

情報・システム研究機構ブックレット

SCIENCE REPORT

031 - 035



デジタルな 人文、 データな社会

大学共同利用機関法人
情報・システム研究機構

情報・システム研究機構ブックレット

SCIENCE REPORT

031 - 035

デジタルな



人文、

データな社会

大学共同利用機関法人

情報・システム研究機構



『サイエンスレポート』について

『サイエンスレポート (<https://sr.rois.ac.jp/>)』は、
大学共同利用機関法人情報・システム研究機構が
運営するウェブサイトです。

2016年12月公開以来、
当機構の大学共同利用機関法人としての役割を踏まえ、
自機構のみならず他機構および
広く大学・研究機関等の話題を採り上げ、
広く一般・マスコミの方々へ向けて、
学術の成果や取り組みに関する情報を提供しています。

その中心的なコンテンツとして、ウェブサイトと同名の連載記事
を定期的に（月1回程度）掲載し、
継続的に新しい記事を追加しています。
先端的な取り組みをわかりやすく紹介し、
人々の疑問に科学者コミュニティが答えるような構成で、
学術の役割と活動を広くご紹介しています。

本書は、その記事が身近にお手にとれるよう、
冊子にまとめました。

目次

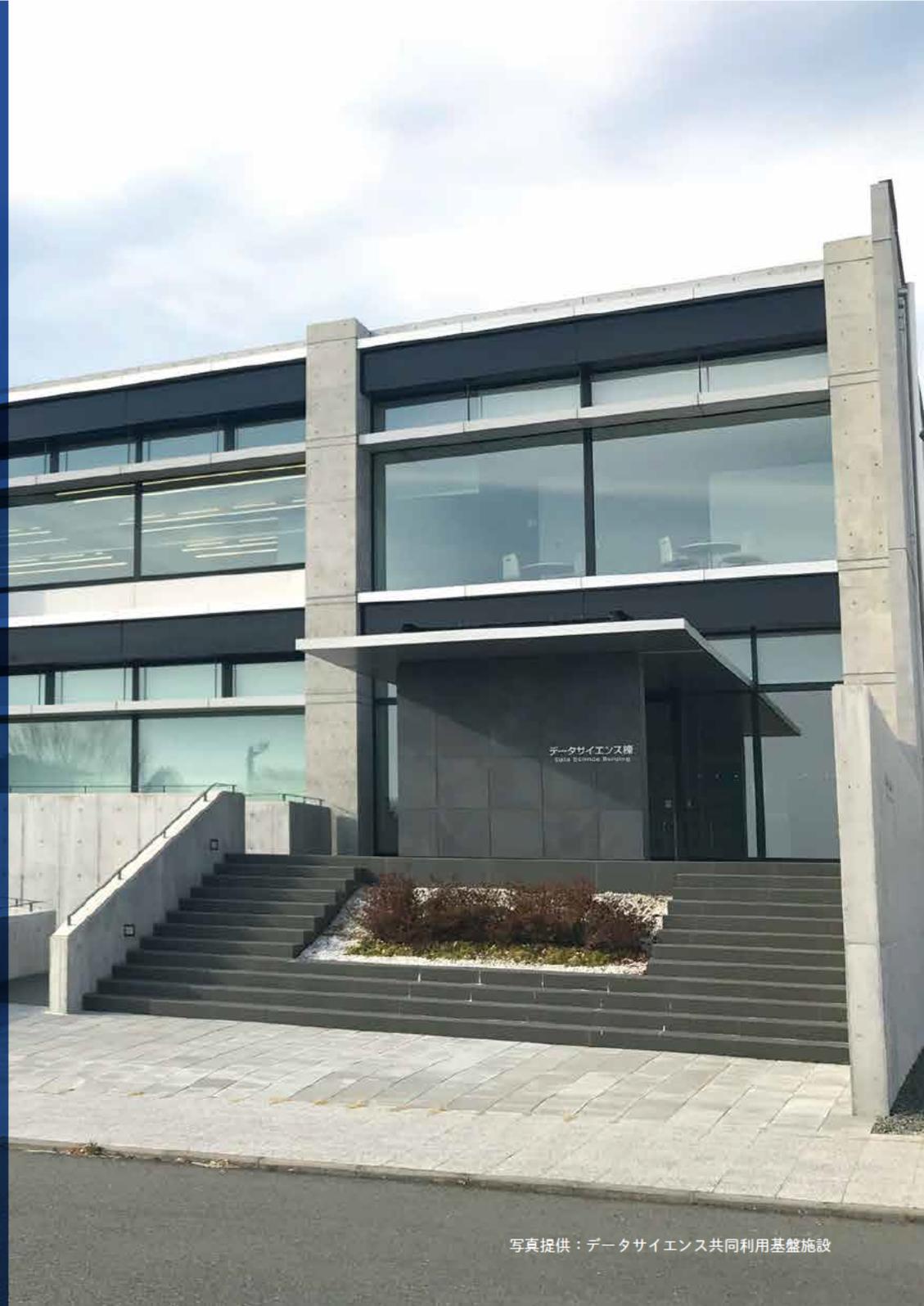
- 03 『サイエンスリポート』について
- 08 Science Report 031
データの中にあるプライバシーをどう守る？
答える人 | 南和宏教授・高部勲センター長
- 18 Science Report 032
データが語る、路地のロジック！
答える人 | 岡檀特任准教授・久保田貴文准教授
- 30 Science Report 033
日英が協力して、日本人の生活時間を解析する。
答える人 | 白川清美研究員・エカテリーナ ハートグ研究員
- 40 Science Report 034
データサイエンス時代の人文学はどう進む？
答える人 | ロバート キャンベル館長・北本朝展センター長
藤山秋佐夫施設長
- 52 Science Report 035
ウィズコロナ社会の諸課題はどう解決できるだろう？
答える人 | 遠藤薫教授・椿広計所長
- 63 大学共同利用機関法人 情報・システム研究機構について
機構長 藤井良一

所属・役職・掲載記事の内容は公開日当時のものです。

SCIENCE REPORT

031 - 035

デジタルな 人文、 データな社会



データの中にある プライバシーを どう守る？

ビッグデータ時代といわれる今日、データの利活用が、さまざまな価値を生み出すことが期待されている。しかし物質、天文といった自然科学のデータはともかく人間・社会のデータの場合、データの中に個人情報が含まれていないかどうか、その安全性を確保しなければならない。国・地方自治体、企業等が公開するデータの中にあるプライバシーの保護は、ビッグデータ時代の最重要課題のひとつと言えよう。そこでこれを秘匿技術によって解決し、利活用を促進する研究が進められている。また国が保有する「公的統計マイクロデータ」について、その安全性を確保した上での利活用促進が、官学連携によって進められているという。

こちらからインタビューの
動画を見ることができます。



答える人

南 和宏 教授

[統計数理研究所、データサイエンス共同利用基盤施設]

みなみ・かずひろ。統計数理研究所、データサイエンス共同利用基盤施設教授。専門は、プライバシー保護技術という分野のうち、特にデータを利活用する際に機密情報を守る匿名化技術。博士(コンピュータ科学、米国ダートマス大学、2006)。イリノイ大学講師、国立情報学研究所特任准教授等を経て、2020年より現職。



答える人

高部 勲センター長 ※2021年より立正大学データサイエンス学部教授

[総務省統計データ利活用センター]

たかべ・いさお。総務省統計データ利活用センター長および統計研究研修所教授。早稲田大学理工学部卒、博士(統計科学、総合研究大学院大学)。ビッグデータに関わる名寄せ技術の研究などを背景に、公的統計マイクロデータの提供、地方自治体や地域の大学等と連携したデータサイエンスの推進、証拠に基づく政策立案(EBPM)の支援に取り組む。





データの中にある個人情報を守るには？

現代の社会では、スマホから取得されるユーザの行動履歴や、オンラインショッピングの購入履歴などの情報が、企業等に自動的にどんどん蓄積されていく。このようなビッグデータは、2017年の個人情報保護法の改正により、きちんと匿名化という処理をすれば第三者への提供が認可され、利活用の道が開かれた。統計数理研究所、データサイエンス共同利用基盤施設の南和宏教授は、統計的な手法を用いて、このようなデータ利活用の際に個人情報を守る匿名化技術に取り組む。

「匿名化というと、名前やIDを削除すればいいように思われるかもしれませんが。しかし実は、それだけでは不十分であることが分かってきました。例えば性別と住所という属性を組みわせたり、別のデータと組み合わせたりすることによって、個人を特定したり、その人の属性を推測できる場合があるんです」。

このことは「歴史的にも証明」されており、「1997年、米国マサチューセッツの州の医療データから州知事の病名が分かってしまったのはその一例」と南教授は言う。「不特定の方々に公開するので、中には悪意ある攻撃者や、組み合わせる

と有用な情報を持っている人がいるかもしれません。そのような人がデータを見ても個人の機密情報を知ることがないように、しっかりとデータを加工するのが匿名化の目的です」。

内部者(調査対象者)による攻撃

度数表 (職業・地域が一致するレコード数)		職業				
		P ₁	P ₂	自分の地域・職業は知っている		
地域	M ₁	20	15			
	M ₂	72	20	1	10	133
	M ₃	38	38	15	40	133
	合計	130	73	46	90	22

数量表 (収入の総和)		職業					合計
		P ₁	P ₂	自分の収入を引けばもう1人の収入が分かる			
地域	M ₁	1440	540	22	70	320	2290
	M ₂	722	1178	375	800	363	3438
	M ₃	2522	2168	1117	1770	1043	8620
	合計						

度数表のセルの値が2の場合、数量表から個人の収入を直接知ることはできない。しかし、調査対象となった人が悪意を持って、別の調査対象者の情報を得ようとする状況も考慮する必要がある。もし、そのような攻撃者が地域M₃、職種P₅の属性をもつ場合、自分が住んでいる地域に同じ職種の人がもう1人いることが度数表から分かる。この場合、攻撃者は該当する数量表のセルの値から自分の収入を引くことで、自分と同じ属性のもう1人の個人の収入を知ることができる。

国が保有する「公的統計マイクロデータ」をつながりやすくする

日本には質のよい社会調査データがあるといわれるが、総務省統計局は、国勢調査をはじめとする公的統計の豊富なマイクロデータ(個票データ)を保有する。和歌山県にある総務省統計データ利活用センターの高部勲センター長は、このような「公的統計マイクロデータ」の利活用促進と、地方自治体や学術機関とのデータサイエンスに関わる共同研究に取り組む。

「公的統計マイクロデータを十二分に活用するための環境の整備として、研究者の方々が情報セキュリティが確保された施設内で、探索的なデータ分析を行うことができ、新たな価値を生み出すことができるオンサイト施設の設置を進めています。現在12拠点に広がっており、日本全国の地域ブロックへの展開を進めています」。

中でも高部センター長が担うのは、統計数理研究所と共同研究を進めている、異種のデータをつなぐ「データマッチング」や「名寄せ」と呼ばれる技術開発だ。「各自治体が保有する行政記録、あるいは企業が保有するPOSデータなどをオープンに利用できるようにしていこうという動きがあります。これらとわれわれの公的統計のマイクロデータを連携させていく場合に、各記録、各個人のデータを、何を手がかりにどう連結させたらいいか、いわゆる「名寄せ」と呼ばれる技術的な課題に取り組んでいます」。

さまざまなデータをつなぐことによって、新たな価値創造が起こるのが、ビッグデータの醍醐味だ。「経済学はもちろん都市工学の観点からの分析、あるいは企業のマーケティングに資する分析、また国・地方自治体の政策立案に資する分析等が行われており、こうした研究は今後も増えてくると考えています」と高部氏は言う。



データサイエンス共同利用基盤施設の オンサイト施設

東京・立川市にあるデータサイエンス共同利用基盤施設のオンサイト施設。統計センターの定めるセキュリティ基準に従って、極めて高い安全性が確保された部屋となっている。



研究者は、ここに一定時間滞在して、設置された端末からリモートアクセスでデータ管理サーバにアクセスし、調査票情報を解析する。受け入れを担当する情報・システム研究機構岡本基主任URAによれば「研究者が利用する調査票情報はすべて統計局・統計センターが運営する中央管理施設のデータ管理サーバに格納されており、端末にデータをダウンロードすることはできない「シンクライアント」のしくみが構築されています」という。さらに画面を撮影すること等による漏えいを防ぐため、監視カメラも設置されている。「利用者も少しずつ増えてきていますので、公的統計マイクロデータ研究コンソーシアムの運営とも連動して、公的統計の利活用と研究の発展に貢献していきたい」と、岡本主任URAは言う。

情報・システム研究機構

データサイエンス共同利用基盤施設 オンサイト施設

https://ds.rois.ac.jp/center3_micro/





分析結果を持ち出す際の安全性基準を作る

オンサイト施設では、訪れた研究者は解析を行い、その結果を追って自分の論文に用いて発表することができる。しかし、その前に(独)統計センターの審査をパスしなければデータを持ち出すことはできない。南教授は、解析結果の持ち出し前にどんな秘匿処理をすべきか、その安全性基準の策定に尽力してきた。

「論文では一般に分析結果をかなり集計したかたちで掲載するため、個人情報の漏えいは起きないのではないかとされるかもしれませんが。ところが意外と難しいのは、集計表のような表データです。集計表の場合、例えば1つのセル(マス目)だけ隠しても、表データには行と列ごとに集計欄があるので、合計からの引き算で簡単に復元できてしまいます」。

「では、実際にどこを隠せば値が守れるのかを決めるのを「二次秘匿問題」と言い、ちょっと複雑な表になると人間が行うのは不可能です。最小限の加工で本当に



隠したいデータを隠すにはどの範囲を隠せばいいのか、隠したセル値に十分な不確実性を担保する「秘匿インターバル」を保証し、かつその安全性を審査するための説明資料も一緒に出力してくれるツールを、R言語を用いて開発しました」。

この二次秘匿自動化ツールは、安全性基準を担うものとして、いくつかのオンサイト施設の協力を得て、現在、評価のプロセスに入っているという。「将来的には、基本的にはこのツールを使って、プライバシーの安全を確保し、機密情報を秘匿処理したデータを、研究者に持ち帰ってもらえればと考えています」。

利用者と連携するコンソーシアムでもっと使いやすく

データサイエンス共同利用基盤施設社会データ構造化センターでは、南教授を中心に、公的統計マイクロデータの学術研究への利用促進のために、産学官が一体となって取り組むことを目的とする「公的統計マイクロデータ研究コンソーシアム」の運営にも取り組む。現在、会員は73名を数え、会員向けの教育・普及活

動や意見・要望の集約などを行っている。計量経済学、社会学の研究者を中心に、最近では都市デザイン、エネルギー関連等の工学系の研究者も増えてきているという。

高部：これまで公的統計のマイクロデータを分析されていた研究者の方は、どちらかというと公的統計だけを対象とされている方が多かったのですが、今や、民間企業等のさまざまなデータがどんどん公開されるという、想定し得なかった事態が生じており、公的統計とつなぐことで大きな価値創造が期待できます。経済センサスなどの公的統計は企業データと比較して、調査事項はそれほど多くはないものの、全数調査であるなど圧倒的に高いカバー率が特徴です。一方、企業データはカバー率は低いものの詳細な項目を持つことが多く、これらをつなぐことによって、利活用の可能性が大きく広がります。

南：つなげればつなげるほど、データは価値が上がるのですが、データが長くなればその分、個人情報や推測できるリスクが高まるため、当然のことながら匿名化は難しくなるんです。まさに高部さんの研究と、私の研究をどうバランスさせるかが非常に難しいところなので、連携して利活用を上げていくことが大切だと考えています。

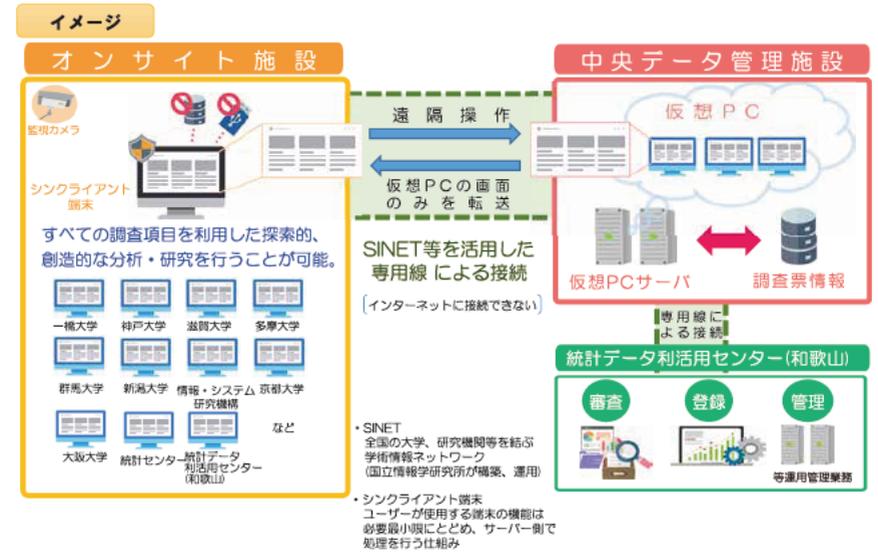
※本インタビューと対談は、オンラインで行われました。
公開日：2020/09/10



miripoは公的統計のマイクロデータを利用するためのポータルサイトである。マイクロデータを利用するための制度、利用実績が紹介されている。



オンサイト施設におけるマイクロデータの活用



オンサイトとは、情報セキュリティが確保された環境で、許可を受けた研究者が調査票情報を用いて、独自の集計・分析を行うことができる専用室をいう。
(出典「公的統計のマイクロデータを使って分析してみませんか？」-総務省統計データ活用センター https://www.e-stat.go.jp/microdata/sites/default/files/share/data-use/onsite_intro.pdf)



左：オンサイト利用室端末全景
下：オンサイト端末の利用室ブース
写真提供：データサイエンス共同利用基盤施設

データが語る、 路地のロジック！

「路地が関係あるんじゃない？」——自殺が少ない地域には、どんな自殺予防の要因があるのかを調査する中で、路地に注目したという統計数理研究所の岡檀特任准教授。どうしたら自殺が減らせるか。自殺には、経済的・社会的な要因があることが知られているが、人が住む場所の物理的な環境に注目し、自殺が少ない町には路地が多いことをつきとめた。ビッグデータ時代の到来以前から、現地へでかけて調査する「質的研究」と、データをコンピュータの中で解析する「量的研究」を組み合わせ、手作りでモデルを作り上げる。岡特任准教授と、その共同研究者である多摩大学の久保田貴文准教授に、データを駆使して路地と自殺の関係を捉える研究について聞いた。

こちらからインタビューの
動画を見ることができます。



答える人

岡 檀 特任准教授

[統計数理研究所、慶應義塾大学] ※2021年より一橋大学経済研究所客員教授兼務

おか・まゆみ。統計数理研究所 特任准教授、慶應義塾大学大学院 特任准教授。博士（慶應義塾大学大学院、2012）。自殺、介護などの社会問題に注目したコミュニティ研究で知られる。専門はこの他、健康社会学、社会疫学。社会調査において質的・量的アプローチを用いて、丹念な実地観察と、社会的な要因の数値化、さらに指標化を目指す。著書に『生き心地の良い町 この自殺率の低さには理由がある（2013年）』ほか。



答える人

久保田貴文 准教授

[多摩大学]

くぼた・たかふみ。多摩大学経営情報学部 准教授。専門は応用数学、統計数学。場所・地域に関連する空間統計データの解析や、医薬・健康に関連するリスク解析のほか、統計データの視覚化も手がける。博士（学術、岡山大学、2012）。岡山大学卒業後、統計数理研究所等を経て2014年より現職。同年第2回藤田利治賞受賞。日本計算機統計学会評議員・理事。





自殺が少ない町には、路地が多い

「私の場合はまず地域の観察から始めます」という、統計数理研究所 医療健康データ科学研究センターの岡檀特任准教授。そもそもはアンケート調査などから、自殺が少ない地域に特有の思考傾向や行動様式に「援助希求（周囲に助けを求める）能力が高い」点があることがわかってきたのだという。そこでさらに観察を続けていくうちに、そのような町は密集していて、路地が多いことがわかってきた。「すると、『もしかして、これ自殺率に関係している？』というところを調べてみたくなるわけです」。

自殺を含めて、身体的・精神的健康に関わるさまざまな要因が、かつて考えられていたように細菌やウイルスだけではなく、今や社会、経済、貧困といったさまざまな要因が関係していることは広く知られている。しかし、「『路地が関係あるんじゃない？』と思ったところで、それだけでは自殺を防止する提言にはつながりません。路地というものをどう数値化して、自殺率データとどう連結させるかというところが肝心」と岡檀特任准教授は言う。

「路地」とは何か、まずその定義からして「すごく曖昧」だという。「路地が少ない地域と多い地域があるのは、誰もが感覚的にはわかっているし、やはり険し

い山間部には非常に少ないし、海岸部の密集したコミュニティには多いわけですね。しかし例えば道路幅何メートル以内の道といった決まりは何もないし、地図にも載っていたり載っていなかったりします。一般的に考えられているのは、自動車は通れないけれども、人は通れて、日常的に利用されているといったところでしょうか」。

そこで客観的に数量化していくために、統計学、都市工学、土木の研究者の意見を聞いたり、中でも地図会社の方々の協力を得られたことが、指標作りを大きく進展させたという。ちなみに、車道に接していない建物には、アクセスするための路地が存在する可能性が高いことに注目したこの路地推定法は、現地でのきめ細かな検証を通じて82%の判定率を達成し、公益財団法人日本測量調査技術協会令和元年度「優秀技術論文会長賞」を受賞した。

質的×量的アプローチで社会的な要因をつきとめる

では具体的に、どのように指標化するのか。岡檀特任准教授の場合は、まずは足を使った入念な地勢調査から始まる。「現地を見て、体感的に路地ってこういうことだよねということを書き出して、その条件を兼ね備えるためのロジックは、どうしたらできるだろうということの研究グループで話し合います」。

「『路地のロジック』って、駄じゃれみたいなんですけど（笑）、次に、路地を抽出するロジックを実際につけて計算します。そして計算が本当に合っているかを、現地に戻って検証するんです。的中率が80%を超えるまで、これを繰り返して精度を上げていきます」。この質的・量的という両方の手法を行き来する混合アプローチが、研究を始めた当初からの岡檀特任准教授の手法だ。「現地へ行かなければ、指標に入れるべきものをやはり見落とすし、計算してから再び現場に戻って合致しないのが見つければ、やはり修正を重ねていきます。私は、データについてはぴったりフィットするものを探し求めるし、これだけでは見たいものが見られない時は、とことん探すし、なければ作る——これは、自分のこだわりですね」。

このようにして東京都、大阪府、三重県を対象に検証した結果、3つとも路地の割合が高いコミュニティほど自殺率が低いことがわかった。ではなぜ路地が多いと自殺率が低いのだろうか？「路地では人と人のコミュニケーションが自然と促されているとか、周囲の異変に気付きやすいなど、いろいろな理由が考えられます。コミュニティをよい状態に維持するには、やはり問題の早期発見が大事なんですけれども、それがなかなかできないために家庭内暴力や児童虐待が重症化して、また非常に悪化してからでないという表面化しないというようなことが、自殺が少ない町では起きにくいわけです。路地に注目することが、このような社会的な要因の礎になる物理的な構造として役に立つのではないかという仮説を立てているわけですが、実際今、このような要因との有意な相関が現れてきているところです」。



「研究の世界に入る前は、データと聞くと、無機質な数字の羅列というイメージがあったんです。でも実際に研究してみて、社会に散らばっているあまたの事象というものを、データというかたちで蓄積し分析することによって、ある人間像が浮かび上がってくるんだなということを実感しています。無意識に繰り返される思考パターンとか、言われてみれば納得するような行動様式とかがあぶり出されて、データというのはある意味ものすごく人間くさい材料なんだなということを感じ、興味が尽きません」。

空間統計学の視点で、路地の指標化を検証する

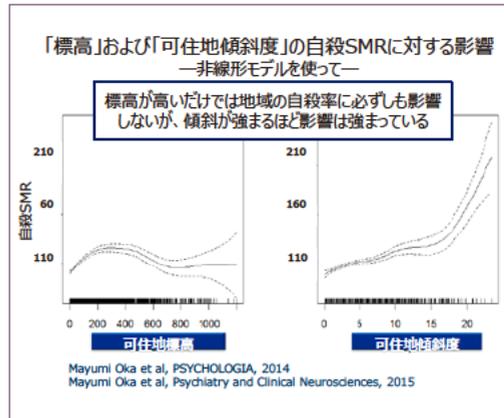
ところで、岡特任准教授がこのような研究を始めることになったのは、約10年前にさかのぼる。そもそもは自殺の要因と住地傾斜度（人が居住し得る土地だけを対象にその傾斜を求めた指標）に注目し、「統数研の藤田利治教授（当時）が構築した3,318市区町村の自殺率のデータセットを使わせていただいたのがそもそも始まり」と、振り返る。「当時はビッグデータもオンサイト室もなかったので、このデータ加工し標準化する作業には、ほんとうに苦勞しました。藤田先生が『食材と一緒にだ』と、『データが傷んでいたら、どんなに腕のいい名シェフが調理しても、そこまでの結果しか出ない。ここで頑張っておきなさい』と繰り返しおっしゃっていたのを、今も肝に銘じています」。

空間統計学を専門とする、多摩大学久保田准教授との共同研究もその当時にさかのぼる。久保田准教授は「『そもそも傾斜度ってどういうふうに計算しましょうかね』といったご相談を岡さんから受けたのが始まりですね。椿広計 現・統計数理研究所所長と、傾斜度と自殺の関係性をモデル化する時に、非線形というモデルを使うと、よりわかってくるんじゃないかといった議論になったことを憶えています」と言う。

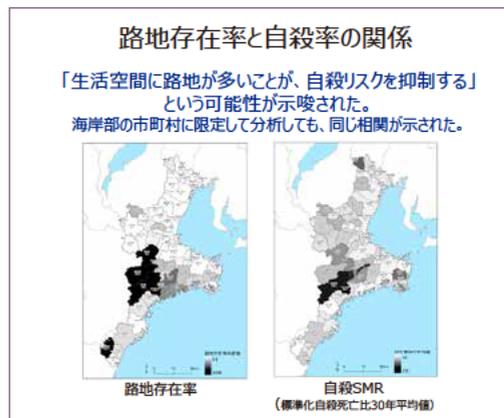


その「非線形」が画期的な結果をもたらした。まず下図にある「自殺SMR」だが、「SMRというのは自殺地域研究によく使われる指標で、標準化死亡比(Standardized mortality ratio, SMR)のことです。実際にその地域で人口構成や年齢分布などから予想される自殺者の数に対して、実際の自殺者が何倍いるかを指標化したもので、値が大きくなるほど自殺が多いことを意味します。予想される自殺者(=100)と比べた数値なので、例えば30ならば自殺希少地域ですし、300ならば、3倍の自殺者が実際にいる自殺多発地域であることを示しています」と久保田准教授は言う。

「この非線形モデルからは、単に標高が高いだけでは必ずしも地域の自殺率に影響しないのですが、ここに傾斜という要素が加わることで強い影響が生じて、自殺率が上がっていることが読み取れます」と岡特任准教授。それは線形モデルでは現れなかったのだという。



議論に基づいて岡特任准教授がモデルを作り、久保田准教授がモデルとその解釈などについてさらに検討を加える。このような共同研究により、続いて、路地の存在と自殺SMRの関係を分析したのが右図である。左は路地、右は自殺率で、路地の少なさと自殺の多さが一致しているのがわかる。



未曾有のコロナ禍社会と自殺について考える

では、自殺をテーマに地域特性を比較する研究から、現在全世界が見舞われているコロナ禍を見ると、どのようなことがわかり、また予見されるのだろうか。

「歴史的に戦後2回ぐらい大きな経済危機があつて、その時にもう非常に分かりやすく自殺率もボンと上昇しています。しかし、これをもっと細かく地域別に見ていくと、実は2回ともに自殺がもともと少ない地域ではそれほど跳ね上がりませんでした。それに対して自殺が大変もともと多かった地域は、その経済危機の時にバンと跳ね上がり方が大きかったことが既にわかっています」と岡特任准教授は言う。「つまり、全国が同じ経済危機にさらされていても、自殺希少地域ではその危険を緩和する何らかの自殺予防因子が働いたことによって、自殺率の上昇を抑制した可能性が考えられるのです」。

岡：日本においては経済問題が自殺と非常に大きく関わっていて、諸外国と比べてもその影響がすごく強いんですね。しかし、今後経済が悪化したらまた自

岡特任准教授が調査で足を運んだ徳島県旧海部町の鞆浦(ともうら)港 写真提供：岡 檀



殺が急増するというのでは困るので、そこをちゃんと研究していかなくちゃいけないという教訓を残していると思うんです。新型コロナウイルスの世界的な感染拡大が人間社会に甚大な影響を及ぼして、今まさにそこを考えなくちゃいけない時期にあります。これから自殺率が跳ね上がるというような事態になったら、そこに地域差はあるのか、その差が生じる要因は何なのかといった分析を、ぜひまた一緒にやりたいと考えています。

久保田：私も同じ考えです。コロナ禍に限らず大災害が起きた後には、一般に、最初の数ヶ月間は自殺者、自殺率は下がる傾向にあります。今回でいうと緊急事態宣言が出て、みんなが非常に不安を持っている4月、5月でさえ、自殺率は前年に比べても低い状況でした。一方で警察庁の発表による速報値では、7月、8月の自殺者数は、去年に比べて増えているんです。これは「いのち支える自殺対策推進センター」とも一緒に進めている分析でわかってきたことなのですが、まず漠然とした不安があって、次にそれが顕在化されて経済面、健康面の不安

が来て、その後に「希死念慮」、すなわち自殺をしたいという考えが生まれてくるというのが大きな流れです。この流れをどこかで断ち切っていくことが必要になります。

岡：そうですね。

久保田：経済的な要因を基に、ひとつには地域の物理的特性、さらに路地といったところまで落とし込んでいく。また私は時系列的に見ていくことにも関心を持っていて、どういうところでどういう人の自殺が増えてくる可能性があるのかを予測し、対策に役立つエビデンスを求めていく必要があると考えています。例えば10万円の交付金支給がアナウンスされた後には、不安が減少するような傾向が見られるだろうか？——といったことを、現在、ツイートの分析を通じて進めているところです。さらに、コロナ禍の状況では例えばワクチンがいつできるか分からないとか、何年後までこの状況が続くのか分からないといった今



まで経験したことのない状況があるため、それらがどのような影響を持つかについても見ていかないといけないと考えています。

社会人を経験後、大学院へ進学したという経歴を持つ岡特任准教授。「研究を始めた頃は数学が苦手なこともあって、質的なフィールド調査中心で行くと考えていたんですけども、すぐにビッグデータが必要だ、それでこそフィールド調査が活けると気づかされました」という。一方、久保田准教授が所属する多摩大学では、3年前にオンサイト施設を設置。久保田准教授はこの運用から高齢者データのプラットフォームづくり、さらにデータサイエンティスト育成プログラムの策定などへ向けて尽力する。「どんな入口からでもいいからデータをちゃんと見ることができて、最終的なゴールに向けて分析し、結果を出していくことが必要で、そういう人材を育てたい」という。

※本インタビューと対談は、オンラインで行われました。
公開日：2020/10/12

江戸時代から続く建築様式「みせ造り」。路地における“ベンチ”の役割を担っている。旧海部町には現在も「みせ造り」が点在し、住民らのコミュニケーションを促す仕掛けとなっている。生活空間にこうしたハブ（情報の集散）スポットがあることによって、住民らは無意識のうちに自らの困りごとを小出しにし、コミュニティにおける問題の早期開示、早期介入に結びついていると考えられる。

写真提供：岡 檀



日英が協力して、 日本人の生活時間を 解析する。

サイエンスレポート031で紹介した「オンサイト施設」は、東京・小平市にある一橋大学小平国際キャンパスにもあり、ここでもデータを利用したさまざまな研究が進められている。中でも一橋大学経済研究所の共同利用共同研究拠点プロジェクトによる英国オックスフォード大学との共同研究では、年2回、英国から研究者が訪れ、日本におけるジェンダーギャップや子供たちの生活時間に関する研究が進められている。残念ながら今年はコロナ禍により来日が実現できていないとのことだが、海を越えて、データ利用に関わるサポートや共同研究が続けられているという。

こちらからインタビューの動画を観ることができます。



答える人

白川清美 研究員 ※2021年より立正大学データサイエンス学部教授および統計数理研究所客員教授兼務

[一橋大学経済研究所]

しらかわ・きよみ。一橋大学経済研究所准教授を経て、現職。青山学院大学経済学部卒、芝浦工業大学大学院 博士課程修了。世界各国の統計局と連携して社会調査データの保全、統合、比較、平等な利活用の促進などに尽力。また共同研究により、公的統計マイクロデータ（個票データ）の二次的利用を促進する秘密分散・秘密計算技術にも取り組む。



答える人

エカテリーナ ハートグ 研究員

[オックスフォード大学]

Ekaterina Hertog。オックスフォード大学社会科学科研究員、Doctor of Philosophy in Sociology。専門は家族社会学、社会政策。現代東アジア、特に日本の家内分業におけるジェンダーバランスに焦点を当てた研究を推進する。また一橋大学経済研究所と、日本における世帯収入、親の教育、子供の日常生活の間の関連を探る共同研究が進行中。



海外の研究者が日本のデータを使えるようにする

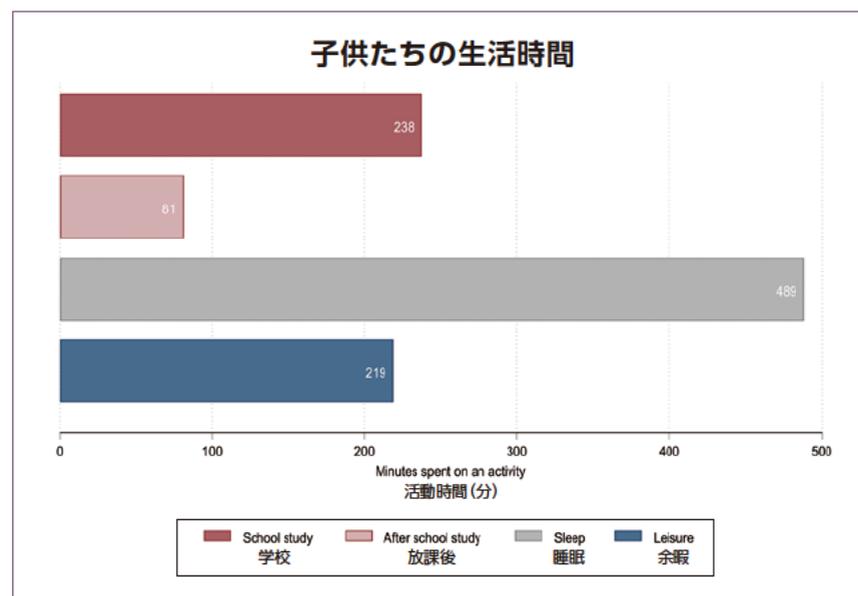
一橋大学小平国際キャンパスにあるオンサイト施設は、徒歩1分の距離にゲストハウスが設けられ、時差を気にせず24時間利用できるなど、海外から訪れる研究者にも利便性の高い施設となっている。およそ2週間の滞在期間に加え、その前後に利用手続きやデータ持出申請などさまざまなサポートが必要だ。

一橋大学経済研究所の白川清美研究員は、公的統計のデータの作成・提供・利用の研究に取り組むかたわら、オンサイト施設を利用する海外からの研究者の支援や共同研究にも尽力する。「実際に使う人の立場になってどのようにデータを使うか」は主要な研究テーマのひとつだ。オンサイト施設のデータを守る秘密分散・秘密計算と呼ばれる技術の研究開発を行う一方、「持ち出し不可であるデータを見られない状態でモデル式を作ったり、計算結果を求めたりするためにどうしたらいいか」などを研究しているという。



オックスフォード大学と一橋大学の共同研究

東アジアの家事や子供の教育について研究するオックスフォード大学社会学部との共同研究が始まったのは、約3年前にさかのぼる。共同研究者の1人で家族社会学などを専門とするエカテリーナ ハートグ研究員は、日本の社会生活基本調査のデータを使って、東アジア、中でも日本の家内分業におけるジェンダーバランスや、世帯収入や両親の教育と子供の日常生活の間の関連について研究を進める。

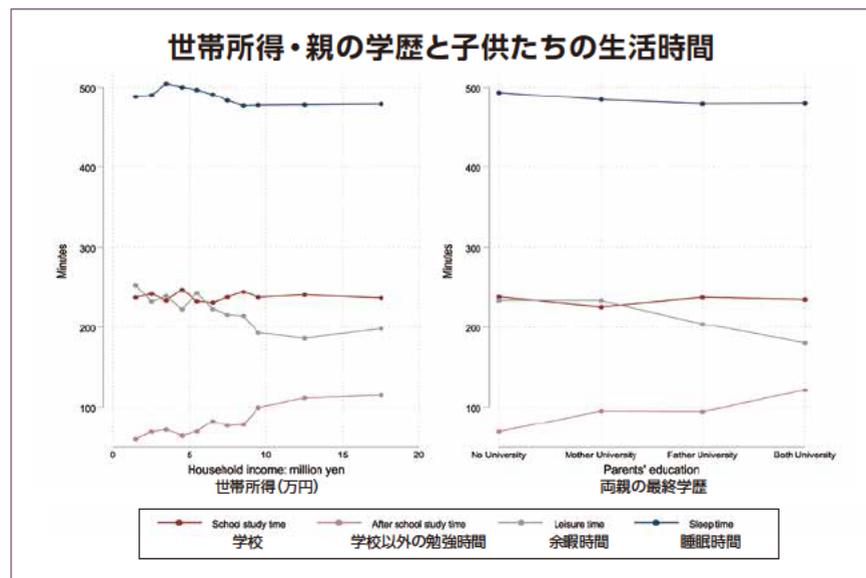


「1991年から2016年までの25年間の社会生活基本調査の調査票情報を使って、日本の子供たちの生活時間のデータ分析をしています。子供の時間を学校の勉強時間、学校以外の時間、睡眠、暇な時間の4種類に分類し、時間の使い方を分析したところ、まず睡眠時間と学校の勉強時間が長く、次に余暇時間となりました。残りは学校以外の勉強の時間で、平均で約90分程度です」。

ハートグ研究員はさらなる解析により、この学校以外の勉強の時間が、子供が



属する家族によって異なることをつきとめた。「子供が属する世帯の所得が大きくなると、学校以外の勉強時間が長くなってしまいます。また親の最終学歴との関連で子供の時間の使い方を分析すると、両親の学歴が高い家族の子供は学校以外の勉強時間が長くなり、余暇時間が短くなってしまいました」。



生活時間データから ジェンダーギャップを明らかにする

一方、同じデータを使った共同研究でも、白川研究員は家庭内の家事・育児などのジェンダーギャップに注目する。「6歳未満の子供を持つ夫婦がどのような時間の使い方をしているか、その無償労働時間に関心を持っています」と、白川研究員は言う。「ポイントは平日と休日の違いで、男性が会社へ行って仕事をして遅く帰ってくると想定される平日と、そうでないはずの休日とでどのような違いがあるかを見てみると、実は休日についても夫婦の就労差は変わらないんです。特に食事の用意については休日であってもまったく変化がなく、夫はもともとやらないということになります。逆に育児には平日の2倍の時間を割いているといったことなどがわかってきています」。

曜日別6歳未満の子供を持つ夫・妻の無償労働時間(平成28年)
-夫婦と子供の世帯

	夫				妻			
	週全体	平日	土曜日	日曜日	週全体	平日	土曜日	日曜日
無償労働	1.35	0.56	3.12	3.13	8	8.07	7.59	7.26
家事	0.22	0.17	0.33	0.34	3.42	3.48	3.23	3.36
食事の管理	0.08	0.07	0.11	0.11	1.57	2.01	1.48	1.46
住まいの手入れ・整理	0.06	0.05	0.07	0.11	0.44	0.44	0.37	0.5
衣類等の手入れ	0.04	0.04	0.03	0.04	0.41	0.42	0.35	0.38
その他	0.03	0.00	0.12	0.09	0.21	0.2	0.23	0.2
育児	0.45	0.29	1.27	1.23	3.21	3.38	3.06	2.17
乳幼児の身体の手入れと監督	0.11	0.09	0.15	0.18	1.4	1.5	1.25	1.01
乳幼児と遊ぶ	0.2	0.09	0.42	0.48	0.57	0.57	0.53	0.58
子供の付き添い等	0.05	0.00	0.22	0.12	0.14	0.12	0.26	0.1
子供の送迎移動	0.05	0.06	0.05	0.03	0.25	0.32	0.1	0.02
その他	0.04	0.04	0.03	0.03	0.06	0.07	0.11	0.04
買い物・サービスの利用	0.17	0.08	0.37	0.45	0.35	0.25	0.54	1.02
家事関連に伴う移動	0.08	0.02	0.25	0.22	0.19	0.15	0.32	0.28
ボランティア活動関連	0.03	0.00	0.1	0.09	0.03	0.02	0.05	0.04

※総務省統計局「平成28年社会生活基本調査」平成29年12月22日より

また、女性の労働問題では「M字カーブ」と呼ばれる現象が知られている。高校や大学卒業後の女性の就職率は高いが、結婚・出産の期間に仕事を辞め、30、40代で復帰するという変遷を示すものだ。「現在はM字の凹みが浅くなってきて

いるのですが、依然として、管理職へつながらない要因にはなっていると考えられます。日本の女性の就労率や管理職率の低さが海外から指摘されていますが、そのような原因をさぐることによって、日本の女性の就業労働がどのようになっているかを明らかにしています」と白川研究員は言う。「またウィズ・コロナの時代になって、在宅ワークが推奨され、夫婦が共存できるような環境づくりが大切であることをつくづく感じます。ネットワークや人工知能なども含むデジタルトランスフォーメーションは、女性の地位向上のひとつの鍵になるのではと期待しています」。

日本のデータは歴史があって「質が高い」

近年、欧米に続き日本でも、エビデンスに基づく政策立案（EBPM）への関心が高まっているが、社会データの場合、自説に都合のいい結果が出たらそれを使いたいという流れになる可能性もある。「それがどういうデータなのか、本質を押さえて使わなくてはなりません」と白川研究員は言う。「また国際的な連携を



通じて、各国のデータと比較できるようにしたり、同じ課題に携わる海外の研究者とどう改善していけばいいかを議論したりすることも重要です」。

白川：日本のデータを使ってみたいかがですか？

ハートグ：一橋大学のオンサイト施設は便利で、研究を進める上でとても感謝しています。日本のデータはサンプル数が多く、また歴史が長くトレンドやパターンを見ることができるデータが1991年からあるといった点でも、とても質が高いです。そのうえ回答率が9割を越える高さであることも日本のデータの特徴で、他の国々ではまずあり得ません。

白川：ありがとうございます。ところで社会生活基本調査は5年に1度行われるのですが、秋の連続した2日において行うため、週末と平日の違いの比較ができないのが難点だと私は考えています。





2018年の生活時間に関する研究会。白川研究員が座長を務める。

ハートグ：社会生活基本調査が10月頃に限定してデータが集められているのは、確かに不便ですね。イギリスには2014～15年に18か月間にわたって集められたデータがあります。これだと春と冬の時間の使い方の比較ができるし、冬になると寝る時間が長くなる（笑）といったこともわかります。

白川：2020年1月にオンサイト施設のコンピュータを入れ替えたので、次回はより高速にお使いいただけると思います。

ハートグ：コロナが収束し次第、ぜひまた訪問したいです。どうぞよろしくお願いします。

※本インタビューと対談は、オンラインで行われました。
公開日：2020/11/10



2019年の生活時間に関する研究会で発表をするハートグ研究員。

データサイエンス時代の 人文学はどう進む？

現代科学がいま、データサイエンスという大きな潮流の中にいることは間違いない。人文学においても、多様な原資料を電子化し、利活用していく「デジタル・ヒューマニティーズ」などの新しい動きが活発だ。そこで今回は、情報・システム研究機構 データサイエンス共同利用基盤施設 (ROIS-DS) と、多数の貴重な古典籍を所蔵する人間文化研究機構 国文学研究資料館 (国文研) との間で共同で進められている研究について、国文研のロバート キャンベル館長 (取材当時)、ROIS-DS 藤山秋佐夫施設長、同・人文学オープンデータ共同利用センター (CODH) の北本朝展センター長に話を伺った。



こちらからインタビューの動画を見ることができます。

答える人

ロバート キャンベル 館長 ※2021年より早稲田大学特命教授

[国文学研究資料館]

Robert Campbell. 人間文化研究機構 国文学研究資料館 館長。ニューヨーク市生まれ。ハーバード大学大学院東アジア言語文化学科博士課程修了、文学博士 (Ph.D. 1992年)。専門は近世・近代日本文学。1985年に九州大学文学部研究生として来日後、東京大学教授などを経て、2017年より現職。専門分野と関連が深い文芸ジャンル、芸術、メディア、思想などに関心を持ち、テレビ等への出演も多数。



答える人

北本朝展 センター長

[データサイエンス共同利用基盤施設]

きたもと・あさのぶ。情報・システム研究機構 データサイエンス共同利用基盤施設 人文学オープンデータ共同利用センター (CODH) センター長、国立情報学研究所教授。東京大学工学系研究科修了、博士 (工学)。デジタル・ヒューマニティーズと呼ばれる研究分野を推進し、特に日本古典籍の魅力あるさまざまな画像データ、くずし字データなどを対象に、AIや機械学習を駆使した活用を進める。



答える人

藤山秋佐夫 施設長

[データサイエンス共同利用基盤施設]

ふじやま・あさお。情報・システム研究機構 データサイエンス共同利用基盤施設 施設長、国立遺伝学研究所特任教授。理学博士 (名古屋大学)。専門は分子生物学、ゲノム科学。1990～2003年にかけて行われた国際ヒトゲノム解読計画に参加し、その後チンパンジーの全ゲノム解読等を達成。2016年より施設長として、分野を超えてデータサイエンスのさまざまな分野の研究促進や若手育成に注力。



「データサイエンス」時代の人文学へ向けて

「データサイエンスという言葉、最近よく聞きますね」と言うのは、2016年度に設置されたROIS-DSを担う藤山施設長だ。「でもデータサイエンスという特定の学問分野があるわけではありません。データというのも、絵とか文書とか、あるいは数値であるとか、実にいろんなものが含まれます。データサイエンスとは、そういういろいろな種類の、中でも特に大量のデータを取り扱い、さらに実際の研究者が使えるようにしたり、また社会の人たちが理解できるようにしたりする、そのための考え方だと思っていただければ結構です。そしてそれはやはり統計学や計算機科学といった学問に基づいた、物事の考え方だということです」。

藤山施設長自身の研究のバックグラウンドは、生物学だ。1990年頃からは、歴史的な研究プロジェクトとなった国際ヒトゲノム計画にも参加・推進した経歴を持つ。「ヒトのゲノムを文字にすると全部でざっと30億文字あります。これを全部解読して計算機の中に入れ、その情報を解析して皆さんに使ってもらおうというヒトゲノム計画は、2004年に一応の完成を見ました。これをもって、生物学もデータ中心科学の仲間入りをすることができたこととなります」と、藤山施設長は言う。「人文学における古文書やくずし字も、古くは偉い先生方が本に埋まるようにして、いろいろ難しいことを考えられていたのですが、データサイエンス時代においては計算機を使うことによって、より広く使ってもらおう

という、そういう進め方になってきたと言えるでしょう。ROIS-DSは、さまざまな学問分野におけるデータサイエンスを推進していますが、デジタル・ヒューマニティーズはその一例なのです」。



デジタル化が難しい、日本の古典籍4つの特色

国文学研究資料館（東京都立川市）のロバート キャンベル館長は、近世・近代日本文学の研究者として知られる。まずは館が所蔵する資料等について聞いた。「当館は文学や歴史に特化した図書館で、原資料の閲覧・複写などが行え、多くの研究者にお使いいただいています。当館が所蔵する書籍・書物は江戸時代までに日本列島で制作されたものが中心で、長い歴史の中で巻物、卷子本、袋とじ等、形態がさまざまに推移しており、これらを全部ひっくるめて「古典籍」と呼んでいます。まさに千年以上にわたる古典史の結晶と、言うことができるでしょう」。



キャンベル館長によれば、日本の古典籍にはユニークな特徴がいくつかあるという。「まず絵がとても多い、つまり視覚要素と文字の要素が不可分であることです。次にさまざまな書体があり、漢文、片仮名混じりなど表記体もさまざまである点です。それから歴史上の人々が知識や情報をどのように積み重ねていったのかは大変興味深い問題ですが、江戸時代までの間に人々がさまざまな時代、さまざまな人々の考えを、注記というかたちで1つのテキストに併存・蓄積させることによって知識を作り上げていったことも、古典籍から知ることができます」。

「さらにもう一つ、日本の古典籍では本の大きさ、材質、表紙の模様といった姿形とその内容が関連しているという特徴があります。大きいものほど古典的で価値が高い、小さくなればなるほど、今日的、あるいは応用性の高い情報が盛

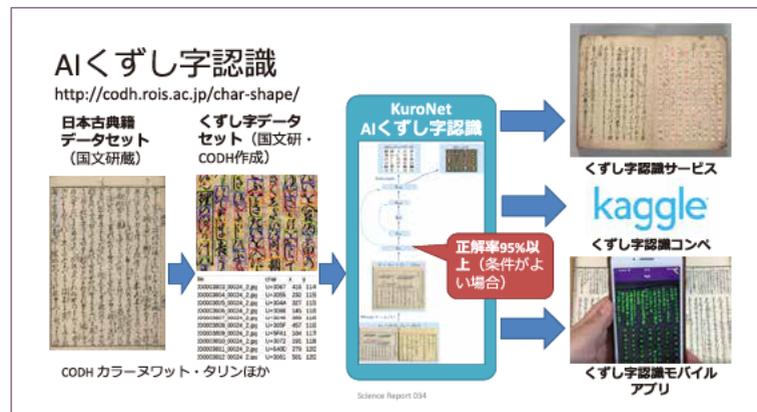
り込まれていると期待されるということです。——こういった文字の電子変換だけでは回収できない多様な示唆に富んだ資料をどうデータ化していくか、蓄積された情報をどうもれなく整理していくのか、これは大きな挑戦であり、課題になります」。

古典籍がデータ公開され、情報学の対象になった

このような課題に対して ROIS-DS は、原資料を信憑性のある汎用性の高いデータとして、永続的、持続的に実際に利活用可能なものにしていくという研究開発において「本当に不可欠なパートナー」とキャンベル館長は言う。実際、国文研との共同研究は、ROIS-DS の設置以来継続的に進められ、資料のデジタル化や、くずし字を読むなどのさまざまな成果を生み出している。

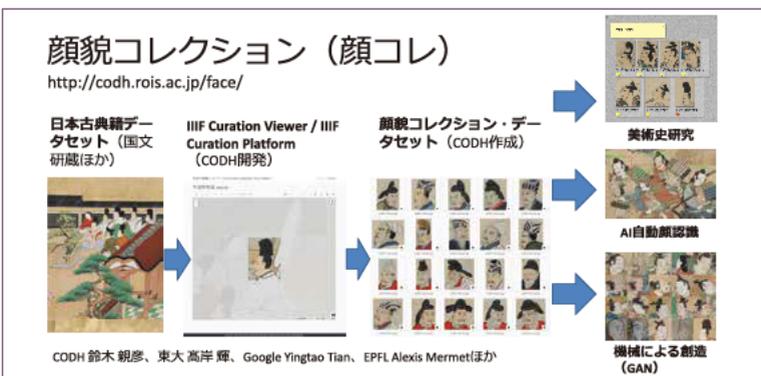
ROIS-DS の中でデジタル・ヒューマニティーズを担う CODH の北本センター長は、情報学をバックグラウンドに持つ研究者だ。「実は国文研と共同研究するまでは、おそらく多くの日本人がそうであるように、私も古典籍にはあまり馴染みがなく、遠い世界のものだと思っていました」と振り返る。「現在のような研究が可能になっ

たきっかけは、まず国文研が大量の本をデジタル・データとして公開してくださったことにあります。これによって古典籍は情報学の研究対象となり、情報学的なアイデアを試せるようになりました。しかも AI 技術の発展とデータ公開がちょうど重なったため、くずし字認識のような研究成果も生まれ、私にとっても非常に意義深いことでした」。



共同研究は、実際にどのようなかたちで進められているのか、「大きく3つの進め方があります」と、北本センター長は言う。「まず国文研で作られたデータを、情報学研究などに使いやすい形式に整備して公開するという方法があります。たとえば、国文研が構築したくずし字のデータセットを、CODHがさらに整理してAI研究者が使いやすい形式に変換し、「日本古典籍くずし字データセット」として公開しています」。

「2つ目は、国文学と情報学の研究者が話し合いながら、新しい活用のアイデアを考えるやり方です。一般の方々にも利用者が多い「江戸料理レシピデータセット」



顔コレは日本の絵本、絵巻などから顔の部分だけを切り出してデータベース化したもの。美術史の分野では以前からある特定の物だけを切り出して集め、比較するという研究が行われてきた。従来はコピーを取り、紙を切り貼りして比較していたものを、マウスで画像を囲みドラッグするだけで、特定のコレクションが出来上がるようにした。メタデータを使えば同じ種類の顔だけを比べて比較することも可能だ。美術史を始め様々な分野への応用が期待される。

は、実は国文研主催のイベント（アイデアソン）に参加した時に、江戸時代の料理本のレシピを現代の調理法で作ってみるとどうなるか、というアイデアから実現していったものです」。

「3つ目は、情報学の側から新しい技術を提案するという方法です。技術から発想することで、国文学の研究者が気づかないような手法を提案します。たとえば『武鑑』という江戸時代の大名や幕府の役人の名前などを集めた資料がありますが、コンピュータ・ビジョンの技術を用いて異なる版の本を自動的に比較するというアイデアを試しているところです」。

古典籍のマイニングで、日本文化の遺伝子を突き止める

北本：現在のわれわれの一番大きな目標は、古典籍を全文検索できるようにするという事です。日本の古典籍は、1ページずつ読まないと何が書いてあるかわからないのが現状ですので、これを検索可能にすることには最も大きな価値があるのではないのでしょうか。さらに将来の目標にはなりますが、人文学に新しい方法を導入することによって、日本の文化の非常にコアな部分を探ってみたいですね。たとえば、キャンベル館長が指摘された日本の古典籍に文字と絵が混ざっているという特徴は、おそらく現代の日本のマンガやアニメにも受け継がれているのだらうと思います。このような日本文化の特徴を、世界にどう伝えていくかも含めて、情報技術がいかに活用できるかを考えていきたいです。



左：書庫 右：図書室 写真提供：国文学研究資料館

「古典籍」は9世紀から19世紀に積み上げられた日本列島の古典籍を大量に含んだ文化資源。左：『画本虫撰（えほんむしえらみ）』（国文学研究資料館所蔵，CC BY-SA 4.0）右：『菓子話船橋（かしわふなばし）』（味の素食の文化センター所蔵，CC BY-SA 4.0）



キャンベル：『鬼滅の刃』の源が、実は18世紀の読本であるとか、あるいは黄表紙であるとかといった面白いことが、情報系の方々の投げかけによって生まれるに違いない(笑)と考えています。

それからもう1点、当館ではギャラリーも併設しているのですが、非常に層の厚い文化資源を、さまざまなパートナーと研究以外の用途で活用いただくという試みも同時進行させています。具体的にはアーティストや翻訳家を招へいして、コラボレーションを通じて活用いただく「ないじえる芸術共創ラボ」を4年ほど前から行っています。例えば芥川賞受賞作家の川上弘美さんに通っていただいて『伊勢物語』についていろいろと学んでいただき、これを元にした『三度目の恋』という全く新たな小説の世界が創造されたといった例があります。それぞれの成果を一堂に集めた特別展示〈時の束を披く〉を2021年2月15日から開催する予定です*。

*この展示は終了しています。



AR（拡張現実）を利用した展示 伊勢物語図屏風に専用アプリがダウンロードされたタブレットをかざしてスキャンすると、写真のように章段ごとに図会が拡大され、場面の解説を表示。現代語訳を読むことも音声で聴くこともできる。

写真提供：国文学研究資料館

藤山：私の研究分野で言うと、ゲノムとは生物の遺伝情報のセットであるわけですが、これらをひとまとめに研究する方法論と考え方を中心としたものが「ゲノミクス（ゲノム科学）」という分野になっています。今のお話を伺っていると、ヒューマニティーズのゲノミクス版としてヒューマノミクスとか、そういう総合的な新しい分野が資料、分析技術、さらに社会へのアウトプットまで含めて発展していけるようなビジョンが見えてきました。

北本：実は私自身も結構ゲノム解析の歴史を参考にしています。ゲノムの世界で起こったのと同じようなことが、これから人文学一般に対しても起こってくる——それがデジタル・ヒューマニティーズという新しい時代の研究方法なのではないかと考えています。（参考：北本朝展「歴史的典籍の検索機能の高度化、そしてスクリーン解析に向けて」https://www.nijl.ac.jp/pages/cijproject/images/fumi_6.pdf）

キャンベル：日本の古典籍のデータ化を通じて、現代においても脈打つ、まさに日本列島に生きた人々の行動のパターン、世界観、死生観、あるいは市民生活におけるさまざまな人々のチョイスというものが、実は歴史的にさまざまな模様としてあぶり出されてくるということを日々実感しています。たとえばこの1年というもの、新型コロナウイルスの脅威が世界中を襲っていま

す。日本ではロックダウンを強制的に行うことなく、業界や地域ごとに行動自粛や何らかの行動パターンを要請するという形での対応が進められてきましたが、これ実は、世界的には特異ともいえる取り組み方なんですね。江戸文学専門の者から見ると、江戸時代の幕藩体制の中で、業界や地域ごとにさまざまな統治能力、すなわちガイドラインを当事者の集団（＝仲間）へ下ろしていったしくみを、まさに彷彿とさせるわけです。これは自然環境に対しても、地域の人々の働きかけによって維持されてきた「里山」のような例を挙げることができ、古典籍の推移とも、照らし合わせて見ることができます。

このようなさまざまな形で、まさに日本人の心、事柄、あるいは言葉のゲノムというものを、古典籍が持っている——1人ひとりで読んでいたのでは発見できないことを、これから情報学との協働の中で見出していけるのではないかと考えています。今日の鼎談で、非常に勇気をいただいたと思います。

藤山：やはり歴史的にも、文書を収集するというのは文化の根源だと考えますし、これに最新の情報技術を適用して研究が進められ、さらに社会に向けて発信されているということに、今日は改めて感銘を受けました。日本（人）文化論といったさらに大きな議論に発展していけばいいなという思いが、共有化できたこともたいへん有意義でした。

※本インタビューと対談は、
オンラインで行われました。
公開日：2021/01/12





林屋子十七萬卷
海山城井平町西福寺 柳菴藏書資料

ウィズコロナ社会の 諸課題は どう解決 できるだろう？

2020年は、新型コロナウイルス感染症が世界的に猛威をふるう年となった。2021年に入っても、国内の感染者数は、緊急事態宣言の発出とも関連して、上昇と下降を繰り返している状況だ。情報・システム研究機構 統計数理研究所では、長年にわたる感染症を対象とした数理モデリング研究を踏まえ、2020年3月よりCOVID-19に挑むデータサイエンスの取り組みを強化してきた。しかしこのような研究を駆動するのは、データだけではなく「ウィズコロナの時代にまずどういう社会を形成していくべきかという指針やビジョン」なのだとして統計数理研究所 椿広計所長は指摘する。そこで今回は、社会科学の分野で長く活躍され、椿所長と協力体制の下、研究プロジェクト「ポストコロナ未来社会と横幹知（横幹連合）」に取り組む学習院大学 遠藤薫教授とともに、学術による解決の道を探る。

こちらからインタビューの
動画を見ることができます。



答える人

遠藤 薫 教授

〔学習院大学〕

えんどう・かおる。学習院大学法学部教授。1977年東京大学卒、博士（学術）（東京工業大学）。専門は理論社会学、社会情報学、社会シミュレーション。メディアと社会／文化変容、世論形成、社会変動などの広い関心に基づき、社会科学の視点から特にソーシャルメディアに関わるデータを駆使した研究で知られる。2003年より現職。『ソーシャルメディアと公共性（2018年、編著）』ほか、著書多数。



答える人

椿 広計 所長

〔統計数理研究所〕

つばき・ひろえ。工学博士（東京大学）。独立行政法人統計センター理事長を経て、2019年より現職。自ら「応用統計家」を任じて公的統計、品質管理、医薬品許認可、環境計測、ビジネス科学などの諸領域で科学の文法としての統計的方法の普及啓発とデータサイエンティストの育成に努める。総合研究大学院大学（SOKENDAI）参与。筑波大学名誉教授。



経済か、感染防止か、価値観の分断

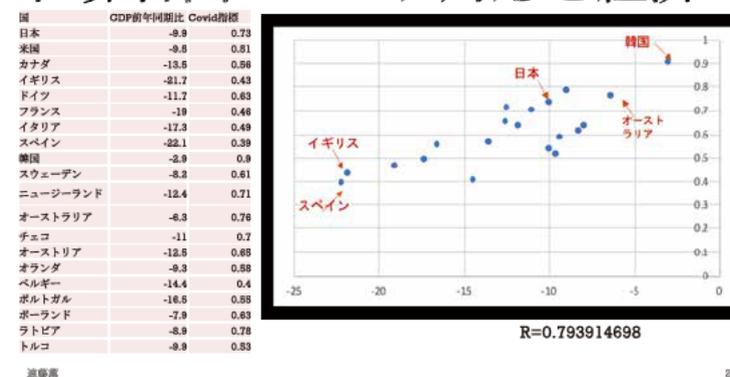
遠藤教授によれば、新型コロナウイルスの感染拡大は「グローバリゼーション、気候変動など、既に世界の未来に不安を与えていた諸問題の帰結」でもあるという。その解決に向けて国連はSDGsを提案しており「持続可能な社会にするために今必要とされる社会倫理を検討するには、社会科学の方法論と、ネットワーク理論、制御理論等のデータサイエンスの方法論をつなぐ「横幹知」という視座がたいへん重要」と遠藤教授は言う。

「コロナの感染拡大に関連して、私が特に着目しているのは、価値観の分断と人々の孤立という問題です。中でも目立った価値観の分断は、経済重視か、感染防止の重視かという二者択一の議論です。人間の命を守るため経済活動を一時停止するのはやむを得ないという議論の一方、経済活動が停止すれば人々の生活や仕事が立ち行かなくなり、かえって命が失われたり、社会全体に混乱が生じる、という主張がある——しかし、感染防止と経済活動の維持は、それほど相反するのでしょうか？」



各国のGDP前期比とコロナ対応指標（国連SDG'sレポート2020年版）を照らして見ると、経済の状況がある程度うまくいっている国はコロナ対応も優れている、反対にコロナ対応がよくない国は経済も悪化という関係があるようだ。「つまり、私たちは今、感染防止か経済かではなく、感染防止も経済も、ということを考えなくてはならないのです。この目標を達成するには、多くのデータを多面的に分析し、適切な政策決定に生かす必要があります。この点に関して、社会科学と自然科学は、これまで以上に強く連携していく必要があると考えています」。

世界各国のコロナ対応と経済



コロナ禍で自殺者はどう推移したか？

また、コロナの感染拡大に伴い、自殺者数の動向も注目を集めている。コロナ感染拡大の初期から注目していた遠藤教授は「当初はむしろ昨年に比べて自殺者数は少ない傾向」があったと振り返る。「ところが、7月頃から自殺者数は急増」し始めた。「きっかけとして有名人の自殺が続いたこと、またそれが大きく報道されたことなどが指摘されました。いわゆるウェルテル効果と呼ばれるものです。しかし、人が自分の人生を諦めるとき、その背後にはさまざまな要因が潜んでいるでしょう」と遠藤教授は言う。

そうした一人一人の状況と死にたい気持ちになることとの関連を示したのが下の棒グラフだ。自殺者数の推移の背景には、さまざまな要因が影響しており「コロナ禍という短期的な状況だけではなく、もっと長期的、世界的な動向も射程に入れて考える必要がある」と遠藤教授は指摘する。「こうした変動について、世界中の研究者たちが議論を重ねています。国連が、誰一人取り残さない世界を目標とするSDGsを発表したのも、この危機意識があればこそです。しかし、現実には貧困層、若年層、女性など、相対的に弱い立場の人々が社会の中で孤立しがちです。自殺者数の動向にもその傾向が表れています」。

「自殺問題は、それ自体重大な社会問題であるだけではなく、世界全体の諸問題を端的に表現するインデックスの役割を果たしていると考えられます。その解決には、短期、長期、地域、国内、世界、個人、コミュニティ、メディアなど、さまざまなレベルのデータを収集分析し、統合的なモデルを作る必要があります。そのために横幹知、すなわち多様なバックグラウンドを持った研究者たちが互いの強みを生かしつつ、新しいデータサイエンスを駆使しつつコラボレーションすることが不可欠なのです」。



新しい社会科学×データサイエンスの可能性

椿：私たちはふだん、学術というのはおよそ一枚岩だと思ってきたわけですが、先ほどのお話のように経済学と、疾病を予防しようとする公衆衛生の価値観が対立している。ウィズコロナの時代といわれる今日ほど、文理融合研究のあり方が問われている時期はないと思います。より長期的な展望に立ってコロナが起きた後の将来の社会をどうデザインするか、そのためにデータサイエンスをはじめ学術システム全体がどういうふうに関与しなければいけないかが問われています。

遠藤：経済学の先生方の中にも、経済を単に利益重視と捉えない方たちもたくさんいるし、また社会倫理、理論的な哲学などの分野の方々ともお話をしていく必要があります。そしてたとえば弱者を切り捨てた場合と切り捨てない場合について、椿先生のご専門のデータサイエンスをベースにモデルをつくって、さまざまな条件でシミュレーションしていく。これによって、科学におけるコミュニケーションをどう行うかについての共通の理解が生まれていくことが重要です。社会学内部でも、たとえば、長い間マクロ社会科学とミクロ社会科学という葛藤がありました。2つの間をつなぐ動的な一般モデルを作れないかという課題に対して、今日ではシミュレーションというツールへの期待が高まっています。しかもシミュレーションは可視化の技術によって、



誰にでも比較的納得しやすいかたちで見せることができます。異なる分野の研究者が同じツールを使うことで、分野ごとの言葉の壁をある程度低くし、さらには産学官、また一般の方々ともつながることができるのではないかと期待しています。

椿：その背景には、数理科学においてエージェント・ベース・モデルのような実験のツール（コンピュータ上に自律的に行動する個体やグループを生成し、システム全体への影響を評価するシミュレーション手法）が開発・発達してきたことがありますね。これが社会科学を、これまでの認識科学から、社会をデザインする設計科学という方向に急速に動かしている気がします。

遠藤：ええ、私はその方向に大きな期待を持っています。人間って非合理的なことをいっぱいしますよね。個人の動機付け、心理、それから世界観といった自然科学では捨象されてしまうような気持ちの部分、これをモデルの中に入れ込



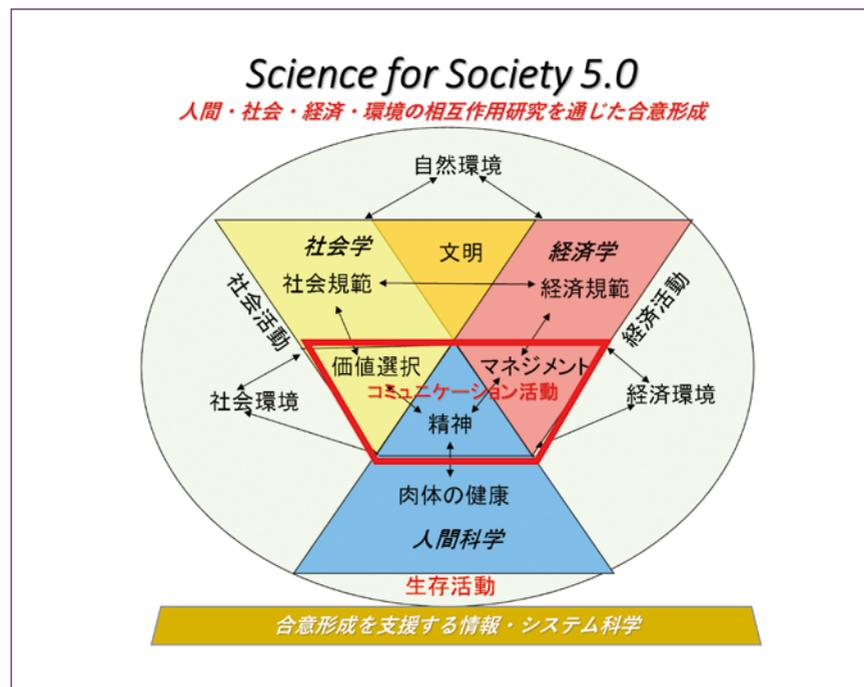
んでいくことによって、新たな発見や、より厚みのある社会の記述というのがシミュレーションで可能になるのではないかとこのように思っています。

ポストコロナの学術の構成とは？

椿：私は人文学や社会学は本来、人間がどんな価値を追求すべきかといった倫理や規範を、まず与える学問だと考えています。検討のための社会実験は簡単にできるものではないので、やはりシミュレーションを用いて、選択すべき価値を検討していくのが妥当だと思います。そこから新しいタイプの人文学や社会科学が生まれてくる。さらにわれわれデータサイエンスや数理、あるいはエンジニアリングの分野の研究者が参集して、ポストコロナの問題解決に寄与していく——社会科学による価値の選択を頂点として、学術全体がシステムとして働くべきだということは、昔からわかってはいたけれども、まさに今こそ、日



本の研究力を支援する立場のわれわれ大学共同利用機関が適切な役割を果たして、実現できればと考えています。



「情報・システム研究機構では、2010年から2015年までの第2期中期目標・中期計画期間において、融合研究「データ中心人間社会科学」を推進しました。私たちを取り巻く様々な環境の作用について、「情報」と「システム」という視点から捉え直し、解明するのが当機構のミッションです。価値選択などに有用な合意形成に資する文理融合研究の在り方について、今回の対話で再確認しました」と榎所長。

現在統計数理研究所は、データサイエンティストの育成、日本で統計数理という学術の何を教えなければいけないかといった教育の課題にも取り組んでいます。今日伺ったことから考えると、シミュレーションとは基本的にどんな考え方なのかという一歩踏み込んだ理解に基づいて、エージェント・ベースのような方法を含めて、自ら社会のためにシミュレーションを設計しようという自発的な研究者に統計数理の知を教えるのも大事な役割だと考えました。研究所内にある統計思考院では、このような役割を意識した研修の場として、2021年度

から大学教員の育成事業を始めたいと思います。社会科学の先生方との共同研究において、きちんとコミュニケーションができるような研究者育成ができれば、先ほどからご議論いただいている横幹知につながるのではと考えています。

遠藤：はい。価値観の対立や分断は、どれか1つが正しいというふうには決まりません。では何をよとするのかを、シミュレーションなどで状況の理解を共有して、みんなで議論することによって、社会としての一つの意味決定ができるかもしれません。しかし現実には、全員が将来にわたって1つの価値観でまとまるとは限らないですよね？ そこで議論を積み重ねることによって、暫定的な答えを出しながら社会を動かしていく。そこにはすべての人たち、つまり「誰も取り残さない」みんなの意見が反映されていく——そのような場づくりを私は1つの目標と考えています。その合意形成に、理解の架け橋としてのデータサイエンス、数理科学が本当に重要だと考えています。

そしてまた決定した結果どうなったか、社会はどう変化したのか、常に次のステップで検証し、それによってさらに次の意思決定を重ねていくというプロセスがたいへん重要です。静的な「エビデンス」ではなく、逐次状況を確認しつつ、ダイナミックに変化していけたら素晴らしい！——と考えております。

榎：若手の数理の研究者たちが、共同研究を通じて、本当に必要な数理的構造やシミュレーション・モデルに気付くチャンスになるのではないかという期待もしているんです。やっぱり人文学、社会学は深いなあと……コロナ禍を経験して、ますますその重要性が顕在化したと思うので、ぜひ先生にいろいろご指導いただければと思います。

※本インタビューと対談は、オンラインで行われました。
公開日：2021/03/10

SCIENCE REPORT



SC ENCE REPORT 001 006

北極を知って地球を知る。

2018年



SC ENCE REPORT 007 012

僕らはゲノムできている。

2018年



SC ENCE REPORT 013 018

データサイエンスで
ここが変わる。

2019年



SC ENCE REPORT 019 024

オーロラは語りかける。

2020年



SC ENCE REPORT 025 030

データは誰のもの？

2020年



SC ENCE REPORT 031 035

デジタルな人文、
データな社会

2021年(本書)

大学共同利用機関法人

情報・システム研究機構について



平成16年、すでに大学共同利用機関として活動していた国立極地研究所、国立情報学研究所、統計数理研究所、国立遺伝学研究所の4研究所が結集し、大学共同利用機関法人情報・システム研究機構が誕生しました。全国の大学等の研究者コミュニティと連携して、極域科学、情報学、統計数理、遺伝学についての国際水準の総合研究を推進する中核的研究機関を担うとともに、21世紀の重要な課題である生命、地球、人間、社会など複雑な現象に関する問題を、情報とシステムという視点から捉え直すことによって、新たな研究パラダイムの構築及び新分野の開拓を目指しています。

平成28年度、当機構は4研究所に「横串」を貫く組織改革を行い、「データサイエンス共同利用基盤施設」を設置しました。これにより、データ共有支援、データ解析支援、データサイエンティスト育成の取り組みを一層強化し、社会のイノベーションにつながるデータ駆動型科学の発展を推進しています。そして研究者コミュニティの要請に応える共同利用・共同研究により、大学等における研究の発展に貢献するとともに、産業界との連携や、市民が参加するオープンサイエンスも進めています。また総合研究大学院大学の基盤機関として、もうひとつの重要な使命である人材育成にも取り組んでいます。

情報・システム研究機構は、各研究所の学理の追究に基づき、データサイエンス時代の新しい研究パラダイム構築を通じて、現代の課題解決や超スマート社会構築等の社会の要請に応じてまいります。皆様の一層のご支援、ご協力を心よりお願いいたします。

機構長 藤井良一
2021年12月

『情報・システム研究機構ブックレット』について

本機構は、新たな研究パラダイムの構築と新分野の開拓を推進し、
また大学共同利用機関法人として、大学等の研究の発展に貢献しています。
『情報・システム研究機構ブックレット』は、
その研究と貢献をわかりやすく紹介していく、シリーズ小冊子です。

情報・システム研究機構ブックレット6

SCIENCE REPORT 031-035

デジタルな人文、データな社会

著 者 大学共同利用機関法人 情報・システム研究機構
監 修 椿広計(統計数理研究所 所長) (Science Report031-033、035)
藤山秋佐夫(データサイエンス共同利用基盤施設 施設長) (Science Report034)
取材・文 池谷瑠絵(情報・システム研究機構 URA広報)
写 真 飯島雄二(p9-10、12-15、19下、20、23、26-27、31-32、34、36-37)
画 像 株式会社らくだスタジオ(p40-43、45、48-49、54、57-59)
デザイン ヤマノ印刷株式会社
シリーズデザイン hata design
発 行 大学共同利用機関法人 情報・システム研究機構
〒105-0001 東京都港区虎ノ門4丁目3番13号
ヒューリック神谷町ビル2階
TEL : 03-6402-6200 FAX : 03-3431-3070
<https://www.rois.ac.jp/>
発行日 2021年12月10日

©Inter-University Research Institute Corporation
Research Organization of Information and Systems, 2021
ISBN978-4-909638-17-5 C0330
Printed in Japan



大学共同利用機関法人

情報・システム研究機構

Research Organization of Information and Systems

ISBN978-4-909638-17-5 C0330