



地域ビッグデータと トにやさしいマッピング・サービス

川 正 俊 有

(秋田大学大学院 理工学研究科 人間情報工学コース 教授)

1 クラウドコンピューティングとパーソナラ イゼーション

近年の社会のデジタル化の勢いは凄まじく、 あらゆるシステムと情報がデジタル化されるよ うになってきている。デジタル化による利点は 多々あるが、サービス提供者側の観点からは、 人件費の削減が第一目的である。従来、人が行っ ていた作業を自動化、半自動化することの経済 性は大きい。ユーザ側にとっては利便性の向上 が第一要因だろう。スマートフォンやパソコン を持ってさえいれば、いつでもどこでも、しか もほとんどのサービスやソフトウェアツールが 無料で便利に利用できる状況になってきた。

仕事も、プライベートも、娯楽も、日常活動 の多くをデジタルシステムを介して楽に行える 時代となった。撮ったばかりの写真も、GPS で測位した自分の位置も、伝えたいテキストの メッセージも、リアルタイムで、特定の人、特 定のグループ、あるいは不特定多数で、情報を 共有できるのは当たり前となった(図表1)。

フェイスブック、インスタグラムやブログの ように、デジタル日記を基本に情報発信できる メディアも一般化し、さまざまな情報をタイム ライン上で管理でき、カレンダーを介して20 年前の写真を自分のデジタル日記に加えること さえできる。

秋田市と秋田大学の協力で開発している羽州街道まちあるきアプリの例 図表 1 (旅行の思い出を知人と共有可能)







クラウドコンピューティング (※1)上では、 E-mailやスケジュールなどユーザ自身の 個人情報を使ってさまざまなパーソナライゼー ション (※2)のサービスが実現されている。 たとえば、出張予定の仙台市あたりの地図を見 ると、自分が予約したホテルが強調表示され、 チェックインの日までラベルとして示されてい る。訪問する予定の組織や場所も地図上で強調 表示され、個人の情報が小気味よく地図の上に 統合化され、便利であるが、少し気味が悪い。 デジタル秘書がクラウドから支援してくれてい るとポジティブにも解釈できる。

- (※1) インターネット上のサーバーにあるコンピュー タが提供している機能を、インターネット経由 で利用する仕組み
- (※2) データを活用して利用者一人ひとりに適した情報を提供すること

これらの個人情報は、E-mailなどのクラウド上の個人データから抽出されたものを利用しており、クラウドコンピューティングやスマートフォンを利用開始する際に、このような個人データの利用に同意していたわけだ。将来的には、地図を開いたときに、自分の情報や好みの情報、今後の移動計画などがきちんと載っていないと、腹を立てるユーザが出てくるようになるのだろう。

2 ビッグデータと個人情報

ここで気になるのは個人情報の漏洩である。 もちろん、漏洩すると困るのではあるが、漏洩 しないのなら、自分のサービス向上のため、み んなのサービス向上のため、公共のため、危機 管理のため、医療進歩のため、そして学術研究 の進歩のためなら、個人データを利用しても構 わないという考え方がある。これが今注目のビッ グデータの公の説明である。 ビッグデータの定義は、普通のコンピュータでは処理できないような膨大な量のデータを意味するらしいが、それならば図書館データ、地球観測データ、地図データ、銀行の口座データなどの大規模データはビッグデータではないのだろうか。これらはビッグデータというよりは、ラージデータという範疇に入るらしいが、ビッグデータと解釈しても問題はない。ビッグデータの解釈は、人さまざまで年々拡大解釈が進んでいる。

今日の日常生活のあらゆる場面でIT環境、特にクラウドコンピューティングをわれわれは利用しており、その副産物として、コンピュータの操作命令・記録データがリアルタイムにあらゆるところでいつでもどこでも大量に生じており、これらは捨てずに蓄積され続けている。これらの副産物データは人間活動の情報そのものでもあり、これらをビッグデータと呼んでいる場合が多い。

ビッグデータのひとつの例として、オンラインのカーナビやスマートフォンの地図サービスを考える。オンラインの地図サービスで、現在地の地図を表示するという操作は、ユーザ自身の位置、進行方向、速度、注目範囲などのデータを、地図の問い合わせとしてサービスプロバイダーのセンターへリアルタイムで常に送り続けていることを意味する。つまり、われわれは自ら個人情報をネットワークの向こうにあるコンピュータへ送り続けることにより、快適な日常環境を維持しているわけである。

多くのユーザの行動記録データが膨大な規模 でリアルタイムでクラウド上に集積されビッグ データとなる。個人情報で支払って、サービス を購入しているという解釈もできる。

3 位置情報サービスとプライバシー

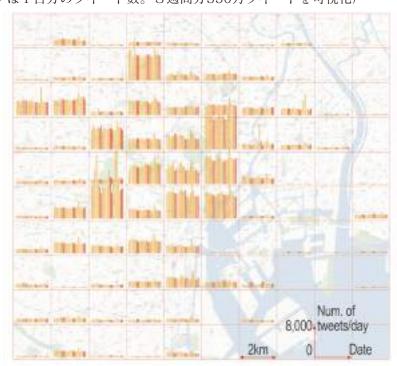
ビッグデータといっても、ツイッターのような公開データ(図表2)から、E-mailのように公開してはいけないものまで個人情報のレベルはさまざまであり、その種類も多様である。GPSによる個人の移動情報は、基本的には個人情報であり、ユーザが同意していない目的での利用は禁止されている。ただし、スマートフォンユーザの多くは、GPSを使った位置情報サービスの利用を開始するときに、サービス向上のために、プライバシーを侵害しない範囲で自分の位置データを利用してよいという契約条件に同意している。

この枠組みを使って、たとえばスマートフォンユーザがGPSを利用していると、同時に、その場所で認識できたWi-Fi発信機のID (MACアドレス)も取得し、それらの両方のデータをサービスプロバイダーのセンターへアップロードし、Wi-Fi発信機の位置データベースの作成にわれわれは無意識のうちに貢献させられているという実態がある。

このWi-Fi発信機の位置データベースの成果は、ユーザに還元されており、地下や建物内のGPS電波が届かないところ、GPS電波の状況が悪い場所では、このWi-Fi位置データベースを使って、Wi-Fi電波を使った位置測位補償が行われており、以前に比べたら格段に、位置測位可能範囲が大きく広がっており、また精度も上がっている。

別の還元の例としては、ある場所で今何が注目されているかを、場所ごとの検索キーワードのランキングを提供するサービスがある。また、位置の精度を下げることにより、個人を特定できなくし、移動データを利用するという例もある。たとえば5キロメートルメッシュで区切り、メッシュ間の移動と携帯電話の匿名IDの組み合わせとして、低い位置精度の移動データに加工して広く利用できるようにし、市場予測などの応用向けに携帯電話キャリアから販売されている。これが最近注目されている「人流」の解析の元データである。

図表2 ツイートのマッピング (各棒グラフは1日分のツイート数。3週間分350万ツイートを可視化)



出典:Fujita 2013

4 ビッグデータと学術研究の将来

人間の多くの行動と活動がデジタル化されるようになると、人間や社会の研究に、このビッグデータは不可欠な存在となっている。情報学の切り口から行う社会科学などの研究は、従来、アナログの情報をデジタル化するところから始まっていた。しかし、これからは、すでに存在し日々成長し続けている、詳細で膨大過ぎるビッグデータを使って、人権を侵さない範囲で如何に意味のある研究結果を導き出すかが研究の主流となる可能性もある。

ビッグデータ自体は、システムの副産物的データ、つまり主目的として設計されていないデータである場合がほとんどである。実際に、何か

の研究に使う場合には、必要な情報が含まれていない、あるいは必要な精度と品質を持っていない場合も多い。プライバシー保護のために除かれたり、精度を落とされたりしている場合も多い。ビッグデータは、データの量は多いが、質は必ずしも高いとは限らない。それゆえビッグデータがあれば、研究ができるというわけではない。データ量が膨大な割には、個々のデータの価値が低いという意味から、ゴミの山から価値のあるものを見つけ出す作業(データマイニングとも呼ばれている)となる場合がほとんどであり、コンピュータはその探索作業を自動化・効率化する(AI)ツールとして重要である(図表3)。

図表3 東京大学・柏キャンパスのオープンキャンパス(2015)のガイドアプリのユーザ移動ログデータ の可視化の例





ニーズ研究かシーズ研究かにもよるが、ニーズ研究の場合、まず研究の目的があり、その目的を達成するためにデータを収集するのであるが、それがたまたまビッグデータの仕様と合う場合もあるだろうが、まったく合わない場合も多い。むしろ、研究に必要な情報の設計をきちんと行い、これらの情報を、ユーザの同意のもとデータを寄付していただく枠組みをクラウドコンピューティングやユビキタスコンピューティング(※3)の中にきちんと組み込むための体系化が今後は重要になる。つまり、副産物としてのビッグデータから、意図したビッグデータへの転換が必要である。

(※3) 社会や生活の至る所にコンピュータが存在して おり、ユーザがコンピュータの使用を意識する ことなく、いつでもどこでも情報にアクセスで きる環境

プライバシー保護を優先すると、集めたい情報を集められない可能性が大きい。危機管理や学術研究では、むしろ積極的に個人データを、精度の高いデータを効率よく集めることが重要である。これは危険もともなうものであり、コンピュータ上で事故が起きるとその被害の規模も膨大なものとなり、安全対策はきわめて重要である。

今後、人類の研究のために、個人の一生分の E-mailやパソコンのデータを寄付する動きも出てくるかもしれない。そして、生きている間は解析してはいけないが、死後あるいは死後10年経てば研究に利用してよい、という時効つき個人データやライフログデータの枠組みが一般化する可能性はある。

5 人にやさしい地図インタフェースデザイン

ビッグデータを作成するためには、使い勝手 の良いサービスやユーザインタフェースが重要 であるが、一方無料であるので多く使われてる 例も多い。ウェブマッピング(※4)も使いや すいというよりは、無料であるために使われて いるというのが現実である。

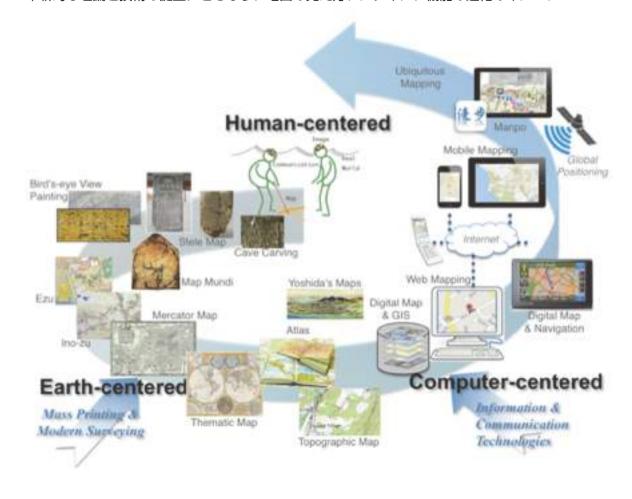
(※4) インターネットを通じて地図データや航空写真 等を取り込み、パソコン上でその表示などを行 うこと

一方、紙の観光ガイドブックや「まちあるき」のパンフレットの中で使われている地図では、ウェブマッピングの表現はまだ主流ではない。ウェブマッピングは、そもそも多目的で何にでも利用できるようにデザインされている。それゆえ、ある特定の目的で利用する場合は不要な情報が入っていたり、ページの全体的デザインに合わない場合もある。

ウェブマッピングは、地図の表現として、真位置(真に正しい位置)を基本としている。このために、ユーザにとっては必要以上に細かいところまで正確に表現してある。日常生活では、絶対位置はほとんど意味がなく、相対的位置が重要である。相対的位置の関係だけを保持すればよい場合、図は大きく単純化できる。むしろ、重要な情報と重要でない情報の「めりはり」をはっきりとさせ、ユーザにやさしい図形表現を実現できる。

このような理由から、観光ガイドブックや「まちあるき」のパンフレットでは、コンピュータで自動可視化したものではなく、グラフィックデザイナーが作成したイラスト地図が使われることが多い。このような人が作った地図は、アート作品にも近いものであり、いろいろな表現を入れることで遊びや味わいが出てくる。ウェブマッピングでは多様性が後退している一方、旅行ブームの現在、紙の地図では、むしろ多様化が進んでいる(図表4)。

図4 革新的な理論と技術の誕生にともなう、地図の見え方、メディア、機能の進化のイメージ



6 アナログ地図画像への経緯度マッピング

この問題を解決するために、私の研究室では、 人間が作成した歪んだアナログ地図を、GPS の位置データなどと連携させる枠組みを研究し ている。アナログ手書き地図画像の全ピクセル (画素) への経緯度の付与を、画像データの経 緯度マッピング(あるいは、地理参照)と呼ぶ。 この手法を使って、だれでもが簡単にこの画 像への経緯度マッピングが可能なプラットフォ ーム「漫歩」を開発した(図表5)。

これは、既存のウェブマッピングをベースに、 新たに位置マッピングを行う層を追加すること により実現した(図表6)。これに類似するプロダクトやシステムはすでに存在するが、本研 究の特徴は、これらのアナログ地図の経緯度マッピングを簡単に行える利用環境の実現と体系化にある。

具体的には、アナログ地図上のPOI(Po int of Interest、興味の地点) に経緯度を与えることだけにより、経緯度マッ ピングを実現している。POI以外の各ピクセ ルの経緯度は、一番近い2点のPOIから算出 する手法をとっている。最低、2点だけでも経 緯度マッピングは可能である。POI群がもし 「お勧めルート」上にあり、屋外で、ユーザが 「お勧めルート」だけを歩いている場合は、G PS誤差が大きくない条件下では問題なく利用 できることを確認している。より精度を上げる ためには、追加的地理参照点という新しい基本 要素を用意しており、これを交差点などに配置 すると、よりよい経緯度マッピングが可能とな る。つまり、アナログ地図上の経緯度マッピン グは、より多くの地理参照点 (POIも含む) があったほうがより位置精度は上がる。

寄稿

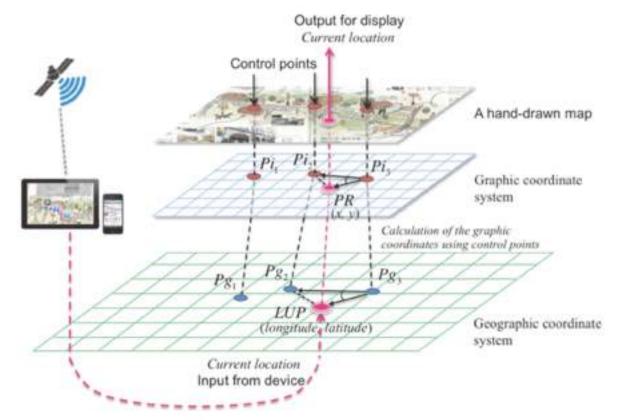


図表5 イラスト地図に経緯度マッピングを付与するツール「漫歩」の i P a d 版ビューアの画面 (イラスト地図:昭文社 ことりっぷ 海外版 ミュンヘン・ロマンチック街道)



出典:Lu 2014

図表6 ウェブマッピングを基本にしたアナログ地図画像への経緯度マッピングのプロセスの概念図 (イラスト地図:昭文社)



7 地域ビッグデータへの展開

寄稿

近年、GAFAに代表される巨大IT企業がビッグデータを占有することが世界的な問題となっている。個人情報の観点からは、EUのGDPR(一般データ保護規則)のように、消費者の権利が保証され、巨大IT企業が保有するユーザ自身のデータの見える化や削除権利が企業の義務となる動きがある。また、独占禁止法の観点からは、データ寡占に対する規制やデジタル課税の強化が重視されている。消費者が銀行を選ぶように、消費者自身が自分のデータをどこで管理してもらい、さらに運用してもらうかを選択できる時代が来る未来も遠くはないかもしれない。

ビッグデータは、「未来の石油」と言われるように、すでにデジタル経済の中心的役割を果たしており、地域においてもデジタル経済を回すことができる、IT基盤の強い足腰を持つこ

とが重要である。つまり、ビッグデータは巨大 I T企業だけのものとは考えずに、地域において ビッグデータを運用できる組織、人材、社会文 化の育成は喫緊の課題と言える。単に I T基盤 だけがあれば良いというのではなく、ビッグデータを地域で独立して回し、それを活用する多様なサービスを産み出すといった、地域ビッグ データの健全な循環システム(図表7)を実現できる能力があるかどうかがこれからの地方創生の成功の鍵になってくるだろう。

(参考文献)

Hideyuki Fujita 2013 Geo-tagged Twitter collection and visualization system,

<u>Cartography and Geographic Information</u>

Science 40 (3): 183-191.

Min Lu 2014 Human-Centered Mapping in Mobile Environments, Doctoral Thesis, the University of Tokyo.

図表7 地域ビッグデータの循環システム構成の観光応用の例

