セマンティックWebと現象学的視点
The Semantic Web and Phenomenological Viewpoints

小坂　武　（諏訪東京理科大学）
Takeshi Kosaka　(Tokyo University of Science, Suwa)

要　旨：セマンティック Web は社会科学の転回と整合的である。そのため、当事者が意味を問いながら徐々に記述していけるが、その方法論は確立されていない。本稿は、リレーショナル DB との対比のもと、RDF の記述とは何であるかを探究する。RDB の開発は合理的視点で、第三者がトップダウンでスキーマを開発し、ユーザは集合操作する。対して、セマンティック Web では各主体がインスタンスを RDF に徐々に記述し、ボトムアップで結果的に大きな知識ベースとなる。この徐々に世界が開示されていく方式は現象学の考えに通じる。このことから、当事者向けの一人称的アプローチの開発が期待される。
Abstract: The Semantic Web is consistent to the turns of social science. Therefore, with the gradual understanding of the world, users themselves can write down RDFs. However, the methodology for it is not yet established. We investigate what RDF descriptions are in contrast to the relational DB. RDB development is practiced in the rationalistic way by the third party in a top-down approach and users practice set operation. In contrast, in the Semantic Web, users gradually describe RDF for each instance, and happen to have a large knowledge base as a result. This gradual description of the world is parallel to the idea of phenomenology. It is expected therefore that a first-person perspective approach be developed.

１．はじめに

セマンティック Web (SWeb)の発展が 2010 年代後半から著しい。個別に各種機関が保持してきたデータが Linked Open Data(LOD)として次々と公開されてきている。例えば、美術館等の文化遺産にかかわるデータを公開する Europeana、米政府関連のデータを公開する DATA.GOV。日本版 DATA.GOV も 2013 年から登場した。

Berners-Lee は 1998 年に、「SWeb はリレーショナル DB (RDB)の大量表現を機械処理可能にすることが、主要な動機だ。」と指摘している。このことから、今日の LOD の発達は彼の想定内のものと言えよう。

言語やサービスは想定内の利用で終始するとは限らない。それが有するアフォーダンスのもと、個人に利便性が認められれば、よりダイナミックに予想外に拡大することがある。Berners-Lee(1999)は次の様にいう。「HTML とその Web が文書群を一つの巨大な書籍のようにオンライン化したことに例えれば、SWeb 技術(RDF、スキーマと推論言語)は世界にあるデータ群を一つの巨大なデータベースにする。」HTML は当初、組織の情報公開に使われたが、今日では個人的な文書の共有利用にも途をあけた。SWeb が公共機関や会社という組織だけのものに限定され続けるのであろうか。

既存データを LOD として公開する方法はほぼ確立されており、特別な困難はない。それらデータは共同利用のものとして精緻化されてきた。それらデータの開発は主に伝統的システム開発方法論、すなわち第三者的視点の方法論を利用して行われてきた。

本稿は、SWeb が単に組織のデータの公開を超え個人のデータ公開に関わるとき、既存の方法論ではなく当事者のデータ開発方法論が必要ではないかと考える。なぜそれが必要であるのかを探求し、しかもそれが大凡どのようなものでありうるのかを探索する。

組織で使われる方法論は三人称の客観的視点にたつものである。個人がデータ公開するとき、そのデータ開発はその個人、当事者が行う。当事者が開発するための方法は、それぞれの人に有意味なもの取り上げる一人称の方法論であると予想される。

一人称の方法論の必要性は経済性の側面では状況依存であるが、システム論的側面、組織的抵抗側面、心理的側面や認知的側面等の知見から望ましいことが知られる (Kosaka,2009)。SWeb は、この一人称の方法論の利用を促す IS 基盤の登場とみなすことが出来る。

組織のデータは利用者を想定し整備されてきた。しかし、個人が関心を有するデータは当事者の利用に実質的に限定されてきた。それが公的空間に開放されることは、当事者が当然視している知識成立・データ成立の条件・信念が他者に理解できるものでなければ、その公開は意味をなさない。この知識成立・データ成立の条件・信念を当事者に意識させることが方法論の核になる可能性があると筆者は考える。

本稿はこの方法論の必要性とその中核的考えを、RDB およびその開発方法論を参考に、RDB と RDF の言語上の差異、およびその基本概念の差異等から考察する。

２．先行研究

SWeb というデータ公開は、組織の場合、公共的な性格を有するが、個人の場合、主観的な性格も有する。そのため、それにかかわる研究は自然科学や工学ではなく社会科学の範疇に入る。

Berners-Lee (1998)は、誰が何について何を言っても許されることで、Web は動くものとした。また、W3C も、「RDF の設計は誰もがどんな資源についても言明を作り出すことを許す」とする (W3C RDF, 2004)。そのため、個人が SWeb に社会的参画することに特別な制約は本来ない。しかしながら、Pinto ら(2009)が指摘するように、オントロジー開発方法論は非専門家の参画を促進すべきだが、そうなっていない。

セマンティック技術はこの 10 年間で進歩したが、合意形成や要求分析が挑戦的タスクであるため、その

応用範囲は極めて限られているとSimperlら（2010）は指摘する。Simperlらがいう要求分析は、単にメタ・データがあれば充分だとする考えに疑問を呈するもので、我々と一部問題を共有する。

オントロジー開発に必要な概念モデリングを構築主義で理解するには利害関係者の世界観の違いをまず理解することが必要となるとEsseinら(2013)はいう。世界観の違いで、データの表現や内容が異なることを指摘し、我々の問題意識を支持する。

これら先行研究は、個人が有するデータをどのように開発・定義するかという点には、方法論については言及していない。当事者が他者にも彼の信念や条件がわかるようにするには、彼自身がその信念や条件を意識する必要があるのではとする我々の考えと接点がない。そのため、本稿は改めて、ゼロの状態から方法論の必要性とその中核概念を考察する。

RDF では二つのモノの関係をそのモノにかかわる情報から離れて記述できる。当事者にとって存在を当然視している二つの存在の間に関係を記述できる。当事者はそれで充分であろうが、他者はその存在そのものが何であるかの、なぜそれが存在するようになったかを理解出来ない。これが当事者が行うデータ公開で起こる問題の一つである。存在措定が暗黙の内になされているが、その当事者を知らない他者はそれを実質的に共有利用できないという問題である。

SWebは従来の人工研究とは異なり、開世界仮説等に基づくため、社会科学と哲学のこの百年間の転回に整合的である（Kosaka,2014）。そのため、社会科学に不整合なものを社会科学領域に持ち込んで検討する危険を冒すことがない。この観点からも、本稿の問題意識を社会科学としての IS 研究で取り上げることに意義があると考える。

３．記述形式の相違と方法論

SWeb の言語群はレイヤーケーキとして定義され、その中核言語は RDF と OWL である。本研究には適切な先行研究がないことから、本稿ではRDFに限定して議論する。RDF は通称トリプルと呼ばれる形式で言明を作り出すためのフレームワークである。なお、以下の説明ではRDFとトリプルを互換的に使用する。

SWebは、AはBの一部で、YはZの一メンバであるかのように、モノの関係を記述し、モノの性質、IS分野では属性と呼ばれるものを記述する。RDF は〈主語〉〈述語〉〈目的語〉で二つのモノ・資源の間の関係を表わす。関係性は主語から述語へと方向性を持った方法で表現される。資源を表すために URI が用いられる。例えば、「ＡはＢを知る」や「Ａは人である」は、非形式的に表現すると、〈Ａ〉〈knows〉〈Ｂ〉、〈Ａ〉〈is a〉〈person〉となる。

トリプルはノードとアークで構成される連結グラフとして視覚化される。主語と目的語はノードであり、述語はアークである。このようなグラフが SWeb として得られれば、SPARQL を用いて、問合せができる。

ここでは、RDBとの比較から、RDFの特徴を抽出する。RDB では、スキーマが先に宣言され、そのものにタップルというレコードが記述・入力されていく。スキーマは組織が有するデータの種類を記述ものである。スキーマの記述には、他の組織や会社との関係に留意する必要はなく、設計者が自在に名前を付けられる。ユーザはスキーマに沿ってデータを投入し、SQL でそれらを検索する。RDB では正規化が施されるため、一般に複数テーブルが生成され、テーブル間の関係は外部キーで識別される。

ここで、留意すべきは、スキーマは誰かによって組織全体を見渡して作られるのに対し、ユーザは既に存在するスキーマに沿ってデータ処理を行う。個々のユーザ特性や意味は捨象される。スキーマは組織のトップダウンで設定される。

RDBのこの考えの範囲でRDFを述べる。RDFは、二つのモノの関係を述語で述べたものであり、それぞれは完結している。RDB であれば、一つのスキーマ、そして各テーブルに複数のタップルがある。RDFでは、RDBの一タップルそのものもが複数のトリプルになる。RDB風に表現すれば、各RDFはそれぞれが一行でしかも二つのフィールドのみがあるものである。

RDBのフィールド名は、RDFでは述語である。一つひとつのトリプルで論理的には自在な名前をつけることができるが、実際は共通的なフィールド名を付与する。RDF は既存スキーマに自分のデータを合わせるのではなく、一人ひとりが有するデータをその人の関心のもと必要な述語を使って記述される。例えば従業員について記述したRDFで、ある従業員には多様な述語を用いて詳しく記述するのに対し、一方では簡素な記述で終わることもある。RDB のように事前に指定されたフィールドを全て埋める必要はない。このことから、RDF作成者の思いを反映したトリプルを設けることができる。RDBとの対比のでは、RDFはボトムアップ方式の記述と言える。

RDF では、コンピュータ上で、あるモノの属性もオブジェクトとして記述・蓄積されるが、他のモノとのリレーションシップもオブジェクトとして記述・蓄積される。そのため、記述者によってはモノとモノ、人と人との関係だけを記述することで満足する可能性がある。属性を持たない個物も生まれる。これは文法的に許され、しかも本人にとっては問題なく有意味である場合もあるが、公的空間に開示されるトリプルとしては意味をなさない。

SWebにはRDBに見られるテーブル概念がない。それぞれのトリプルがバラバラに存在し、検索時に関係付けられるだけである。

本節では、RDB との対比のもと、その文法、記述の相違からくる知見として、トップダウン・アプローチとボトムアップ・アプローチの相違を識別した。RDBでは全体的な分析を先に行った後に、スキーマを固定しインスタンスを投入する。それに対して、RDF では各インスタンスを、全体を考慮することなく、関心のあるところから記述し始めることができる。

これは方法論について、RDB では客観的な方法論、第三者的な方法論、すなわち個人の違いを無視してロールで人間を捉える方法論が利用される。それに対し、ボトムアップの RDF 記述では関連世界にいるその当事者が自身の関心を記述する方法論であり、それは第三者的な方法論とは異なることが示唆される。

４.外部関係の相違と方法論

本節ではデータの記述者は誰であるかを探索する。SWeb と RDB の本質的な差は、しばしば前者が開世界仮説（OWA）に、後者が閉世界仮説（CWA）に依拠することとされる（Bergman,2009)。

SWeb は OWA に基づく。「…システムのすべてのデータがどこにあるかを知ることは現実的ではなく、事実上不可能になります。…いくつかのデータが欠落しているように思われる場合には、一般的に、データがシステムの他のどこかに存在するかもしれないと推測しなければなりません。」(W3C OWL-Guide, 2004)　一方、述語論理に基づく RDB などでは、ある事実が存在せず、ある命題が真であることの証明に失敗した場合、その命題は偽であるとする「失敗による否定」(Negation as failure）を採用している。

OWA と CWA との相違が直接、方法論にどのような影響を与えるかをここで吟味するのではなく、それが教える意義をここでは参考にする。OWA は世界が閉じているのではなく、内外に起こる知識の進化を取り入れることを意味する。CWA は外部で起こっている進化や変化を考慮することなく、その時点で存在するデータだけでものごとを判断する。そこで、ここでは外部に開いているか閉じているかを、データ作成者と関係付けることで方法論に関する知見を導出する。

RDB は組織の関係や挙動を観察することを通じて、スキーマを作成する。例えば従業員情報の DB を作成するとき、一人一人の違いを捉えるのではなく、大きく見渡しどのような人々がいるかに関心がある。実存的な関心は RDB 開発には見られない。そして、データのフィールド名を設定するに当たっても、その組織で使われている言葉を用いて記述する。その結果、各組織で独自のスキーマが定義され、世界に無数のスキーマが存在するようになり、互いに情報交流できない状況を生み出している。また、一度定義したスキーマの変更は、アプリケーションだけではなく、既存データの容易でない作り直しを意味する。そのため、外部との接触を断つだけでなく、外部の進化や変化を取り入れられない状況を生み出す。IS が環境変化に取り残される運命となりがちであることを、これは意味する。

一方、RDF では最初から公開することを目的としているため、外で通用している語を用いて記述しようとする。利用されるオントロジーは、世界で共通的に利用される FOAF、Dublin Core、WordNet 等で、そこにある語彙を使って自分にかかわる世界をトリプルに記述する。関心を有する世界を記述する語彙等が不足する場合、独自のオントロジーを作成し公開することもできる。

この状況は、人々が会話をするために、既存の言葉で述べることができるのであれば既存の共有語彙を使い、既存の語彙が世界を記述するに不足するのであれば独自に定義することと同等である。そのため、SWeb は最初から、外部世界との交流が意図されている。この結果、外部で定義される語彙を取り入れ、環境の変化を取り入れ、SWeb が進化していける。これが、OWA だけではなく、共通の語彙、その進化を受け入れていることによって、IS の進化を可能にすると考えられる。

本節では、誰が記述するかを考察するために、RDB と RDF における外部との関係をもとに議論してきた。RDB はデータ作成者とはしばしば異なる第三者が客観的な観察によって、個々の違いや意味を捨象して、組織に役立つと考えられるスキーマを準備する。対して、SWeb では、データを準備する人が共通的に利用される語彙を用いて、自分にかかわるデータを記述する。そのため、記述されたトリプルはその時点で他者やコンピュータにとって理解できる可能性がある。

要約すれば、RDB は組織内部語、RDF は共通語という相違がある。データ作成者がデータ利用者そのものである可能性が RDF では一般的と考え得る。対象世界を意味ある存在として捉える当事者が、その世界についてトリプルを作成する。それは、対象世界を観察に依拠する方法論ではなく、意味ある世界として当事者意識で使う方法論であろう。筆者は当事者が当事者の世界を間主観的に把握するための方法を一人称システム分析（1ppSA）として提唱してきたが、これがこのデータ開発に、トリプル開発に利用できる可能性がある。

５.方法論のスケッチ

以上の検討で、RDF の開発方法論はボトムアップで当事者が使うものであることが理解される。そして、トリプルは当事者が記述するものの、それは公開される性格を同時に有する。意味を表すメタ・データがあるだけでは個人が提供するデータは他者に理解できるとは限らない。

人は世界を通常当然視している。最も簡略化してそれを RDF で表現すれば、「あの人はあれが好きだ」といった言明になる。当事者にとってあの人も存在するし、あれも存在する、当然視された世界である。そして、URI を利用して、そのまま<A><likes><B>と記述可能である。しかし、これは他者に意味をなさない。また時間が経過したとき、本人にも分からないものとなり得る。

人間が当然視しているものを、なぜそのように当然視しているかを明らかにするのが現象学である。メルローポンティ(1967,p.23)はいう。「現象学的世界とは、先行しているはずのある存在の顕在化ではなくて存在の創設であり・・・」。また、フッサールによれば「その主観性がその隠された内的な方法において世界を所有し、世界を成立させ、形成しつづけているその仕方へと立ち帰るのである。現象学者の関心は、出来上がった世界を目指すことでもなければ、その世界において外的に意図されていた行為を目指すことでもない。」

(1995, p.323) これら知見から、我々は当然視しているものを顕在化、記述化するのではなく、それをもう一度存在措定することが必要であることを我々は学ぶ。どのような意味で、どんな条件でそれが存在化するのか、当然視しているかを明らかにせよと。

例えば、「あの人はあれが好きだ」は極端な簡略化であるが、極く普通にみられるのは「ロメオは聖母子像が好きだ」のように名前がつくことが多い。しかし現象学では名前を付けるだけでは世界を顕在化させたとはしない。「ロメオはジュリエットを知っている」、「ロメオはシェクスピアの作品だ」と記述すれば、ロメオなる存在を措定したことになり、他者にもその存在が共有されるであろう。同様なことが、聖母子像についても言える。「ラファエロの作品だ」、「その聖母はフォルトリーナがモデルだ」と記述すれば、小椅子の聖母子像が存在措定され、他者と共通理解を生む。現象学は知る、製作する、モデルだとかを意味と言うが、関連物や属性に述語で結ぶことで存在の創設となり、存在が確となり、公的空間で通用することが可能になる。

現象学は第三者が行うのではなく、当事者である人間が一人称で行うものである。我々は現象学に基づく実存的システム分析を 1ppSA と呼ぶが、RDF 作成で必要となるであろう方法論はこれに通じると考える。以下に、その方法論を現象学にそって略述する。

まず、当事者が存在しているとする確信をまず棚上げする、すなわちエポケー（括弧入れ）する。そうすると、改めて存在を創設しなければならない状況になる。なぜ自分はそれを当然視しているかを改めて考え、創設しようとする方法が現象学でいう超越論的還元である。しかも、意味や条件を意識的にいろいろ入れ替えて、確信を高めていく。それを形相的還元という。

この存在措定は止まるところを知らないが、ある段階で他者にもその存在が通用する段階に到達するであろう。時代の変化とともに、さらなる意味や条件を追加する必要が起こるが、追加・改訂を随時行うことを SWeb は許す。RDF は現象学をするための記法といっても過言ではない。しかし、具体的な現象学の方法論はまだ一般的ではない。Kosaka(2013)は、図的方法に活動理論を使い、現象学を方法論とした 1ppSA を提案している。それが一つの参考となるであろう。

６．おわりに

SWeb は組織が有する既存データを公開する方式として広く有効活用されつつあり、これは SWeb の意図された目的を達成しつつある。HTML が個人の文書の公開を促進したが、SWeb も個人が有するデータの公開を促進する可能性がある。組織が有するデータは他者の利用を想定して用意されてきた。一方、個人が有するデータは他者が使うことを想定されていなかったが、SWeb 上ではパブリックな性格を持つようになる。このギャップを埋め、健全に SWeb が発達するようにするには、データ開発方法論が必要であると本稿では想定した。そしてその方法論の性格を言語の差異と、外部との関係という構造上の問題から探求した結果、当事者が使える一人称の方法論が浮かび上がった。そしてそれは現象学がいう存在措定をするものになることが理解され、方法論のおおよそのスケッチを提示した。この分野の研究は、オントロジー開発方法論の必要性が指摘される程度で、まだ本稿が指摘する内容の段階に至っていない。そのため、この種の議論を起こすことがまず必要と筆者は考える。

最後になるが、SWeb にはスキーマがなく、データ構造とデータそのものが世界の認識や世界の変化に呼応して当事者が変更していけることが分かった。単に既存データの公開ではこのような議論にはならないものの、当事者視点を入れることで、変化への対応、進化が自然であることが理解される。これは本研究の副産物であるが、SWeb が今後の IS 基盤として有効なものであることを示唆する。すなわち、Truex ら(1999)が指摘する時代と共に進化する IS を、SWeb を基盤として構築できる可能性がみえる。

主要な参考文献

- Bergman, M. (2009). The Open World Assumption: Elephant in the Room. http://www.mkbergman.com/852/the-open-world-assumption-elephant-in-the-room/, Accessed Sept. 19, 2014.
- Berners-Lee, T. (1998). Relational Databases on the Semantic Web. http://www.w3.org/DesignIssues/RDB-RDF.html, Accessed Sept. 19, 2014.
- Berners-Lee, T. (1999). Weaving the Web. Orion Business, London, UK, 1999.
- Esswein, W. and Lehrmann, S. (2013). About the Need for Semantically Enriched Referecne Models. AMCIS 2013 Proceedings.
- フッサール, E.『ヨーロッパ諸学の危機と超越論的現象学』中央公論新社, 1995.
- Kosaka, Takeshi. (2009). Theoretical Investigation into Systems Analysis, PACIS 2009 Proceedings.
- Kosaka, Takeshi. (2013). A Foundation of a First-Person Perspective Systems Analysis. AMCIS 2013 Proceedings.
- Kosaka, Takeshi. (2014). The Semantic Web and Turns of Social Science. MCIS2014 Proceedings.
- メルローポンティ, M. (1967).『知覚の現象学 I』みすず書房.
- Pinto, H. S., Tempich, C. and Staab, S. (2009). Ontology Engineering and Evolution in a Distributed World Using DILIGENT. Handbook on Ontologies. 2nd Ed. Springer, 153-199.
- Simperl, E., Mochol, M. and Burger T. (2010). Achieving Maturity: The State of Practice in Ontology Engineering in 2009. International Journal of Computer Science and Applications, 7(1), 45-65.
- Truex, Duane P., Baskerville, R., and Klein, Heinz (1999). Growing Systems in Emergent Organizations. Communications of the ACM, 42(8), 117-123.
- W3C OWL-Guide, Recommendation (2004). http://www.w3.org/TR/2004/REC-owl-guide-20040210/, Accessed Sept. 19, 2014.
- W3C RDF: Concepts and Abstract Syntax (2004). http://www.w3.org/TR/2004/REC-rdf-concepts-20040210/, Accessed Sept. 19, 2014.