

プラットフォームとの関係性を踏まえた 日本企業のとるべきデータ利用戦略の展望

筑波大学ビジネスサイエンス系

立本博文

エコシステム型の産業構造への変化

産業環境の変化

1950年代

- 完全垂直統合企業の時代
- 中央研究所

1980年代

- 企業ネットワークの時代
- ・キャッチアップ国産産業の台頭
- ・クローズドな企業ネットワーク
- 戦略的アライアンス

1990年代

- 企業ネットワークの時代2
- ・シリコンバレーの台頭
- ・オープンな企業ネットワーク
- オープン標準

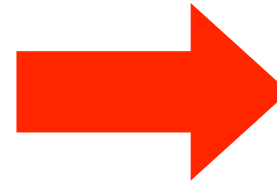
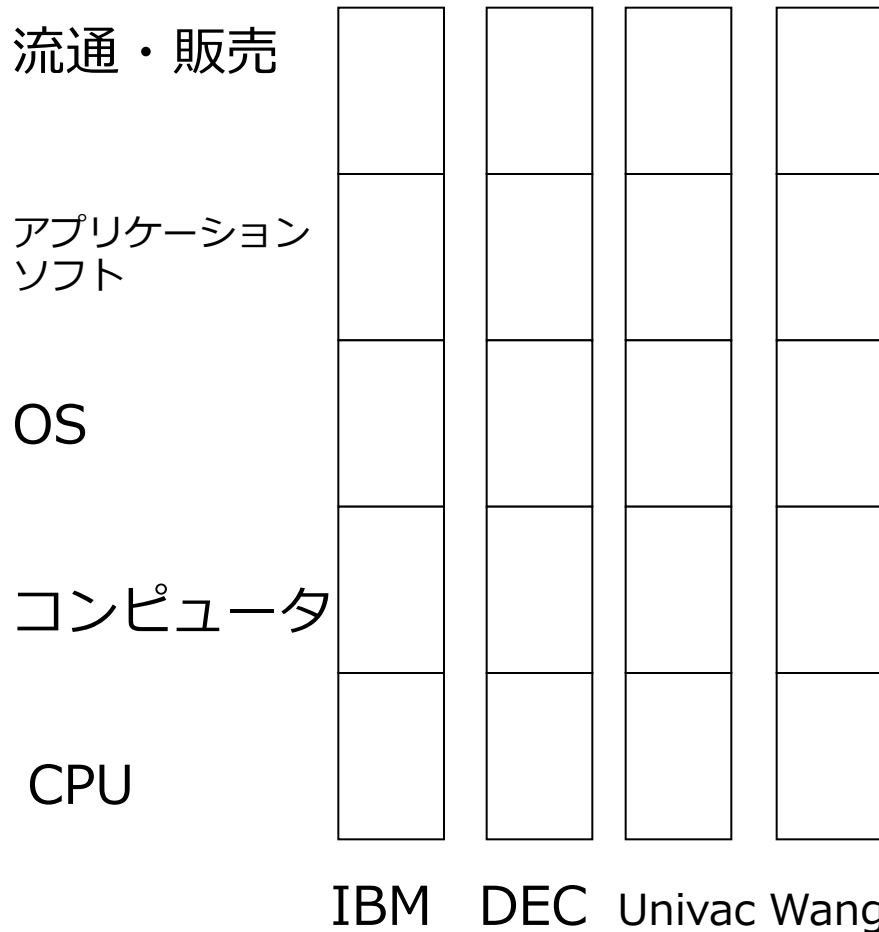
2000年代

- ビジネスエコシステムの時代
- ・オープンイノベーション
- プラットフォームビジネス

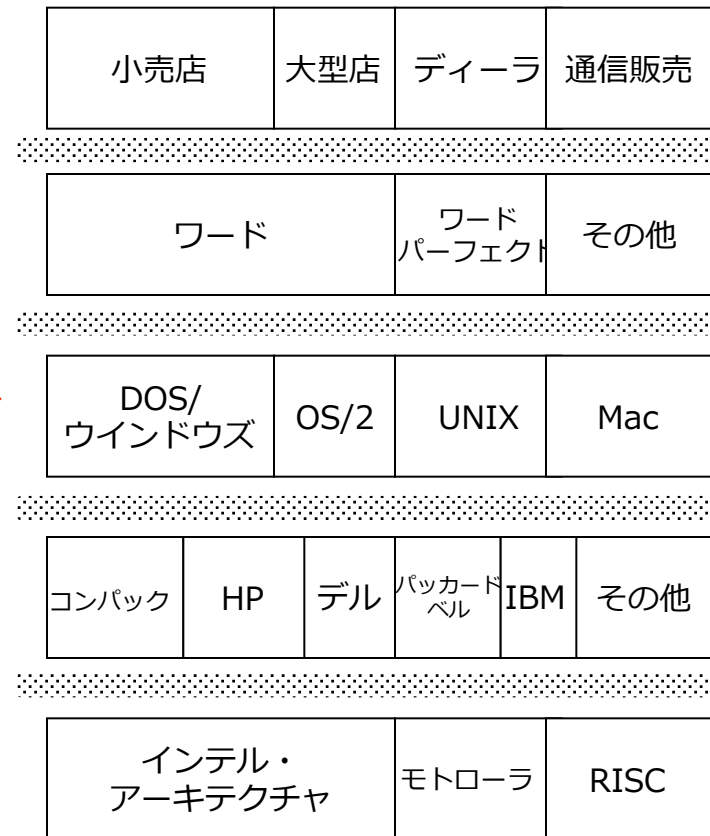
エコシステム型の産業：レイヤー型の産業構造

縦割り型コンピュータ産業
(1980年頃)

レイヤー型のコンピュータ産業
(1995年頃)



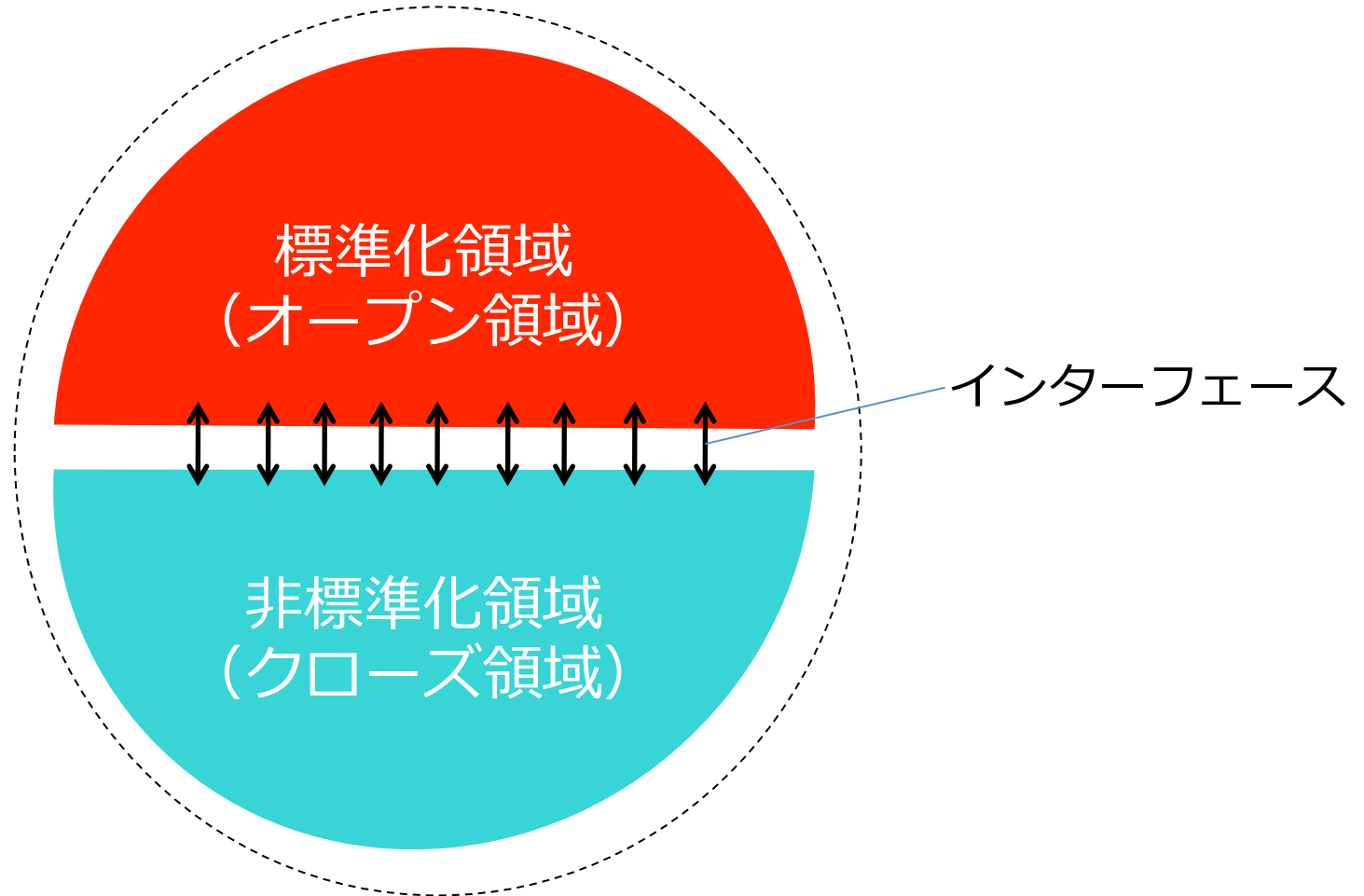
・水平分業化
・オープン化



..... オープン化

引用：Grove(1996)に筆者追記

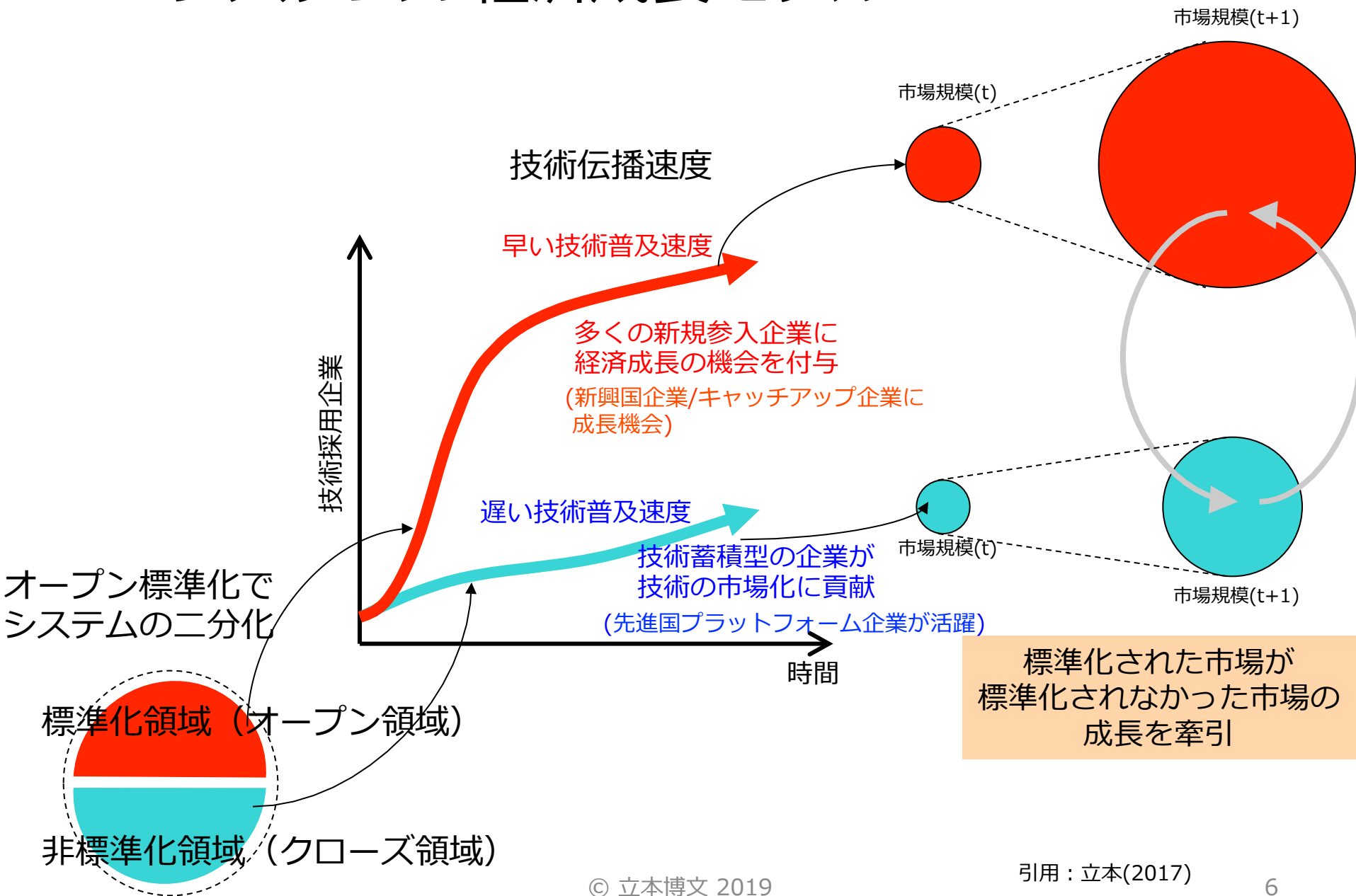
オープン標準化



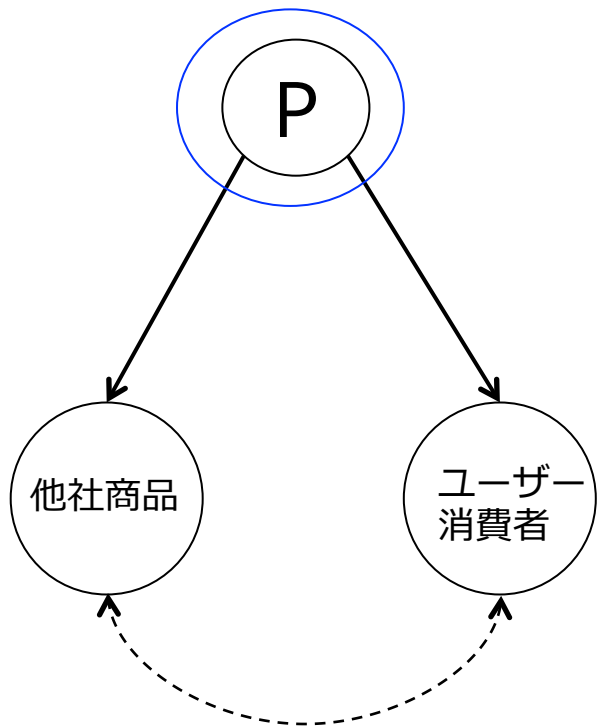
標準化 = 1つのシステムを2つに分割

エコシステムの経済成長モデル

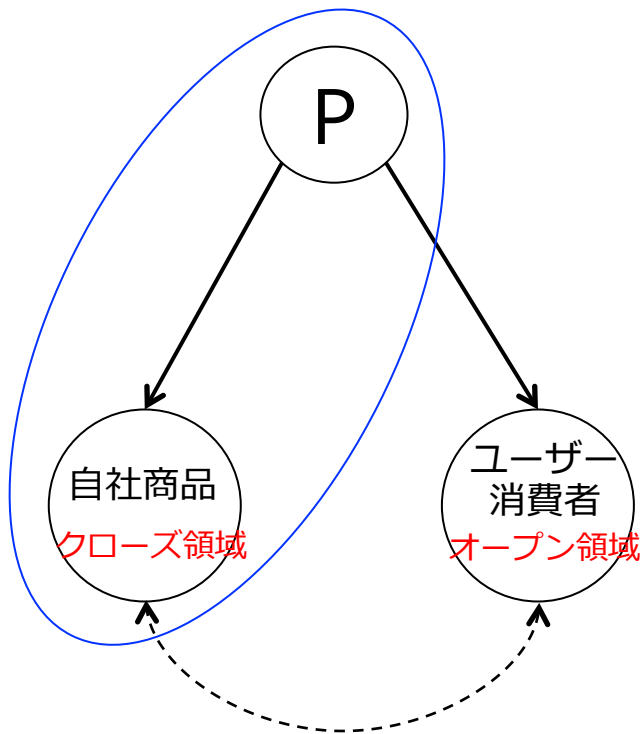
先進国企業と新興国企業の協業による急速な市場拡大



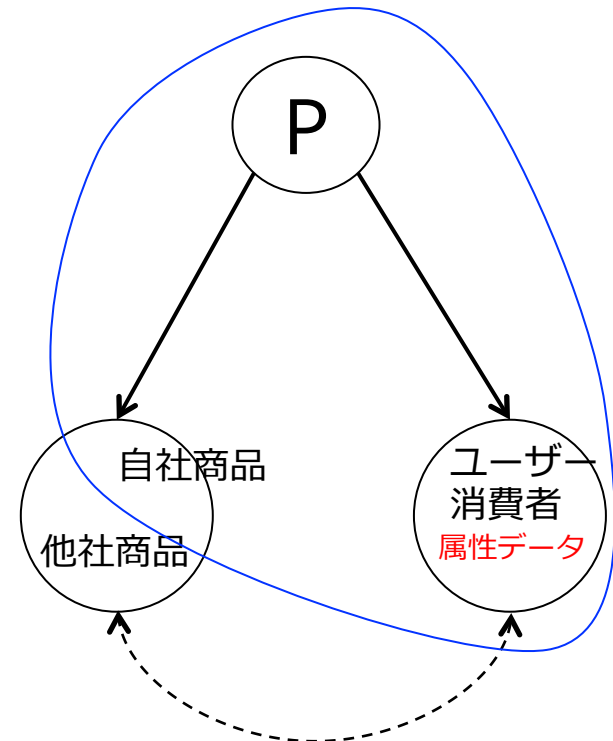
①第1形態
マッチング



②第2形態
オープン&クローズ



②第3形態
データプラットフォーム



Ⓟ プラットフォーム企業

↔ ネットワーク効果

○ プラットフォーム企業の収益源

ネットワーク効果
の発生メカニズム

メーカーと消費者の
組み合わせによって発生

オープン標準 (オープン化)
によって発生

消費者の
属性データによって発生

引用 : 立本(2019)

IoT/Big Data/AI

データでつながることを可能にした汎用技術フレームワーク

IoTデバイス

大量のデータを獲得する

Big Data

大量のデータを保存・抽出する

AI（機械学習）

大量のデータを使って高精度の
予測モデルを作成する

機械学習の典型的なシナリオ

【第1段階】 学習済モデルの作成

- 【テニス】 対戦相手データ
+
戦績データ
- 【店舗】 顧客属性データ
+
購買ブランドデータ
- 【タイタニック】 乗船客データ
+
生死データ

データセット
(教師データ)

学習済モデル

- 【アルゴリズムの選択】
 - ロジスティック回帰
 - SVM
 - ランダムフォレスト
 - ディープラーニング

【パラメタの推定】

アルゴリズム

【第2段階】 学習済モデルの利用

- 【テニス】 今期の対戦予定
- 【店舗】 来店客の属性
- 【タイタニック】 運行中の客船

データセット
(実際のデータ)

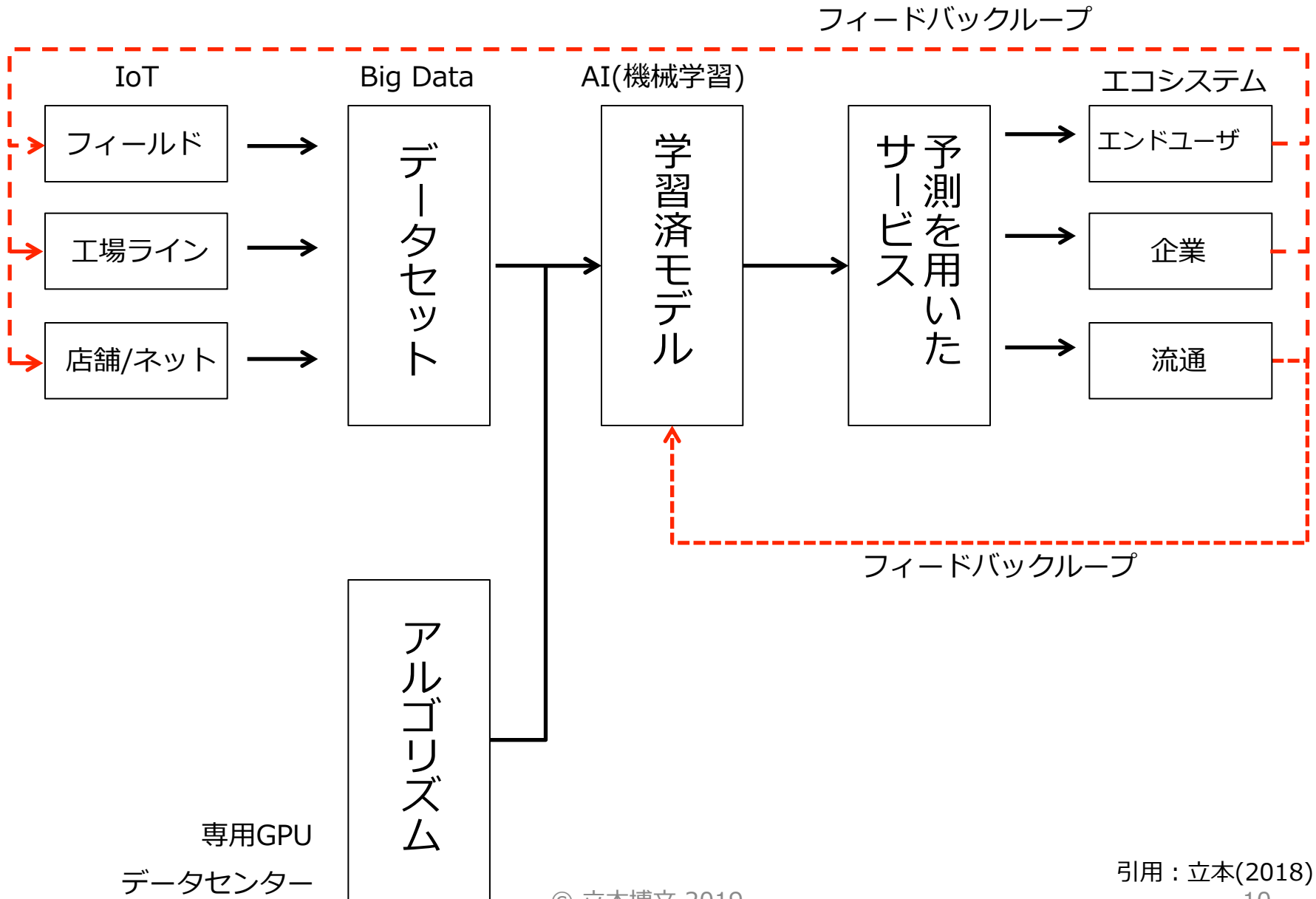
学習済モデル

高精度の予測

- 【テニス】 勝敗予想
- 【店舗】 購買ブランド予測
- 【タイタニック】 海難事故の生存確率

引用：立本(2018)

データ資源から見たバリューチェーン



プロジェクトの中でMLコードはごく一部

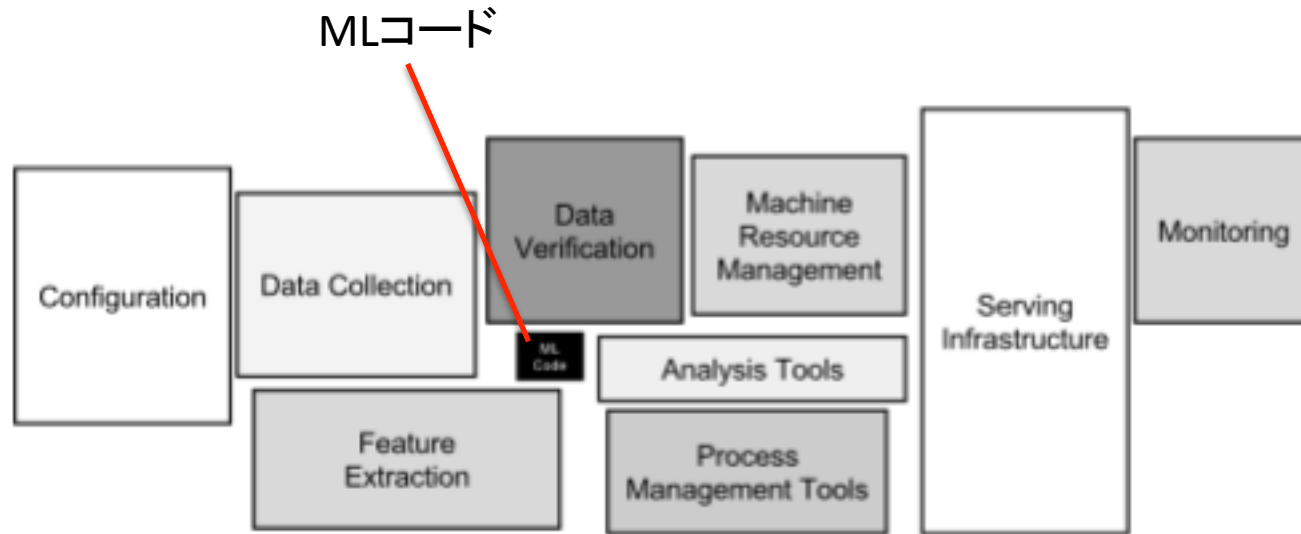


Figure 1: Only a small fraction of real-world ML systems is composed of the ML code, as shown by the small black box in the middle. The required surrounding infrastructure is vast and complex.

Sculley D et al.(2015) “Hidden technical debt in Machine learning systems”
In Proceeding NIPS'15 Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2
Pages 2503-2511

- 機械学習(ML)システムの中で、機械学習コードはごく一部
- 構成/データ取得/特徴量抽出/インフラ/分析/モニタリング

データの種類

- ①当該データが何を表すデータであるのかに着目した整理
 - ・個人データ（年齢、性別、住所、嗜好等の個人に関するデータ）
 - ・産業データ（機械の稼働状況、人や物の流れ等の事業活動に関するデータ）
 - ・公的データ（公的統計、観光、公共交通分野等に関するデータ）
- ②当該データが示す行動等が行われた場所に注目した整理
 - ・バーチャルデータ（Webサイト等のオンラインで集められたデータ）
 - ・リアルデータ（現実空間に設置されたセンサー等から集められたデータ）
- ③データの管理や提供の方法に着目した整理
 - ・**オープンデータ**（天気予報、公的統計等の公開されているデータ）
 - ・**クローズデータ**（個人データ、産業データ等の秘匿されているデータ）

データは原材料だけれども**オープンなデータは少ない**

データチェーンの整備

- IoT/Big Data/AIの文脈では、データは**原材料**
- データは「当然に知的財産権の対象」にはならない
(データは特許権,著作権等のような存在ではない)
- データは有体物ではなく、所有権等物権の対象にならない
(データを所有する、というのは法的にはいえない)
- データに関する債務関係について契約していると考えられる

(例) 工場に導入した工作機械に付随するセンサー由来のデータ
→ そのデータが、直ちに工場所有者のものである、とは言えない

データチェーンの整備：

データへのアクセスは（多くは）**当事者間の契約**による。

契約によってデータへのアクセスを確保する

例外はオープンデータへのアクセス

参考：阿久津(2019)

データ利活用に関する質問票調査を用いた 産業別比較

立本博文（筑波大学ビジネスサイエンス系）
平井祐理（東京大学政策ビジョン研究センター）
渡部俊也（東京大学政策ビジョン研究センター）

RIETIディスカッション・ペーパー 18-J-032

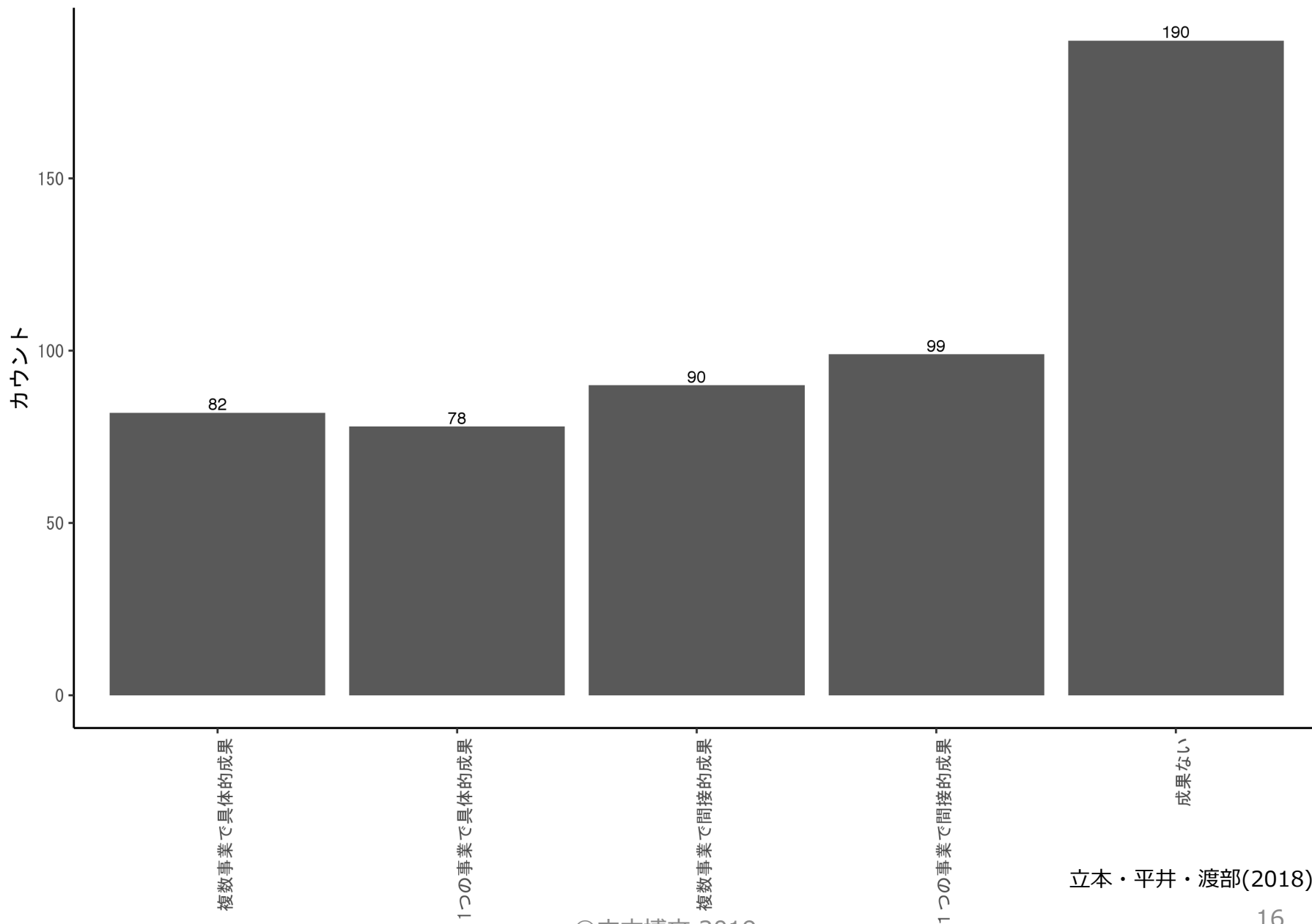
<https://www.rieti.go.jp/jp/publications/nts/18j032.html>

調査概要

- 日本企業のデータ利活用の実態把握。
 - どのような企業がどの程度データ利活用を行っているか。
 - どのようなデータを利活用に用いているか。
 - データ利活用を円滑にする要因は何か。
- 調査期間：2017年9月15日～2017年11月6日（11月27日回収分まで追加）
- 2016年度の状況について調査。

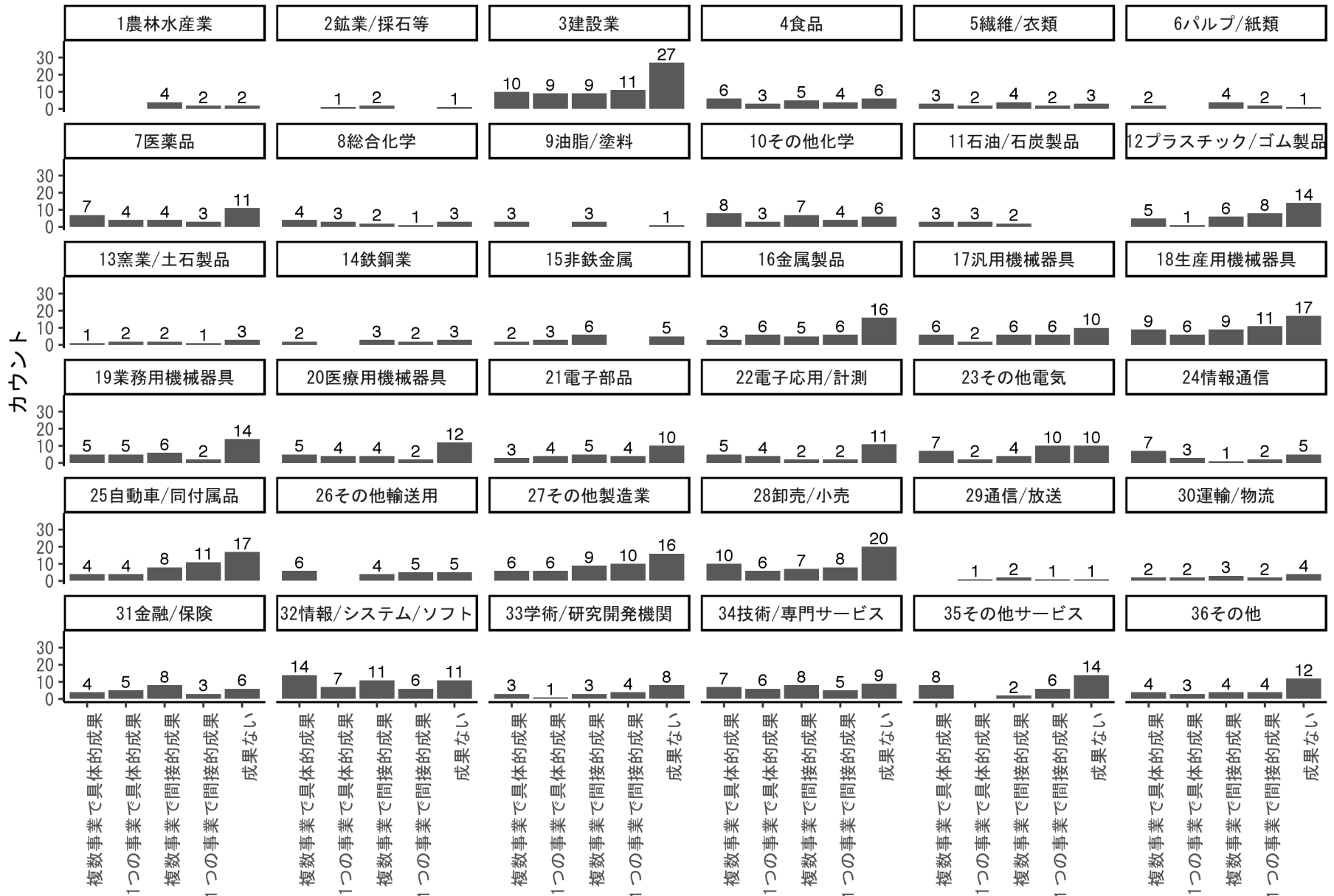
	調査対象	回収数	回収率	
全体	6278	567	9.0%	※うち無回答5件
特許出願件数上位	4621	461	10.0%	※有効回答のみ
東証一部	1170	70	6.0%	※有効回答のみ
WANTEDLY	487	31	6.4%	※有効回答のみ

問2-7:データ活用の成果がありますか？（全社レベル）

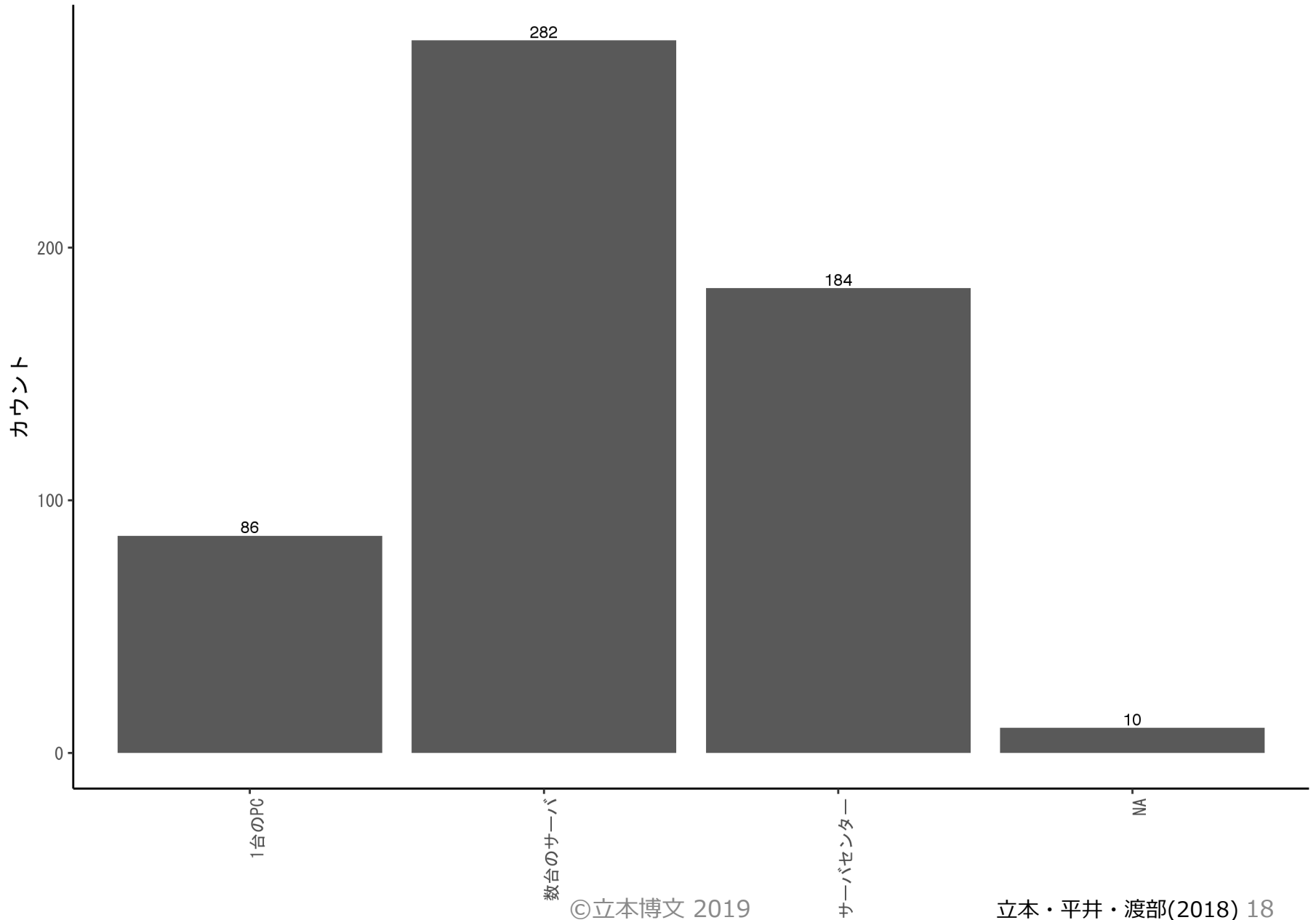


立本・平井・渡部(2018)

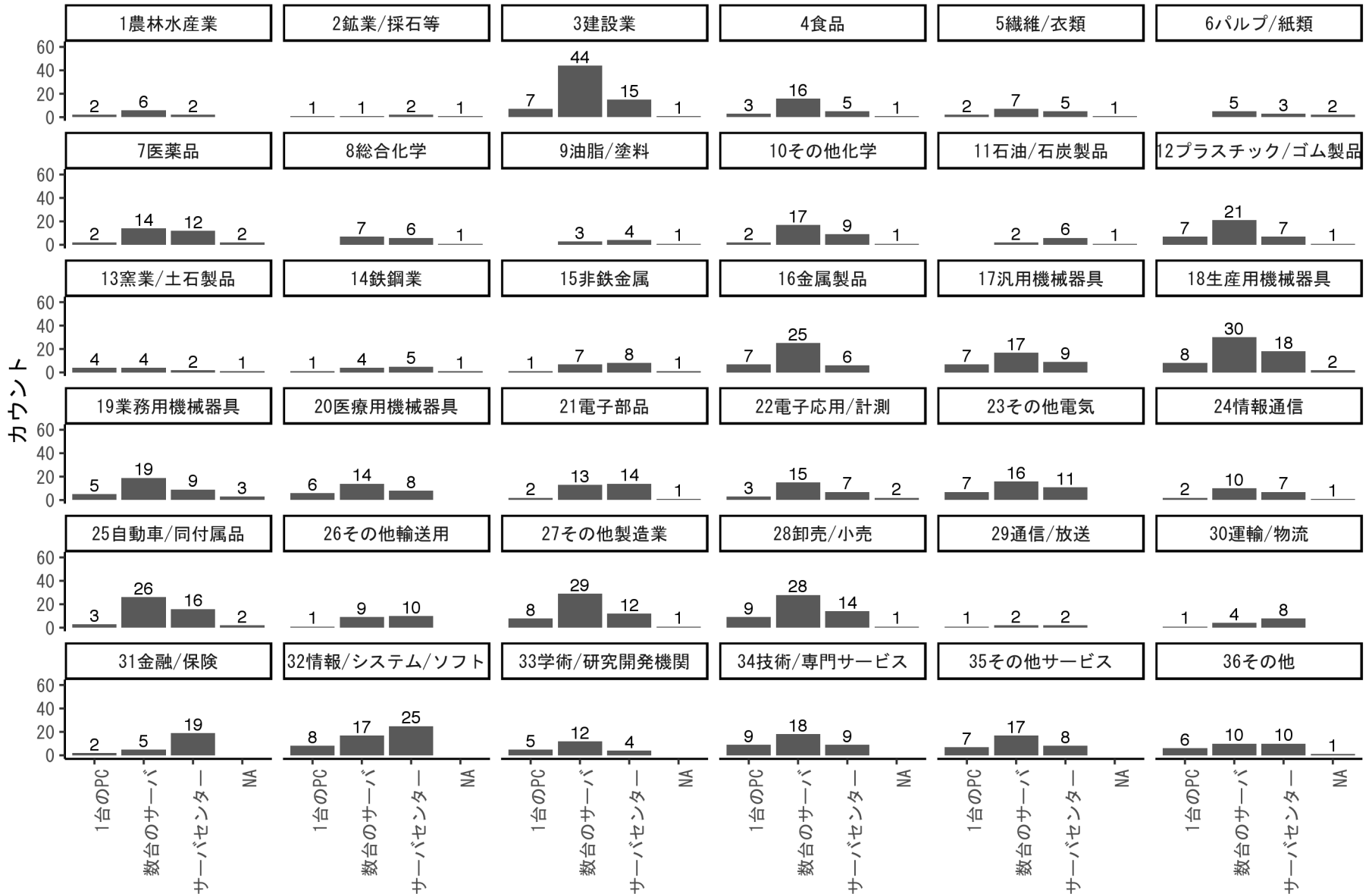
問2-7:成果について(産業毎)



問2-1:利活用を期待しているデータの総容量

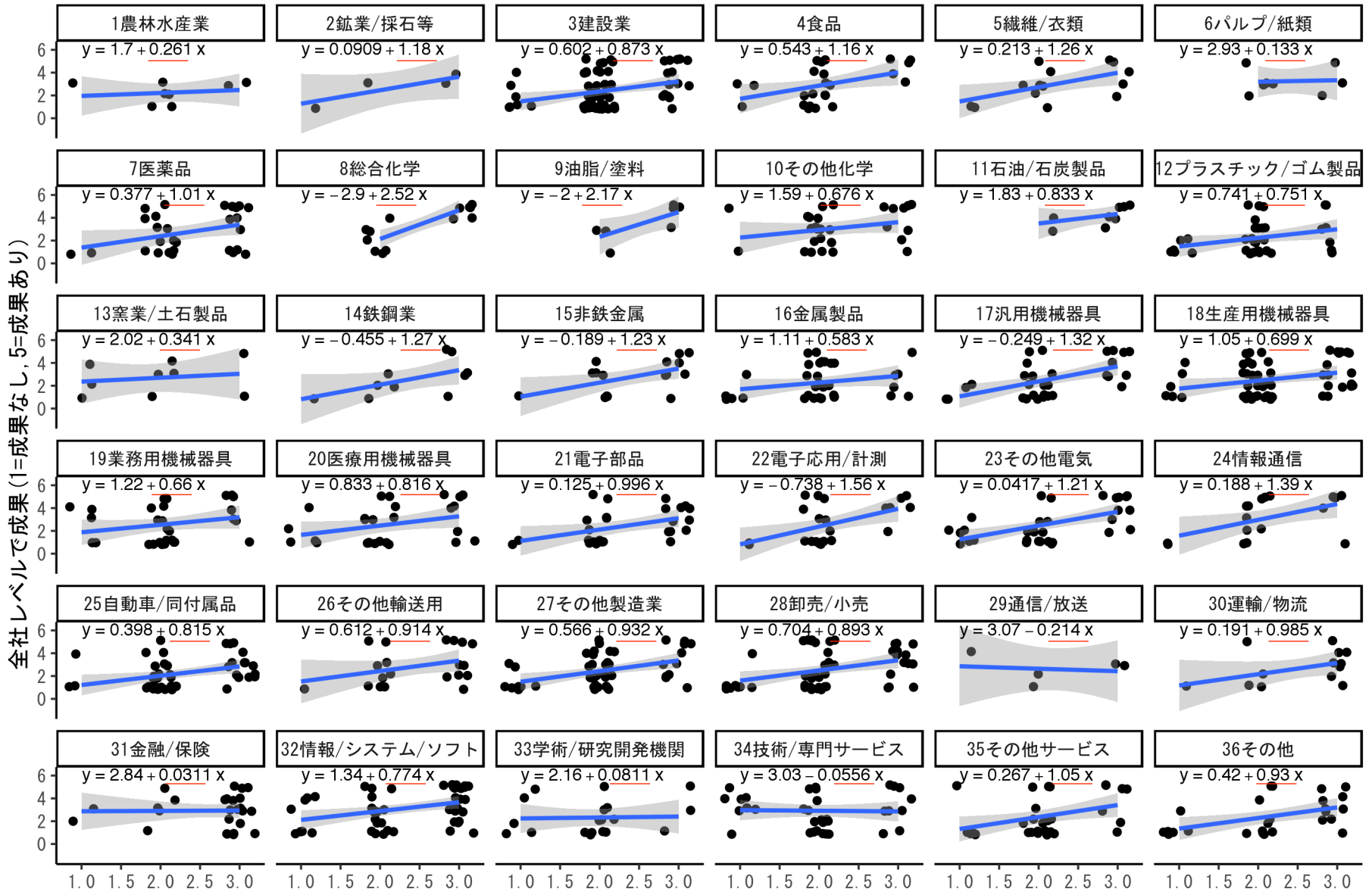


問2-1:利活用を期待しているデータの総容量(産業毎)



全社レベル成果へのデータ総容量の影響

— 各産業毎の影響係数



影響係数

出力

入力

$$Y = \beta_0 + \beta_1 X$$

β₁
影響係数

産業ごとに各設問の影響係数を算出

Yとして、全社レベルの成果項目である以下の1つの成果項目を用いた

問2-7(データ利活用成果)

Xとして、以下の9つの設問の回答を用いた

問2-1 (全社データ総量)

問2-2 (データ利用率)

問3-10 (事業データ総量)

問3-11 (ビックデータに該当)

問2-3 (担当者合計)

問2-4 (契約書)

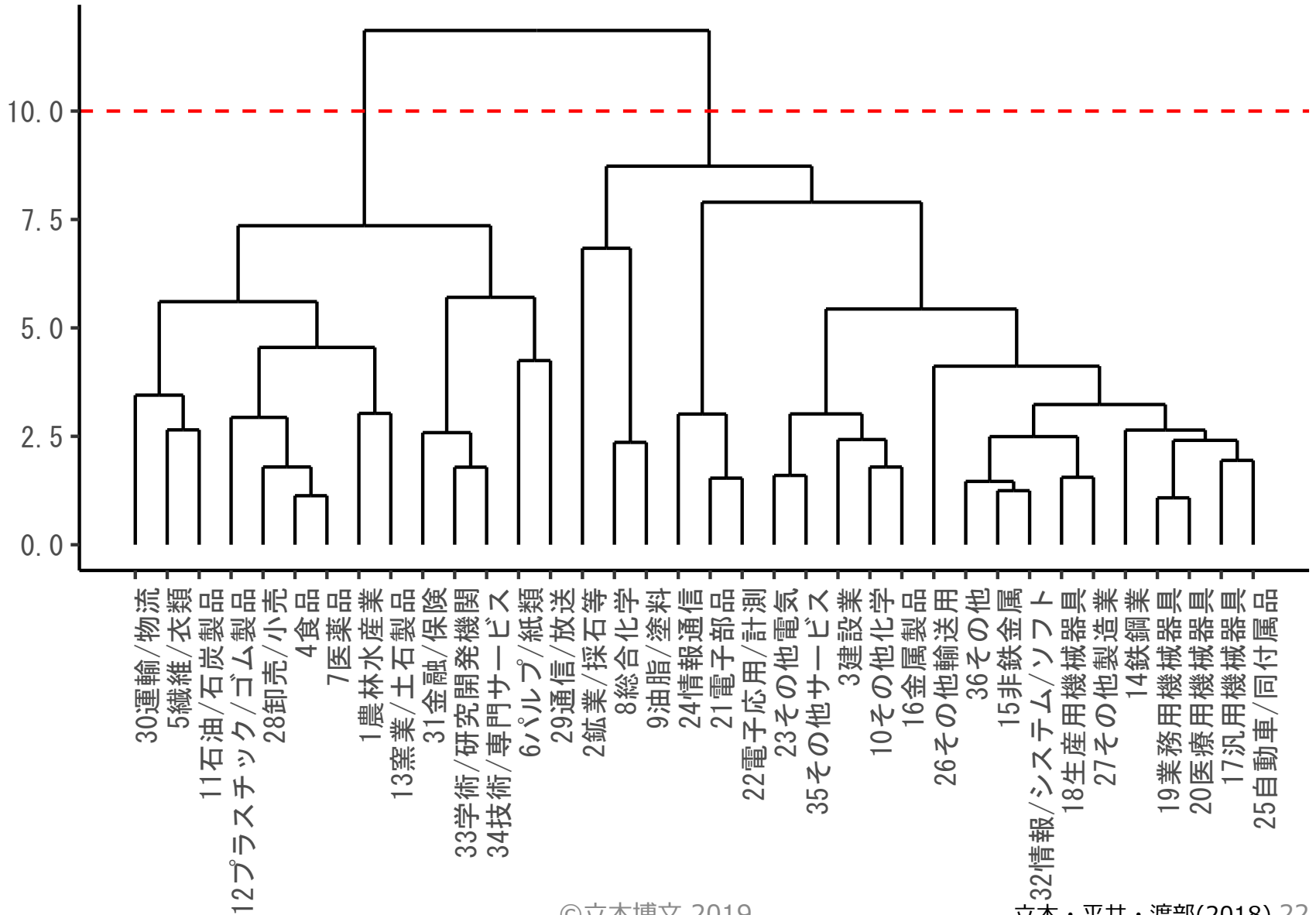
問3-14 (データ利活用経験年)

問3-17 (イニシアティブの割合/【自社】)

問3-18 (高度なデータの処理・解析)

クラスター分析の結果

9つの設問の影響係数（産業毎）に関してクラスター分析を行い、類似グループを推定



各クラスターに属する産業リスト

No.	業種ラベル	clustID	No.	業種ラベル	clustID
1	1農林水産業	1	1	2鉱業/採石等	2
2	4食品	1	2	3建設業	2
3	5繊維/衣類	1	3	8総合化学	2
4	6パルプ/紙類	1	4	9油脂/塗料	2
5	7医薬品	1	5	10その他化学	2
6	11石油/石炭製品	1	6	14鉄鋼業	2
7	12プラスチック/ゴム製品	1	7	15非鉄金属	2
8	13窯業/土石製品	1	8	16金属製品	2
9	28卸売/小売	1	9	17汎用機械器具	2
10	29通信/放送	1	10	18生産用機械器具	2
11	30運輸/物流	1	11	19業務用機械器具	2
12	31金融/保険	1	12	20医療用機械器具	2
13	33学術/研究開発機関	1	13	21電子部品	2
14	34技術/専門サービス	1	14	22電子応用/計測	2
			15	23その他電気	2
			16	24情報通信	2
			17	25自動車/同付属品	2
			18	26その他輸送用	2
			19	27その他製造業	2
			20	32情報/システム/ソフト	2
			21	35その他サービス	2
			22	36その他	2

©立本博文 2019

クラスター毎の影響係数のパターン

	クラスタ1	クラスタ2
産業数	14	22
企業数	299	681
(成果) Q2-7:データ利活用成果の平均(クラスタ毎)	2.722	2.624
Q2-1:全社データ総量の影響 ^{*1}	0.533	1.117
Q2-2:データ利用率の影響	0.373	0.461
Q2-3:担当者合計の影響	0.120	0.095
Q2-4:契約書の影響	0.410	0.510
Q3-10:事業データ総量の影響	0.311	1.061
Q3-11:ビッグデータに該当の影響	-0.139	1.202
Q3-14:データ利活用経験年の影響	0.240	0.327
Q3-17:イニシアティブの割合/【自社】の影響	-0.006	-0.013
Q3-18:高度なデータの処理・解析の影響	-0.114	1.774

*1 全社データ総量のデータ利活用成果への影響のこと

立本・平井・渡部(2018)

- ①クラスタ1,2ともに成果（「Q2-7.データ利活用成果の平均」）は同水準
- ②クラスタ1,2ともに「Q2-1.全社データ総量」「Q2-2.データ利用率」は正の影響
ただし「Q2-1.全社データ総量」の影響はクラスタ2の方がクラスタ1よりも大きい
- ③クラスタ2では「Q3-11.ビッグデータに該当」「Q3-18.高度なデータの処理・解析」
は正の影響。クラスタ1では負の影響

引用文献

Grove, S. A. (1996) *Only the Paranoid survive:How to Exploit the Crisis Points That Challenge Every Company and Career*, New York: Currency/Doubleday.

阿久津匡美「データ時代の研究・開発契約等のあり方ーその契約内容でデータをりよう
できますか」『研究 技術 計画』Vol.33, No.4,324-344, 2019.

経済産業省(2017)「第四次産業革命に向けた競争政策の在り方に関する研究会報告
書」<http://www.meti.go.jp/press/017/06/20170628001/20170628001-2.pdf>

立本博文『プラットフォーム企業のグローバル戦略：オープン標準の戦略的活用とビジ
ネス・エコシステム』総頁数414ページ, 有斐閣, 2017.

立本博文「IoT エコシステムの将来像」『研究 技術 計画』Vol.32, No.3, 279-292,
2017.

立本博文・平井祐理・渡部俊也「データ利活用に関する質問票調査を用いた産業別比
較」18-J-032, 経済産業研究所, 2018. [https://www.rieti.go.jp/jp/publications/
summary/18120001.html](https://www.rieti.go.jp/jp/publications/summary/18120001.html)

立本博文「エコシステム型の産業環境と知財マネジメント」知財管理, Vol69. No.4,
2019. (近刊)

ご静聴ありがとうございました