

特集

Web上の大量のデータから 人間・社会活動を知る

データ中心人間・社会科学の創出を目指して

個人情報開示に対する心理的障壁と
漏えい防止技術

Webデータで探る観光政策

情報学と統計学の融合がもたらすもの

情報通信技術と法制度をめぐる課題

NII Today

National Institute of Informatics News



NII Interview

データ中心人間・社会科学の創出を目指して

データ中心の人間・社会科学を創り出そうという「新領域融合研究プロジェクト」が2010年4月に始まった。国立情報学研究所（NII）と統計数理研究所（ISM）がタッグを組み、Web上にある膨大なデータを活用する手法を編み出す。それで政策決定を支援し、大災害時でも復旧力の強い社会を作ろうというプロジェクト。目標は大きいのだ。

高橋 NIIと統計数理研究所は、2004年に国立極地研究所と国立遺伝学研究所とともに大学共同利用機関法人「情報・システム研究機構」という組織の中に入ったのです。

曾根原 ええ、それ以来、研究機構は地球環境・生命・情報の3本柱でやってきましたが、昨今のインターネットやWebの普及などを考えて、第4の柱として「人間・社会」という研究領域ができたんです。WebやSNSなどの空間にはデータがたくさんある。もともと、デジタルもアナログも関係なくデータはいたるところにあったわけですが、それがネットワークでセンシングできるようになった。そのデータをうまく解析してモデル化して、実社会の困っているところにフィードバックして人々の生活を良いものにしていくというスパイラルを作らなければいけないと思っています。これを私たちは、情報循環基盤と呼んでいます（図）。

リアルタイムで人の動きがわかる

高橋 具体的にはどういうご研究を？

曾根原 例えば、Webから宿泊予約データや交通予約データを集めています。飛行機と電車とホテルの予約サイトからデータをとると、毎日、リアルタイムで人の動きがわかります。

高橋 Webを使っていないホテルもあるのでは？

曾根原 今はほとんどありません。最初は京都市観光部の人に手伝ってもらってデータを出してもらったんですが、どんな小さなホテルでもWeb予約サイトを使っていることがわかった。もちろん、電話予約もあります。それで満室になると旅館から予約サイトに「部屋が埋まったよ」と電話が来る。お盆など正確に取れない時期もあるのですが、誤差は9%以下です。公的統計機関がとった実態調査と私たちのWeb分析データはほとんど一致しています。統計データは、結果が出るまでには3~4カ月かかりますが、Web分析データは瞬時に得られます。

曾根原 登

Noboru Sonehara

国立情報学研究所
情報社会相関研究系教授

高橋 なるほど、人手をかけずに瞬時に情報が集められるわけですね。でも、これをどう活かすのですか？

曾根原 基本的にはデマンド（需要）コントロールができるということです。ホテル側からすれば、料金設定や集客をうまくできるようになる。

高橋 でも、それは今でも大体わかるような気がする。お正月とお盆は宿泊料金が高くなっているし…。

曾根原 料金設定だけでなく、閑散時にマラソンや国際会議などのイベントを企画したり、さらにはどこに新しい空港や観光地を造るかといったことにもこうした科学的な根拠データが活用できます。

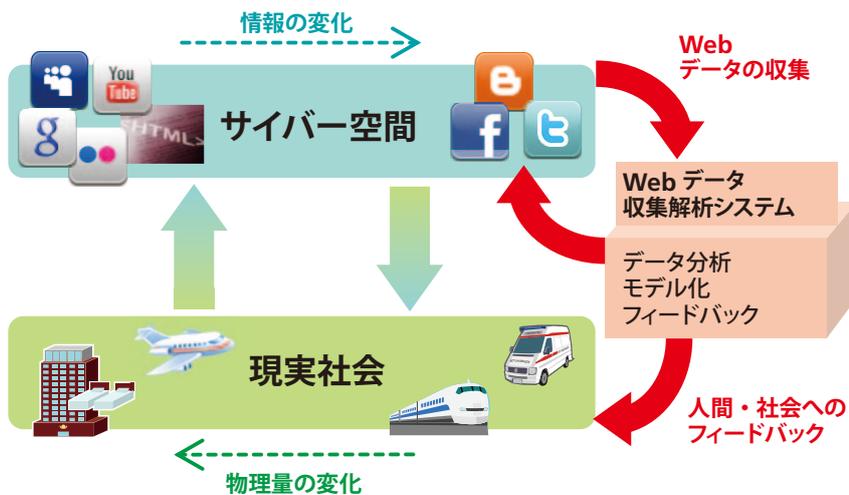
高橋 そうですね。政策決定はエビデンス（証拠）に基づくべきだという点は同感です。

制度改革につなげたい

曾根原 災害時に役立てることもできます。例えば、関東で大地震が起きたら、「渋谷や羽田に何人いるか？」を把握する手段が今はない。本当はパスモやスイカ、携帯のデータを使えばかなりの精度で「何人」と出るはずなんです。そういう社会システムになっていない。電気通信事業法によって、事業者は通信の秘密を守らなくてはならないという規則があって、携帯電話会社は位置情報を把握していても、それを安否確認や



Web Data Driven Information Flow



Webデータ駆動型情報循環システム

サイバー空間で取得・収集された多様で多量なデータは、新しい価値を持った情報として現実社会にもたらされる。現実社会で発生する新しい情報はサイバー空間に写像される。このように、サイバー空間と現実社会の間における情報の循環により、より良い社会、効率的な社会が実現されることになる。

災害直後の復旧支援などの政策決定に簡単には利活用できないようになっています。

高橋 携帯電話の位置情報が外に出してしまうとなると、相当プライベートな所までわかってしまうので、誰でも警戒感を持つと思います。個人データがわからないような形で人数だけ開示するのならいいんじゃないですか。

曾根原 難しい問題です。今回の震災で日頃から使い慣れていない情報システムは、緊急時などいざという時には簡単に使いこなせないことも分かりました。例えばWebの安否確認システムでも、ずっとスクロールして下の方にある「同意します」というボタンを押さないと使えないようになっていた。緊急時の政策決定では、個人データを瞬時に活用できるような危機に強い社会基盤が必要だと思います。

高橋 災害時に限るという条件つきであれば、皆さん情報開示に賛成するのでは。

曾根原 はい。いま、タクシーにはビデオカメラとセンサがついていて、

ぶつかったときの前後の数十秒の映像だけ取っておくようになっています。あれと同じことをやればいいわけです。普段は携帯電話の位置情報はわからなくて、地震が来たときにはその直前の情報を出してもらう。そのときは全部の携帯電話会社に協力してもらわないといけない。電車にも協力してもらおうと、瞬間的に全国マップができるわけです。それだったらいいかなと思って、今クライシスに強い社会システム設計をしています。

高橋 そういうことを実現させる技術はすでにあるんですね。それを阻んでいるのが電気通信事業法など

の制度だとすると、問題は制度であって、技術ができることは限られているんじゃないですか？

曾根原 それは違う。技術によってデータに新しい価値が見いだせるようになるわけです。その結果、制度に影響が及ぶ。こんな価値があるんだというのを見せないと、制度は変わってこない。

高橋 じゃあ、このプロジェクトはそれを見せるのが目標ということになりますか？

曾根原 そうです。災害時や緊急時に、どこに誰がいるかがすぐにわかるようになれば、避難や救援など支援にも活用できます。モバイルやSNSなどICTが世界で一番進んでいる日本だからこそできることがあるはず。これは、情報学に課せられた使命のひとつだと思います。プライバシーの問題も「なんとなく気持ち悪い」といったレベルで片づけるのではなく、科学的なデータに基づいて判断できるような社会にしたい。

今は新しい情報サービスを開発することに対して、コンプライアンスコストが非常にね上がっています。コンセンサスを得るのは結構難しいですが、そういうところをデータに基づいてやっていくのがこれからの政策支援だと思います。

インタビューの一言

やりたいことは山ほどあって、でも現実にできることは限られていて、だからもどかしくてしょうがない。曾根原さんの話からは、そんな焦燥感がびんびんと伝わってくる。「危機に強い社会システム」は誰もが望むこと。でも、民間企業に政府・自治体、そして政治家と多岐にわたる関係者の協力なくしては実現しない。従来の研究プロジェクトの枠を飛び出ることが最初から運命づけられたプロジェクト。そう理解した。

高橋真理子 Mariko Takahashi

朝日新聞編集委員



個人情報開示に対する心理的障壁と漏えい防止技術

近年、個人のプライバシーに関する情報は、公開を拒否する人も多く、個人を特定できないよう、匿名化が必須である。しかし、過度の匿名化によって情報の質が低下し、情報を有効に活用できないという問題も生じている。どうすれば情報を保護しつつ、活用を促すことができるだろうか。

匿名化と統計データ活用のジレンマ

——現在、お2人は新領域融合プロジェクトの1つ、「人間・社会システム」分野において、それぞれどのような研究をされていますか？

越前 私は、これまで大学や研究機関において、閉じて利用することが前提だった個人調査データなどの統計情報を、一定の匿名性を確保することにより、組織の枠を超えて学術的に活用する研究を行っています。

たとえば、ある医療データに、氏名、住所、年齢、病名、投薬に関する情報が記載されていたとします。そのまま公開すれば個人が特定されてしまうので、氏名や住所を削除し、茨城県を関東に、32歳を30～39歳といった具合に、属性をぼかすことにより、同じ属性をもつ人を複数人存在させ、個人を特定できないようにするのです。

ところが、このように匿名化を進めれば進めるほど、データの正確さや価値は損なわれていきます。いわば、“データの匿

名化の度合い”と“データの学術的有用性”はトレードオフの関係にあるということです。

そうしたなか、2009年に新統計法が施行され、調査票データの学術利用が可能になりました。それにより、データの匿名化よりも、学術的有用性に重点が置かれるようになった。従来、匿名化されたデータは誰でもアクセスできるという前提で、データの匿名化の度合いを高めることに重点が置かれてきましたが、今後は、匿名化の度合いを弱めて利用しやすくする一方で、利用者によるデータの漏えい防止に重点が置かれることになったというわけです。

小林 確かに、従来、二次分析用に公開される統計データは、使いづらいものが多くありました。例えば、

都市と村落の比較分析をしたくても、県レベルの居住地域情報しか付与されていないなかったり、個票ではなく集計データしか提供されていなかったり——。そもそも統計法が対象としているデータの中でも、私が扱うような社会調査データというのは、個人を特定できないものがほとんどなので、医療データなどとは異なり、これ以上の個票レベルでの匿名化の必要はないと思います。実際に、統計上、頻度が稀なデータは、トップコーディングとって、欠損値扱いにしていますし、匿名化を強固にする方法論自体は、すでにかなり確立されているのです。

そこで、私のほうは、いかにして人々に個人情報を提供してもらうか、という観点で研究をしています。なかでも現在、学術に限らず期待されているのが、人間行動やコミュニケーションに関するライフログの活用です。しかし、ライフログというのは、それこそプライバシーに関わるデータなので、その提供には抵抗を

越前 功

Isao Echizen

国立情報学研究所
コンテンツ科学研究系准教授

小林哲郎

Tetsuro Kobayashi

国立情報学研究所
情報社会相関研究系准教授

Web Data Driven Information Flow

もつ人が多い。そこで、何が心理的障壁になっているのか、いかにして社会にとって有用なデータを提供してもらうことができるのかについて研究を進めています。

情報漏えい防止の技術と、心理的障壁を押し下げる工夫 ——具体的な研究内容をお聞かせください。

越前 匿名化データ向けのフィンガープリント手法とって、データの匿名化プロセスの多義性（同一の匿名性を確保するのに多様な匿名化プロセスが存在すること）を用いて、個々の匿名化プロセスと利用者の識別情報を紐づけすることで、漏えいした匿名化データから、漏えい元を特定できる手法を考案しました（図）。たとえば、生年月日と性別からなるデータであれば、利用者Aには、「1971年、男性」と書いたものを渡し、利用者Bには「1971年8月10日、性別不明」といったデータを渡して、同等の匿名化レベルを確保しながら、利用者ごとに異なる匿名化プロセスで生成したデータセットを用意するのです。その匿名化プロセスと利用者IDを紐づけすることで、万一漏えいした場合、どの利用者からデータが漏えいしたのか、特定することができるというしくみです。このルールが適用されていることを知ると、利用者は情報の管理に慎重にならざるを得ません。つまり、匿名化プロセスそのものをデータ漏えいの抑止力に使うわけですね。

これを SNS やブログに応用すれば、利用者が属するグループごとに匿名化のレベルを変えながら、利用

SNSへのフィンガープリント手法の適用

投稿者が投稿したメッセージを、利用者が属するグループごとに異なる匿名化レベルで匿名化する（この例では、投稿者の所属や出身地情報を匿名化している）。グループ内では、同一の匿名化レベルで異なるメッセージを生成する。個々の匿名化プロセスと利用者の識別情報を紐づけしておけば、漏えいしたメッセージから漏えい元が特定できる。

者ごとに異なる匿名化プロセスで生成したテキストを表示することで、漏えいしたテキストから漏えい元を特定できるようになるのです。

小林 私のほうは、スマートフォンを使用している人を対象に、スマートフォンから得られるライフログ（①位置情報／② Web 閲覧履歴／③通話・SNS・Gmailによるコミュニケーション）の提供を依頼する被験者実験を行いました。その際、3種類の謝金（①1000円／②5000円／③1万円）を設け、提供してもらうライフログの種類と金額を組み合わせた条件を設け、各要因が提供率に及ぼす効果について調べました。その結果、やはり謝金が多いほど、情報提供率が上がるのがわかった。一方で、いくらお金を積まれても、個人情報は提供したくないという人が、3割程度いることもわかりました。なかでも、通話記録などコミュニケーションに関わる情報の提供には、内容が記録されないとしても抵抗がある人が多いようです。逆に、GPS データは比較的ハー



ドルが低いようです。位置情報の提供にはリスクを伴うはずですが、そのことはあまり認識されていないのかもしれない。翻って、GPS データのように、比較的ハードルの低い情報から徐々に間口を広げ、心理的障壁を取り払い、提供してもらう情報の種類を増やしていくこともできるのではないかと考えています。

個人情報の提供には高い障壁があり、またその漏えいはあってはなりません。一方で、皆が提供すれば、それだけ全体としての情報の質が向上し、「公共財」として利用価値が高まっていくと考えられます。従って、情報提供を促進するインセンティブの設計については、今後いっそうの研究が必要だと思います。

越前 同感です。データの匿名性と有用性という、相反する要素を扱うのは難しいことですが、それをうまくバランスさせる技術を開発することで、情報の有効活用に貢献できるよう、今後も研究に取り組んでいきたいと思っています。

（取材・構成 田井中麻都佳）

Webデータで探る観光政策

近年、日本政府は、これから有望な産業の1つとして観光業に力を入れている。観光業の活性化対策にも、データ中心科学が重要な役割を果たすことになるだろうと、NIIでも研究を始めた。Webデータの有効活用を目指す一藤裕・融合プロジェクト特任研究員に、研究の現状と将来の展望を聞いた。

Ⅰ 観光立国を目指して

2003年、日本政府は観光立国を宣言した。以来、観光業は重要な産業の1つとされ、活性化対策が強力に推し進められている。

2008年には観光庁が発足し、同庁が3カ月ごとに公表する「宿泊旅行統計調査報告」は、自治体が観光政策に役立てることを目的としている。国の政策に応じて、各自治体も努力しており、例えば、大規模なイベント開催時には調査会社に依頼して経済効果を評価し、今後に生かそうとしている。

このような現状に対して、一藤特

任研究員は「本当に効果的な活性化対策が行えているのか」と疑問を投げかける。「宿泊旅行統計調査報告」も「経済効果評価」も、調査に時間とコストがかかり過ぎると感じているからだ。一藤研究員は、Web社会にあふれる膨大なデータの有効利用を研究する曾根原教授のグループの一員で、2009年からWebデータの観光業への活用を研究している。

Ⅱ 玉石混交のWebデータ

観光業における活気の“ある・なし”の指標といえば、まず観光客数が増えられる。観光客数を知ることができるWebデータとして、一藤研究員が最初に注目したのが、宿泊施設の予約サイトだ。

最近では、多くの宿泊施設が予約サイトに登録しており、利用客がいつでも空室の状況を確認できるようになっている。そこには、宿泊プランの料金設定なども細かく掲載されており、情報が充実している。そこでまず、京都市内の宿泊施設について調査・研究を開始した。

「Webデータを利用する場合、それが信頼できるデータであるという“前提”が必要です」。予約サイトは、宿泊施設が公式にアップした情報で、玉石混交といわれるWebデータの中では信頼性が高い。ただし、宿泊施設は多くの客に利用してもらえる

ように、ダブルブッキングにならないようにしながら複数のWeb予約サイトに予約可能プラン情報などを提供している。そのため、単体のサイトを見るだけでは観光客数を把握できない可能性があり、複数のWeb予約サイトを対象として情報収集を行い観光状況の把握を行う必要がある。さらに、そこから観光客数を正しく見積もるためには、京都市内のすべての宿泊施設が予約サイトに登録していること、さらにアップされている予約状況が実際の宿泊者数に即していることが求められる。

これを評価するために、Web上で集めた宿泊施設予約データと、京都市が公式に把握している宿泊施設情報との比較を行った。

その結果、Web上にある京都市内の従業員10人以上の宿泊施設数が196、部屋数が約2万なのに対して、2011年の「宿泊旅行統計調査報告」に記載されている宿泊施設数は191、部屋数は約2万だった。

また、予約可能プラン数から空室数を推定し月別の宿泊施設利用者を算出した結果と統計調査報告との比較では、すべての月で、Web予約データと実績の間に最大9%程度の差しかなかった。

こうして、「京都市内の宿泊施設の予約データからは、京都市の観光客数を見積もることができる」とい



一藤 裕

Yu Ichifuji

新領域融合研究センター
国立情報学研究所
融合プロジェクト 特任研究員

Web Data Driven Information Flow

う前提が確立された。現在、この予約データの具体的な活用方法の検討が始まっている（図）。

「今」使える データを提供する

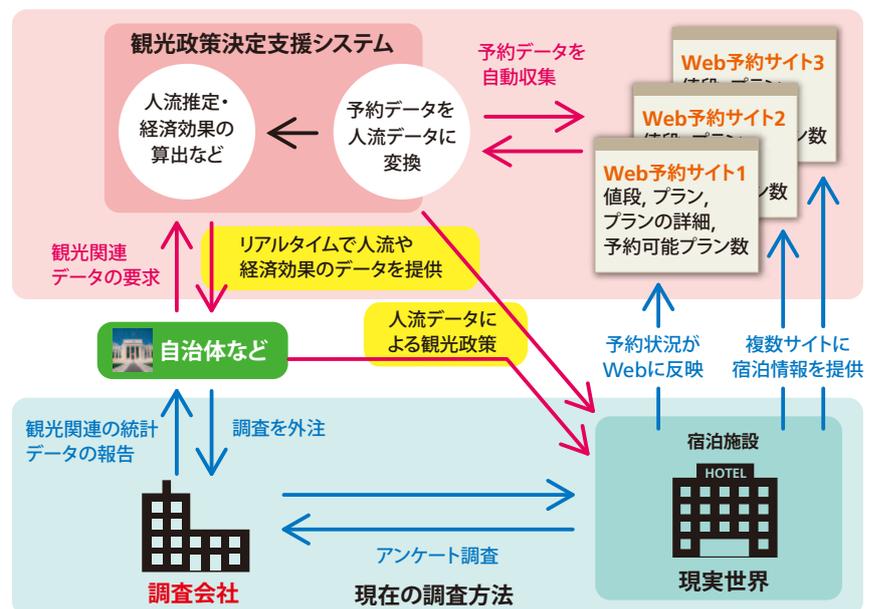
Web データから観光客数が予測できると、何が可能になるのだろうか。研究に協力している京都市は、調査会社から出された経済効果評価報告などを補足する情報として、Web データを使いたいという。

一方、一藤研究員は「Web 上でデータを集めるのはコンピュータプログラムです。ですから、調査会社が調査員を使って調べるのに比べて、データ収集に要する時間もコストも少なくてすみます」と Web データの利点を説明する。特に“データ収集に時間がかからない”ことが重要で、観光業活性化のカンフル剤になるかもしれないという。例えば、イベント開催の直前の集客状況がわかるので、悪ければ広告を打つなどの対策を講じることができるのだ。

従来の調査報告が過去の情報なのに対して、“ただ今・現在の状況”を知ることができる Web データは、このようなリアルタイム制御を可能にする。また、観光業を効果的に活性化するために、観光客の少ない時期を狙ったイベントなども開催しやすくなるだろう。

一藤研究員は、新幹線の乗客数データの研究も始めている。宿泊施設予約データと合わせて分析すれば、観光客の動きがより詳細にわかると予想されるからだ。

さらに「現時点では信頼性を評価することが難しいので利用できませんが、旅行先での SNS への投稿内



観光政策決定支援システムの概要

観光政策決定支援システムでは、コンピュータを使ってWebから自動的に観光関連データを収集し、観光客数などの観光用データに変換する。これにより、低コストかつリアルタイムな観光データの収集・分析が可能になった。

容などが利用できれば、“これまでにない面白いツアー”を企画できるかもしれないと話し、将来は、さまざまな Web データを組み合わせることで、旅行会社や宿泊施設、さらには個人に対していろいろなサービスを提供したいと考えている。

緊急時にも生きる情報

「このようにして集めたデータを、災害時にも活用したいと思っています」。2011年3月の東日本大震災では、多くの緊急システムが機能しなかった。その理由を「ある大学が送った安否確認メールは、学生たちにスパムメールだと勘違いされました。混乱した状況で、知らないメールアドレスから安否確認メールが送られてきたら、誤解しますよね」と指摘し、「このようなことが繰り返されないためにも、日常利用しているシステムを、緊急時に

も利用できるようにしておくことが大事です」と結論づける。

この考えのもと、仙台市の宿泊施設 Web 予約データを使って、緊急時に何ができるかを検討した。そして、“営業再開日”の情報を集めてグラフ化すると、被災地の復興状況が可視化できることを示した。宿泊施設の営業再開は、電気・ガス・水道などライフラインの復旧や、シーツなど備品の供給再開を暗に示しており、このデータから読み取れる情報は思いのほか多いのだという。

玉石混交の Web データだが、丁寧に選り分けて慎重に評価すれば、観光業の活性化や緊急時の状況把握に活用できることが明らかになってきた。将来的には Web データから、観光政策への提言が出されるかもしれない。Web データは、これまで以上に重要な情報になることだろう。

(取材・構成 池田亜希子)

情報学と統計学の融合がもたらすもの

大学共同利用機関法人 情報・システム研究機構は、2005年、「新領域融合研究センター」を設立し、NII、統計数理研究所、国立遺伝学研究所、国立極地研究所から得られる多種多量なデータをもとに、融合研究を推進してきた。2010年から始まった第2期では、従来の生命・地球環境・情報基盤分野に、新たに「人間・社会システム」分野を加え、NIIと統数研による融合研究に取り組んでいる。

情報学と統計学の 融合が必要な理由

——なぜいま、情報学と統計学の融合研究が必要なのでしょう？

曾根原 これまで、人間・社会に関するデータというのは、5年に1度実施される国勢調査のように、調査員による調査やアンケート調査など、オフラインで集計された静的データがほとんどでした。そこで私たちは、WebやSNSなどのインターネット情報やモバイル情報から得られる動的データの活用を進めています。Webデータというのは信憑性^{しんぴようせい}に欠ける面もありますが、これをキャリ

ブレーション（補正）し、刻々と変化する動的データとオフラインデータを融合することで、より現実に近い情報を収集・分析することができますと考えています（図）。

椿 おっしゃるように、国が行っている基幹統計はあくまでも定点観測なので、ダイナミックな動きまで見ることはできません。一方で、2005年に施行された個人情報保護法以後、調査協力が得られにくくなっています。そうしたこともあって、従来の定点観測を補うべく、リアルタイムかつダイナミックなWeb情報の活用を進めているのです。

産業界ではすでに、Webデータを活用した動きが始まっていますね。たとえば、建設機械メーカーのコマツでは、保守の目的で、遠隔地の機械をモニタリングしており、稼動状況を見ることができるという。このデータから、いまだどこで経済活動が活発に行われているかがわかるのです。つまり、サイバー空間上の情報から、国が求める基幹調査データに近い情報を取得することが可能だということです。

曾根原 一方で、サイバー空間から取得できるデータと、基幹データでは、性質が違います。それらを突合して、より価値の高い情報に高める

ことで、安心・安全や災害・減災、環境などの分野に貢献すべく、研究を進めています。

科学的根拠データを 政策決定に役立てる必要性

——基幹データは、どのように活用されているのでしょうか？

曾根原 そもそも基幹データというのは政策決定に活用されるものですが、最近では、教育研究にも活用され始めています。立川にある新領域融合研究センター・オンサイトデータ解析拠点では、学術目的であれば、基幹データの閲覧利用ができるよう



椿 広計

Hiroe Tsubaki

統計数理研究所 副所長



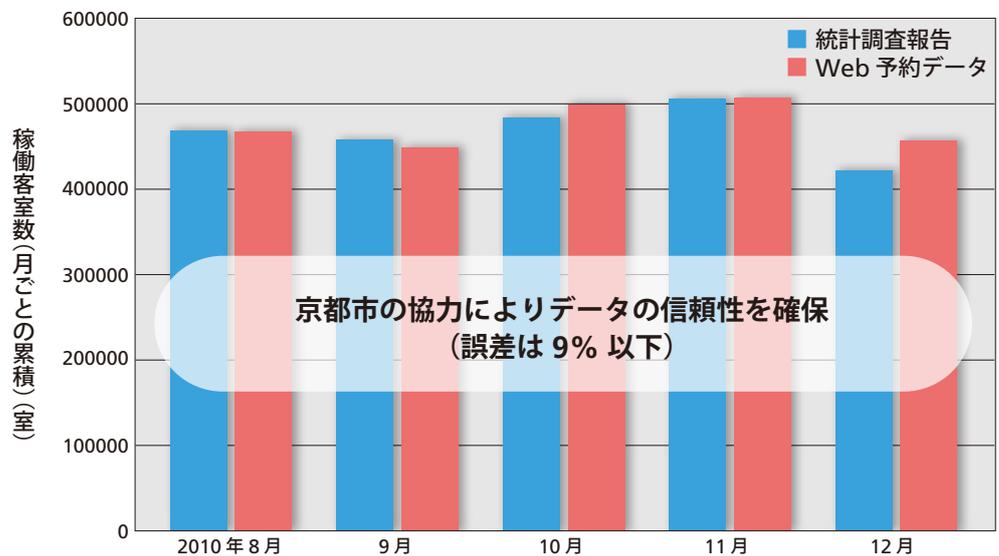
曾根原 登

Noboru Sonehara

国立情報学研究所
情報社会相関研究系教授

Web 予約データによる推定と統計調査データとの比較

Web 予約データによる稼働客室数と宿泊旅行統計調査報告による稼働客室数が、一致することを明らかにした。このため、Web データにより観光政策の決定支援が費用対効果よく実現できる。



になりました。ただ、やはりネットワーク化が不可欠ですし、こうしたデータは産業界でこそ活かされるものではないでしょうか。

椿 同感です。個人情報保護法施行以来、住民基本台帳の閲覧ができなくなり、民間のマーケティング調査がたいへん難しくなっていると言います。確かに、一定の制限は必要かもしれませんが、基幹データを利用して価値の高いものにして共有し、民間でモデル化や分析の技術をもつことができれば、日本の競争力にとって大きなプラスになると思います。

曾根原 問題になるのは、基幹データは誰のものか、という根本的な問題でしょうね。本来、基幹統計の作成に協力するのは国民の義務であり、そうして集めたものは、国民に広く還元すべきでしょう。研究目的で集めたデータなども、どこかに死蔵されたままでは、非常にもったいないと思います。

椿 いまや、国政にとってもっとも重要なものの1つが情報ですからね。そう考えると、情報は税と同等の価値があり、皆の共有財産なのではないでしょうか。それを国民や企業が秘匿してしまったら、国は何もできなくなってしまいます。一方、国は集めたデータを正しく活用する義務がある。情報をきちんと国政に活かし、さらに、分析された情報を民間に還

元すべきでしょうね。

曾根原 現代社会では、ちょっとしたことが機会損失につながってしまうので、今後はますます、科学的な根拠データに基づく政策決定や制度設計をしていく必要があると強く感じています。

データ中心科学が異分野をつなぐハブになる

——具体的には、データの融合でどのようなことが可能になるのでしょうか？

曾根原 現在、私たちはWebデータを用いることで、リアルタイムの観光政策支援を行ったり、産業環境情報をWeb上で統括して閲覧できるシステムの開発などを手掛けています。また、救急車から医療センターまで患者の映像や位置情報などを伝達することで、救急搬送の迅速化を図るシステムの開発も手掛け、成果をあげつつあります。

椿 ただし、課題もあります。たとえば、統計データから、2種類の薬を併用している患者さんに、重篤な副作用が出ていることがわかることがある。ところが匿名性が障壁となって、患者さんに直接コンタクトをとることができないのです。匿名化することで情報の質を落とすことがないよう、工夫が必要ですね。

また、分析結果の活用も大きな課

題です。たとえば、自殺に関する慶應大学の研究調査で、坂道のある街での自殺率が高いという結果が得られたそうです。一方、海岸部での自殺率は低い。それなのに、震災後の復興では、高台移転が進められようとしています。そう考えると、やはり、どこかの機関が統括的に情報をウォッチして、それを活用するしくみをつくっていく必要があるでしょう。

曾根原 私たちは情報学と統計学の専門家なので、制度設計まで手掛けることは困難です。そうした意味では、情報と統計に限らず、法律、政治、教育など、より広い意味での分野融合が必要なのだと思います。その核となるのが、まさにデータ中心科学ではないかと。データというのは、どの分野においても必要不可欠なものですし、そもそも分野に関係なく存在するものですからね——。

椿 まさに私たちなど、統計学、情報学の人間が、異分野をつなぐ、ハブの役割を担うべきなのかもしれませんね。

曾根原 大学ではなかなか難しいと思いますので、やはり大学共同利用機関がその役割を果たすべきなのでしょう。使命を果たすべく、今後も融合研究に邁進していきたく思います。

(取材・構成 田井中麻都佳)

情報通信技術と法制度をめぐる課題

東日本大震災の際、SNSが重要な役割を果たしたことは記憶に新しい。ICTの進化により、誰もが自由に世界規模で簡単に情報を発信できるようになった。その一方で、プライバシーの侵害や個人情報の漏えいなど深刻な問題も起こっている。そこで、東京大学大学院の宍戸常寿准教授に法律の専門家の立場から、現在のICTの問題点と対応策について話を聞いた。

大震災で

浮き彫りになった課題

東日本大震災時、携帯電話の通話やメールなどが通信制御により十分な機能を果たせず、安否確認や緊急避難指示が滞った。他方で、TwitterやFacebookなどSNSが情報発信サービスとして非常に有用であることが分かった。

「これは、大震災といった非常時であっても、情報発信の自由が電気通信事業法によって守られたことの好例でしょう」と宍戸准教授は語る。電気通信事業法とは、電気通信事業者やインターネットサービスプロバ

イダー（ISP）に対し、通信の秘密を守り、ユーザーの情報発信や情報交換の自由を侵害してはならないという法律だ（図）。

通信事業者やISPが、インターネットを通してやり取りされる情報を蓄積、管理できるという特別な立場を利用して、その内容を勝手に取得したり、自分の都合の良いように利用したり、情報をコントロールしたりといったことがないように、法律で定めているのである。

加えて、個人情報保護法により、ネットワーク上で管理されている個人情報も厳重に保護されており、個人情報を取得するには本人の同意が必要となっている。個人情報とは、氏名、生年月日、住所などを含み、個人を識別できる情報一般をさす。

ところが、東日本大震災により、ICTの果たす役割の重要性が再認識された一方で、緊急時になればなるほど、社会全体として、ICTを介した個人情報を利活用することの有用性も叫ばれ始めたのである。

例えば、被災地のどこに誰が住んでいるのか。そして、現在、被災者は一体どこにいるかといった情報を迅速かつ正確に把握したい場合、被災者が所有しているGPS機能搭載の携帯情報通信端末に関する情報からたどることは技術的には可能だ。

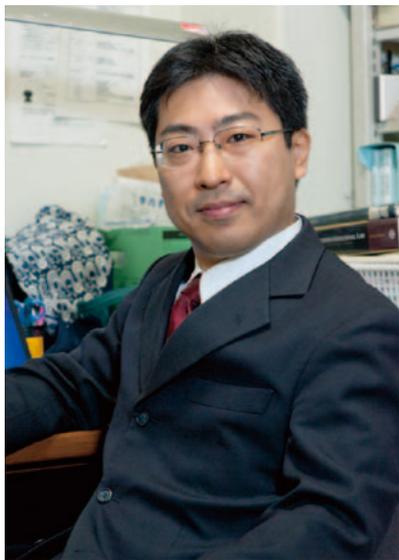
また、避難所などで被災者に対して適切な医療を施したいといった場合も、ICTを介してその人の病歴や持病を調べることができれば、いつも飲んでいる薬や投与してはいけない薬などを知ることができる。

しかしながら、これらは個人情報であるため、現在のところ、電気通信事業者やISPを含め第三者が原則として本人の同意なしに取得し、利用してはならない。こういったことから、緊急時における個人情報の取り扱いに関する見直しの必要性が言われ始めているのである。

必要な新しいルール作り

実は、個人情報保護法16条3項2号で、本人の同意を得るのが困難で、なおかつ生命、身体、財産が脅かされている場合に限っては、本人の同意を得ることなしに個人情報を取得してもよいということになっている。しかし、実際問題として、大震災などの緊急時に個人情報の取得の是非を、誰がどのような判断に基づき、決定するかは非常に難しい。なぜなら、一度流出してしまった個人情報を回収するのは不可能だからだ。

「そのため、緊急時、誰がどのような手続きにのっとり、個人情報を取得し、利活用してよいかについて、国家をはじめ地方公共団体、電気通



宍戸常寿

George Shishido

東京大学大学院法学政治学研究所 准教授

信事業者、法律家など関係者が早急に集まり、徹底的に議論し、すぐにも新たなルールを検討すべきでしょう」と穴戸准教授は指摘する。

そして、その際の重要ポイントを2点あげる。

1点目は、災害の規模など、法律の特例として認めることができる基準を定めること。2点目は、仮に特例として認められ、個人の同意を得ず、個人情報の取得が許された場合であっても、平常時に戻れば、取得した個人情報の読み取りができなくなる、あるいは自動的に消去されるなど、新たな仕組みを導入する必要があるということだ。

穴戸准教授は強調する。「大災害時、技術的には可能だからといって、安易に個人情報の取得と利活用を許してしまうと、あとで取り返しのつかないことになりかねない。とはいえ、それを恐れてこのまま何もしなければ、元来、最適に扱うことができれば非常に有用な情報が、全く活かされないことになる。このような現在の状況を打破するためには、1日も早い新たなルール作りが必要で

浮上するライフログと個人情報保護法の問題

一方で、個人情報保護法の観点から、ライフログに関する問題も浮上している。ライフログとは、インターネットユーザーの生活や行動に関する記録のことだ。

現在、パソコンやスマートフォンなどの携帯情報端末を介して、インターネットユーザーのWebサイトの閲覧履歴やネットショップでの購買履歴といったライフログが、電気通信事業者が管理するサーバーなどに蓄積され、保存され続けている。また、GPS搭載の端末であればユーザーの位置情報や移動情報も取得可能となっている。そのため、現在、ライフログをもとにした、観光や交

1947	日本国憲法施行 ・表現の自由、通信の秘密を規定（21条）
1950	電波法・放送法成立
1964	「宴のあと」事件判決 ・プライバシーの保護をはじめて認めた判決
1984	電気通信事業法成立
1985	NTT 民営化
1988	行政機関の保有する電子計算機処理に係る個人情報の保護に関する法律成立
1999	住民基本台帳法改正 ・住基ネットの導入決定
2000	IT 基本法成立 ・IT 戦略本部の設置
2001	プロバイダ責任制限法成立
2002	電子政府・電子自治体推進のための行政手続オンライン化関係三法成立
2003	個人情報保護法、行政機関個人情報保護法成立 住基ネット本格稼働
2007	新統計法成立 ・統計データの匿名化と匿名データの提供が可能に
2008	青少年インターネット環境整備法成立
2010	放送法改正
2011	東日本大震災

情報に関連した法整備の略年表

情報に関連した法律は、情報通信技術（ICT）の進化・普及に合わせて新しい法律が制定されたり、既存の法律が改正されたりして法整備が行われてきた。とくに東日本大震災では、ICTの果たす役割の重要性が再認識された。その一方で、緊急時における個人情報の取り扱いについてはその見直しクローズアップされている。また、個人情報保護の観点からは、ライフログに関する問題も浮上している。現状では、ICTの進化の速度が速く、それに対応する新たな法整備やルール作りが緊急の課題となっている。

通に関するサービス、個人の嗜好に即した広告宣伝などが、増加の一途をたどっている。

今後、ユビキタス社会が進めば、このようなサービスはさらに増えると同時に、高度化していくことになるだろう。

これが、ユーザーの同意に基づくものであり、ユーザーにとって有用なものであれば、何ら問題はない。しかしながら、実際にはライフログが勝手に取得され、利活用、あるいは悪用されているケースは少なくない。それがユーザーの不安感や不快感をおおっているだけでなく、犯罪の温床にもなっている。

ところが、閲覧履歴や購買履歴の場合、使われている情報端末の特定はできるものの、個人情報を特定することはできない。そのため、現状では個人情報保護法の対象にはならない。その結果、無法状態となっているのである。

これに対する解決策として、穴戸准教授は、「まずは、ライフログを利活用している業者は、情報の収集目的などを公表したり、自社の透明性を高めていく必要があるでしょう」と提言する。

加えて、SNSなどの普及により、誰もが自由に、世界規模で、しかも簡単に情報を発信できるようになったことにより、名誉毀損やプライバシーの侵害といった問題も深刻化している。このように、現在、ICTの進化の速度に、法制度が追いついていない状況なのである。穴戸准教授は語る。「ICTは両刃の剣です。私たちの生活をより便利にしてくれる反面、深刻な問題を引き起こす可能性も秘めています。今後、法整備やガイドラインの策定、技術的な対処などを早急に進めることで、私たちはICTを基盤とした安心、安全な未来を築いていく必要があるでしょう」。（取材・構成 山田久美）

多様な状況で 個人の判断が求められている

片山紀生

国立情報学研究所
コンテンツ科学研究系准教授

年賀状でおめでとうを使うべきか、それとも避けるべきか。もちろん東日本大震災を踏まえての悩みである。震災の影響がまだ大きい今、おめでとうという気持ちになりきれない自分がいた。Q&Aサイトなどを見ると、過度の自粛は必要ないという意見もあれば、「迎春」などですませてはどうかという意見もあった。どちらか一方が正しいわけではなく、一人ひとりの置かれている立場と状況で判断すべき問題なのだと思う。

振り返ると、東日本大震災は、個人に難しい判断を強いてきた。被災地の方々におかれては、難しい判断がさまざまな形で必要だったのではないかと思う。心からお見舞い申し上げたい。一方、東京近郊に住む私たちにとっても原発による放射能汚染の影響は少なくなく、どのような対策を取るべきかは個人の判断に委ねられている。今回の震災の大きな特徴に、被災者が多く、かつ、被災地域が広いことがある。そのため、状況は千差万別で、そのような中で個人が判断を下さなければならぬという難しい状況に立たされている。私は、テレビ放送を蓄積・分析する研究グループにおり、複数のテレビ局がどのように震災を報じているか注視してきた。震災前には、複数の局で類似したニュース映像が使われることが少なくなかったが、震災直後には、共通して使われる映像は官邸などの記者会見および陸上自衛隊が撮影した映像ぐらいとなり、大部分は各局が取材

した地震や津波の映像であった。一方、数日して原発の状況が深刻化すると、今度は現地の取材映像は減り、政府の記者会見や図表などを使った解説が中心となった。

このように各放送局は、震災直後から特別編成を組んでさまざまな情報を発信してきたが、今回の震災では、不明確な情報が多過ぎることと、一人ひとりの置かれている状況があまりにも千差万別であるために、個人の多様性に十分対応できていないように思う。その結果、個人が判断を下すには、テレビで放送される情報にも目をやりつつ、インターネットなどからも自分自身の手で情報を集めることが不可欠となっている。インターネットの情報は誤ったものや不完全なものも見られるものの、テレビでは伝えられない詳しい情報もある。私たちは今、これまで以上に、自分自身で情報を集める人で判断することが求められており、情報学としての新たな課題を突き付けられているように感じる。すなわち、私たち学術研究者は、情報の発信と活用について改めて見詰め直し、それらをさらに発展させることによって、復興に貢献することが求められている。

さて、私の年賀状の文面であるが、結局、「迎春」と書いた隣に「明るいよい年でありますように」と書き添え、おめでとうは使わないことにした。これもまた個人の判断であるが、よい年を願う気持ちは誰しも同じだと思う。一日も早い復興に貢献したい。



情報から知を紡ぎだす。

NII

国立情報学研究所ニュース [NII Today] 第55号 平成24年3月

発行：大学共同利用機関法人 情報・システム研究機構 国立情報学研究所
http://www.nii.ac.jp/

〒101-8430 東京都千代田区一ツ橋2丁目1番2号 学術総合センター
編集長：東倉洋一 表紙画：小森 誠 写真撮影：由利修一／大塚 俊
制作：日本印刷株式会社 デザイン：GRiD

本誌についてのお問い合わせ：企画推進本部広報普及チーム
TEL：03-4212-2131 FAX：03-4212-2150 e-mail：kouhou@nii.ac.jp

表紙イラスト

農業では、種々のセンサーで集められた温度、水、風、肥料などのデータをもとに管理すれば、立派な作物ができる。同様に、センサーネットワークで集められた大量のデータを活用すれば、人間・社会活動がよく分かり、より良い社会の実現に役立てることができる。