

日常生活におけるユーザ支援に向けた 複数情報源の統合手法の基礎的検討

沼 晃 介・栗 芝 正 臣 (専修大学ネットワーク情報学部)

Multiple Resource Integration for User Support in Everyday Life

Kosuke NUMA, Masaomi KURISHIBA (School of Network and Information, Senshu University)

In this research, we describe our fundamental studies on multiple resource integration for user support in everyday life. For this purpose, in the vision of the Semantic Web, resources are linked by RDF. There, however, are still problems on handling contradictions and on protecting privacies. We propose a description method in which private information are personally “tagged” instead of describing physically accurate data and in which context information on data generation are attached.

キーワード：セマンティックウェブ，ソーシャルウェブ，情報統合，RDF/RDFS，ユーザ支援

Key words : Semantic Web, Social Web, Information Integration, RDF/RDFS, User Support

1. はじめに

近年，ウェブやモバイル端末，さらにはIoTという流行り言葉に現れるようにさまざまな情報機器がネットワーク化され，我々が日常的に生み出すデータはいよいよ莫大なものとなっている。これらのデータを適切に活用することができれば，これまで以上に高度に個人化したきめ細やかな情報サービスの提供が期待できる。本稿ではこのような目的のための情報の取り扱い手法について検討する。

今日のデータの増加は，ユーザ個人が日常的に触れるウェブのサービスや情報機器など，実に多様な情報システムに支えられている。ユーザを日常生活におけるさまざまな状況下で支援するためには，これら多数の情報システムから得られるデータを適切に結びつけていくことが求められる。本稿で議論するのは，このような複数情報源の統合手法である。

2. 日常生活の支援と文脈情報

2.1 日常生活におけるユーザ支援

本稿がスコープとする情報システムによる日常生活の支援とは，例えば以下のようなものを想定している（図1）。

現在徐々にIoT環境が整備され，ユーザが日常的に触れる電子機器は今後ますます増加が予想される。これらの機器はネットワークで相互に接続され，ある機器を用いたユーザのアクションが他の

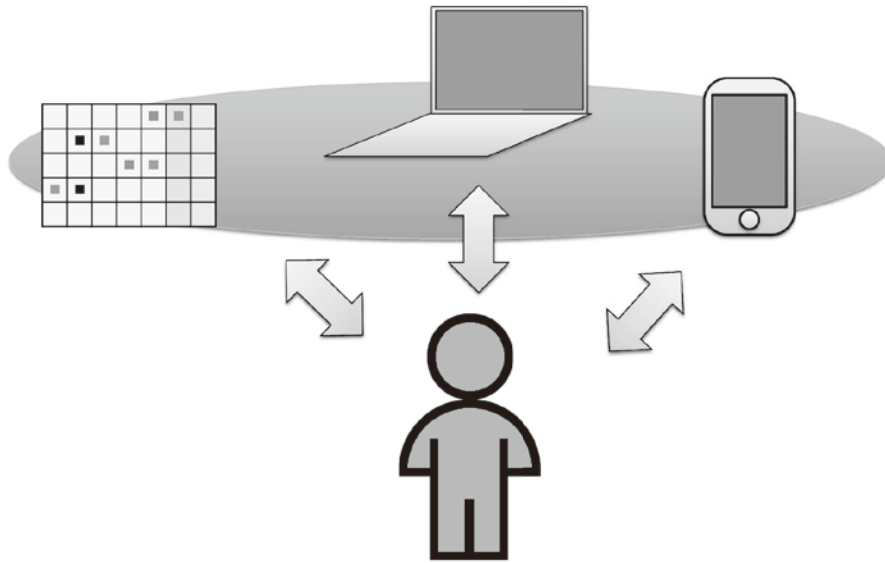


図1 日常生活において利用されるさまざまな情報システム

機器にも伝播し、こうした機器のネットワークが総体としてユーザの活動を捕捉することが可能になるだろう。またパソコンや携帯電話などのモバイル端末を介してユーザが利用するウェブ上のサービスにも多様なユーザの情報が蓄積される。ソーシャルネットワーキングサービス（SNS）やスケジュール、電子メールのほか、店舗検索や乗換案内など、ユーザの興味や活動を示す情報が各種のサービスにまたがってさまざまに送信されている。

これら多岐に渡る機器やサービスは現在多くの場合、それぞれの目的に沿って必要な情報入力をユーザに求めている。システムが相互に独立していることで仕組みが単純になる側面もあるが、一方で状況によっては他のシステムに入力したのと似たような入力を何度も求められるようなこともある。それぞれのシステムが連携し情報を共有することで、ユーザに対しても適切な情報提供を行うことができるようになるだろう。

2.2 ユーザの文脈情報と情報統合

ここで期待されるユーザ支援システムの役割は、ユーザの入力に対して適切な返答を出力する従来の一対一の情報提供ではない。多数の情報システムがネットワーク化することでユーザの置かれた状況を極力正しく察知し、そうしたユーザの文脈に応じて個々のシステムがその時その場で必要な機能を提供する、いわば一対多の関係を結んでいくことである。あるいは中には複数のユーザが同じシステムを共有することもあるから多対多の関係といてもいいかもしれない。

この時システムは、利用するユーザについて、多くのことを知っている。少なくとも、多くのことを知っているかのように振る舞う。この時の両者の関係は、ユーザがシステムを使いこなすような主と従の関係というより、ユーザとシステムが対等に対話をするようなパートナーのような関係に近い。

このような仕組みを実現するためには、種々の情報システムが持つユーザに関する情報を適切に統合し、複数の情報源からもたらされるデータに基づいてユーザの置かれた文脈を推論する枠組みが必要である。

複数の情報システムにまたがってユーザの活動にまつわる状況を統合する試みは、例えば Google

や Apple などの複数の製品やサービスにおいて一部で実現されている。例えばメール内に記載されたスケジュールを抽出してスケジューラに追加したり、通知された住所を別の地図アプリで開いたり、ひとつのサービスにおいて取得されたユーザの情報を他のサービスで活用してユーザの利便性を高めている。

3. 複数の情報システムの間での情報統合手法

3.1 API 通信

複数の情報システム間における情報統合の素朴な手法は、当該システム間でやり取りされる情報の呼び出しと提供方法を定めておく方式である。アプリケーションプログラミングインタフェース (API) と呼ばれるシステム間の通信の規約に沿って情報の授受が行われる。

ウェブサービスとして情報のリクエスト手法とデータの返却手法を定義することで、他のシステムからの情報利用が可能となる。API の整備されたウェブサービスは、IFTTT に代表されるようなウェブサービスの連携を目的としたシステムを介し容易に接続される。

この他 CSV や iCal などの一般的なフォーマットでデータを出力することで外部でのデータの再利用性を高めているサービスも見られ、これらも広義には API に基づく情報連携と呼ぶことができよう。

3.2 セマンティックウェブと RDF/RDFS

一方で情報を機械可読な形式で記述し、データの汎用性を高めながら自由な推論を実現しようとする取り組みがある。セマンティックウェブとは、WWW における情報の記述と共有の理念を深化させ、より厳密な論理の記述を目指すものである。

セマンティックウェブにおいては、リソースは XML と RDF を用いて記述される。RDF とは Resource Description Framework を表し、リソースを記述する枠組みを定める。RDF ではリソースは Subject (主語)、Predicate (述語)、Object (目的語) という 3 つ組 (RDF トリプル) を用いて記述する。この関係は図 2 に表されるようなグラフ構造を用いて表現することができ、こうした記述の積み重ねにより知識を大きなグラフとして表す。また RDF の記述に用いられる語彙の定義方法は RDFS (RDF Schema) によって定められる (W3C [1999, 2014])。

RDF/RDFS を用いて記述されたリソースは一意の ID (URI) を用いて指し示され、こうしてあらゆる物事に関わる情報が一意に記述される。分散して記述された RDF を統合することで、より大きな知識を描き出すことができる。セマンティックウェブの目標は、こうして機会可読な形式であらゆる知識を表現することである。

RDF の実践的な応用として、ウェブコンテンツに関するメタ情報の記述が行われたが、まさにこうした記述こそがユーザに関する文脈情報の記述そのものといえるだろう。そのコンテンツを誰が、いつ、どこで書いたのかといったメタ情報に関する語彙セットが、例えば Dublin Core などで整備された。



図 2 RDF におけるトリプルの構造

4. より精緻なユーザに関する情報の記述と共有

セマンティックウェブの思想に則れば、適切に記述、共有されれば、ユーザに関する情報も統合的に扱い、その上での推論も可能となる。確かにさまざまなシステムを用いた際のユーザの状況について、システムが検知したあらゆる情報を記述し、共有すれば、システム間での情報統合は可能となろう。

しかしユーザ個人の日常に深く関わる情報をこのように記述する場合、大きく2点の課題があると考えられる。すなわち (1) 情報の矛盾への対応, (2) ユーザの個人情報やプライバシーの保護、である。本章では以下それぞれについて順に説明する。

4.1 情報の矛盾への対応

複数の情報源を統合する際には常に付きまとう問題であるが、特にユーザの活動の文脈を記述する場合、情報の正確さや厳密さに幅があることが想定される。ユーザが利用する情報システムにおけるアクションを元に取得できる情報からユーザの文脈を推定することになるが、これらの判断は各システムに依存する。さまざまな情報システムを連携させるにあたっては、これらの間での記述に矛盾が含まれる可能性がある。

4.2 ユーザの個人情報やプライバシーの保護

ユーザの精緻な情報、例えば厳密な時刻や位置情報をすべて記述しシステム間で共有することには問題があるだろう。ウェブ上に全公開することにはもちろん危険が伴うが、仮に何らかの制限を加えたとしても、情報はあらゆるシステムの間で利用できてこそきめの細かいサービスの実現につながると考えられる中、接続されるすべてのシステムが果たして信頼に足るかという問題は解決できない。悪用されうる不必要なユーザの個人情報やプライバシーに関わる情報を扱う際の適切な仕組みが求められる。

5. 解決手法と議論

前章の課題に対して、以下のような情報共有手法が一定の解決法となるだろう。

5.1 個人情報のタグ化

例えばユーザの位置情報を扱う場合、大きく2つの意味合いがあると考えられる。第一は絶対的な場所であり、コンテンツが生み出された場所の物理的な近さや位置関係が意味を持つ場合である。これについても特に日常の中でリアルタイムに共有されることには問題がある場合があると考えられ、共有することのメリットとデメリットをユーザ自身が把握し、公開範囲を選択していく必要があるだろう。

しかし第二には、場所を相対的に捉える場合があり、例えば「自宅」などというようにユーザとの相対的な関係に意味があるケースも少なくない。このような場合、実際の空間的な位置を共有する必要はなく、ユーザ自身が位置に付けた「タグ」にこそ意味がある。

タグはユーザごとの世界の捉え方を反映するため、例えばあるユーザの「自宅」は別のユーザの「自宅」とは物理的な位置は大きく異なることもある。このことをあえて利用し、実際の位置情報はシステムの内部処理に用い、他のユーザにも共有される場面においてはタグを用いるという手段が考えら

れる。

5.2 文脈の明記

システムで生み出される情報にはそれぞれのシステム上でのユーザの文脈がある。それを理解した上であえてつなげることに意味があるというのが情報統合のモチベーションだが、あえて無視することと本当に無視することとは異なる。統合や推論の過程で矛盾を検知した場合、その文脈を辿ることができれば、よりユーザの感覚に沿った情報提供が可能となると期待される。

例えば SNS のようなシステムが複数利用される場合、ユーザは意識的にそれらを使い分けしているはずである。一方はある友人層との交流に、もう一方は他の友人層との交流にと、扱うコンテンツもつながっている友人も異なるかもしれない。このような場合に、ある情報を「ユーザの友人」に共有してよいのかどうか。このようにそれぞれのシステムの使い分けがユーザの文脈を表しており、システム間での人間関係にはデータ上は一見矛盾があるかもしれない。

しかしユーザが厳密にどのような目的でそれぞれのシステムを利用しているのかというと、やはりそれも確定的に明示しにくい。なぜならば多くの場合ユーザは、直感的にシステムの使い分けを行っており、明確な基準が存在しないことが多い。また同一のユーザの中でも、時間経過とともに使い方が変化することがある。

ユーザの文脈とはこのように刻一刻と変化する状況のことであり、かつその状況をユーザがどう捉えているかという解釈の問題でもある。従って正確にこれを記述することは、実質的に不可能といえる。

ここでは哲学的に状況を理解することは避け、工学的に文脈が異なることが充分区別できるだけの情報を付加することを提案する。具体的には、いつ（時刻）、誰が（ユーザ）、どのシステムを用いたかという記述を加える。

6. おわりに

近年人工知能分野が改めて注目を集めている。こうした流行の背景には深層学習と呼ばれる格段に進化した機械学習の技術の登場がある。機械学習は、データを元に訓練し、新規なデータに対しても適切な分類を行うような技術である。今日のビッグデータ時代に、こうした多量の情報をさまざまに活用する試みといえる。しかし一方で本稿で扱ったセマンティックウェブ技術は、古典的な記号論理の記述と推論をウェブに適用したものである。ウェブを取り巻く人工知能のアプローチとしては、両極に位置する。

重要なのはこれらの二極の技術は対立するものではなく、相補的なものであることである。深層学習によるトップダウン的な分類アプローチと、推論に基づくボトムアップ的な論理展開を組み合わせることで、さらなる技術発展が期待される。

本稿では、こうした視野のもとに、セマンティックウェブ技術をより広いユーザの日常的な情報支援に適用するための基礎的な議論を行った。今後の課題は、機械学習によるアプローチとの接続、ならびに実利用可能なアプリケーションの開発と実験である。

謝辞

本研究は専修大学情報科学研究所の共同研究助成を受けて行った。

参考文献

W3C Recommendation, “Resource Description Framework (RDF) Model and Syntax Specification”, <https://www.w3.org/TR/1999/REC-rdf-syntax-19990222/>, 1999.

W3C Recommendation, “RDF Schema 1.1”, <https://www.w3.org/TR/rdf-schema/>, 2014.