

モバイル環境における構造化データ作成支援システムの提案

Structured Data Making Support System in Mobile Environment

後藤 孝行*¹ 濱崎 雅弘*² 武田 英明*^{1*3}
 Takayuki Goto Masahiro Hamasaki Hideaki Takeda

*¹国立情報学研究所 *²産業技術総合研究所
 National Institute of Informatics (NII) National Institute of Advanced Industrial Science and Technology (AIST)

*³総合研究大学院大学
 The Graduate University for Advanced Studies (SOKENDAI)

The role of CGM(Consumer Generated Media) is one of the greatest contributions in creating Linked Data. We propose a camera application called Kotoli, that supports a creation of structured data. Kotoli makes structured data by taking a picture with a template written about the subject. Kotoli also describes the relation between structured data by combining templates. Then, users can create a structured data without a burden almost. There is an editor which can make these templates easily on Kotoli. In this paper, we describe our proposed system and show a prototype of Kotoli.

1. はじめに

Linked Data など、機械可読な Web(Web of Data)において、複数の異なるデータの組み合わせ(マッシュアップ)による創発的なデータ活用が期待されている。例えば、オハイオ州ゼンズビルでは人種によって水道インフラに明確な差があったことを、住んでいる人種の情報と、水道のある家の情報を地図にマッピングすることで明らかにした[Suddath 08, Lee 10]。一見関連なさそうな異なるデータ集合を組み合わせることで、意外な発見があるかもしれないところにマッシュアップの可能性がある。

意外性のあるマッシュアップの実現には、様々なデータの組み合わせが可能で、そして、そのデータの多様性が重要である。多様な構造化データを作り出すには、ユーザ生成コンテンツの役割が大きく、数多くのシステムが提案されている。SemanticMediaWiki[Völkel 06] や OntoWiki[Hepp 06], Freebase[Bollacker 08], SocilalInfoBox[Hamasaki 11] は、Wiki 的アプローチにより構造化データを作成する。また、近年のスマートフォンの普及にともなってモバイル環境で利用するシステムもいくつか提案されており、OntoWiki Mobile[Ermilov 11] は、スマートフォン上で OntoWiki が利用でき、DBpedia Mobile[Becker 09] はユーザの現在の所在地や写真、レーティングを DBpedia のリソースと共に投稿し公開できる。FixMyStreet*¹ は町の落書きなど地域課題を投稿するウェブサービスで、位置情報や写真、コメントなどを地域課題というテーマでデータを収集できる。市民が行政に積極的に関わっていくオープンガバメントの実現という点でも興味深い試みである。

モバイル環境でのデータ作成は、人々が実世界において出会う様々な事柄をその場でデータ化できることで、位置情報などデバイスのセンサ情報を自動的に付与したり、撮影対象を観察しながら記録可能なことから、今後ますます重要になると考える。一方で、モバイル環境では画面が小さいなどの物理的制約が存在し、また移動中は落ち着いてデータ入力ができない、

などといった時間的制約も存在する。このことから、自由に属性を設定することが可能な反面、データセット管理に関する知識を要求されながらデータを記述していく OntoWiki Mobile のようなシステムは、誰もが使える手軽なデータ作成環境とは言えない。DBpedia Mobile は、ユーザが入力できるのはレーティングだけなので手軽に利用できる一方、DBpedia に存在するリソースに対しての写真とレーティングなので、多様なデータ作成環境とは異なる。FixMyStreet は、入力項目が限定されているので利用し易く、また地域課題という全体のテーマが存在するため特徴あるデータ収集が可能である。しかし、当然ながら地域課題というテーマ限定なので、多様なデータ作成環境には使えない。そして、このような特定のテーマに特化するようなデータ収集を実現するには、専用のウェブアプリケーションや、カメラアプリケーションを用意する必要があり、誰でも気軽に独自のテーマに基づいたデータ収集ができるわけではない。

そこで我々は、簡単にユーザ独自のテーマ(視点)に基づいたデータ作成・収集が行えるウェブアプリケーション Kotoli を提案する。Kotoli は、ユーザが指定する被写体をユーザの指定した視点(入力項目)で記録するためのテンプレート作成環境(テンプレートエディタ)とそのテンプレートを利用したカメラ(テンプレートカメラ)、記録したデータを保存するデータベースを提供する。また、テンプレートカメラは Web ブラウザ上で動作し、自分の Web ページに、テンプレートカメラへのリンクと、このカメラで撮影した写真一覧を埋め込むことができる。これにより、ユーザ独自のテーマに基づく写真共有サイトを気軽に作成でき、より特徴のあるデータ収集が可能になる。

我々はデータ入力項目のカスタマイズだけでなく、どのようなアングルで写真を撮ればいいのかなど撮影視点を指示することができるテンプレート撮影も提案する。どのアングルで撮影するのか、どの要素に注目して撮影するのか、を視覚的に指示することで統一感のある写真データを記録できる。

連絡先: 後藤孝行, 国立情報学研究所, 東京都千代田区一ツ橋
 2-1-2, tygoto@nii.ac.jp

*¹ <http://www.fixmystreet.com/>

2. Kotoli

2.1 アプローチ

ユーザ参加型のデータ作成においては、ユーザの自発的な貢献が必要となる。そのため、(1) 参加の動機付けとなる短期的・長期的なメリットが示されていること、(2) ユーザの能力の範囲内で過度な負担なく貢献ができること、の二点が重要となる [江渡 12]。Kotoli は、参加の動機付けとしてユーザ独自の写真共有サイトの構築、また、負担の軽減のためシンプルなテンプレート作成ルールになっている。

Kotoli は、ユーザ独自の視点に基づいた記録を行うテンプレートが作れることで、テンプレート作成ユーザが望むテーマの写真を集めることができる。また、自分のページにカメラへのリンク、写真を埋め込むことで写真共有サイトを作ることができる (図 1(1,2))。データ作成者は、気に入ったテーマの写真共有サイトに参加することで、興味ある写真閲覧が可能になる (図 1(3))。

Wiki 的アプローチをとっているアプリケーションのように、クラス定義や、その継承を考えるのはユーザにとって負担が大きい。Kotoli は、テンプレートには一つの事柄 (一枚の写真) についてのみ記述するというルールだけを用意した。全体を撮影して、次に一部だけを撮影する場合、その一部だけを撮るテンプレート (例えば、手だけ、足だけなど) を作り、テンプレート同士をリンク付けする。一部だけを撮るテンプレートを他のテンプレートも利用することで、横断的な探索が可能になる。このように、複雑な構造を記述しなくても、ある程度の情報の再利用性を高めることができる。LOD として公開する場合、テンプレートに対して構造化データ (外部データへのリンクやプロパティの語彙など) を追加して、蓄積したデータを RDF に変換する (図 1(4,5))。構造化する段階を二つに分けることで、ユーザ独自視点のデータ収集環境を手軽に作れるようにした。データ作成ユーザにとっては、テンプレートを利用して撮影することは、事前に撮影対象を限定することになるが、あらかじめ被写体についての情報が付与されていることは、データ作成時の入力負担の軽減につながる。

2.2 テンプレートエディタ

テンプレートエディタ (図 2) は Web ブラウザ上で動作し、撮影に関する指示を記述するカメラテンプレートと属性と属性値を設定できるプロパティテンプレートを作ることができる。プロパティテンプレートとカメラテンプレートは対になっており、プロパティテンプレートで入力された値は、カメラテンプレートで撮影された事柄に関連づけられる。カメラテンプレートは、カメラ、画像、テキスト、リンクの 4 種類の要素をテンプレートにツールパレットからドラッグアンドドロップで配置していくことで、被写体の指示やアングルの指示などが行える。一つのテンプレートに一つの事柄だけを記録するために、カメラ要素は一つしか配置できないようになっている。リンク要素において、すでに作ってあるテンプレートを呼び出すことができる。テンプレートを流用することは、属性を共有することにつながり、蓄積したデータを横断的に探索することが可能になる。例えば、動物の前足だけを撮影するテンプレートを作り、特定の動物を撮影するテンプレート (猫だけ、犬だけを撮影) においてそのテンプレート使い回すことで、前足という視点で異なる動物を横断して写真を探索できる。

プロパティテンプレートは、属性と属性値の入力フォームを作ることができ、属性値のフォーマットとして、テキスト、数値、値のリスト、レーティング (5 段階) を設定できる。フォーマットを事前に設定しておくことで、文字入力を極力なくし

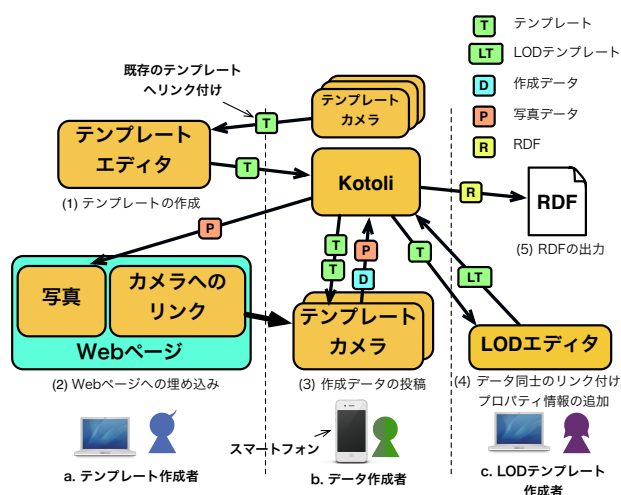


図 1: システム構成。最初に (1) テンプレート作成者が、テンプレートエディタを利用してテンプレートを作成する。このとき、既にあるテンプレートにリンク付けできる。テンプレートを保存するとテンプレートカメラとして利用できるようになる。次に (2) テンプレート作成者は、自分の Web ページ (Blog など) にテンプレートカメラへのリンクを張り、また作ったテンプレートで撮影された写真一覧を表示できるコードを埋め込む。 (3) データ作成者がテンプレート作成者の Web ページを閲覧し、カメラを気に入るとそのカメラを用いて撮影・プロパティの入力を行いデータを投稿する。撮影された写真は、テンプレート作成者の Web ページに反映され、共有される。 (4) 蓄積されたデータを LOD として公開したいとき、LOD テンプレート作成者は、テンプレートに対してさらに構造化データ (外部データへのリンクやプロパティの語彙など) を追加した LOD テンプレートを作成する。 (5) LOD テンプレートを利用して蓄積データを RDF ファイルに変換しダンプする。

ユーザの負荷を軽減することができる。

2.3 テンプレートカメラ

作成したテンプレートは、スマートフォン上の Web ブラウザでひらくことでカメラとデータ入力フォームとして利用できる (図 3)。

テンプレート上のカメラ要素をタップすることで、(1) カメラが起動し写真を撮影することができる。 (2) 撮影後に写真をトリミングすることができ、 (3) トリミング後はテンプレートに埋め込まれた形で表示される (図 3(1,2,3))。

写真撮影において、撮り方は様々である。Kotoli は、テンプレートによって撮り方を指示することができる。事例を提示することによって、適切なトリミングが期待できる。また、撮り方が提示されることで、今まで知らなかった撮影視点 (ものの見方) を共有できる。新たな撮影視点を知ることで、新たな撮影体験を得ることができる。

プロパティモードにすると、テンプレートが裏返し、属性値入力フォームの表示に切り替わる。これによって、属性値が入力可能になる。

Submit ボタンを押すと、投稿する写真、プロパティ値が確認でき、問題なければ投稿する。

2.4 Web ページへの埋め込み

Blog などの自分の Web ページに、作成したテンプレートカメラへのリンクと、そのテンプレートカメラで撮影した写真

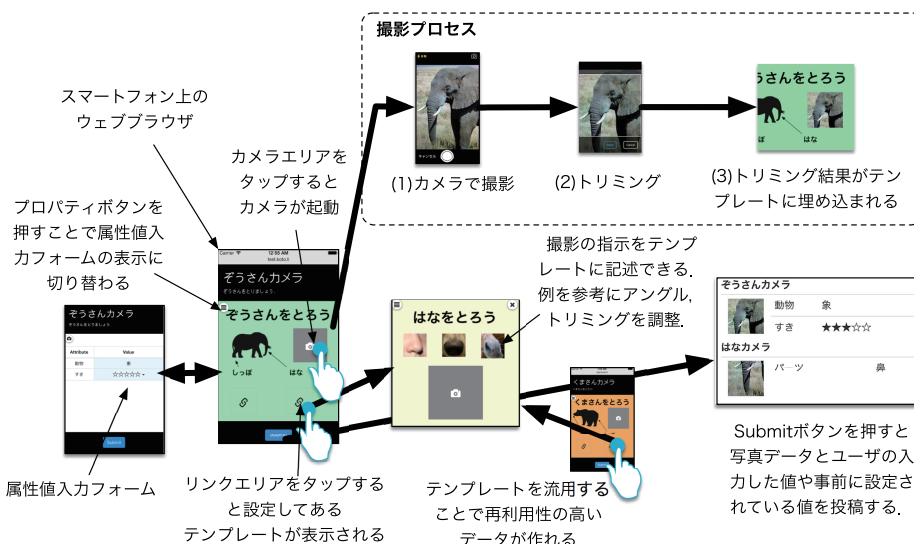


図 3: テンプレートカメラ



図 2: テンプレートエディタ

図 4: Web ページへの埋め込み

一覧を表示するコードを埋め込むことができる (図 4)。これによって、独自のテーマに基づく写真共有サイトをほとんど手間をかけずに作るができる。そして、データ作成者は、Web ページの閲覧の過程で写真を撮ることができる。

Web ページの閲覧過程で写真が撮れることによって、観光案内ページをみながら、そこに埋め込まれているカメラで名所を撮影するといったことが考えられる。また、動物園で様々な動物について調べながら写真を撮るといったこともできる。撮影するカメラに、共通する器官 (例えば、耳、鼻など) の専用テンプレートがリンクされていると、あとから特定の器官だけをすべての動物を横断して閲覧するといったことができる。

2.5 LOD エディタ

テンプレート同士のリンクだけでは、テンプレートの繋がりが表現できていない。複数のテンプレートにリンクされるテンプレートがあることによって、横断的な検索・閲覧が一部可能になっているが、データのマッシュアップには利用しにくい。例えば、テンプレートのプロパティの語彙も複数のテンプレートに共通するものがあるが、最初の段階では別のものとして扱われる、といった、語彙の統一がなされていない。また異なるテンプレート自体が同じ事柄を指しているものが存在する

可能性もある。

そこで、Kotoli では、LOD エディタが用意されている。LOD エディタは、テンプレートに対して、外部データへのリンクや、プロパティの語彙などを追加した LOD テンプレートを作成する (図 5)。そして、LOD テンプレートを通して、蓄積されたデータを RDF に変換しダンプする。ダンプしたファイルを RDF ストアに格納し、SPARQL で問い合わせができるようにすると、マッシュアップサービスからデータが利用できるようになる。

3. 実装

Kotoli のブラウザ上で動作するクライアントは、HTML5+javascript で作成した。カメラやマイクなどのデバイスの機能を利用できる HTML Media Capture API を利用し、ブラウザ上からカメラデバイス呼び出している。サーバ側は、Python で作成した。認証システムには、Twitter API*2 を利用した。データベースには、MongoDB*3 を利用した。

*2 <https://dev.twitter.com>

*3 <https://www.mongodb.org>

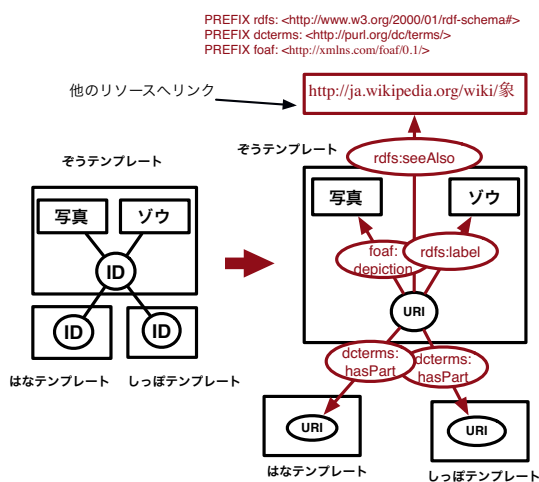


図 5: LOD エディタによる情報の追加
(赤文字が追加される情報)

4. 議論

4.1 運用経験

Kotoli を 2 週間ほど共著者を中心に運用した。ここでは短い期間ではあるが、その運用経験について述べ、考察する。

テンプレートカメラは、事前に撮影するものを限定するため、突発的な出会い、出来事を対象にした撮影にはあまり向かない。ただし、撮影行為において、タスク (例えば、名所を撮影する、動物を撮影する、など) が有る場合や、習慣的に撮影するものにおいて有効に感じる。ただ、食事の撮影において、撮影したあと、食べるというプロセスが間に入るため、食事の評価データ入力まで、間が空く。Web フォームにおいて、submit を押さないまま放置することに慣れない、といった意見があった。ブラウザ上ですべてできる反面、ブラウザ上の習慣に影響されるので、撮影、データ入力、投稿が連続した流れの状況でない、使いにくいのかもかもしれない。

追加で写真を撮りたいときがあり、テンプレートが対応していないと撮影を諦める必要があった。同様に、任意の情報を記述したいときに、入力プロパティが設定されていないと不便である。リンクされていない別のテンプレートを呼び出したいときがある。このように、テンプレートが想定していない行為を望んだとき、不満が生じる。もうすこし、テンプレートの設計において柔軟な要素を取り入れる必要があるかもしれない。例えば、プロパティは自由に追加できる、なにも情報が付与されていない空のテンプレートを自由にリンクさせて写真を追加できるようにする、などが考えられる。また、撮影に応じて、テンプレートを推薦するような仕組みも必要かもしれない。

4.2 構造化の粒度について

Kotoli は、テンプレートには一つの事柄 (一枚の写真) についてのみ記述するという制約だけで、とてもシンプルなルールを採用している。記述データの再利用性を考えるならば、テンプレートをクラスとして定義して、それを継承して利用することで、テンプレート作成段階で複雑な構造を記述できるようにするのが、適当だと考える。しかし、ユーザ独自のテーマ、視点で多様なデータを集めようとするとき、テンプレート設計段階で高度なデータ管理知識の要求は手軽さを損なわせ、適切ではない。気楽にテンプレートを作らせ、あとからデータを統合していくほうが、多くのユーザが参加するシステムには合っていると考える。

効率的なデータの統合が可能か、またテンプレートを事柄ごとに分けたことがデータ統合にどう影響するのかを検証する必要がある。

5. おわりに

モバイル環境における構造化データ作成支援システム、Kotoli を提案した。Kotoli は、事前に撮影対象について記述したテンプレートを利用して写真撮影することで、撮影時のデータ入力の手間を軽減した。また、Web ページに、カメラのリンクを張り、そのカメラで撮影した写真一覧も埋め込むことができることで、写真共有サイトを手軽に作るができる。これにより、ユーザ独自の視点による写真データベースを作ることができ、多様性のあるデータ集合の蓄積に役立つ。

今後は、運用経験でわかった問題点を修正し、Kotoli を公開して外部の人も交えた運用を行いたいと考えている。

参考文献

- [Becker 09] Becker, C. and Bizer, C.: Exploring the Geospatial Semantic Web with DBpedia Mobile, *Web Semantics: Science, Services and Agents on the World Wide Web*, Vol. 7, pp. 278–286 (2009)
- [Bollacker 08] Bollacker, K., Evans, C., Paritosh, P., Sturge, T., and Taylor, J.: Freebase: a collaboratively created graph database for structuring human knowledge, in *Proc. of SIGMOD '08*, pp. 1247–1250 (2008)
- [Ermilov 11] Ermilov, T., Heino, N., Tramp, S., and Auer, S.: OntoWiki Mobile – Knowledge Management in your Pocket, in *8th Extended Semantic Web Conference (ESWC2011)* (2011)
- [Hamasaki 11] Hamasaki, M., Goto, M., and Takeda, H.: Social Infobox: Collaborative Knowledge Construction by Social Property Tagging, in *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work, CSCW '11*, pp. 641–644, New York, NY, USA (2011), ACM
- [Hepp 06] Hepp, M., Bachlechner, D., and Siorpaes, K.: OntoWiki: Community-driven Ontology Engineering and Ontology Usage based on Wikis, in *Proc. of WikiSym '06* (2006)
- [Lee 10] Lee, T. B.: The year open data went worldwide, http://www.ted.com/talks/tim_berniers_lee_the_year_open_data_went_worldwide (2010)
- [Suddath 08] Suddath, C.: Making Water a Matter of Race, <http://content.time.com/time/nation/article/0,8599,1822455,00.html> (2008)
- [Völkel 06] Völkel, M., Krötzsch, M., Vrandečić, D., Haller, H., and Studer, R.: Semantic Wikipedia, in *Proc. of WWW '06*, pp. 585–594 (2006)
- [江渡 12] 江渡 浩一郎, 濱崎 雅弘: 集合知による Linked Data の構築 (特集) Linked Data とオントロジー, *人工知能学会誌*, Vol. 27, No. 2, pp. 181–188 (2012)