

**科学研究費助成事業 研究成果報告書**

平成 29 年 6 月 7 日現在

機関番号：32675  
研究種目：基盤研究(C) (一般)  
研究期間：2014～2016  
課題番号：26330349  
研究課題名(和文) リンクト・オープン・データをものづくり分野で活用するための情報連携方式の開発

研究課題名(英文) Mash-up Design for Linked Open Data

研究代表者  
藤井 章博 (FUJII, Akihiro)  
法政大学・理工学部・教授

研究者番号：40241591  
交付決定額(研究期間全体)：(直接経費) 2,500,000円

研究成果の概要(和文)：「リンクト・オープン・データ」として、構造化したデータを相互にマッシュアップすることで新しいサービスを生み出す研究を行った。まず、東大阪市のネジEDI向け製品コード体系「N研コード」をリンクトオープンデータとして公開した。その成果は、電子情報通信学会の論文誌で出版した。また、J-POP分野に関して、電子商取引サイト、DB-Pedia等を連携づけるアプリケーションを開発し、LODチャレンジにおいて奨励賞を受賞した。さらに、Twitterのデータから即時性のある重要キーワードを抽出するサービス等を実装し、情報処理学会において発表し奨励賞を得た。今後の展開として、成果をIoT分野へ活用する。

研究成果の概要(英文)：As LOD (Linked Open Data), EDI code for screw distribution is used and made it into RDF type data base in 2015. We have published this result in an academic journal. In 2016 J-POP application is made with several mash-up data such as E-eCommerce site and DB-pedia. This results won a minor priers at LOD challenge Japan. Twitter data and public transportation data are utilized for creating new services. The Information Processing Society Japan gave a student award for the presentation made in March 2016.

研究分野：分散システム

キーワード：リンクトオープンデータ Web サービス マッシュアップ RESTful Web API Open Street Map Twitter

## 1. 研究開始当初の背景

様々な分野で、LOD(Linked Open Data, リンクト・オープン・データ)を構築し新しいアプリケーションに利用しようとする試みが広がっておりつつある。

国内では、RDF形式の日本語 Wiki である DBpedia Japanese や日本語 Wikipedia オントロジー、理化学研究所の「Link Data」、中部大学の「SparqlEPCU」が挙げられる。データセットの作成や、LOD を活用したアプリケーションの開発を補助するアプリケーションも開発されている。また 2011 年から「Linked Open Data チャレンジ Japan」も開催されている。これらにより、LOD のためのデータセットや、LOD を活用したアプリケーションが加速度的に増加している。

ソフトウェア開発上の学術的課題は、複数の Web API を組み合わせて新たなアプリケーションを開発するための開発環境と手法の確立が急務であることである。

## 2. 研究の目的

本提案では、LOD 形式で提供される情報リソースを活用するマッシュアップアプリケーションを開発し、ものづくり等様々な分野での業務支援を目的とするソフトウェア開発環境の支援ツールを構築する。

## 3. 研究の方法

RESTful Web API によって RDF を公開し、後者は、Web ブラウザ上で SVG(Scalable Vector Graphics) の DOM(Document Object Model) を HTML5 機能で操作し実現する。

## 4. 研究成果

はじめに、成果の概要を述べる。「リンクト・オープン・データ」として、構造化したデータを相互にマッシュアップすることで新しいサービスを生み出す研究を行った。まず、東大阪市のネジ EDI 向け製品コード体系「N 研コード」を LOD として公開した。その成果は、電子情報通信学会の論文誌で出版した。また、J-POP 分野に関して、電子商取引サイト、DB-Pedia 等を連携づけるアプリケーションを開発し、LOD チャレンジにおいて奨励賞を受賞した。さらに、Twitter のデータから即時性のある重要キーワードを抽出するサービス等を実装し、情報処理学会において発表し奨励賞を得た。今後の展開として、成果を IoT 分野へ活用する。以下、各項目の詳細を述べる。

### 1) ねじ LOD

「N 研コード」と呼ばれる EDI 向けのねじ製品のコード体系は、東大阪市に本拠を置くねじ企業間情報処理研究会による共通コード策定への取り組みに端を発する。同研究会で

は、大阪、京都、奈良にあるねじ商社約 30 社が集まり、共通の EDI システムを構築することを目的に、研究会を設立し、1997 年(平成 9 年)以来、ねじ統一名称および関連事項のコード付け、プロトコルの制定、通信手段の決定、EDI 導入に伴う基幹システム構想の具体化などを手がけてきた。現在、ねじ分野の商品名のレベルで約 3400 種類、25 万アイテムの業界統一コードを制定して活用している。平成 25 年には、この研究会は発展的解消され、N 研コードは、大阪鋳造卸商組合の管理のもとで運用されている。

N 研コード体系は、総合品種マスターと総合商品マスターからなり、構造化のもとで表形式にまとめられている。我々は、ものづくり分野に LOD を普及させることを目標とし、民間企業とコンソーシアムを作り、2015 年 7 月にもものづくり分野の LOD を提供する Web サイトを公開した。  
( <http://monodzukuri.lod.org/> )。

### 2) 機械図面とのマッシュアップアプリケーション

ねじ LOD の利活用のあり方を検討する目的で、ねじ LOD と機械図面のデータをマッシュアップしたアプリケーションを開発した。機械図面には、それを構成する部品の名称や、寸法情報が記入されていることが多い。そこでそれらのテキストを抽出し、ねじ LOD に対して検索を行う。これにより、機械を構成する部品の商品名や、寸法の合う部品の商品名を、簡単に調べることが可能になる。

LOD を活用したアプリケーションを開発する時、その LOD の構造や、リソースに関する理解が必要になる。SPARQL エンドポイントを利用することにより、LOD の構造やリソースの検索は可能であるが、SPARQL の文法に関する知識が求められる。そこで、より多くの開発者にねじ LOD を活用してもらうため、キーワードによりねじ LOD 内のリソースを簡単に検索できるアプリケーションを実装した。

### 3) J-POP マッシュアップアプリケーション

電子書籍市場の拡大はまだまだ途上であり、電子書籍の利用率は決して高いとはいえない。この利用率が低い要因として、「電子書籍のコンテンツ数が紙媒体の書籍に比べてまだまだ少ない。」、「紙媒体の書籍と価格が変わらないため、わざわざ電子書籍に切り替える必要性を感じない。」といった理由が考えられる。これらの理由には、電子書籍の作成に必要な電子データ作成の手間やコストが大きく関わっている。「手間がかかるからコンテンツを作成しない。」、「コストがかかるから価格を安くできない。」といった問題を公開されている既存の LOD を活用することで解決することを考える。そこで本研

究では、LOD を電子書籍化するサービスを提案し、電子書籍の普及率と利用率の向上を図ることを目的とする。

DBpedia Japanese とは、日本語版の Wikipedia から構造化データを抽出して LOD として再公開するプロジェクトである。主に、Wikipedia 各ページのタイトルや概要、Infobox の情報を変換して作成されている。

本研究では、国内の LOD のデータセットの中から、この DBpedia Japanese をデータセットとして使用した。DBpedia Japanese は、SPARQL を用いて問い合わせを行うことができ、結果を取得する際のデータ形式は様々な形式を指定することが可能である。例えば、HTML 形式を指定して、Web ブラウザで閲覧することや、JSON (JavaScript Object Notation) 形式や XML (Extensible Markup Language) 形式を指定して、コンピュータで処理できるようにするといったこと等が可能である。

本研究では、LOD に制約が無いという特性を活かすために、様々な Web API (Web Application Programming Interface) とマッシュアップさせてオリジナルのサービスを作成した。Web API とは、Web サイトに外部のサイトの提供する機能や情報を組み込んだり、アプリケーションソフトから Web 上で公開されている機能や情報を利用したりする際に用いられるものである。Web API で機能を公開しているサーバに対して、インターネット等の通信ネットワークを通じて依頼内容を HTTP リクエストの形で送信することで、処理結果が HTTP レスポンスの形で送られてくる。送受信されるデータの形式は API によって様々だが、Web でよく用いられる JSON や XML 等が用いられることが多い。

本研究において、Web API を利用したサービスの概要としては、「YouTube API」や「Dailymotion API」を利用した動画共有サイトでの動画検索、「楽天商品検索 API」を利用した通販サイトでの商品検索をそれぞれ、楽曲の検索結果を表示するページで行えるようにした。上記に挙げた API 以外にも様々な API が現在公開されている。本研究では、DBpedia Japanese と 3 つの Web API を組み合わせたサービスを作成した。

#### 4) 市バス運行情報と観光情報のマッシュアップ

本事業は、恵那市バスの「停留所と路線の情報」「時刻表の情報」を GTSF データとして整備することから始めた。これまで述べたように、Google GTSF だけに依存するデータ作成手順は、その後の運用において Google の事業展開だけに依存せざるを得なくなるという懸念がある。そこで、OSM を利用したデータ構築も併せて行ってきた。OSM は地図情報を運用するための情報基盤である。そこで、市バスの運行情報のような定期的に刷新を

伴うデータの運用に関しては、オープン化を意識したデータ作成ツールが不可欠であると考えた。28年度事業では、観光課の担当者と恵那市バスとの間の業務フローを分析し、その結果に基づいていくつかの表計算ファイルとデータ処理用のプログラムを開発した。これらのファイル・プログラムを利用することで、GTSF/OSM 双方に対応したデータの作成が効率化できる。また、最新の時刻表データをオープンデータとして (pdf 形式でなく) 公開できれば、GTSF 事業と直接関連しないサイトであっても、検索エンジン上でヒットし、利用者が参照できるようになるはずである。

#### 5) 行政情報とのマッシュアップ

本研究では、オープンデータと地図情報サービスである Open Street Map を関連付け、地方自治体の情報を視覚的に提供するサービスを開発した。OSM (オープンストリートマップ) は、利用者の誰もが自由に編集でき、自由に利用することができる。そのため、個人の意思をより反映することができる。2016年5月には登録者数は約260万人にも上っている。Google マップの場合、利用規約の複雑さやアクセス量が一定以上になった場合、有償となるデメリットがある。したがって、自由に編集をすることができる OSM を使用することで、自治体情報に沿ったサービスを開発することができることから、本研究に用いるのに適すると考えられる。

OSM の地図を自分の Web ページに組み込み、情報提供を行うサービスの開発を行った。まず、OSM の地図をベースマップとして用いるため、表示させた。JavaScript で、OSM の地図の URL とクレジットを表記するソースコードを記述した。GeoJson とは、JSON (JavaScript Object Notation) を用いて、空間データをエンコードし、非空間属性を関連付けるフォーマットである。

市と町の地図情報、人口総数、世帯総数を示す画面を作成し、これらのメンバーを用いてヒートマップを作成した。市町村をクリックするとその市町村の詳細情報をポップアップで表示する JavaScript と、色分けを行う Java アプリケーションを作成した。

#### 6) Twitter API の活用に関する研究

Twitter などのマイクロブログが多くの人々に利用され、情報の配信や共有が行われている。Twitter を利用するユーザは 140 字以内の tweet と呼ばれるテキストをやりとりすることで情報の配信が可能になる。本研究では、Twitter の利用方法の中でハッシュタグに注目する。これは、テレビ番組に対する感想や内容を投稿する場合に、検索を容易にするためのキーワードとして付与される。付

与されたハッシュタグによって、ユーザは関心のあるイベントを特定して tweet を閲覧することができる。このようなある特定のイベントに関して投稿された tweet は " イベントストリーム " と呼ばれ、イベント実況状況を如実に表現している。しかし、イベントストリームは移り変わりが頻繁であり、ユーザが関心のある tweet を閲覧し、イベントストリーム内の全ての内容を理解することは容易ではない。そこで、イベントストリームの解釈を適切に行えばリアルタイムに状況の把握が行える。また、テレビ番組をリアルタイムに視聴していないユーザが過去の tweet をハッシュタグから特定し、閲覧する場合にも問題が起こる。そこで本研究ではイベントストリーム中の特定の状態を抽出することによってこの問題に対処する。イベントストリームでは、視聴者の関心の高まりにより tweet 数が急激に上昇する現象がある。この現象は " バースト " と呼ばれている。バーストの生起要因は様々考えられるが、特にニュースやスポーツ番組、ラジオ放送といった速報性の高いメディアとの関連が強い。本研究では、イベントストリームに関してバースト検出とバースト検出内の解析の2つに分けることで、時々刻々と変化するある特定の出来事に対する投稿から要約生成を行うことで実況状況を分かりやすく可視化し、その評価を行った。

#### 5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[雑誌論文(査読有)](計 2 件)

江上周作、藤井章博、川村隆浩、大須賀明彦「BOM エージェントの実現に向けた LOD の構築」、電子情報通信学会論文誌 D Vol. J 98 D/ No. 8、 pp.992-1004 2015/06

藤井章博、IEEE 論文に基づく IoT 研究動向の計量書誌学的調査 科学技術動向 149、19-24 2015/03/25

[学会発表](計 5 件)

Rishabh Upadhyay、Akihiro FUJII Semantic Knowledge Extraction from Research Documents, , Proceedings of the 2016 Federated Conference on Computer Science and Information Systems, ACSIS, 8, 439-445 2016/09/14, (ポーランド、グダニスク)

Rishabh Gyanendra Upadhyay, Akihiro FUJII, Extracting Knowledge from Technological Research Papers in Application of IoT, 2016 Proceedings of PICMET '16: Technology Management for Social Innovation 2645-2651

2016/09/08, (米国, ホノルル)

赤木文彦, 小城卓也, 田邊佑貴, 藤井章博, バースト検出に基づく要約手法の検討 情報処理学会第 78 回全国大会 2016/03/11, 慶応義塾大学 (神奈川県, 横浜市)

Daiki MURAYAMA, Shouki MIZUNO, Akihiro FUJII, Distributed Synchronization over RESTful Web API Distributed Synchronization over RESTful Web API NBIS-2015 The 18-th International Conference on Network-Based Information Systems 2015/09/04, (台湾, Taipei)

子安貴宏, 藤井章博, 技術文書に見るインターネット要素技術の動向に関する研究 研究・技術計画学会第 29 回年次学術大会 2014/10/18, 立命館大学 (滋賀県, 草津市)

[その他]

ねじ LOD ホームページ:

<http://monodzukurilod.org/nejii/>

Music Sparql: アプリケーション

Linked Open Data チャレンジ Japan2014, アプリケーション部門審査員特別賞, 2015/03/07

<http://musicsparql.appspot.com/>

#### 6. 研究組織

(1) 研究代表者

藤井章博 (FUJII, Akihiro)

法政大学・理工学部・教授

研究者番号: 40241591