を設定していること等に鑑みると、その主眼は、今後の産業発展の構造が AI やその前提となるビックデータの獲得に依存しており、現状のままでは、将来の競争関係が固定化するおそれがあるとの客観的な予測と現状認識にあるとみられる。既に欧州域内や欧州の国民のデータを米国や中国の大企業に取得されている現状において、デジタル主権は、米国や中国主導の AI 原則を妨げる方向性と相まって、今後の競争関係を決める重要な概念となる可能性も秘めている。

7. 2. 各国のデータ越境フロー規制と国際動向

7. 2. 1. マルチレベルの規制

グローバルなデータのガバナンス、データの越境移転に関し、まだ国際的ルールは存在せず、OECD、G20、WTO、APEC などでもこれまでルールメイキングに向け検討が進められてきた。

OECD は、プライバシーの尊重を基本的な価値として、また国境を越えた個人データの自由な流れを推奨する形で、1980 年に世界で初めてデータの国際流通についてのガイドラインを公表し、その後、グローバルな取組みをリードしてきている。

G20 でも、2016 年の杭州サミットにおいて、デジタル経済に関する政策議論を開始し、2019 年、大阪において、データフリー フローウィズトラスト及び越境データ流通の重要性が確認された。APEC は、APEC 越境プライバシールール(CBPR)システムにおいて、アジア太平洋地域のデータ流通を推奨し、WTO は、グローバルな貿易制度にかかる、デジタル製品への非関税措置の実施、データの自由流通を推奨することで情報技術製品の自由貿易の実現を目指してきた。その他、情報技術協定 (Information Technology Agreement) が情報技術製品の関税撤廃することでデータの越境流通促進を間接的に推奨してきた。

OECD プライバシーガイドラインは、その後、2013 年に改訂されており、その 6 条では、加盟国にプライバシーに関するミニマムな基準を定めるが、具体的で明確な基準が定められておらず、ガイドラインにとどまって法律的な効力はない。APEC の APEC 越境プライバシールール(CBPR)システムも、G20 で示されるフレームワークもいずれも強制的なものではない。WTO 下の GATT は加盟国を拘束するものの、各協定の法的効力について、例えば、GATT 2 条は最恵国待遇を規定するが、7 条では、加盟国の差別的扱いを許容するような規定もあるなど、統一的なルールを定めるものではない。

このように、マルチな取組みにおいて、データの越境移転の国際的統一ルールの制定には未だ至っておらず、各国のこの点の規制の状況も様々である⁵³。国境を越えたデータの流れについて、OECD が開発した Digital Services Trade Restrictiveness Index によると、デジタルサービスへの制限が少ない指数を示したのは、米国、スイス、オーストラリアなどで、逆に、中国、インド、ロシア、インドネシア、南アフリカ、ブラジルなど非 OECD 諸国は高い制限指数を示しており興味深い⁵⁴。

くなり、それは強いサイバー主権のスタンスに現れている。サイバー主権とそれに伴うデータ主権は、データローカライゼーションを説明し、言い換えれば、データローカライゼーションは、サイバー主権の主張によって実現されるともいえる。

⁵² https://doc-research.org/wp-content/uploads/2020/09/Sovereignty-in-a-digital-era____.pdf

⁵³ <https://www.unilu.ch/en/faculties/faculty-of-law/professorships/managing-director-internationalisation/research/the-governance-of-big-data-in-trade-agreements/>

⁵⁴ https://stats.oecd.org/Index.aspx?DataSetCode=STRI_DIGITAL

7. 2. 2. EU

EU の政策は、市民権保護の観点から個人情報の適切な保護の仕組みを積極的に推進することで、データの自由な流れに制限的な姿勢をとる一方、非個人データについては、原則として国境を越えたデータの流れを奨励している。EU 域内では加盟国に対して、公共の安全の理由に基づき正当化されるものを除き、ローカライゼーションの規制を、2021 年 5 月 30 日までに廃止することを義務付けている。個人情報に関するものは、非個人情報とのパッケージであれ、GDPR の適用を受ける。

一般データ保護規則（GDPR）は、個人データは、個人データを EEA 域外に移転することを原則禁止する。EU 域内から域外へ個人データを移転するには、①十分な個人データ保護を保障する国として認定されるか、②BCR（Binding Corporate Rules：拘束的企業準則）の締結、③SCC（Standard Contractual Clauses：標準契約条項）の締結、④明確な本人同意を得るという条件を満たさなければならない。日本も、2019 年 1 月①に認定されている。

2019 年 5 月、EU は、個人情報、非個人情報が混合するデータセットの処理方法に関するビジネス向けの実務ガイダンス⁵⁵を公表した。本ガイダンスは、制限的な個人情報の流れと自由な非個人情報の自由な流れを調整することにあり、特に以下の点に焦点を当てている。

- 個人データと非個人データの概念と、いわゆる「混合データセット」におけるそれらの組み合わせ
- データの自由な移動の原則とデータのローカライゼーション要件の禁止
- データのポータビリティ

なお、この様に、EU 域内においては、この様にデータの自由な流れを推奨するような動きがあるものの、これが EU 域外に適用されるわけではない。

域外へのデータ移転に関し、非個人情報に限定された情報の移転についても、前述のとおり、それらを包含するものとして「デジタル主権」概念を提唱し、域外のデジタル企業への対抗のため、域外への移転に制限を設ける措置が設けられ、域内でのデータ管理を推奨する形で方向性がシフトしてきている。域外のデジタル企業への対抗に関しては、EU 対内直接投資審査規則を 2019 年 4 月に発効し、その中で、加盟国に対して、外国投資の審査時に検討すべき要素を例示する。審査対象とすべき分野・業種としては、重要インフラの他、重要技術や個人情報を含む機微な情報へのアクセスなどが列挙されている（同規則 4 条 1 項）⁵⁶。

2020 年 12 月、欧州委員会は、「デジタルサービス法（DSA）」と「デジタル市場法（DMA）」の 2 つの法案を発表した。DSA は、「メガオンラインプラットフォーム」を含むあらゆる種類と規模の仲介サービス事業者を対象としており、DMA は、主に「ゲートキーパー」の地位にあるコアプラットフォームサービスプロバイダを対象とする。両法は、EU が重要視する個人の権利を保護し、メガプロバイダーと個人ユーザーの間の不平等な関係を調整することはもちろん、産業政策的な観点では、小規模なプラットフォームや中小企業、スタートアップ企業のデジタル経済における成長のために、大規模なデジタルサービスプロバイダーの市場での影響力を制限することにもあるといえる。米中の台頭への懸念、欧州の影響力の復興という政治的な意図も前提にあるとみられる。

7. 2. 3. 米国

EU の統一法制モデルとは異なり、米国では、連邦レベルには統一されたデータ保護の基本法は

⁵⁵ <https://ec.europa.eu/digital-single-market/en/news/practical-guidance-businesses-how-process-mixed-datasets>

⁵⁶ <https://eur-lex.europa.eu/eli/reg/2019/452/oj>

存在せず、業界ごとの分散型法制モデルが採用されている。電気通信、金融、健康、教育、および子供のオンラインプライバシーの分野には、特別なデータ保護法がある。連邦取引委員会（FTC）は、個人データの収集と使用を包括的に規制する連邦法ではないものの、不公正で欺瞞的なデータ保護の実践に対する禁止を規律している。連邦法レベルでは、外国にあるデータへのアクセスについて、民間企業が米国外で保持する電子通信データへアクセスを簡易化することを目的とした CLOUD 法（Clarifying Lawful Overseas Use of Data Act）⁵⁷が、2018 年 3 月にトランプ大統領が署名して成立した。CLOUD 法は 1986 年に米国で導入された電子通信プライバシー法（Electronic Communications Privacy Act）を改正したものである。CLOUD 法は、米国が Microsoft 社のアイルランドのサーバに格納された通信データの開示を求めた裁判が、連邦最高裁に係属している中で成立しており、同法が成立したことにより、データが米国内に存在するか否かに関わらず、米国政府が自国の企業に対して、自社が所有、管理するデータの提供を求めることができることと定められ、その結果、連邦最高裁は、上記の訴訟は争訟性を喪失した（moot）と決定し、両当事者の合意の下、手続きが打ち切られた⁵⁸。

現状では、サイバーセキュリティとデータプライバシーを規律する米国の法令は複雑で、統一性を欠いている状況において、近時、連邦レベルの個人データの保護規制の制定を目指し「2019 年アメリカデータ普及法」⁵⁹や個人識別データに関してオンラインサービスプロバイダーの義務の強化を目的とした「データ・ケア法」⁶⁰等が上院に提出されているが、いずれも廃案となっている。ただ、これらの議員立法の動向は、米国内の立法ニーズを反映しているものとみられる。議会調査局（CRS）においても、2019 年 3 月に、「データ保護法：概要」という報告書⁶¹を出し、コモンロー上のプライバシー保護について、合衆国憲法におけるプライバシー保護の沿革について概観したうえで、現行の連邦レベルでのデータ保護にかかる法令⁶²の内容を調査している。州レベルでの立法については、カリフォルニア州消費者プライバシー保護法（CCPA）制定されて以降、他の州でも、州レベルのプライバシー保護法の立法が提案されている。それらについては、IAPP（International Association of Privacy Professionals）が法案の内容を、消費者の権利と企業責任に分類し整理している⁶³。国家安全保障にかかるデータ保護についても、連邦レベルでの法案について未だ成案とはなっていない⁶⁴。

対外的には、オバマ政権で締結された TPP 第 14.13 では、データ設備の現地での設置を義務付けたり、設置するのは現地で経営に従事する条件を付することはできないと規定した。トランプ政権で TPP からは脱退したものの、本内容は、2018 年に締結されたアメリカ・メキシコ・カナダ協定（USMCA）で踏襲され、米国が締結する自由貿易協定では、この内容が踏襲される様な交渉を米国は行っている。

ただ、米国が他国との二国間・多国間協定においては、データの自由な流れを奨励し、世界貿易

⁵⁷ <https://www.congress.gov/115/bills/hr1625/BILLS-115hr1625enr.pdf#page=866>

⁵⁸ https://www.supremecourt.gov/opinions/17pdf/17-2_1824.pdf

⁵⁹ <https://www.govtrack.us/congress/bills/116/s142>

⁶⁰ 2018 年に初めて上院に提出(S. 3744 (115th))されたが廃案になり、2019 年 12 月に再度提出（S. 2961 (116th)）され再度廃案となっている。

⁶¹ <https://crsreports.congress.gov/product/pdf/R/R45631>

⁶² 「Graeme-Ridge-Billiley Act (GLBA)」、「Health Insurance Circulation and Accountability Act (HIPAA)」、「公正信用報告法 (FCRA)」、「ビデオプライバシー保護法 (VPPA)」、「家族教育の権利とプライバシー法 (FERPA)」、「連邦証券法」、「児童オンラインプライバシー保護法 (COPPA)」、「電子通信プライバシー法 (ECPA)」、「コンピューター詐欺および虐待法」、「連邦取引委員会法 (FTC 法)」、「金融消費者保護法 (CFPA)」等。

⁶³ <https://iapp.org/resources/article/state-comparison-table/>

⁶⁴ 2019 年 11 月、共和党のジョシュ・ホーリー(Josh Hawley)上院議員は、データ保護と移転に関する「2019 年国家安全保障・個人データ保護法」(National Security and Personal Data Protection Act of 2019 (NSPDPA)) を上院に提案したが廃案となっている(National Security and Personal Data Protection Act of 2019 (2019; 116th Congress S. 2889) - GovTrack.us)。法案の目的について、データセキュリティ要件の強化と海外投資の審査を強化することにより、外国政府による脅威から米国民のデータを保護することとされる。

における米国の支配的な地位を維持することを中心とした貿易上の利益によって推進されている。この様なアメリカの強い反データローカライゼーションの方針は米欧間にも摩擦を生じさせている。

7. 2. 4. 中国

中国では、技術産業のエコシステムが不完全で、自主的な形でイノベーションの能力は未だ未成熟で、技術的自立性を実現すべきであるとされ、その戦略として、ビッグデータ、モノのインターネット、クラウドコンピューティング、人工知能などの先進的なコア技術を、強力なネットワークを通じて「デジタルチャイナ」を推進するための手段として取り入れることが提案されている。IDC (International Data Corporation) によると、中国は世界のデータ生成量のほぼ 5 分の 1 を占めており⁶⁵、中国にとっては、越境移転措置を強化し、国内のデータ関連産業を重点的に発展させていくことは、現実的で、戦略的な選択肢となっている。

王毅國務委員兼外交部長は、2020 年 9 月、グローバルデジタルガバナンスシンポジウムのハイレベル会議において、「グローバル・データセキュリティ・イニシアティブ」⁶⁶を公表し、多国間主義の堅持、安全と発展の双方への配慮、公平と正義の堅守という三原則のもと 9 つの項目を公表した。中国政府は、データを「国の重要な基本的戦略資源」として、民法典及び刑法修正案（十一）⁶⁷で規定を設けるとともに、個人情報保護法（草案）とデータセキュリティ法（草案）を全人代向けに提案している。更に、2020 年 11 月には「14 次 5 か年計画」を公表し、データ資源の権利化、取引流通、越境移転と安全保護などの基本制度と標準規範を検討することを明確にしている。

中国のデータ越境移転に関し、デジタル主権とデータセキュリティを確保するために、重要データという概念を核にそのデータ越境移転を規制している。業界ごとに、例えば、金融、保険、信用管理、ネット予約タクシーの経営サービス管理、地図、インターネット出版など個別の法令でデータのローカライゼーション（現地での保存）の規定を置いているほか、統一的な法律として、2016 年にサイバーセキュリティ法、先に述べたように、2020 年にデータセキュリティ法（草案）および個人情報保護法（草案）が審議入りし、統一的な中央政府の規制⁶⁸の方向性が徐々に明らかになっている⁶⁹。

⁶⁵ <https://www.seagate.com/files/www-content/our-story/trends/files/idc-seagate-dataage-whitepaper.pdf>

⁶⁶ <http://new.fmprc.gov.cn/web/wjbzhd/t1812945.shtml>

⁶⁷ 2020 年 12 月 26 日、中国刑法修正案（十一）が全人代を通過し、生産安全設備・設備の閉鎖、破壊にかかるデータ・情報の改ざん・隠蔽・毀損の罪が追加された。

⁶⁸ 中央政府の動きに合わせ、地方政府は、「天津市データセキュリティ管理措置（仮）」、「貴州省ビッグデータ安全・セキュリティ条例」、「深圳経済特区データ条例（意見募集草案）」、「中国新臨港地区（上海）パイロット自由貿易区全体計画に関する通知」、「海南自由貿易港建設全体計画」など、越境データ、データ安全・セキュリティ、データオープン性、データ権利などが実施されており、地域ごとの産業の発展状況なども勘案して、パイロット的に様々な立法が提案・実施されている。

⁶⁹ 中国において、のデータ越境フロー規制は、サイバー主権の代表的な実践の一つであり、サイバー境界という関係からも立場を明確にしている。サイバー主権の原則に導かれて、国境を越えたデータの流れの規制に関する法律は、国家内のインターネット管理の問題とみなされ、データ主権の主張として具現化されている。国家はデータを基本的な戦略的国家資源として捉えており、当然ながら主権の原則を適用している。したがって、データフローの管轄境界を定めるために国家の境界を用いることは、国家主権の正しい意味とされている。データローカライゼーションはデータ主権のルール表現であり、データ主権はネットワーク主権によって理論構築を完成させることが試みられている。中国は EU や米国のロングアーム管轄とは異なり、管轄範囲に基づいて規制している。これは防衛主義の論理であり、サイバー主権の論理と一致している。サイバーセキュリティ法第 2 条では、「中国国内におけるネットワークの構築、運用、保守、使用、サイバーセキュリティの監督と管理は、同法に準拠する」と規定されているが、これは明らかに領土的な意味合いを持つ規定であり、自己制限的な性格を持ち、中国法

サイバーセキュリティ法は、中国がはじめてサイバー領域に関する基本法として制定されたものであり、データの越境的な流れに関する明確な法的規定を提供する。同法第 37 条では、重要情報基盤の事業者の中国国内での個人情報及び重要データの保管を義務付け、国外に移転する際の安全評価システム定めた。その後、2020 年、サイバーセキュリティ審査弁法⁷⁰が実施され、サイバー製品やサービスの調達に関連する潜在的な国家安全保障リスク（重要なデータの盗難、漏洩、毀損のリスク）を重点的に評価することを定めた。また、2019 年 5 月に公表されたデータセキュリティ管理弁法（意見募集稿）⁷¹では、重要データの定義を提案しており、意見募集稿 38 条（五）では、未公開の政府情報、大規模な人口、遺伝的健康、地理、鉱物資源など、漏洩した場合に国家安全保障、経済安全保障、社会の安定、公衆衛生及び安全に直接影響を及ぼす可能性のあるデータ、と規定したうえで、企業の生産・運営・内部管理情報や個人情報を重要データから除外している。

これまでに用いられている重要データの定義としては、2017 年、重要データ識別ガイドライン（意見募集案）⁷²において、初めて重要データの詳細な定義や、28 業種の重要データの種類と範囲が列挙され、判断基準が示されたとされ、同ガイドラインについては、2020 年、情報セキュリティ標準化委員会により、重要データ識別ガイドラインの起草が提案されている。2020 年版のガイドラインは、重要データの定義を単純化し、重要データを、国家経済運営類、安全保障類、天然資源・環境類、公共衛生類、機微技術類、ユーザー類、政府の秘密類などに分類する方向で検討が進んでいるとされる。

この様な中央政府の統一的な取組みのほかに、各業界において、例えば、通信業界では、事業者のデータ資源及び重要データのマッピング調査したうえで、業界標準として「基本電気通信事業者の重要データ特定ガイドライン」を策定し、重要データの定義、重要データ特定の原則、ルール、ワークフロー等を定めるなどの動きもみられる。

2020 年には、データセキュリティ法（草案）が提案され、5 条ではデータの自由流通を奨励すると規定され、10 条では、国家がデータセキュリティに係る国際規則や標準の制定に参加し、越境データのセキュリティ、自由流通を促進する旨を定めるとともに、国家管理型の制度設計を再度明文で定めた。同法案では、重要データ保護については、目録制度を採用し、重点的に保護をする旨を定め（19 条）、また、重要データ処理者に対しては、通常のデータ処理者に比して追加の措置や報告義務を課している（25 条、28 条）。

7. 2. 5. 韓国

個人情報の保護とデータの利活用のバランスを政策目標する韓国は、個人情報保護法・「情報通信網法（情報通信網利用促進及び情報保護等に関する法律）」・「信用情報法（信用情報の利用及び保護に関する法律）」の改正案を 2018 年 11 月、国会に提出していたところ、2020 年 1 月の審議により本会議を通過し、8 月 5 日より施行された。

これら 3 つの法律を韓国ではデータ 3 法というが、今回の改正による所管省庁別に分けられた不要な重複規制を取り除き、第 4 次産業革命の変化を受け、個人や企業のデータ利活用が促進されるよう改正された。

また、データの利活用と言う意味での取組みは、2020 年 7 月 14 日、Covid-19 による深刻な経済不況の克服と構造大転換対応という課題に直面する中、「韓国版ニューディール」の一環として、「デジタルニューディール」政策を発表し、2020 年 12 月 24 日、関係省庁が合同で、「AI 法・制度・規制のロードマップ」を公表した。第 4 次産業革命をけん引する汎用技術である AI はデータ

の治外法権を制限することができるが、その結果、規制当局がデータの現地化措置を採用することを奨励することになるとされる。

⁷⁰ http://www.cac.gov.cn/2020-04/27/c_1589535450769077.htm

⁷¹ http://www.cac.gov.cn/2019-05/28/c_1124546022.htm

⁷² 同ガイドラインは、情報セキュリティ技術 データ出境安全評価ガイドライン（草案）の付録として、2017 年 5 月に意見募集がされた。

<https://www.tc260.org.cn/front/postDetail.html?id=20170527173820>

結合してこそ国、社会のデジタルフォーメーションを加速するとして、本ロードマップでは、①AIの特性に応じた「時代を先駆ける法・制度」②グローバル議論を考慮した「総合的法・制度」③民間自律を尊重する「市場に優しい法・制度」④社会的合意を通じた「共生・包容ベース法・制度」という4つの目標をかかげ、データに関しては、以下のような内容の法令・課題の実施を目指している。

課題名	関係部署	実施予定
データ産業振興のための基本法の制定	科技情報通信部	21 年上期
データ管理業のための法的根拠を用意	科技情報通信部・個人情報委員会	21 年上期
個々の産業や需要を考慮したデータ活用支援	産業通商資源部・中小ベンチャー企業部	21 年上期
自動化された個人情報の処理に依存した意思決定への対応権導入	個人情報委員会	21 年上期
データマイニングのための著作物利用許可	文化体育観光部	21 年上期
結合手順と仮名の情報安全対策の規制見直し	個人情報委員会	21 年下期
データの排他的と不当な利用行為に対する規律案の設定	特許庁・公取委	22 年下期

図表 75 韓国における関連法令・課題の実施

7. 2. 6. その他

上記以外の国についても、個人情報保護法制や、データ利活用の促進に向けた動向は活発である。特に、2019 年から 2020 年にかけて個人情報保護法について、ASEAN 地域では関連法令の導入が進められている。インドネシアでは、2020 年 11 月に、個人情報保護法 (RUU PDP) の実施を目指し立法手続のため、通信情報部 (Kemkominfo)、内務省 (Kemendagri)、司法府 (Kemenkumham) 3 機関で実務会議を実施している。ベトナムでも、2020 年 1 月、公安部が、個人情報に関する法令草案 (Draft Decree on Personal Data Protection) に対応するメールや通信分野の行政違反処罰を強化した施行令を作成し、2 月 3 日 (Nguyen Xuan Phuc) 首相は、本法案に署名し、4 月 15 日に施行されている。タイでも「個人情報保護法 (Personal Data Protection Act、BE 2562 (2019))」が、全面施行に向けて動いている。シンガポールは個人情報保護法 (修正) の意見募集を 2020 年 5 月に実施し、南アフリカは 2013 年に制定された個人情報保護法が 2020 年になり施行され、2020 年 12 月、ニュージーランドではプライバシー法が発効した

インドについては、政府がデータローカライゼーションの取り組みを行っており、国内企業からも支持を受けているとの報告がある⁷³。ただし、データローカライゼーションをめぐる地政学的な問題が続く状況下において、欧米の大手テクノロジー企業とインドや中国の大手テクノロジー企業

⁷³ 例えば、インドビジネス特派員連盟の会長であり、インド準備銀行の金融包摂諮問委員会のメンバーでもあるアナンド・シュリバスタフ博士は、「データは主権的な権利であり、すべての国は国民の機密データを保護する権利を持っている」と主張しました。同氏は、インドのフィンテック企業がデータのローカライズを支持しているのは、データをインドの国境内に保管されるべき重要な資産と考えていると説明する。実際、Paytm (中国の大手テック企業アリババの支援を受けている) は、「インドの新興企業のエコシステムがグローバルなテック企業と比較して公平な競争の場を確保する」ために、データローカライゼーションを強く支持する姿勢を公にしている

(以下 URL の記事「データグローバル化 v. データローカライゼーション」参照。
<https://www.centerforfinancialinclusion.org/data-globalization-vs-data-localization>)。

が対立するも熾烈となっている。インドネシアやベトナムなどの国々は、データローカライゼーションの規定を緩めるよう動いたり、2018年に初めて提示されたインドのデータ保護法案の草案が、インド国外への個人データの転送と保管に関する制限の撤廃など、多くの修正を加えて議会に再提出された等の経緯もあり、ソフトな形でローカライゼーションを緩和する動きもみられる。なお、法案当初の草案にあった「センシティブな」個人データと「クリティカルな」個人データの両方に対する制限は残っている。また、2020年7月に、インド電子情報技術省（委員長：クリス・ゴパラクリシュナン氏）が設置した専門委員会は、非個人情報に関する諸問題に関しての検討をふまえ、非個人データガバナンスにかかるフレームワーク⁷⁴を公表している。関係する報道⁷⁵によると、現在国会で審議中の個人データ保護法案に、非個人データ共有の内容も盛り込まれていると言われている。データ共有の規制は公平な競争の場を作り、インド国民とコミュニティへの経済的利益を全体的な福祉の最大化に転換でき、政府の政策立案とサービス提供を改善すると委員会は分析するが、さまざまなステークホルダーから賛否両輪あり、インド政府による、（企業の）匿名の非個人データ利用の実現に向けた検討が行われている。

8. 諸外国における企業のデータ利活用と越境移転にかかるモデル契約

8. 1. 外国における企業のデータ利活用に関する調査

8. 1. 1. EU: 「欧州における企業間のデータ共有に関する調査（Study on data sharing between companies in Europe）」⁷⁶と「B2B データ共有の原則とガイダンスに関する SME のパネル協議の結果」⁷⁷

【欧州における企業間のデータ共有に関する調査】

2017年の「欧州データ経済の構築」に関する諮問結果⁷⁸において B2B データのアクセスと再利用および責任に関する検討の重要性が確認された後、2018年4月26日、欧州委員会より「欧州における企業間のデータ共有に関する研究」と題する報告書が公表された。同研究では、データの統計分析を用いて欧州経済領域（EEA）内での企業間データの共有と再利用の規模の定量的に分析し、EU 域内の企業間でのデータ共有が相当の規模と実用化を達成している現状を紹介するとともに、企業のデータ共有のケーススタディを通して、データ共有の成功要因と課題を紹介し、最後に、欧州経済圏における企業間のデータ共有と再利用の分野における今後の政策展開を支援するための提言を行っている。

EEA における企業間のデータ共有と再利用の具体的な定量的な規模に関して、B2B データ共有という概念は、一般的に企業間でのデータ交換を指す言葉として使われているところ、データ共有概念には、データ交換の需給両面が含まれており、データを外部で共有する企業と、データへのアクセスを許可された企業が主体となる。本報告書では「データ共有」と「データ再利用」は、実際には需給の異なる視点から見た同じ行動の記述であり、前者はデータ供給者の視点に限定され、後者はデータ需要者の視点に限定して「データ共有」（企業 B がその事業目的のために必要なデータ

⁷⁴ フレームワークの概要は、以下参照。<https://www.prindia.org/report-summaries/non-personal-data-governance-framework>

⁷⁵ <https://www.globalgovernmentforum.com/india-considers-non-personal-data-sharing-regulation/>

⁷⁶ European Commission; Study on data sharing between companies in Europe (<https://dl4ld.nl/KK0118016ENN.en.pdf>)

⁷⁷ <https://ec.europa.eu/digital-single-market/en/news/sme-panel-consultation-b2b-data-sharing>

⁷⁸ <https://ec.europa.eu/digital-single-market/en/news/synopsis-report-public-consultation-building-european-data-economy>

を、企業 A が生成または収集したデータを、競合他社や下請け企業ではない他の企業 B に無償または有償で提供すること」と「データ再利用」（企業 B が、直接の市場競争相手ではない他の企業 A のデータを、企業 B 自身の事業目的のために再利用するプロセス（請負・下請関係を除く）。再利用の対価は無償である場合もあれば、サービスの提供を含む何らかの報酬等に基づくものである場合もある。）の概念を定義している。

B2B データ共有の EU 域内の経済価値は巨大だが、B2B データ共有はまだ発展途上の段階にあるとされ、域内の企業の多くが自社のビジネスにとって重要と考えているが、実際のデータ導入は進んでいない。100 社を調査した調査では、取引関係のデータを取り引きしている企業は 11%に過ぎず、データを共有するためにオープンデータポリシーを採用している企業は 2%に過ぎないと結論づけている。EU のデータ市場は、2016 年の欧州のデータ企業の収益総額は約 620 億ユーロに達しており、2015 年には 560 億ユーロ、2014 年には 510 億ユーロと推定され、急成長を遂げているが、データを多用する企業は 6.3%にとどまっている。B2B データ共有に取り組んでいない企業の 3 分の 1 は、今後 5 年以内に他社とのデータ共有を開始することを検討しているとされる。データの供給側と同様に、ほとんどのデータ・ユーザーは、データ・サプライヤーからのデータへのアクセスがビジネスにとって重要であると考えている。データを再利用したことがない企業のほぼ半数が、今後 5 年間で B2B データを再利用する可能性あるとする。また、2016 年の EU の GDP に対するデータ経済の貢献度は 1.99%に達し、最も楽観的な予測では 2020 年には 4%に達する可能性がある」と結論づける。

欧州の企業間データ共有産業の主な形態に関する主要な形態について、①データの収益化（より典型的で一般的なアプローチは、企業が他の企業とデータを共有することで追加的な収益を得る一方的な方法を指し、サービスへのアクセスの結果としてのデータの収益化も含まれている。このような多くの企業の主な事業目的は、（データ対象者からの有効なライセンスに基づいて）彼らが生成したデータから収益を得る）、②データ取引マーケットプレイス（データ供給会社とデータ需要会社は、信頼できる仲介者を介して、データ供給会社が設置または管理する安全なオンラインプラットフォーム上でデータを取引し、プラットフォーム上でデータ取引ごとに手数料を受け取る）、③産業用データプラットフォーム（一部の企業は、戦略的パートナーシップを締結し、自主的にクロードで安全な専用プラットフォームに参加して、データ交換から相互利益を得ることを選択します。このような共同プラットフォームに参加している企業は、その結果として新製品やサービスを開発したり、既存の製品やサービスのパフォーマンスやレベルを大幅に向上させたりすることができるとされる）、④テクノロジーイネーブラー（データ共有の「テクノロジーイネーブラー」または「テクノロジーイネーブラー」とされる企業もあり、データ共有から直接収益を得るのではなく、データ共有のためのテクノロジーソリューションの開発と実装に特化して取り組むとされる。例えば、「テクノロジーイネーブラー」は、データ共有から直接収益を得るのではなく、データ共有のためのテクノロジーソリューションの開発と実装に取り組み、Web ベースのソフトウェア製品、アプリケーション、またはプラットフォームから収入を得る。）、⑤オープンデータ戦略（あまり一般的ではないが、新製品や新サービスの開発を促進するためにデータを全面的にオープンにすることを選択する企業もある。この戦略を採用している企業の多くは、特にエネルギー分野を中心にデータのオープン化を法的に義務付けられている）、という 5 つの分類が提示され、ケーススタディの企業はそのいずれかに分類されている。

企業間データ共有や再利用の成功要因について、データ利用者やデータ提供者との信頼関係の構築、データの需要を理解してパートナーシップを確立し、データを使って何ができるのか具体的なユースケースを特定し、シンプルで使いやすいツールを導入することが成功要因として挙げられている。

他方で、欧州企業間のデータ共有が直面する主な障壁として、技術的、コスト的、法的障壁とされ、①技術的な障壁としては、相互運用性の欠如、セキュリティ上の懸念、その他のサイバーセキュリティ要件など、②コストの障壁としては、主にデータを長期的に安全に保存する必要性に基づくインフラ投資、③法的障壁は、主に「データの所有権」の不確実性、データの法的利用の境界線、データ保護に対する企業の法的ニーズを満たすことの難しさなどが挙げられる。また、データの再利用の面では、データの再利用の拒否、差別、高コスト、相互運用性の欠如、データの標準化など

が、現在企業が直面している主な障壁とされる。

欧州では、2013年10月の早い段階で、EU理事会⁷⁹により、デジタル単一市場の統合、経済のあらゆる分野におけるデータ駆動型イノベーションの促進、欧州のデジタル空間における大規模な投資経済、経済全体でのデータへのアクセスと共有を改善するための戦略的技術の促進などの分野での行動を呼びかける戦略的アジェンダを策定し、インテリジェントな輸送システム、食糧安全保障、気候、エネルギー、健康など、特定の分野におけるデータ交換について、データの自由な流れ、企業間のデータへのアクセスと転送、新興技術の責任問題、データのポータビリティと相互運用性の理解を促進するとともに、両者の標準化を進めてきたことが紹介されている。

本報告では、B2Bデータ共有の概念について、従前、重要な規制や指令の範囲が明確に理解するために定義が不明確だった点の問題点を指摘し、EEAの企業や利害関係者と協力してB2Bデータ共有の定義を策定するよう勧告するとともに、欧州におけるベストプラクティスの特定と共有を支援し、技術的なガイダンスを提供するためのデータ共有支援センターの設立等を勧告している⁸⁰。

また、具体的な施策として、欧州委員会において、①欧州委員会がB2Bデータ共有と再利用への参加のための基本的な概念、原則、条件を定めた汎用的なB2Bデータ共有フレームワークを作成し、データ転送と他社へのアクセスを成功させるための推奨事項を提供すること、②対話型データの操作性やデータの相互運用性や標準の発展を支援すること、③中小企業を含む企業を標準化プロセスに参加させ、企業のニーズと課題を考慮し、部門別基準を考慮すること、④小規模な零細企業に対し、データ共有を推進するためのB2Bデータ共有ファンド事業への企業の参加を支援すること等が提案されている。

【B2Bデータ共有の原則とガイダンスに関するSMEのパネル協議の結果】

「欧州データ経済の構築」に関するステークホルダー対話のフォローアップとして、欧州委員会は「共通のヨーロッパのデータ空間に向けて」とそのスタッフ作業文書を公表し、合わせて、2018年10月から2019年1月にかけて、B2Bのデータ共有に関して中小企業へのアンケートを実施し979回答を得ている。

その回答としては、中小企業の33%が他の企業から積極的にデータを取得しており、30%が接続されたデバイスを使用している（または使用する予定がある）、87%は、IoTオブジェクトが新しい課題であり、54%がそれらが産業文脈における公平性の観点から、現在、法律によって十分に対処されていないと考えているとされる。

回答者は、欧州委員会が提示した原則が、契約慣行に適度にプラスの影響を与えると概ね考えており、他方で、約30%は、企業が「競争を歪め、特に企業を1つのプロバイダーとの取引にロックインすることを回避することには効果的でない」と回答している。

8. 1. 2. 中国:「工業企業データ資産管理現状調査報告」

2018年12月に中国情報通信研究院(CAIST)等は、工業企業データ資産管理現状調査報告を公表している。

ネット上、WeChat上で、①データ資産管理の組織、②情報技術力、③データ管理に関してアンケートを実施し、120回答の内、74回答が有効として整理されている。回答企業としては、北京が30%、2000人以上の大企業が半数とされ、コンピュータ、通信、その他の電子機器製造業(24.3%)、鉄道、船舶、航空宇宙、航空宇宙産業、その他輸送用機器製造業(16.2%)、自動車製造業(10.8%)、一般機器製造業(6.8%)、電力・電力業熱生産・供給業(5.4%)、その他の選択肢を選択した企

⁷⁹ 同協議会の結論は、データアクセスやデータ転送を改善するための重要なイネーブラーとしてクラウドコンピューティングを挙げている。<https://ec.europa.eu/digital-single-market/en/cloud-computing>

⁸⁰ その後、データ共有支援センターが立ち上げられ、データ利活用の各種支援を実施している。<https://eudatasharing.eu/about-us>

業（5.4%）は、工業企業に関連するサービスを提供している企業が中心であったとされる。

①データ資産管理について、98.6%の工業系企業がデータ管理業務に投資価値があるものとしており、工業系企業では、データ資産管理の重要性について一定の認識が形成されていること、データ管理関連の投資開始または計画について 87.8%の企業が行っていること、データ管理業務のための専任組織を設置について、55.4%の企業が行っているが、人員の確保についてはまだ不十分で、47.3%の企業が 0～5 人のチームにとどまること、33.8%の企業において自社のデータ管理モデルをいまだ開発していないこと、25.7%の企業しか、データに関する管理方法や企業標準を定期的に更新していないこと、78.4%の企業は社内データについてある程度認識しているが、継続的なデータの整理・把握作業は改善が必要としていることが確認されている。

②情報技術力に関して、企業が収集・管理しているデータの総量は 66.22%の企業において総データ量が 20TB 以下で、87.84%の企業の 1 日のデータ増分が 100GB 以下の企業であること、企業のデータ資産管理の取組みについて、60.7%は、社内のビジネスニーズに牽引されていること、データ管理の課題について、71.62%がデータ連携の難しさ、60.81%が方法論の不足、59.46%が短期的に結果が見えないことを挙げていること、0.54%の企業がデータ管理のために文書記録方式を使用していること、ニーズのある技術ツールとして、ビッグデータ処理（44.59%）、データ管理（44.59%）、報告書分析（43.24%）があること、データ管理システムの構築方法としては、自社開発 31.08%、外部調達 44.59%であること、企業のデータ活用に対するニーズは、生産・運用監視（68.92%）と資源利用の効率化（62.16%）であることが確認されている。

③データ管理について、クラウドストレージについては、51.34%の企業は慎重であること、企業側のデータ流通のニーズは高く、データ流通を志向しないという企業は 2.7%にとどまったこと、データセキュリティ管理では、従来の物理的分離とネットワークの分離を採用している企業が 41.3%を占め、厳格なフロー管理システムデータセキュリティを確保している企業は 26.1%を占めていたことが確認されている。

また、その後、2020 年に、中国情報通信研究院（CAIST）が公表した「ビッグデータ白書」では、①情報技術基盤の面で、トップ企業で構築されているビッグデータプラットフォームとデータインフラに比べて、一般の企業はまだ ERP、CRM、SCM などの運用管理ソフトウェアを使用している、データ管理コストが高いこと、②企業レベルのデータ収集・保存・分析コストが高く、データ管理の入力と分析をする人員の確保が課題であること、③専門性について半数近くの企業が専門的なデータ管理チームの設置がないこと、④データ資産管理の推進についてデータのほとんどが生産オペレーションや設備の故障を監視するために使用、データアプリケーションのシナリオは多様とは言えず、データ資産がビジネス開発を推進するには至っていないことが報告されている。

8. 2. 越境データ移転のモデル契約の参考例

8. 2. 1. EU：越境的データ移転のための標準契約条項

EU では、データ移転について、個人データと非個人データの両方で構成されるデータセットになっている可能性が高いとの認識を前提に、非個人データの自由流通と GDPR の規制の両方が含まれる場合のガイドラインを発出している⁸¹。両社のデータが不可分である場合は、GDPR のデータの適用が妨げられないとされ、非個人データの場合は、自由流通が奨励されていることから、データの移転のレベルでは、非個人データであっても、GDPR の標準契約条項が参考となる。

欧州委員会は、2020 年 11 月 12 日に、規則（GDPR）2016/679 に基づく第三国への個人データの移転のための標準契約条項に関する実施決定（案）と、その新版である「越境的データ移転のための標準契約条項」（以下「新 SCC」）を付録の形で公表した。これら 2 つの文書は、まだ正式に

⁸¹ 2_EN_ACT_part1_v4.docx (europa.eu)

は採用されていないが、Schrems II 一事件⁸²を受けて、国境を越えたデータ転送における SCC の利用に関する欧州委員会の姿勢と指令を反映している。

新 SCC は、幅広いシナリオに適用可能な汎用条項を含むだけでなく、(1)コントローラ間の転送、(2)コントローラからプロセッサへの転送、(3)プロセッサ間の転送、(4)プロセッサからコントローラへの転送の 4 つの異なるシナリオに対応した、よりターゲットを絞ったテンプレート条項を提供している。第 1 項の総則では、データ主体が第三者受益者として SCC 関連の権利を主張できることや例外を明記し、SCC の適用の目的と制限を明確にし、また、一定の条件を満たした場合に第三者が権利と義務を伴うオプションのアクセス条項を提供する。新 SCC の第 II.2 項では、現地法が国境を越えた移転に与える可能性のある影響を考慮し、有効な評価に基づいて、データ輸入者が SCC に基づく義務を履行することを、移転先の国の法律が妨げると信じる理由がないことを保証することを締約国に要求している。このような評価は、移転の具体的な状況（移転先の種類、処理目的、移転した個人データの性質など）、移転先の関連法（開示要件の有無など）、および SCC が提供する以上の保護を考慮することとなる。これらの一般的規定に加えて、新 SCC では、データ対象者の権利の実現など、様々な移転シナリオに応じたより具体的な適用規定が定められている。

8. 2. 2. ASEAN : 「データモデル契約条項」と「クロスボーダーデータフロー認証」

ASEAN では、世界有数のサイバーセキュリティインフラを持つシンガポールや、ラオス、カンボジア、ミャンマーなどの同盟国がサイバーセキュリティインフラの構築途上にあるなど、ASEAN 内のサイバーセキュリティや個人情報保護のレベルにはばらつきがあるところ、サイバーセキュリティとデータ保護のグローバル化が進む中で、ASEAN 全体のサイバーセキュリティとデータ保護のレベルを高め、よりグローバルな競争優位性を獲得するために、ASEAN は統一的なサイバーセキュリティとデータフローの基準を取得することが急務となり、ASEAN は、2021 年 1 月 22 日に「ASEAN Data Management Framework (DMF)」⁸³や「ASEAN Model Contract Clauses for Cross Border Data Flows (MCC)」⁸⁴を公表した。

従前の経過として 2016 年には、2015 年 ASEAN 経済共同体ブループリント 2025 および ASEAN ICT マスタープラン 2020 の勧告に基づき、アジア太平洋経済協力 (APEC) プライバシーフレームワーク (2015 年) やその他の国際的に認められた個人データ保護のための標準またはフレームワークの内容を重く引き継ぎ、ASEAN 閣僚会議は「個人データ保護に関する ASEAN フレームワーク」を採択した。ただ、加盟国レベルでフレームワークを適用する経済圏は、自国の状況に適用される例外を採用することができ、フレームワークは国内的にも国際的にも拘束力はないとされた。

2018 年には、2015 年の ASEAN 経済共同体ブループリント 2025 と個人データ保護に関する ASEAN フレームワークの呼びかけに基づき、ASEAN は「デジタルデータガバナンスに関する ASEAN フレームワーク (ASEAN Framework on Digital Data Governance)」⁸⁵を発表した。

そして、2019 年 11 月にラオスのビエンチャンで開催された第 19 回 ASEAN 貿易大臣会合で採択された「ASEAN クロスボーダーデータフローメカニズムへの主要アプローチ」⁸⁶は、ASEAN がこれらのメカニズムのうち、「ASEAN モデル契約条項」と「ASEAN クロスボーダーデータフロー認証」の 2 つの開発に注力することを提言した。

以上のような開発背景を踏まえ、2021 年 1 月 22 日、ASEAN は「デジタルマスタープラン 2025」

⁸²

<http://curia.europa.eu/juris/document/document.jsf?text=&docid=228677&pageIndex=0&doclang=en&mode=lst&dir=&occ=first&part=1&cid=12312155>

⁸³ https://asean.org/storage/2-ASEAN-Data-Management-Framework_Final.pdf

⁸⁴ https://asean.org/storage/3-ASEAN-Model-Contractual-Clauses-for-Cross-Border-Data-Flows_Final.pdf

⁸⁵ https://asean.org/storage/2012/05/6B-ASEAN-Framework-on-Digital-Data-Governance_Endorsed.pdf

⁸⁶ <https://asean.org/storage/2012/05/Key-Approaches-for-ASEAN-Cross-Border-Data-Flows-Mechanism.pdf>

を発表した⁸⁷。第1回 ASEAN デジタル大臣会合では、「ASEAN デジタルデータガバナンスフレームワーク」の4つの主要イニシアティブと関連する主要文書への対応と実施として、地域のデジタル経済とデジタル貿易のための ASEAN 内の個人データフローを実施するための具体的なイニシアティブとして、「ASEAN データ管理フレームワーク (DMF)」と「越境データフローのための ASEAN モデル契約条項コンポーネント (MCC)」のリリースが承認された。国境を越えたデータフローの ASEAN 認証の詳細は現在開発中で、2021 年までに完成する見込みとされる。

越境データフローのための ASEAN モデル契約条項(MCC)は、ASEAN Framework on Personal Data Protection (ASEAN Framework on Personal Data Protection)⁸⁸に基づき、ASEAN が公表している国境を越えてデータが流れる際に遵守すべき最低限の義務を定めたもので、個人データの移転に関する基本的な考慮事項のガイダンスを提供するための自主基準とされる。MCC はすべて、定義セクション、データ輸出者の義務セクション、データ輸入者の義務セクション、商業条項で構成されている。

具体的には、MCC は① データ・コントローラからデータ・プロセッサへ：データ・プロセッサから下流へのデータ処理を含む、データ・インポーター（この関係では「データ・プロセッサ」とも呼ばれる）にデータを送信する契約者または供給者であるデータ・エクスポート業者のために、データ・プロセッサから下流へのデータ処理を含む、送信企業に代わってデータを処理するデータ・インポーター（「データ・コントローラ」とも呼ばれる）にデータを送信する場合 データ・プロセッサ（「サブプロセッサ」とも呼ばれる）、② データ管理者からデータ管理者へ：データ輸出者が、データ輸入者の使用のためにデータを送信する場合、データ輸入者は送信されたデータを独自の目的のために処理し、データを受領した時点でデータを完全に管理するという、2つの場合を分けている。

ある ASEAN 諸国の管理者または処理者は、適切な保護措置が講じられている場合、および ASEAN 諸国の法律で要求される場合に限り、データ対象者に強制力のある権利および有効な法的救済手段が利用可能であることを条件に、個人データを他の ASEAN 諸国に移転することができる。このようなセーフガードは、ASEAN MCCs の下で ASEAN 加盟国が採択したモデルデータ保護規則によって規定されることがあります。さらに、データ輸出者は、その所在する各 ASEAN 加盟国の法律に従って、コントローラまたはプロセッサとしての一般的な義務を履行することとされる。

また、移転の当事者間の実質的に同様な保護を提供するために、ASEAN 加盟国の当事者間の MCC は、これに基づいて移転された個人データが、各 ASEAN 加盟国が保証するものと実質的に同等の保護を受けることを保証する必要性を考慮に入れるべきである。データ・インポーターは、他の ASEAN 加盟国のインポーターが MCC を遵守し、その他の方法で保護の継続性を確保する場合、またはデータ対象者の同意に基づく場合に限り、データ・インポーターが他の ASEAN 加盟国のインポーターに転送することを許可しなければならないとされる。

基本的なデータ保護の前提条件としては、データ輸出者は、データが適用される ASEAN 加盟国の法律に従って収集、使用、開示、送信されることを保証し、このような法律がない場合、データ対象者は、その目的が合理的かつ実行可能な場合には、その目的を通知し、同意することとされる。

さらに、データ輸入者は、本契約に基づく個人データの紛失、不正使用、コピー、変更、開示、破壊、アクセスがあった場合、直ちに、または当事者が指定した合理的な時間内に、不当な遅延なく、関係当局およびデータ輸出者に通知が求められる。

⁸⁷ <https://asean.org/storage/ASEAN-Digital-Masterplan-2025.pdf>

⁸⁸ <https://asean.org/storage/2012/05/10-ASEAN-Framework-on-PDP.pdf>

■Appendix

2020 年度「データ利活用に関するアンケート調査」調査票

データ利活用に関するアンケート調査

独立行政法人経済産業研究所
理事長 矢野 誠

■本調査の目的

本調査は、経済産業省経済産業政策局知的財産政策室、及び、経済産業省商務情報政策局情報経済課の協力のもと、企業の競争力の源泉と期待されているデータについてその利活用の実態を把握し、データ利活用から成果を得るために重要な要因は何かについて明らかにすることで、企業におけるデータ利活用促進の一助とすることを目的としています。

IoT（モノのインターネット化）、ビッグデータ、AI（人工知能）などを軸とした第4次産業革命の進展に伴って、企業の事業活動に関連するデータが日々大量に生み出されており、あらゆる産業の企業にとってデータの効果的な利活用は不可避の課題となっています。また、データの利活用は、自社のみで行う場合だけでなく、他組織からデータを取得したり他組織にデータを提供したりするなど、取引や連携の形態も多様化しています。本調査は、このような状況を踏まえ、どのような企業がどの程度データ利活用を行っているのか、どのように他組織と連携してデータ利活用を行っているのかなどについて調査し、データ利活用によって有意義な成果を得るために重要な要因について明らかにすることを目的としています。

■ご回答にあたってのお願い等

1. 本調査は、独立行政法人経済産業研究所が（株）東京商工リサーチに委託して実施しているものです。
2. 本調査におきましては、調査結果は統計的に処理し個別の回答は絶対に公表されることはありません。
3. 本調査で集められた情報は機密を厳守し、独立行政法人経済産業研究所の機密データセット保護管理規程のもと機密事項として扱われ、研究目的以外で使用されることはありません。
4. 本調査票は2部構成です。

【第Ⅰ部】：貴社の概要について（1～3 ページ）

問1：貴社の概要について

【第Ⅱ部】：貴社のデータ利活用について（4～19 ページ）

問2：貴社のデータ利活用の全般的な状況について

問3：貴社のデータを貴社内で利活用する場合について

問4：データ取得プロジェクトについて

問5：データ提供プロジェクトについて

ご回答者として、経営企画部門（経営企画担当者）が取りまとめてご回答いただくことを想定していますが、データ利活用の責任部門やデータ利活用を実施している部門が企業によって様々である状況を踏まえ、必要に応じて知財、法務、情報システムなど、適した部門のご担当者様にご回覧いただき、ご回答をお願いできれば幸いです。

5. できる限りご回答いただければと存じますが、回答内容に差し支えのある質問項目があれば、回答をとばしていただいても結構です。
6. 回答は、2020年7月17日（金）までに、専用のWEBアンケートサイトにてご回答いただくか、または調査票に直接ご記入の上、同封の返信用封筒に入れてご返送いただきますようお願い申し上げます。
7. 本調査の結果に関しては、ディスカッションペーパーとして公開するとともに、報告会を開催する予定です。

■ご回答者様の情報

下欄にご記入ください。

貴社名			
ご回答者様のご所属名			
ご回答者様のご所属部門 ※最も近い部門 1 つに○を付けてください。	1. 経営企画系 4. 広報系 7. 研究開発・技術系 10. その他 (	2. 知財系 5. 総務系 8. 事業系	3. 法務系 6. 情報システム系 9. 営業系)
ご回答者様のお名前			
ご回答者様のご連絡先 (電話)			
ご回答者様のご連絡先 (E-mail)			
本調査結果の報告会の案内	1. 要 2. 不要		

第 I 部 貴社の概要について

問 1 貴社の概要について

貴社の概要についてお伺いします。2020 年 3 月末時点 (2019 年度) の状況をご記入ください。

貴社が子会社等をお持ちの場合であっても、子会社等は含めず、貴社単体の状況についてご回答ください。

ただし、貴社が純粋持株会社の場合は、子会社等の傘下の主要なグループ会社を 1 社選び、その企業単体についてご回答ください。

問 1-1 貴社の経営形態について、当てはまるもの 1 つに○を付けてください。「1. 純粋持株会社である」に当てはまる場合は、子会社等の傘下の主要なグループ会社を 1 社選び、その企業名をご記入いただくとともに、以降の設問ではその企業単体についてご回答ください。

- | | |
|---------------------------------|---|
| 1. 純粋持株会社である → 回答する主要なグループ会社名 (|) |
| 2. 事業持株会社である | |
| 3. 純粋持株会社でも事業持株会社でもない | |

問 1-2 貴社の設立年をご記入ください。

(西暦)	年
------	---

問 1-3 貴社の資本金額または出資金額をご記入ください。また、貴社の資本金または出資金に占める外国資本の比率をご記入ください。

資本金額または出資金額	百万円
外国資本の比率※	%

※貴社の発行済株式総数もしくは出資金総額に占める外国投資家による所有株式数又は出資金額のおおよその比率をご記入ください。

問 1-4 貴社の 2019 年度 の売上高、営業利益、研究開発費をそれぞれご記入ください。研究開発費については、総額に加え、社内研究開発費及び外部支出研究開発費もご記入ください。

売上高	百万円
営業利益	百万円
研究開発費（総額）※1	百万円
うち、社内研究開発費※2	百万円
うち、外部支出研究開発費※3	百万円

- ※1 研究開発費とは、研究開発にかかる人件費、原材料費、消耗品費等のその他の経費、有形固定資産購入費（土地、建物、構築物、船舶、航空機並びに耐用年数 1 年以上かつ取得価額が 10 万円以上の機械、装置、車両、その他の運搬具、工具・器具及び備品の購入に要した費用）を指します。研究開発費には、「社内研究開発費」と「外部支出研究開発費」の 2 つがあります。
- ※2 社内研究開発費とは、自己資金、社外から受け入れた資金を問わず、社内で使用した研究開発費のことを指します。なお、研究開発と他の活動とを分けて算出することが困難な場合には、案分した金額をご記入ください。委託研究開発（共同研究開発を含む）等のための外部支出（貴社の海外拠点を含む）は含みません。
- ※3 外部支出研究開発費とは、社外（外部）に委託した研究開発（共同研究開発を含む）等のために支出した研究開発費をいいます。支出名目（委託費、賦課金、奨学寄附金等）は問いません。

問 1-5 貴社の 2020 年 3 月末時点 のおおよその従業員数、正社員数をそれぞれご記入ください。

従業員数※1	人
うち、正社員数※2	人

- ※1 従業員数には、有給役員、常勤職員及びパート、アルバイト、臨時・日雇の者で 1 ヶ月以上にわたって雇用されている方はすべて含めてください。
- ※2 従業員のうち、正社員・正職員として処遇している人の数をご記入ください。

問 1-6 貴社の事業がカバーする業種として、当てはまるものすべてに○を付けてください。また、そのうち貴社の事業を最もよく表す業種の番号を 1 つ記入してください。

1. 農林水産業	13. 窯業/土石製品製造業	25. 自動車/同付属品製造業
2. 鉱業/採石等	14. 鉄鋼業	26. その他輸送用機器製造業
3. 建設業	15. 非鉄金属製造業	27. その他製造業
4. 食品製造業	16. 金属製品製造業	28. 卸売/小売
5. 繊維/衣類製造業	17. 汎用機械器具製造業	29. 通信/放送
6. パルプ/紙類製造業	18. 生産用機械器具製造業	30. 運輸/物流
7. 医薬品製造業	19. 業務用機械器具製造業	31. 金融/保険
8. 総合化学製造業	20. 医療用機械器具製造業	32. 情報/システム/ソフト
9. 油脂/塗料製造業	21. 電子部品製造業	33. 学術/研究開発機関
10. その他化学製造業	22. 電子応用/計測機器製造業	34. 技術/専門サービス
11. 石油/石炭製品製造業	23. その他電気機器製造業	35. その他サービス
12. プラスチック/ゴム製品製造業	24. 情報通信機器製造業	36. その他
		()

うち、貴社の事業を最もよく表す業種の番号 1 つ

問 1-7 貴社のグループ会社数をご記入ください。グループ会社数は、子会社及び関連会社、親会社の総数を国内・海外含めてご記入ください。

社

問 1-8 貴社及び貴社のグループ会社が設置する海外拠点の所在する地域について、当てはまるものすべてに○を付けてください。海外拠点とは、販売・生産・研究開発・地域統括・物流の各拠点や駐在員事務所等を含みますが、代理店は含めません。海外拠点を設置していない場合は「6. 海外拠点はなし」に○を付けてください。また、拠点が所在する海外の国数のおおよその総計をご記入ください。

- | |
|--|
| <ol style="list-style-type: none">1. 中国（香港含む）2. その他アジア3. 米国4. ヨーロッパ5. 上記以外の地域6. 海外拠点はなし |
|--|

拠点が所在する海外の国数の総計

ヶ国

問 1-9 貴社の保有特許件数（国内）について、当てはまるもの1つに○を付けてください。

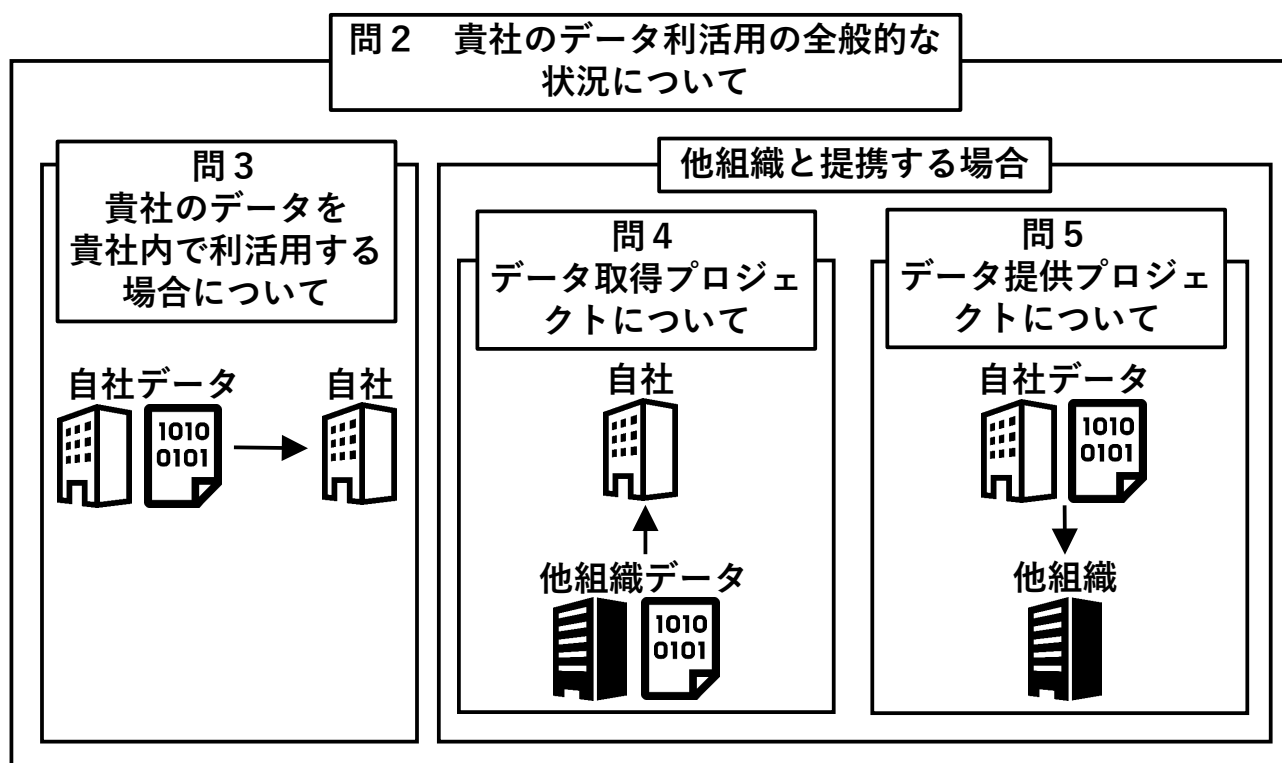
- | |
|---|
| <ol style="list-style-type: none">1. 0 件2. 1～9 件3. 10～99 件4. 100～999 件5. 1000～9999 件6. 10000 件以上 |
|---|

※次の第Ⅱ部（4～19 ページ）も引き続きご回答ください。

第Ⅱ部 貴社のデータ利活用について

■記入上の留意点

- ・第Ⅱ部では、貴社のデータ利活用についてお伺いします。第Ⅱ部は、問2～問5からなります。



- ・2020年3月末時点の状況をご記入ください。
- ・貴社が子会社等をお持ちの場合であっても、子会社等は含めず、貴社単体の状況についてご回答ください。ただし、貴社が純粋持株会社の場合は、問1-1でご記入いただいた傘下の主要なグループ会社単体についてご回答ください。
- ・企業が事業活動を行う中で、販売履歴データ、機器の稼働データ、顧客や従業員のデータ、個人の行動データ、気象データ、交通量データ等、様々なデータが発生します。本アンケート調査では、これらのデータを生成、やり取り・通信、蓄積、分析した後、自社の製品やサービスの開発、生産性の向上、組織の課題解決といった事業活動に活かすことを「データ利活用」といいます。データ利活用は1社単独で行う場合もあれば、複数組織と連携する場合もあり、その形態は事業によって多様です。なお、「データ」とは、電磁的記録（電子的方式、磁気的方式その他の人の知覚によっては認識することができない方式で作られる記録であって、電子計算機による情報処理の用に供されるもの）を指します。



- ・本アンケートにおいて、各用語は以下の通り定義します。なお、明確な数値や状況が把握できない場合は、ご回答者様や事業のご担当者様の推定によりご記入いただいても構いません。

用語	定義
組織についてのデータ	個々の組織において発生するデータで、個人と切り離されたデータを指します。 (例：自社の販売履歴データ、顧客組織の生産機器の稼働データ、提携組織の家畜の体温等)
個人についてのデータ	従業員や個人顧客等、個人に関するデータを指します。 (例：自社従業員の勤務時行動データ、提携組織の個人顧客のウェブサイト閲覧データ等)
自然現象・社会現象・金融・経済等についてのデータ	政府のデータを含め、気象等の自然現象についてのデータ、交通量や住宅の分布等の社会現象についてのデータ、景気や為替等の金融・経済についてのデータ等、「組織についてのデータ」「個人についてのデータ」以外のデータを指します。
個人情報	個人情報保護法上の個人情報を指します。 その個人情報とは、生存する個人に関する情報であつて、当該情報に含まれる氏名、生年月日その他の記述等により特定の個人を識別することができるもの（他の情報と容易に照合することができ、それにより特定の個人を識別することができることとなるものを含む）、個人識別符号が含まれるものを指します。 (例：氏名、顔写真、指紋、旅券番号、免許証番号等)
ビッグデータ	規模の大きさ (volume)、高い発生頻度 (velocity)、多様な形式 (variety) といった特徴から、従来の技術や分析手法では処理が難しいデータを指します。
データサイエンティスト	事業において競争優位性をもたらすために、データの収集・加工・分析に優れた専門性を発揮し、知見を引き出す人材を指します。
限定提供データ	不正競争防止法における「業として特定の者に提供する情報として電磁的方法（電子的方法、磁気的方法その他の知覚によっては認識することができない方法）により相当量蓄積され、及び管理されている技術上又は営業上の情報（秘密として管理されているものを除く）」を指します。
データの来歴管理	データの起源について追跡したりデータの変更が行われたプロセスや行為者、システム等について記録したりすることによって、データが不正に変更されていないか、正確に計算されたか等について管理することを指します。

問 2 貴社のデータ利活用の全般的な状況について

問 2 では、データ利活用に関する貴社の体制や活動、成果等、貴社のデータ利活用の全般的な状況についてお伺いします。

問 2-1 貴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はおおよそどの程度ですか。当てはまるもの 1 つに○を付けてください。実際に管理している方法ではなく、容量についてお答えください。

例えば、社外のサーバセンターで管理している場合であっても、データの総容量が数台のサーバで管理できる程度であれば「2. 数台のサーバで管理できる程度」を選択してください。

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

問 2-2 問 2-1 で回答した貴社に利用権限があるデータの総容量のうち、実際に利活用を行っているデータの容量はおおよそどの程度ですか。当てはまるもの 1 つに○を付けてください。

1. 20%未満
2. 20%以上～40%未満
3. 40%以上～60%未満
4. 60%以上～80%未満
5. 80%以上

問 2-3 貴社では、全社的なデータ利活用を推進する専門部門や担当者を設けていますか。当てはまるもの 1 つに○を付けてください。専門部門や担当者を設けている場合、担当者のおおよその人数をご記入ください。

1. 担当を設けていない
2. 既存部門の中から、全社的なデータ利活用を推進する担当者を定めている
→ 担当者 () 人
3. 全社的なデータ利活用を推進する専門部門を設置している
→ 担当者 () 人

問 2-4 貴社には、データサイエンティストはおおよそ何人いますか。当てはまるもの 1 つに○を付けてください。

1. 0 人
2. 1～10 人
3. 11～50 人
4. 51～100 人
5. 101 人以上

問 2-5 貴社ではデータ利活用に関する活動がどの程度行われていますか。それぞれ当てはまるもの1つに○を付けてください。

		全く そう 思わない	そう 思わ ない	ど ち ら と も い え ない	そ う 思 う	強 く そ う 思 う
戦略・方針 について	①データ利活用を積極的に推進している	1	2	3	4	5
	②社内全体で連携してデータ利活用を行うことを積極的に推進している	1	2	3	4	5
	③社外の組織と連携してデータ利活用を行うことを積極的に推進している	1	2	3	4	5
実施体制・ ツールにつ いて	④個人情報に該当するデータの利活用を行える体制が整備されている	1	2	3	4	5
	⑤ビッグデータの利活用を行える体制が整備されている	1	2	3	4	5
	⑥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	1	2	3	4	5
	⑦データのアクセス権限を共通化する等、社内全体で連携してデータ利活用を行う体制が整備されている	1	2	3	4	5
	⑧データのアクセスを権限に応じて制限する等、社外の組織と連携してデータ利活用を行う体制が整備されている	1	2	3	4	5
	⑨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）	1	2	3	4	5
スキル・知 識・人材に ついて	⑩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）	1	2	3	4	5
	⑪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）	1	2	3	4	5
	⑫データサイエンティスト等、高度なデータの処理・分析を行える人材を育成、雇用している	1	2	3	4	5
	⑬ディープラーニング等の高度なデータの処理・解析結果を理解し、事業活動に活かせる人材を育成、雇用している	1	2	3	4	5
データにつ いて	⑭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）	1	2	3	4	5
	⑮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）	1	2	3	4	5
	⑯事業活動に関連するあらゆる情報のデータ化を進めている	1	2	3	4	5
	⑰事業活動の目的や今後の展開に沿ってどのようなデータが有用か十分に吟味し、データを設計している	1	2	3	4	5
	⑱複数のデータを組み合わせられるよう、データを設計している	1	2	3	4	5
	⑲データの有用性を評価し、必要なデータを取捨選択する等、定期的にデータ設計の見直しを行っている	1	2	3	4	5

問 2-6 貴社におけるデータ利活用のこれまでの成果について、当てはまるもの1つに○を付けてください。

1. まだ成果は得られていない
2. 具体的成果（売上やコストダウン等）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（事業活動に役立つノウハウやアイデアの獲得等）が得られている
3. 具体的成果（売上やコストダウン等）は得られていないが、複数の事業で間接的成果（事業活動に役立つノウハウやアイデアの獲得等）が得られている
4. 少なくとも1つの事業で具体的成果（売上やコストダウン等）が得られている
5. 複数の事業で具体的成果（売上やコストダウン等）が得られている

問 2-7 貴社におけるデータ利活用全般の直近1年間（2019年度）の成果について、それぞれ当てはまるもの1つに○を付けてください。（貴社にとっての成果についてご回答ください。）

	全く そう 思わ ない	そ う 思 わ な い	ど ち ら と も い え な い	そ う 思 う	強 く そ う 思 う
①データ利活用によって、顧客維持が改善した	1	2	3	4	5
②データ利活用によって、売上高の伸びが改善した	1	2	3	4	5
③データ利活用によって、利益が改善した	1	2	3	4	5
④データ利活用によって、投資収益率（ROI）が改善した	1	2	3	4	5
⑤データ利活用によって、全体的な財務業績が改善した	1	2	3	4	5
⑥データ利活用によって、競合よりもはやく新しい市場に参入した	1	2	3	4	5
⑦データ利活用によって、競合よりもはやく新しい製品やサービスを市場に導入した	1	2	3	4	5
⑧データ利活用によって、競合よりも新製品や新サービスの高い成功率を達成した	1	2	3	4	5
⑨データ利活用によって、マーケットシェアが向上した	1	2	3	4	5

問 2-8 貴社におけるデータ利活用の今後の方針について、当てはまるもの1つに○を付けてください。

1. 縮小する見通し
2. 現状維持
3. 拡大する見通し

問 2-9 問 2 でご回答いただきました「貴社のデータ利活用の全般的な状況」について、お気づきの点やご意見、回答に関する補足点等がありましたら、ご自由にお書きください。

問3 貴社のデータを貴社内で利活用する場合について

問3では、自社で保持している（自社で生成あるいは購入して保持している）データを自社内において利活用する場合についてお伺いします。

問3-1 貴社で保持しているデータのうち、直近1年間（2019年度）で貴社内で利活用していたデータの種類として当てはまるものすべてに○を付けてください。いずれのデータも利活用していない場合は、「4. 利活用しているデータはない」に○を付けてください。

- | |
|---|
| 1. 組織についてのデータ
2. 個人についてのデータ
3. 自然現象・社会現象・金融・経済等についてのデータ
4. 利活用しているデータはない |
|---|

問3-2 直近1年間（2019年度）において貴社で保持しているデータを貴社内で利活用していた目的について、それぞれ当てはまるもの1つに○を付けてください。目的としていた場合はその成果の程度として「4. 目標以上に達成した」「3. 目標を達成した」「2. 目標を達成していない」のいずれかに、目的としていなかった場合は「1. 目的としていない」に○を付けてください。

	目的としていない	目的としている		
		目標を達成していない	目標を達成した	目標以上に達成した
①組織改革や経営戦略の策定（付加価値の高い業務への集約、組織運営、全社戦略の策定等）	1	2	3	4
②人材の育成や能力開発	1	2	3	4
③生産性向上や業務効率化（生産量の最適化、生産方法の改良、業務サポート等）	1	2	3	4
④既存商品の改良（既存の製品・サービスの改良や高品質化、低価格化、メンテナンス精度の向上等）	1	2	3	4
⑤新商品の開発（新製品の開発、新サービスの提案等）	1	2	3	4
⑥業界や市場の分析（市場動向の分析、競合他社の調査等）	1	2	3	4
⑦CRM（customer relationship management：顧客関係管理）の向上（顧客満足度の向上、顧客接点の充実、営業支援等）	1	2	3	4
⑧他組織との連携構築（業務連携、業界内や商流内でデータを共有することによるビッグデータ化、標準や規格への対応等）	1	2	3	4

問3-3 問3でご回答いただきました「貴社のデータを貴社内で利活用する場合」について、お気づきの点やご意見、回答に関する補足点等がありましたら、ご自由にお書きください。

--

問 4 データ取得プロジェクトについて

問 4 では、データ活用を行うことを主として、他組織（他社（グループ会社を含む）、大学、研究機関、公的機関等、自社以外の組織）が保持しているデータを受領し、提携契約を結んだプロジェクト（業務提携や共同研究、委託等を指し、単にデータを購入した場合は除く）についてお伺いします。

以下、「データ取得プロジェクト」と呼びます。

問 4-1 貴社では、これまでにデータ取得プロジェクトはありましたか。当てはまるもの 1 つに○を付けてください。

1. ある → 問 4-2 へ
2. ない → 問 5 へ

問 4-2 貴社では、これまでにデータ取得プロジェクトはおおよそ何件ありましたか。当てはまるもの 1 つに○を付けてください。

1. 1～5 件
2. 6～10 件
3. 11～50 件
4. 51～100 件
5. 101 件以上

問 4-3 問 4-2 で回答したプロジェクトの件数のうち、直近 1 年間（2019 年度）で提携契約を結んだデータ取得プロジェクトはおおよそ何件ありましたか。当てはまるもの 1 つに○を付けてください。

- | | |
|-------------|------------|
| 1. 0 件 | → 問 4-10 へ |
| 2. 1～5 件 | → 問 4-4 へ |
| 3. 6～10 件 | |
| 4. 11～50 件 | |
| 5. 51～100 件 | |
| 6. 101 件以上 | |

問 4-4 問 4-3 で回答した直近 1 年間（2019 年度）のデータ取得プロジェクトの件数のうち、相対（貴社と他 1 組織との 2 組織間）で提携契約を結んだプロジェクトの割合はどの程度ですか。当てはまるもの 1 つに○を付けてください。

1. 20%未満
2. 20%以上～40%未満
3. 40%以上～60%未満
4. 60%以上～80%未満
5. 80%以上

問 4-5 問 4-3 で回答した直近 1 年間（2019 年度）のデータ取得プロジェクトの件数のうち、PoC 段階のプロジェクトの割合はどの程度ですか。当てはまるもの 1 つに○を付けてください。

PoC とは、Proof of Concept（概念実証。戦略仮説・コンセプトの検証工程）を指します。

1. 20%未満
2. 20%以上～40%未満
3. 40%以上～60%未満
4. 60%以上～80%未満
5. 80%以上

問 4-6 問 4-3 で回答した直近 1 年間（2019 年度）のデータ取得プロジェクトにおいて、提携契約を結んだ相手組織の属性と所在する地域について、当てはまるものすべてに○を付けてください。

	日本	中国（香港含む）	その他アジア	米国	ヨーロッパ	その他の地域
①グループ会社	1	2	3	4	5	6
②供給業者	1	2	3	4	5	6
③顧客またはクライアント	1	2	3	4	5	6
④競合他社	1	2	3	4	5	6
⑤民間のリサーチ会社等、その他の企業	1	2	3	4	5	6
⑥大学、研究機関、公的機関等、その他の組織	1	2	3	4	5	6

問 4-7 問 4-3 で回答した直近 1 年間（2019 年度）のデータ取得プロジェクトにおいて、受領したデータの種類として当てはまるものすべてに○を付けてください。

1. 組織についてのデータ
2. 個人についてのデータ
3. 自然現象・社会現象・金融・経済等についてのデータ

問 4-8 問 4-3 で回答した直近 1 年間（2019 年度）のデータ取得プロジェクトにおいて、限定提供データとして受領したデータはどの程度ですか。当てはまるもの 1 つに○を付けてください。

1. 全くない
2. 一部のデータ
3. 半分程度のデータ
4. ほとんどすべてのデータ

問 4-9 問 4-3 で回答した直近 1 年間（2019 年度）のデータ取得プロジェクトにおける貴社にとっての目的について、それぞれ当てはまるもの 1 つに○を付けてください。目的としていた場合はその成果の程度として「4. 目標以上に達成した」「3. 目標を達成した」「2. 目標を達成していない」のいずれかに、目的としていなかった場合は「1. 目的としていない」に○を付けてください。

	目的としていない	目的としている		
		目標を達成していない	目標を達成した	目標以上に達成した
①組織改革や経営戦略の策定（付加価値の高い業務への集約、組織運営、全社戦略の策定等）	1	2	3	4
②人材の育成や能力開発	1	2	3	4
③生産性向上や業務効率化（生産量の最適化、生産方法の改良、業務サポート等）	1	2	3	4
④既存商品の改良（既存の製品・サービスの改良や高品質化、低価格化、メンテナンス精度の向上等）	1	2	3	4
⑤新商品の開発（新製品の開発、新サービスの提案等）	1	2	3	4
⑥業界や市場の分析（市場動向の分析、競合他社の調査等）	1	2	3	4
⑦CRM（customer relationship management：顧客関係管理）の向上（顧客満足度の向上、顧客接点の充実、営業支援等）	1	2	3	4
⑧他組織との連携構築（業務連携、業界内や商流内でデータを共有することによるビッグデータ化、標準や規格への対応等）	1	2	3	4
⑨データの蓄積や分析等、データ利活用を行うこと自体を製品・サービスとする売上獲得	1	2	3	4

問 4-10 貴社において、他組織からデータを受領してデータ取得プロジェクトを行うかどうかを判断する上で、リスクとして特に懸念することは何ですか。特に懸念することとして当てはまるものすべてに○を付けてください。当てはまるものがない場合は「15. 特になし」に○を付けてください。

データについて	1. 受領するデータの希少価値が高い（同様のデータは容易に入手できない、代替できるデータが他にない等） 2. 受領するデータが個人情報を含んでいる 3. 受領するデータが重要な営業秘密を含んでいる 4. 受領するデータの来歴管理が十分になされていない 5. データの国際間取引に対する懸念がある
提携契約を結ぶ相手組織との関係について	6. 相手組織との関係において契約書によるガバナンスができない（契約書において詳細に取り決めができない、口頭ではなく書面の契約書を交わすことができない等） 7. 相手組織による投資が十分でない（金銭的負担が見合っていない、相手組織の人員や時間が十分に割かれぬ等） 8. 相手組織にプロジェクトを遂行するのに十分な専門性や能力がない 9. 相手組織が信頼できない（約束を遵守しない、公平でない等） 10. 相手組織と十分な情報交換やコミュニケーションができない 11. 相手組織と過去に取引の経験がない 12. 相手組織の評判が良くない 13. 相手組織の情報セキュリティが十分でない 14. 相手組織との関係において自社の交渉力（バーゲニングパワー）が弱い
	15. 特になし

問 4-11 貴社では、データ取得プロジェクトを行う場合の提携契約を結ぶ相手組織との契約書のひな型がありますか。当てはまるもの1つに○を付けてください。

1. 契約書のひな型はない
2. 契約書のひな型を作成している途中である
3. すでに契約書のひな型はあるが、それを使いこなしていない
4. すでに契約書のひな型があり、それを使いこなしている

問 4-12 貴社では、データ取得プロジェクトを行う場合、データ利活用を行うことを主としない他組織との提携プロジェクトの契約（一般的な業務委託契約、共同研究契約、調達購買契約等）と比較して、その契約書はかなりカスタマイズする必要があり、相当な法律業務が必要です。当てはまるもの1つに○を付けてください。

1. 全くそう思わない
2. そう思わない
3. どちらともいえない
4. そう思う
5. 強くそう思う

問 4-13 貴社では、海外からデータを受領してデータ取得プロジェクトを行う場合、国内のデータを受領する場合と比較して、その契約書はかなりカスタマイズする必要があり、相当な法律業務が必要だと思いますか。当てはまるもの1つに○を付けてください。

1. 国内のデータを受領する場合よりも必要ないと思う
2. 国内のデータを受領する場合よりもやや必要ないと思う
3. 国内のデータを受領する場合と変わらない
4. 国内のデータを受領する場合よりもやや必要だと思う
5. 国内のデータを受領する場合よりも必要だと思う

問 4-14 貴社では、データ取得プロジェクトを行う場合、データの来歴管理やそれに関連する法的責任についてどの部門（担当者）で検討していますか。検討している部門（担当者）として当てはまるものすべてに○を付けてください。どの部門（担当者）でも検討していない場合やわからない場合は「6. 検討していない、わからない」に○を付けてください。

1. 情報システム部門等のデータを処理する部門（担当者）
2. AI 推進部門等のデータ利活用を総括する部門（担当者）
3. 知財・法務部門（担当者）
4. 各プロジェクトを担う事業部門（担当者）
5. その他（ ）
6. 検討していない、わからない

問 4-15 貴社では、データ取得プロジェクトを行う場合に、知財・法務部門（担当者）はどの段階から関与していることが多いですか。最も当てはまるもの1つに○を付けてください。

1. 知財・法務部門（担当者）が関与することはあまりない
2. 知的財産に関連する課題が生じたり契約交渉が必要になったりする段階から関与していることが多い
3. 事業の初期段階（事業計画や提案等の段階）から関与していることが多い

問 4-16 問 4 でご回答いただきました「データ取得プロジェクト」について、お気づきの点やご意見、回答に関する補足点等がありましたら、ご自由にお書きください。

問5 データ提供プロジェクトについて

問5では、データ活用を行うことを主として、自社で保持しているデータを他組織（他社（グループ会社を含む）、大学、研究機関、公的機関等、自社以外の組織）に提供し、提携契約を結んだプロジェクト（業務提携や共同研究、委託等を指し、単にデータを販売した場合は除く）についてお伺いします。以下、「データ提供プロジェクト」と呼びます。

問 5-1 貴社では、これまでにデータ提供プロジェクトはありましたか。当てはまるもの1つに○を付けてください。

1. ある → 問5-2 へ
2. ない → アンケートは以上です。ご協力ありがとうございました。

問 5-2 貴社では、これまでにデータ提供プロジェクトはおおよそ何件ありましたか。当てはまるもの1つに○を付けてください。

1. 1～5 件
2. 6～10 件
3. 11～50 件
4. 51～100 件
5. 101 件以上

問 5-3 問5-2 で回答したプロジェクトの件数のうち、直近1年間（2019年度）で提携契約を結んだデータ提供プロジェクトはおおよそ何件ありましたか。当てはまるもの1つに○を付けてください。

- | | |
|-------------|-----------|
| 1. 0 件 | → 問5-10 へ |
| 2. 1～5 件 | → 問5-4 へ |
| 3. 6～10 件 | |
| 4. 11～50 件 | |
| 5. 51～100 件 | |
| 6. 101 件以上 | |

問 5-4 問5-3 で回答した直近1年間（2019年度）のデータ提供プロジェクトの件数のうち、相対（貴社と他1組織との2組織間）で提携契約を結んだプロジェクトの割合はどの程度ですか。当てはまるもの1つに○を付けてください。

1. 20%未満
2. 20%以上～40%未満
3. 40%以上～60%未満
4. 60%以上～80%未満
5. 80%以上

問 5-5 問5-3 で回答した直近1年間（2019年度）のデータ提供プロジェクトの件数のうち、PoC 段階のプロジェクトの割合はどの程度ですか。当てはまるもの1つに○を付けてください。

PoC とは、Proof of Concept（概念実証。戦略仮説・コンセプトの検証工程）を指します。

1. 20%未満
2. 20%以上～40%未満
3. 40%以上～60%未満
4. 60%以上～80%未満
5. 80%以上

問 5-6 問 5-3 で回答した直近 1 年間（2019 年度）のデータ提供プロジェクトにおいて、提携契約を結んだ相手組織の属性と所在する地域について、当てはまるものすべてに○を付けてください。

	日本	中国（香港含む）	その他アジア	米国	ヨーロッパ	その他の地域
①グループ会社	1	2	3	4	5	6
②供給業者	1	2	3	4	5	6
③顧客またはクライアント	1	2	3	4	5	6
④競合他社	1	2	3	4	5	6
⑤民間のリサーチ会社等、その他の企業	1	2	3	4	5	6
⑥大学、研究機関、公的機関等、その他の組織	1	2	3	4	5	6

問 5-7 問 5-3 で回答した直近 1 年間（2019 年度）のデータ提供プロジェクトにおいて、提供したデータの種類として当てはまるものすべてに○を付けてください。

1. 組織についてのデータ
2. 個人についてのデータ
3. 自然現象・社会現象・金融・経済等についてのデータ

問 5-8 問 5-3 で回答した直近 1 年間（2019 年度）のデータ提供プロジェクトにおいて、限定提供データとして提供したデータはどの程度ですか。当てはまるもの 1 つに○を付けてください。

1. 全くない
2. 一部のデータ
3. 半分程度のデータ
4. ほとんどすべてのデータ

問 5-9 問 5-3 で回答した直近 1 年間（2019 年度）のデータ提供プロジェクトにおける貴社にとっての目的について、それぞれ当てはまるもの 1 つに○を付けてください。目的としていた場合はその成果の程度として「4. 目標以上に達成した」「3. 目標を達成した」「2. 目標を達成していない」のいずれかに、目的としていなかった場合は「1. 目的としていない」に○を付けてください。

	目的としていない	目的としている		
		目標を達成していない	目標を達成した	目標以上に達成した
①組織改革や経営戦略の策定（付加価値の高い業務への集約、組織運営、全社戦略の策定等）	1	2	3	4
②人材の育成や能力開発	1	2	3	4
③生産性向上や業務効率化（生産量の最適化、生産方法の改良、業務サポート等）	1	2	3	4
④既存商品の改良（既存の製品・サービスの改良や高品質化、低価格化、メンテナンス精度の向上等）	1	2	3	4
⑤新商品の開発（新製品の開発、新サービスの提案等）	1	2	3	4
⑥業界や市場の分析（市場動向の分析、競合他社の調査等）	1	2	3	4
⑦CRM(customer relationship management:顧客関係管理)の向上（顧客満足度の向上、顧客接点の充実、営業支援等）	1	2	3	4
⑧他組織との連携構築（業務連携、業界内や商流内でデータを共有することによるビッグデータ化、標準や規格への対応等）	1	2	3	4

問 5-10 貴社において、他組織に自社のデータを提供してデータ提供プロジェクトを行うかどうかを判断する上で、リスクとして特に懸念することは何ですか。特に懸念することとして当てはまるものすべてに○を付けてください。当てはまるものがない場合は「15. 特になし」に○を付けてください。

データについて	1. 提供するデータの希少価値が高い（同様のデータは容易に入手できない、代替できるデータが他にない等） 2. 提供するデータが個人情報を含んでいる 3. 提供するデータが重要な営業秘密を含んでいる 4. 提供するデータの来歴管理が十分になされていない 5. データの国際間取引に対する懸念がある
提携契約を結ぶ相手組織との関係について	6. 相手組織との関係において契約書によるガバナンスができない（契約書において詳細に取り決めができない、口頭ではなく書面の契約書を交わすことができない等） 7. 相手組織による投資が十分でない（金銭的負担が見合っていない、相手組織の人員や時間が十分に割かれぬ等） 8. 相手組織にプロジェクトを遂行するのに十分な専門性や能力がない 9. 相手組織が信頼できない（約束を遵守しない、公平でない等） 10. 相手組織と十分な情報交換やコミュニケーションができない 11. 相手組織と過去に取引の経験がない 12. 相手組織の評判が良くない 13. 相手組織の情報セキュリティが十分でない 14. 相手組織との関係において自社の交渉力（バーゲニングパワー）が弱い
	15. 特になし

問 5-11 貴社では、データ提供プロジェクトを行う場合の提携契約を結ぶ相手組織との契約書のひな型がありますか。当てはまるもの1つに○を付けてください。

1. 契約書のひな型はない
2. 契約書のひな型を作成している途中である
3. すでに契約書のひな型はあるが、それを使いこなしていない
4. すでに契約書のひな型があり、それを使いこなしている

問 5-12 貴社では、データ提供プロジェクトを行う場合、データ利活用を行うことを主としない他組織との提携プロジェクトの契約（一般的な業務委託契約、共同研究契約、調達購買契約等）と比較して、その契約書はかなりカスタマイズする必要があり、相当な法律業務が必要です。当てはまるもの1つに○を付けてください。

1. 全くそう思わない
2. そう思わない
3. どちらともいえない
4. そう思う
5. 強くそう思う

問 5-13 貴社では、海外へデータを提供してデータ提供プロジェクトを行う場合、国内でデータを提供する場合と比較して、その契約書はかなりカスタマイズする必要があり、相当な法律業務が必要だと思いますか。当てはまるもの1つに○を付けてください。

1. 国内でデータを提供する場合よりも必要ないと思う
2. 国内でデータを提供する場合よりもやや必要ないと思う
3. 国内でデータを提供する場合と変わらない
4. 国内でデータを提供する場合よりもやや必要だと思う
5. 国内でデータを提供する場合よりも必要だと思う

問 5-14 貴社では、データ提供プロジェクトを行う場合、データの来歴管理やそれに関連する法的責任についてどの部門（担当者）で検討していますか。検討している部門（担当者）として当てはまるものすべてに○を付けてください。どの部門（担当者）でも検討していない場合やわからない場合は「6. 検討していない、わからない」に○を付けてください。

1. 情報システム部門等のデータを処理する部門（担当者）
2. AI 推進部門等のデータ利活用を総括する部門（担当者）
3. 知財・法務部門（担当者）
4. 各プロジェクトを担う事業部門（担当者）
5. その他（ ）
6. 検討していない、わからない

問 5-15 貴社では、データ提供プロジェクトを行う場合に、知財・法務部門（担当者）はどの段階から関与していることが多いですか。最も当てはまるもの1つに○を付けてください。

1. 知財・法務部門（担当者）が関与することはあまりない
2. 知的財産に関連する課題が生じたり契約交渉が必要になったりする段階から関与していることが多い
3. 事業の初期段階（事業計画や提案等の段階）から関与していることが多い

問 5-16 問 5 でご回答いただきました「データ提供プロジェクト」について、お気づきの点やご意見、回答に関する補足点等がありましたら、ご自由にお書きください。

ご協力ありがとうございました。