

Linked Open Data を利用するアプリケーションの作成支援 —スキーマが与えられていないメタデータの視覚化支援手法の提案—

田中 圭[†] 西出 頼継[‡] 永森 光晴^{††} 杉本 重雄^{††}

筑波大学大学院図書館情報メディア研究科[†]

筑波大学情報学群情報メディア創成学類[‡]

筑波大学図書館情報メディア系^{††}

知的コミュニティ基盤研究センター^{††}

1. はじめに

Linked Open Data (LOD) の普及に伴い、メタデータを地図、年表、グラフといった様々な表示形式に視覚化し、アプリケーションの中で利用する機会が増えている^{[1][2][3]}。LOD として提供されている Resource Description Framework (RDF) 形式のメタデータを利用するにはメタデータの構造情報や記述されているデータ型などを定義した記述規則 (スキーマ) を把握する必要がある。しかし、LOD として提供されているメタデータには記述規則が明らかではないものが多い。

本研究ではメタデータの利用手段として視覚化を取り上げ、スキーマが与えられていないメタデータの視覚化支援手法を提案する。本研究では、各表示形式に必要な属性の組み合わせを定義し、利用するメタデータの記述内容と対応させることでメタデータの視覚化を支援する。

2. LOD を利用するアプリケーション

LOD として提供されるメタデータは 2009 年以降急増している^[3]。近年ではメタデータを LOD として公開することに加えて、メタデータを加工してアプリケーションの中で利用するようになってきた。例えば研究室に所属する学生のメタデータ (図 1-A) がある場合、このメタデータを用いて学生を生年月日や名前に基づきソートし出力する機能や、LOD データセットのメタデータを統合し、地図上に出力する機能 (図 1-B)、学生ごとの論文を年表形式で出力する機能 (図 1-C) の実現などに利用できる。このような、LOD として提供されるメタデータを利用した機能が実装されているアプリケーションを、本論文では「LOD を利用するアプリケーション」と呼ぶ。

3. LOD のメタデータを利用する際の問題

LOD を利用するアプリケーションの作成では、様々な分野のメタデータを組み合わせて利用することが求められる^[5]。筆者らがこれまで行ってきた LOD を利用するアプリケーションの作成では次の問題があった^[5]。

A) メタデータの記述に利用されている語彙や構造が多様である

LOD として公開するメタデータを記述する際、メタデータ作成者は独自のメタデータ語彙を定義し、用いることがある。そのため、公開されているメタデータ内の情報資源がどのように構造化されているかは、メタデータ利用者がメタデータ作成者の公開しているメタデータ語彙の定義を理解し、メタデータの構造を利用者自身が詳細に把握した上で利用しなければならない。

B) メタデータスキーマが付与されているメタデータが少ない

メタデータスキーマは、メタデータの記述に用いるメタデータ項目と、その項目は必須か、任意か、などといった取り決め (スキーマ) を明示的に定義したものである^[4]。メタデータスキーマを参照すれば、構造情報から、利用に必要な情報を容易に取得することができる。例えばメタデータ内の利用者が必要な情報を抽出し、機械が処理しやすいような変換が可能である (図 2 参照)。しかし、LOD として提供されるメタデータはメタデータスキーマが与えられていないものが多い。

本研究ではスキーマが与えられていないメタデータを、メタデータ語彙定義を蓄積したレジストリと、各表示形式に必要な属性、メタデータ項目の組み合わせを定義して、必要な情報を取得するためことを可能にする手法を提案する。

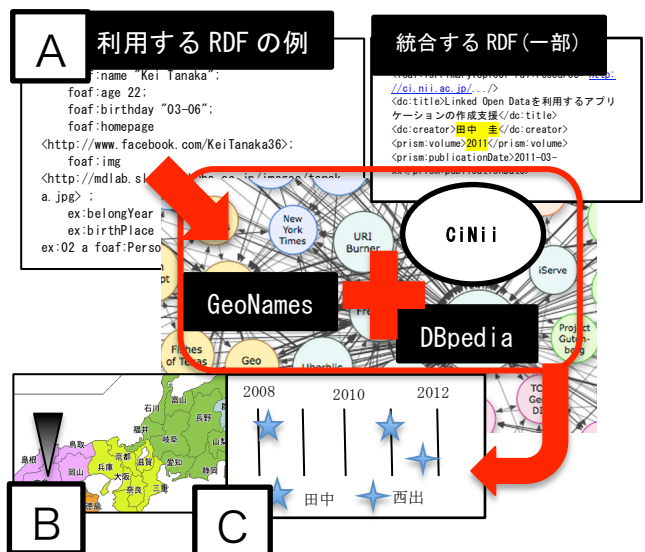


図 1 メタデータの利用例

“A Programming Model to Help Visualization of the Linked Open Data without Metadata Schema”

[†] Kei Tanaka. Graduate School of Library, Information and Media Studies. University of Tsukuba.

[‡] Yoritsugu Nishide. School of Informatics. Univ of Tsukuba.

^{††} Mitsuharu Nagamori. Shigeo Sugimoto. Faculty of Library, Information and Media Science. University of Tsukuba.

^{†††} Research Ctr for Knowledge Communities. Univ of Tsukuba.

4. スキーマが与えられていないメタデータの視覚化支援手法の提案

4.1. メタデータ語彙定義レジストリ

メタデータ語彙の定義はネットワーク上で公開されていることが多い。メタデータ利用者はこれらスキーマを参照しメタデータ項目の意味を理解するが、メタデータ語彙定義の公開方法はコミュニティによって様々である。そのため、機械に特定のメタデータ語彙定義を取得させようとするには手間がかかる。そこで本研究では公開されているメタデータ語彙定義^[6]を収集し、RDF形式で蓄積するレジストリを作成した。このレジストリを利用することで語彙定義情報を取得する手間が省くことができる。

4.2. 表示形式に必要な属性とメタデータ項目の組み合わせ

メタデータスキーマを参照することで利用するメタデータを適切な形に変換する支援ができる。これはスキーマの中にメタデータの構造情報と、メタデータ項目と代表的な語彙や関係のあるメタデータ項目が記述されているためである。このため、スキーマが与えられていないメタデータを利用するには、メタデータ内のメタデータ項目を利用目的に合わせて変換するための変換方針が必要である。

そこで本研究では、地図や年表、グラフといった各表示形式が必要とする属性(例えば地図ならば地名や緯度経度、年表ならばイベント名、開始時間)の組み合わせを定義し、各属性にその属性の記述に用いるメタデータ項目群を結びつけた。(図3参照)これにより、メタデータ利用者が指定した表示形式に必要な属性に結びつけられているメタデータ項目群がメタデータを変換するための方針となる。

4.3. スキーマが与えられていないメタデータの視覚化支援手法

4.2節で述べた各表示形式と属性の組み合わせ、そして属性と結びつけられたメタデータ項目群により、メタデータ利用者がメタデータと視覚化したい表示形式を決定すれば、メタデータ内に記述されている視覚化するために必要なメタデータ項目の有無を判定可能である。このとき必要な属性を記述するメタデータ項目群と同様に扱うことが可能なメタデータ項目があるかどうかを判断するためにメタデータ語彙定義レジストリを用いてスコア化し、メタデータ利用者に提示する。(図3参照)

5. おわりに

本稿ではLODとして提供されているメタデータの視覚化支援手法を提案した。今回提案した支援手法はメタデータ利用者が利用目的を設定し、利用するメタデータ内に記述されているメタデータ項目の中に必要なメタデータ項目があるかどうかを判断しスコア化している。これによりメタデータ利用者は

スキーマが与えられていないメタデータをアプリケーションの中で利用することが容易になった。今後の課題は視覚化だけでなく、様々な利用目的に対応することである。

参考文献

- [1]viewshare. <http://viewshare.org>
- [2]Semantic Web Challenge. <http://challenge.semanticweb.org/index.html>
- [3]Linked Data. <http://linkeddata.org>
- [4]メタデータ情報基盤構築事業. <http://meta-proj.jp/project.html>
- [5] 田中圭他. Linked Open Dataを利用する開発支援環境-メタデータスキーマレジストリを利用した記述規則の多様化への対応-. 第74回情報処理学会全国大会, 2012.
- [6]prefix.cc . <http://prefix.cc>

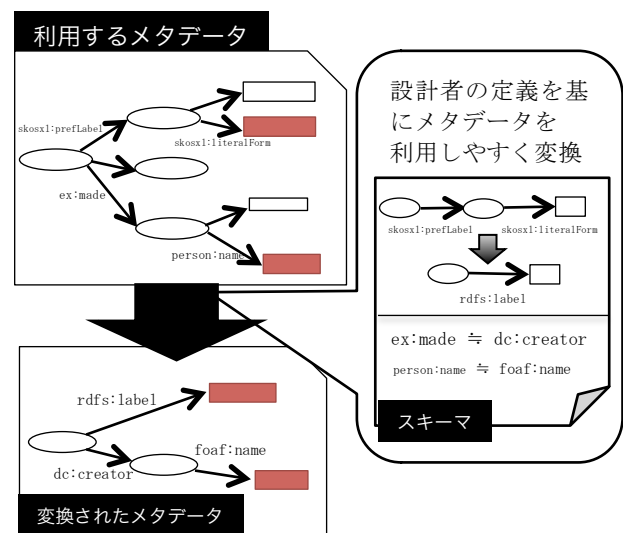


図2 メタデータスキーマを用いた情報取得

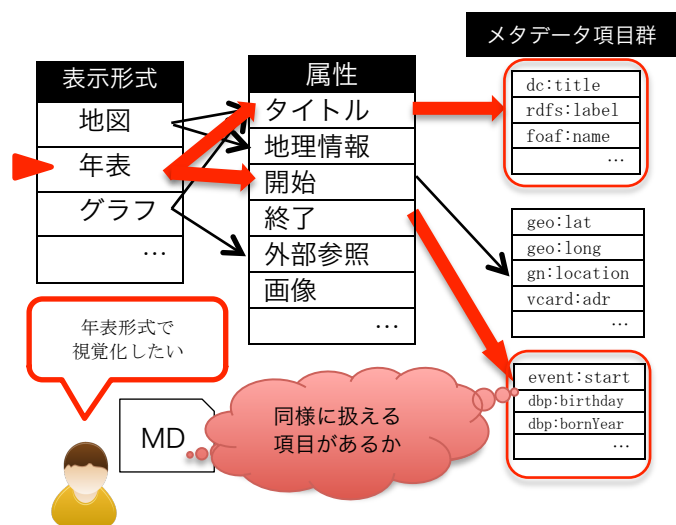


図3 表示形式、属性、メタデータ項目の組み合わせ利用例