

データの多様な相貌

H20システムの中のデータサイエンス

福島真人

バブル、ハイプ、データサイエンス

映画『チューリップ・フィーバー——肖像画に秘めた愛』(1701七年)は、市場の活況に沸く一七世紀アムステルダムを舞台に、修道院で育った孤兎とイケメンの若手画家の恋愛を、チューリップを巡る急速なバブルとその崩壊に絡めて描いた作品である⁽¹⁾。この背景には世界最初のバブル経済と呼ばれる現象がある。

科学技術の分野でも、こうしたバブルが存在するという指摘は、科学技術社会学(STS)や、技術予測の分野等でもすでによくなされている。一般にこれは「ハイプ」「熱狂」「サイクル」という名で喧伝されている場合も多いが、テクノロジー初期段階での急

室の活動の詳細分析)(Latour & Woolgar 1979; 福島 2017; 鈴木 2017)のデータ(サイエンス)版ともいえる流れであるが⁽²⁾、こうしたいわばデータ形成への接写的アプローチとでもいってきものからみえてくるのは、データのエコシステムの複雑な相貌であり、(ビッグ)データ／データサイエンスとどう概念や領域がもつ根本的な曖昧さである。

ビッグデータのアプロトタイプ

心理学者のロッシュ(E. Rosch)らは、人が特定の概念を語る時、最もそれらしいと暗黙に想定する事例(アプロトタイプ)から出発すると指摘したが、これは例えば「鳥」と聞いて、ベンギンやダチョウを想定する人は少ないという意味である(Rosch & Lloyd 1978)。同様に、ビッグデータという言葉のアプロトタイプとして、ビジネス系雑誌のシンクデータ特集等で想定されているのは、スマホの追跡情報を使って人の流れを解析し、それを応用して新店舗の位置を決定するとか、Twitterの情報から新たな流行のトレンドを予測するといった応用例である。より専門的な事例たとえばタンパク構造のデータベースから創薬のための新たなシーザ化合物を探索するといったケースは、専門家用の会合(たとえばバイオ医療関係での講演等)で見聞きすることが多い。更にビッグデータを使って老朽化するインフラを救つたり、農業をビッグデータの力で飛躍させる等、その応用範囲の拡大は留まるところ

を知らないようにみえる⁽³⁾。

しかし、同じビッグデータやデータサイエンスという言葉を使つても、そこで語られているデータの内容、質、およびその解析方法はお互いかなり異なっている。データが殆ど自動的に得られるもの、その内容が、もともと手書きのものから、テキストや画像、特定の学術分野によって厳格に規定されたもの等、様々である。当然、それを解析する方法も実に多様であり、それぞれの分野に特化した処理方法が適用される必要がある。

このデータ関連の言説における、限定的なアプロトタイプと周辺事例の多様性という構造は、STSが注目してきた科学のアプロトタイプと実際の多様性の話と似た面もある。クーン(T. Kuhn)のパラダイム論は理論物理学をそのアプロトタイプとしているが(クーン 1971)、一九七〇年代後半から始まったSTSのラボラトリ研究等では、むしろバイオ系のラボ活動が集中的に研究され(Latour & Woolgar 1979他)、更に環境問題(地球温暖化やエネルギー問題)がそれに加わり、データ関係の問題についても、これらの領域での調査が全体の議論に大きな影響をあたえている。

これと並行して、九〇年代以降科学実践の多様性そのものを強調するアプローチも盛んになった。クノール・セチナ(K. Knorr-Cetina)の「認識的文化」(epistemic Cultures)論がその代表例だが、マウスを使った遺伝子研究と、欧洲原子核研究機構(CERN)における素粒子物理学の研究活動の違いを比較して、対象とされる現象へのアクセスやデータの分析(マウスは研究者が直接操作でき

激な期待の上昇と、失望への急落、そしてゆっくりとした回復という形で定式化されている。もちろん実際の事例では話ははるかに複雑ではあるが(Borup et al. 2006; 福島 2017; 山口 + 福島 2019)。」の図式からいえば、データ化する社会とそれをやるべきはハイプ的現象真っ只中という印象を受ける。

いうまでもなく、ビッグデータ／データサイエンスを巡つて膨れ上がる言説について、その光と影のどちらに着目するかで話はかなり変わつてくる。影の部分への指摘としては、プライバシーの侵害や、みえざる差別の問題といった議論が、最近GAFA批判等でよく知られているが、STSやその関連分野では、(ビッグ)データやデータサイエンスの実際のあり方をより中立的に観察分析するという活動も盛んである。これはラボラトリ研究(実験

るが、素粒子は数量的データとしてしか、その姿は確認できない)、業績の発表や研究者のキャリアパスのパターン(筆頭著者として名前が出るか、数百人のチームの中でアルファベット順の名簿に埋没するか)等、その対照的な姿が、科学実践の多様性の証左として強調されたのである(Knorr-Cetina 1999)。

データの多様なエロシズム

こうしてみると、現在メディアで喧伝されているビックデータ論は、データの多様な相貌の、ある一部の側面で全体を語つくるという印象を受ける。もちろん、ビッグデータ論者たちの中には、分野ごとの意味の違いを強調しているケースもある。例えば、リー(J. Lee)は、ウェブ上にある、大量で質が低いがその相関を瞬時に分析することでそれなりの利点があるデータと、より局所的、階層的で、その分野での固有な特性を十分に反映している必要がある産業関連データを区別し、後者を「インダストリアル・ビッグデータ」と呼んでいる(リー 2016)。こうした弁別が重要なのは、ビッグデータのハイプにともない、新たな専門職としてデータサイエンス(サイエンティスト)という新業種もまた、派手に喧伝されているからである。

樋口(2013)は、このデータサイエンスに関する系譜を、一九八〇—九〇年代のウォール街のクオンツ(数理分析家)、ヴァリアン(H. Varian)の間、一一世紀における最もセクシーな職とし

き関わり方を探るといった内容のシンポジウム、情報交換会は頻繁におこなわれている。しかし、それぞれの分野において、データサイエンスの必要性と同時に、それとの関係のむずかしさが指摘されることも珍しくない。例えば池田・森井(2018)は、天文学はすでに長いことデータ爆発状態にあり、データサイエンスとの協働は必須であると強調している。この著者の一人は、宇宙情報解析関係のシンポジウムで、近年その応用に注目があつまるスペースモーニング(日高他 2017)や機械学習の手法についてその宇宙科学への応用事例を紹介していた。興味深かったのは、宇宙科学者がデータサイエンティストにアプローチしてくる場合、プロジェクトの空白部分、あるいはその最終段階になつて、なんとかしてくれと泣きついてくることが多いが、それではうまくいかないという指摘である。データサイエンスは全体のプロジェクトの構造とかかわるので、やるならプロジェクトの最初から参加させてほしい、と言うわけである⁽⁶⁾。実際、特定領域の研究者たちが期待するデータサイエンス像と、データサイエンティストが具体的に提供できる手法にはギャップがある。解析手法の多様なレポートリーから、何を採用し、それに適した形でどうデータを集めめるかは、プロジェクトの最初から関与して全体を調整しないとうまくいかない。さらには当該分野で從来から用いられている分析手法と新しいそれをどううまく組み合わせるか、といった「技」も必要になつてくるのである。

大声で喧伝されるデータサイエンスの「汎用性」の主張と、そ

ての「統計学者」、そしてこれらに人工知能、機械学習等が合流し、「ハーバードビジネスレビュー」の一〇一四年「ビックデータ特集」におけるデータサイエンティスト論に至るという流れとして簡潔にまとめている。データサイエンティスト論に至るという流れとして簡潔も現在多く出版されているが、本邦ではその育成に関して諸外国に大きく遅れをとつてゐるといった議論も多い⁽⁷⁾。

データサイエンスへの期待の高まりには、過去における別領域のハイプと似た面もある。その一つがナノテクノロジーであるが、

こうした用語が人口に膚浅するようになり、政策担当者のため込みが顯著になつてくると、多くの分野がわれもわれもとの名称を採用するようになる。厳密な意味でのナノスケール(10^{-9} m)とは余り関係のないレベルの技術までナノテクノロジーを自称するようになり、表面的にはナノテクノロジストは急増することになるが、その実態は曖昧模糊としてくるのである(ベルバ 2009)。

実際、様々な研究分野において、データサイエンスとのあるべ

これが現実のデータ構造に適用される場面での状況依存的な性格について、STSの観点からより一般的に論じた議論としては、ラップス(D. Ríbes)らの「ドメイン(特定領域)」論がある。彼らは、人工知能、ソフトウェア開発、そして科学政策という複数の領域において、このドメインという観点が果たす役割を詳細に分析している(Ríbes et al. 2019)。生物学や宇宙科学固有の知識はドメイン知識であるが、こうした汎用性とドメインの複雑な関係は、STSで研究が盛んな、バイオインフォマティックスがその先駆的あり方を示しているともいえる。データ化の荒波は、データサイエンスが多種多様なドメインと接触することを意味し、ちょうど「バイオ・」という接頭語の代わりに、様々な別の言葉がそこにつく可能性を示している。これはデータサイエンス(サイエンティスト)に汎用性と局所的解法の両方を同時に要求する」となるが、言うは易く、おこなうは難しの典型的のような要求である。

知識インフラ論

そこで起こりうる多様な問題を、少し時間を遡つて、筆者が関わってきた事例を中心みてみよう。STSのバイオ系科学研究によつて先導された議論の一つが「知識インフラ」研究であり、データサイエンスの諸問題を考えるための多くの示唆がそこに含まれている。ここでいう知識インフラとは、研究用のデータベースやデータ管理といった、研究の情報基盤、およびその維持保全

のための諸活動を意味する (Star & Ruhleder 1996)。更に分類系統学 (systematics) (Hine 2008)、生物多様性 (Bowker 2005) といった分野についての分析も関連する領域として多く取り上げられてきた。

スター (S. Star) はこの知識インフラに着目したのは、ゲノム研究の発展により、バイオ研究分野への情報科学の参加が加速化し始めた時期である。彼らのインフラ概念は後の STS の関連研究に大きな影響を与えたが、その要点の一つは、インフラはその安定した稼働時では「透明」(transparent) あるいは「不可視」(invisible) であり、それが「可視化」するには主に不具合が発生した時だという点である (Star & Ruhleder 1996)。ここから複数の重要な論点が生じる。その一つは、インフラの維持管理問題である。インフラの不可視性を保つには、それを維持する努力があるが、それはしばしば外からは見えにくく、社会的ステータスに関する問題が生じるという点である。もう一つは、ユーザーにとっての負荷である。インフラがユーザーにとって不可視であるためには、ユーザー側もその使用法に習熟し、うまく使いこなす必要がある。実際、この二つの点は、知識インフラの実際の稼働に関して、様々な問題を引き起してきた。

安定していれば不可視なインフラにおいて、その保全、補修、維持管理といった作業は外からは見えにくく、シャドーワーク化することがある。それをあつかう部門が自律性をもち、特定の職能として社会的認知 (ステータス) を確保している場合もあれば

フォマティックスという一大勢力に成長しているが、バイオ情報の量、およびその守備範囲の爆発的な拡大は、そこに様々な新たな問題を生み出している。例えば、ゲノム解析が 4 基塩基 (ATGC) の配列の解析を中心とするのにに対して、設計図としての遺伝子が作り出すタンパク質は、二〇種類のアミノ酸の連鎖からなり、更にそれは多様な立体構造をもっている。しかもタンパク質の種類によっては、構造そのものがはつきりしなかつたり、あるいは構造解析のための結晶化が困難だつたりと、解析の難易度は飛躍的に高まる。この場合、どの部分の情報をどう集め、解析するかはその目的によって大きく変わってくる (福島 2017 第 6 章)。

あるバイオデータの専門家は、データの構造が単純でしかも終わりのあるもの、つまりどこまでやれば特定の目標が達成できるかが明確なものは、「データベースにとつて気持ちがいい」と指摘する。他方、話がタンパク質になると、構造の複雑さに加え、その全体像が完全に掴めているわけでもなく、難易度が増すという (8)。更にこれが生態学 (例えば生物多様性) といった分野に拡大すると、データ取得や分類を巡る問題も含めて、そのデータ構造の複雑さも指数級数的に増大する。バウカー (G. Bowker) はそれを「データ多様性」(data diversity) とよんで、その相互間の複雑な調整問題を詳細に論じて (Bowker 2005)。

データのライフサイクル

更に、先のスターらの研究 (Star and Ruhleder 1996) からは、データを取得し、それを公開、分析するなど一連の流れにおいて、研究室のお宝データを公的に共有 (公開) することによっての葛藤、といったより広いテーマも読み取ることができる。ここには、研究室が獲得した個別の標本、化合物といったものに対して、公的なライブラリーをつくるといったケースも含まれる (福島 2019 第 7 章)。実際、筆者が参加したデータ関係シンポジウム (この場合は極地科学) でも、研究者間のデータ共有の難しさ、データ供出のための労力 / コストの問題、国別のデータ公開政策の違い、あるいは陸域と海域それぞれの研究におけるデータ公開慣習の差といった多様な問題が活発に論じられていた (9)。

更に知識インフラには、蓄積されたデータの活用問題もつきまとった。ビッグデータ利用の議論には、その鮮度や利用価値があるとう。ビッグデータ利用の議論には、その鮮度や利用価値がある時期一瞬のものから、長期的な蓄積により価値があがる場合まで様々なバリエーションがあり、後者では古いデータそのものに大きな価値がある。しかし、最新の手法を使って古い時代のデータを集めるのは異なり、かつて獲得された科学データは、その手法が古く、現在使うにはその精度が問題になる可能性もある。地球温暖化に関心が高まる現在では、気象に関する正確な時系列データの重要性はますます増大しているが (10)、一部で論争が政

(Henke 2000)、急成長する知識インフラのよう、その維持管理作業の評価が社会的に不安定な場合もある。しかしその評価が安定しないと、新規インフラ開発者はその影の部分は忌避しようとするかもしない。この点が、筆者が別稿で「価値振動」(value oscillation) と呼んだ側面である。これは、「インフラ設計は必要だが、自分はそのシャドーワークにはかわりたくない」という矛盾した価値観を示している (福島 2019 第 7 章)。社会学者ゴフマン (E. Goffman) の演劇論的比喩を使えば、インフラ開発といふ舞台上で脚光を浴びるのはいいが、舞台裏的な作業 (維持管理) は嫌だというのである。実際、STS の知識インフラ研究では、その活動が知識「インフラ」とみなされることへの関係者の複雑な心情が多く記録されている (福島 2019 第 7 章)。

この点は一番田の論点、つまりインフラのユーザー問題にも影響を及ぼすことになる。この「舞台上」の活動としては、例えば斬新なデータ分析用アルゴリズムの開発などが挙げられるが、こうした新技術の絶え間ない開発は、ユーザーにとっては、延々と繰り返される再学習の強制ともいふことを意味し、インフラが常に悪い意味で「可視化」されるという事態を招きかねない。OS の頻繁なアップデートに対するユーザーの困惑がその典型である。

前述したスターらの研究は、バイオ研究現場におけるいわゆるウェット (生物) 系とドライ (情報) 系研究者の協働と相剋の初期の民族誌とみる (11) もやめた。この境界領域は現在バイオイン

治化するにつれ、そのデータの精度については科学的観点を超えた、十分な精査、補正が必要となつてきている(1)。

更にいえば、研究が進み、かつては注目していなかつた問題が浮き彫りになり、初期に記録されたデータを再検討する必要が出てくる場合もある。先程の生物多様性における初期のサンプルなどがその例だが（Boxkeller 2005）、過去の気温や海水温のデータ等、後に必要な再検討を行なうためには、そのデータ環境についてのデータ（いわゆる「メタデータ」）が必要となる。メタデータは詳細な方がいいが、詳細なメタデータがいつも保存されているわけではない。またこれからメタデータを管理するとしても、保存できる情報量や労力の限界のため、無制限に、というわけにはいかないのである。

ここで問題になるのは、後に精査が必要となるとしても、事前に何をもつて必要なメタデータとするか、原則的に決めようがないという点である(2)。例えば、ある時期の観察者はその観察環境をメタデータとするかもしれないが、後にその装置についての詳細があると分つた場合、その装置についての詳細なメタデータ（たとえば製作過程、使用方法、バリエーション、実際の作動様式等）が必要となるかもしれない。しかし問題は、そうしたデータがどこに残つているかである(3)。更に異分野でのデータを統合する場合、そこに多種多様なメタデータが付随するとなると、話は更に複雑になる。いつしたメタデータを巡る煩雜な諸問題を、エドワーズ（P. Edwards）らは「科学の摩擦」（science friction）と呼び、

と、またバイオ系と情報系の抜きがたい文化（本稿でいう認識的文化）の差も継続しているといつた指摘もある（Lewis & Bartlett 2013）。」のようにデータの多様なエコシステムと、多彩なドメインの構造は、汎用性を謳つて今のところ鼻息の荒いデータサイエンスにむずかしい課題を提供しつづけるのである。

おわりに

筆者が関心をもつ別の領域である組織安全問題において（福島 2010）、かつて興味深い経験をしたことがある。テクノロジーがらみの組織における安全問題は、その被害の甚大さもあって関係者の関心も高く、また、現在のデータサイエンティストのように、問題を横断的に検討できるとする専門家も育ちつつある。この分野では特に航空分野での安全管理が充実しているので、ある救命救急センターでも、時折その系列の専門家による安全講習会が開かれていた。しかしそれを受講した看護師らの反応は余り芳しくなかつた。その理由は、こうした専門家の講演が、結局のところ航空業界のそれを暗黙の前提（プロトタイプ）としており、救急医療の現場からみると、かなり違和感のあるもので、殆ど現場では役に立たないと感じられたからである。

現在、ビッグデータ／データサイエンスにまつわる言説はハイブ状態にある。それはかつてナノテクノロジーが、万物の基礎である原子を操作することによって、物質から社会にいたるまで、

この問題が科学全般に起こりうることを示唆している。しかもそれについての一括した解決法はないだろうと主張しているのである（Edwards et al. 2011）。

ドメインを巡る摩擦

実際、バイオインフォマティックスのように領域間の融合が進んでいるようにみえるケースでも、常に変化するバイオ、および情報科学のフロンティアに応じて、両者間の最適解をどうやって見つけるかは、結構むずかしい問題である。この分野でのある先端的な研究者は、こうした困難を認めつつ、「現実的な」解法として、この二つの分野（ウェットとドライ）について違う比率をもつ研究者たち（ウェット中心、ドライ中心、その中間等）をゆるやかに組み合わせ、プロジェクトの性質に応じてその構成比率を変えていくという彼のやり方を紹介している(4)。しかしこれが組織レベルでの協業体制となると、前述した宇宙科学での発言のように、データサイエンス側とプロジェクト側がその都度接触して、相互学習する形で最適解を模索するしかない。

しかしバイオインフォマティックスに関していえば、前述したスターらの研究でもこの両者（バイオ系と情報系）の潜在的な対立は指摘されており（Star and Ruhleder 1996）、更に後続の調査でも、バイオインフォマティックス側とプロジェクト側が、結局バイオ主導でおこなわれ、情報系はその下部単位とされていること

文字通り全てを変えていくと喧伝されたのに似ている面もある（ブルーベ 2009）。しかし現実には、データ解析が必要とされる諸分野の全体構造（エコシステム）はきわめて複雑で、そこには分野固有の特性、慣習や伝統も存在する。それはちょうど生物多様性や、或いはグローバリゼーション／地域といった、いわゆる「グローバル問題」に似た関係性がみられると言つてもよい。

どんなデータにも、その取得から利用にいたるまで、さまざまな個性的履歴があり、データ解析の目的によつては、その履歴そのものが重要な役割を果たす場合も少なくない。その一部はメタデータという形で明示化されるが、別の面は背後の手続の中に埋もれてしまう。そうした見えないプロセスとしては、知識インフラに関わる様々なシャドーワーク、標準化／データ整備への絶えざる調整、またアルゴリズムに潜む様々な政治社会的前提といった多くの興味深い問題が存在するが、紙面の都合上本稿では十分に掘り下げられなかつた。いずれにせよ、これらのシステムの影に隠れた細部の一つ一つに注意を向けることで、ともすればデータの巨大さに圧倒され、ハイブに煽られがちなわれわれも、事態のあり方をもう少し醒めて観察することが出来るかも知れないのである。

註
(1) <https://www.youtube.com/watch?v=sV116qXHsM>. (1)〇1〇年五月
II〇日アクセス)

(27) 「やった議論の結括は紙面の限界で殆ど不可能だが、城田 (2012)、ハルマー (2013)、マヤヤー・ハヨーン・ベルガーナクキエ (2013) 等は紹介、批判のやれやれの傾向のパターンを示してくる。」の問題に特化した園庭シヤートスであることは研究会ノートとして筆者がよく参照するが、「Journal of Big Data (学際的なデータ・クラウド・アーキテクチャ応用研究センター)、Big Data and Society (おおきなデータ記述的、批判的アプローチで筆者もその国際編集部の一員を務める)、Data and Society 研究会 (データワークを掲げただだP.O.研究組織や、人種、ジェンダー、労働問題等、明確に批判的な視点を採用している)」など。またOERの専門誌でも近年投稿が増大している。

(28) 前報の箇を https://www.scopenet.or.jp/main_scope.net/pdf/Vol71.pdf (110110年K田・H田トセバ)。後報は https://www.kantei.go.jp/jp/singi/it2/seminar_bunka/data.ryutsuseibi/detakatsuyou_wg_daif/siryou3.pdf (110110年K田・H田トセバ)。

(4) たゞ日本米の格差はここにの参考 <https://globe.asahi.com/article/12844087> (110110年五月一八日トクセバ)。

(5) 田北野愛子 (弘前大学) を研究代表者とする研究課題として、科学の諸領域における「やったシナリオワーク」の比較研究が110110年度から始まり、筆者以外にも鈴木舞氏などが参加している。

(6) 110110年「宇宙科学情報解説システム」(宇宙科学研究所)。また「学び方—タサイエンスはバフでいる(マム)が、若手が立ち着いて仕事がやる感じでもなる」とこう耳の痛い指摘があった。

(7) やれ以外にもたとえは気象学等のデータ問題はよく参照される (Edwards 2000)。

(8) 筆者は110110年6月のインタビュー。それゆえノム研究の成功を認識し、その手法をタンパク構造解析に応用しようとしたらこねる構造ケーブル研究が、多くの対立と論争をおこした様子は相原 (2019 第6章) 参照。

(9) 110117年開催の極地科学シンポジウム、北極圏での研究協力セッション (110117年K田・H田トセバ)。

（10）たとえば110111年「観測データが語る気候変動—長期観測データの取得・発掘・保存」シンポジウム(於：年代田放送会館ホール)等。

(11) 「やった政治化した对立については、モシニアーナフラー (2010) が反温暖化論者のキヤハーマーンとして、屋外設置の温度計に関して伝統的な調査を、温度計がエアコンの室外機のそばにおいてある、といった事例をウェブ上にアップするといった活動を紹介している。

(12) 状況論じて、「物理学的なアプローチでは、「状況」の状態をもといいで、主体が知覚する範囲での状況」と、第II者が定義する状況の間ではそれがたら、ギャップがある」とが問題化された (Chaiklin & Lave 1983)。

(13) 前掲気候ハハボシウムでの重要なテーマは「(レガリティ)とは別篇で語じる予定」。

(14) 筆者は110110年6月のインタビュー。

文献

- 福島真入 (2010) 「計測と出発符——リスク・実験・高信頼性」 東京大学出版企
福島真入 (2017) 「試験の工場——科学技術の社会的研究」 東京大学出版企
Henke, C. (2000) The Mechanics of Workplace Order: Toward a Sociology
of Repair. *Berkeley Journal of Sociology* 44: 55-81.
田浦昇恵 (2017) 「バーベーチやトコソタのアドバイス……データ構造を解き明
かす先駆技術」 カナルシティ
橋口晃之 (2013) 「データ・サイエンティストがピックルータで私たちの未来
を創る」 [情報論理] 第六卷 | 第 1 — 1 頁
Hine, C. (2008) *Systematics as Cyberscience: Computers, Change, and Contin
uity in Science*. MIT Press.
池田暁報 + 森井幹雄 (2018) 「トータル科学と天文学」 [天文月報] 111 卷十七号、
1110—1127頁
Knorr-Cetina, K. (1999) *Epistemic Cultures: How the Sciences Make Know
ledge*. Harvard University Press.
ペーパー・ヒ (1971) 「総合社会学の構成」 吉川茂記、吉川茂記
Latour, B. & Woolgar, S. (1979) *Laboratory Life: The Social Construction of
Scientific Facts*. Sage Publications.
二二一 (2016) 「トータル・ピックルータ——第4次産業革命に向
けた製造業の挑戦」 日刊工業新聞社
Lewis, J. & Barlett, A. (2013) Inscribing a Discipline: Tensions in the Field
of Bioinformatics. *New Genetics and Society* 32 (3): 243-263.
マヤヤー・ハヨーン・ベルガーナクキエ (2013) 「ピックルータの正体
——情報の産業革命が世界のすべてを変える」 齋藤栄一郎訳、講談社
マハヤー・ハヨーン・ベルガーナクキエ (2010) 「地球温暖化とキャッシュカル
秋川ハイテクハーモニーフラットの激闘」 著別出版社、日本評論社
Ribes, D. et al. (2019) The Logic of Domains. *Social Studies of Science* 49 (3):
281-309.
Rosch, E. & Lloyd, B. (eds) (1978) *Cognition and Categorization*. Lawrence
Erlbaum Associates.
城田真醫 (2012) 「ピックルータの衝撃——巨大なデータが戦略を決める」 東
洋経済新報社
ハルマー・ゼ (2013) 「ハッタル＆ハイズ——天才データナツメイの挑戦」
三編集子訳、五総研P.社
Star, S. & Ruhleder, K. (1996) Steps toward an Ecology of Infrastructure:
Design and Access for Large Information Spaces. *Information Systems
Research* 7 (1): 111-134.
鈴木大輔 (2017) 「科学鑑定のエヌノグラフ——11月11日ハムにおける基
科学ラボボトニーの実践」 東京大学出版企
山口智子 + 福島真入編 (2019) 「予測がいくつ社会——「科学の植葉」の使わ
れ方」 東京大学出版企
（44-444 4430-4431・科学社会誌 [201-202]）