

特集

リンクするデータ

(Linked Data)

～広がり始めたデータのクラウド～

Linked Data

0. 編集にあたって

1. Linked Data の仕組み *Linked Data - The Story So Far*
2. Linked Data とメディア—メディアが Linked Data を活用する理由—
3. 次世代の医薬品開発を支える知識流通
4. オープンガバメントとオープンデータ
5. Linked Data と地理空間情報
6. 日本における Linked Data の現状と普及に向けた課題

0. 編集にあたって

長野伸一 (株) 東芝 研究開発センター
萩野達也 慶應義塾大学 環境情報学部

本特集は、Web 世界の新しい情報インフラとなりつつある Linked Data (リンクするデータ) について、最新の技術動向、産業界での取り組み、および国内での普及に向けた課題について広く紹介するものである。Linked Data とは、個々の情報や事象に対してグローバルな ID として URI を与え、領域を横断して任意の種類のリックを付与した、Web 上のデータの集合を言う。至ってシンプルな仕組みであるが、情報を社会に公開し共有するボトムアップな取り組みが進むにつれて、世界規模のデータの Web を形成するとともに、社会における情報の価値を高めつつある。従来の Web が、人間が読む文書のネットワークであるのに対して、Linked Data は機械が読むデータのネットワークであることから、データの Web やデータクラウドとも呼ばれる。Web 世界で意味情報を共有する新たな取り組みとして、欧米を中心に注目されている。

1999 年に Tim Berners-Lee により提唱されたセ

マンティック Web は、Web 文書にメタデータを付与することにより意味情報(セマンティクス)を導入するという、当時としては画期的なものであった。オントロジーという知識体系を与えることによって、意味情報を理解するエージェントと呼ばれるソフトウェアがユーザの代わりに、チケットの予約を手配したり、最適なドライブルートを発見したりするという壮大なものであった。セマンティック Web は、次世代 Web として現在の Web 世界を一変させる可能性を有していると大きな期待が寄せられた。

しかし、Tim Berners-Lee の提唱から 10 年が経過した昨今、セマンティック Web が広く浸透し普遍的なものになっているとはいいがたい。その理由の 1 つには、オントロジーが持つ形式的、統制的な側面と、Web のオープンでボトムアップな特性との間に大きな隔たりがあり、トップダウンにより Web 上で意味情報を共有することが思うように浸透しなかった点にあらう。実際、オープンな世界を想定した



Web は、不完全な形式や情報の欠如が見られたとしてもある程度は寛容しつつ、大きな発展を遂げてきた。文書の Web と同様に、Linked Data もオープンでボトムアップに生成されたデータの Web である。個別情報（インスタンス）を中心とした意味情報のデータセットを Web 上で共有する仕組みとして広まりつつある。特に欧米では、データのオープン化の流れとも相まって、データセットを Linked Data として公開する取り組みが進められている。学術領域に限らず、産業界や社会での活用も始まっており、ビジネスや社会に大きな変革をもたらされようとしている。

本特集は、大きく分けて3部から構成されている。最初に、「1. Linked Data の仕組み (Bizer 他, 萩野訳)」で、Linked Data の概念とその技術的な原理について説明するとともに、Linked Data が誕生してから現在に至るまでの発展について紹介する。

次に、Linked Data の公開が進んでいる領域（産業界、政府）を取り上げ、各領域が抱えるビジネスや社会の上での課題と、Linked Data による解決に向けた試みの紹介を通して Linked Data の特徴を明らかにし、Linked Data によりもたらされる将来の社会の姿について述べる。報道メディアは、早い段階で Linked Data をビジネスで活用し始めた領域の1つである。昨今、報道メディア各社はオンラインサイトを立ち上げ、保有している膨大なコンテンツのオンライン化を進めつつあり、サービスとしての魅力の向上、および収益の向上へといかに結びつけるかが大きな課題となっている。「2. Linked Data とメディア (乙守他)」では、メディア業界における Linked Data の先進的な活用事例を紹介するとともに、データのリンクからサービスのリンクというメディアの進化の可能性について紹介する。

生命科学分野は、早くから学術情報を中心にデータのオープン化に取り組んできたが、昨今の2010年問題などの医薬品業界を取り巻く状況は、特に創薬における情報共有の重要性を示唆している。「3. 次世代の医薬品開発を支える知識流通 (細見他)」では、W3C の Linking Open Drug Data タスクフォースを中心に、世界中に散在している創薬に関するさまざま

なデータを Linked Data として共有することで統合的な利用の実現を目指す取り組みを紹介するとともに、創薬分野における Linked Data の活用可能性について紹介する。

データのオープン化という点では、米国や英国を始めとする世界各国で、Web の特性を活用した積極的な政府情報の公開や、行政への市民参加を促進するオープンガバメントが急速に進展している。「4. オープンガバメントとオープンデータ (佐藤他)」では、Linked Data として公開される政府データに焦点をあてつつ、オープンデータがもたらす社会と今後の展望について紹介する。また、「5. Linked Data と地理空間情報 (清水他)」では、従来は行政が主体となり整備が進められてきた地理空間情報の現状を紹介するとともに、地理空間情報の Linked Data 化の現状と課題について述べる。

最後に、「6. 日本における Linked Data の現状と普及に向けた課題 (武田)」では、社会においては情報共有を活発化することが情報の価値向上につながるという視点から、その手段として期待される Linked Data の国内での普及に向けた課題について述べるとともに、日本における Linked Data 公開の取り組みを紹介する。

Web は誕生して20年以上が経過し、人々の暮らしに広く浸透して、社会に不可欠な情報インフラとなっている。そして Linked Data は、始まったばかりの新しい Web の取り組みであるが、今後はデータが中心となって、ヒトとモノと情報とを紡いだ社会基盤を形成するとともに、新しい価値を持ったサービスや社会システムが登場するであろう。本特集により、データの提供者と利用者を含めた議論が活発化し、国内のさまざまな分野の技術者・研究者が集まったコミュニティの立ち上げと、Linked Data の多様な分野への展開とが促進されることを期待する。

参考文献

- 1) Berners-Lee, T., Hendler, J. and Lassila, O. : The Semantic Web, Scientific American (2001).
- 2) Shadbolt, N., Hall, W. and Berners-Lee, T. : The Semantic Web Revisited, IEEE Intelligent Systems, Vol.21, No.3, pp.96-101 (2006).
- 3) Berners-Lee, T. : Long Live the Web, A Call for Continued Open Standards and Neutrality, Scientific American (2010).

(平成 23 年 1 月 18 日)