

行政と市民によるオープンデータ共創支援プラットフォーム LinkData.org

LinkData.org: A Platform Supporting Co-Creation of Open Data among Administrators and Citizens

下山 紗代子*1
Sayoko Shimoyama

豊田 哲郎*1
Tetsuro Toyoda

*1 独立行政法人理化学研究所 情報基盤センター 統合データベース特別ユニット
Integrated Database Unit, Advanced Center for Computing and Communication (ACCC), RIKEN

Open Data activity has spread rapidly throughout Japan in recent years. However, data catalog sites are distributed independently and the results of particular events and contests using their data are invisible from one another. In order for Open Data to be used effectively, secondarily processed datasets, applications using particular data and ideas for utilizing data should be combined and shared together. LinkData.org enables administrators and citizens to publish and link together datasets, applications and ideas and to easily realize horizontal development of products by a “Reverse Mashup” method using Linked Data technologies. This model, which supports co-creation and utilization of Open Data amongst administrators and citizens, will lead to the solution of various regional issues.

1. はじめに

組織や個人が持つデータをインターネットを通じて公開し、第三者に利活用の機会を供する「オープンデータ」と呼ばれる取り組みは、2000年代から米国やEU諸国を中心に進められてきた。2013年には日本政府によるデータカタログサイトDATA.go.jp*1が公開されるとともに、オープンデータをテーマとしたハッカソンイベントやコンテスト等が多く開催され、オープンデータは国内においても広がりを見せている。

しかし現状、データカタログサイトは分散して存在し、それらのデータを使ったイベントやコンテストの成果は互いに見えない状態となっている。オープンデータが有効活用され、維持されるためには、一次オープンデータを二次加工したデータ、オープンデータを使ったアプリケーション、オープンデータを活用するためのアイデア等が組織やイベントを越えて結びついた形で可視化され、共有されることが必要である。

本稿では、LinkData.org*2におけるオープンデータ活用支援の取り組みについて実際の活用例を交えながら述べるとともに、今後のオープンデータに必要なプラットフォームのあり方について議論する。

2. 関連研究

データの二次利用性を高めるための新しい情報公開・共有の仕組みとして、Linked Dataが用いられ、データがRDFで公開されるようになった。RDF(Resource Description Framework)とは2004年にW3C(World Wide Web Consortium)によって勧告されたデータの標準モデルであり、Web上で扱う情報を主語、述語、目的語という概念で整理し、この3つの概念(トリプル)を基本単位として表現することで、データの機械可読性を高めている。その一方で、RDFデータ作成の負担を軽減するための支援の必要性が指摘されていた[嘉村 2011]。

そこで我々は表形式のデータからRDFに変換して公開できるシステムLinkData*3を開発し、RDFによるデータの公開の支援を進めてきた[下山 2012]。LinkDataを利用することで、RDF

連絡先: 豊田 哲郎, 理化学研究所情報基盤センター統合データベース特別ユニット, 埼玉県和光市広沢 2-1 情報基盤棟, 048-462-1111, toyoda@base.riken.jp

*1 <http://www.data.go.jp/>

*2 <http://linkdata.org>

*3 <http://linkdata.org/home>

の専門知識を持たない利用者が作成した表形式データをRDFに変換し、公開することが可能になった[Shimoyama 2013]。

3. LinkData.orgによるオープンデータ活用支援

LinkData.orgの基盤では、オープンデータの活用を支援するために、データの作成と公開を支援するLinkDataに加え、アプリケーションの作成と公開を支援するApp.LinkData*4、アイデアの共有を支援するIdea.LinkData*5、及びそれらを市町村毎に分類して提供するCityData*6を公開している。以下、LinkData.orgにおける具体的な支援機能について述べる。

3.1 “逆マッシュアップ”によるオープンデータを活用したアプリケーション作成支援

(1) 逆マッシュアップの定義

“逆マッシュアップ (Reverse Mashup)”とは、データを利用したアプリケーション作成を容易にするための新たなソフトウェア工学手法である。一般的にアプリケーション開発において用いられるマッシュアップ(以降“順マッシュアップ”と呼ぶ)では、API化された複数のデータを束ねるためのプログラムを作成することで、アプリケーションからデータを呼び出して利用することが可能になる。一方本研究において提唱する逆マッシュアップは、既存アプリケーションのプログラムを再利用し、新たなデータをAPI化して挿入することで、プログラミングをせずに新たなアプリケーションを作成することを可能にする手法である(図1)。



図1: 順マッシュアップと逆マッシュアップの比較

*4 <http://app.linkdata.org>

*5 <http://idea.linkdata.org>

*6 <http://citydata.jp>

(2) 逆マッシュアップの実現に必要なシステム要件

本研究では逆マッシュアップの実現に必要なシステム要件を以下の通り定義した。

- ① データ API とアプリケーションは独立して存在し、且つ誰でも自由に組み合わせることができる
- ② データの形式が標準化され、API として誰でも利用できる
- ③ データのスキーマを再利用して誰でも新しいデータを API として公開できる
- ④ 既存のアプリケーションを複製して誰でも新しいアプリケーションとして公開できる

上記要件を満たす基盤を構築するために、App.LinkData を開発した。App.LinkData では、LinkData に公開されている RDF を入力データとして利用し、JavaScript アプリケーションを開発する機能を提供する。本システムのオンライン JavaScript エディタを用いてプログラムを記述することで、サーバを持たないユーザでもアプリケーションを公開することができる。さらに、App.LinkData では RDF におけるトリプルの概念を用いてデータを取得する JavaScript ライブラリ^{*7}を提供している。これにより、共通のスキーマを持つデータであれば入れ替えても動作するようなアプリケーションの開発が可能となる。

3.2 "引用" という概念の採用によるイベントを越えたアイデアの共有支援

本研究において開発した Idea.LinkData は、これまでイベント内部でしか共有されていなかったアイデアを公開することで、他者の考えを取り入れながらアイデアを発展させ、実現につなげられるよう支援するためのシステムである。Idea.LinkData では、以下の“アイテム”と総称するコンテンツを適宜追加していくことで1つのアイデアを記述し、単一の web ページとして公開する機能を提供している。

- ・ データ
- ・ アプリ
- ・ アイデア
- ・ web ページ
- ・ 論文
- ・ 画像
- ・ 動画 (YouTube^{*8} の動画コンテンツに対応)
- ・ スライド (Slideshare^{*9} のスライドコンテンツに対応)
- ・ テキスト
- ・ コード

Idea.LinkData では学术论文における“引用”の概念を採用し、公開されているアイデアを引用して新しいアイデアが作成されると、引用元のアイデア及び引用先アイデアにおいて相互リンクが自動生成される。これによりアイデアの引用関係が可視化され、アイデアがどのように発展したか辿ることが可能となる。

3.3 データ・アプリケーション・アイデア・ユーザの対応関係をを用いた評価と利用度の計測

LinkData.org に登録されたデータ・アプリケーション・アイデア・ユーザの対応関係を記録するために、4つのエンティティ(表1)と13のリンク(表2)で構成される依存グラフ(図2)を定義した。この対応関係の依存グラフに基づき、あるノードへのリンク数を集計することで、これをユーザからの評価と他からの利用度の合算値とみなし、評価指標として算出することができる。

*7 <http://app.linkdata.org/tutorial/api/2.0/script2.js.html>

*8 <https://www.youtube.com>

*9 <http://www.slideshare.net/>

表1: LinkData.org におけるエンティティの定義

エンティティ	定義
データ	LinkData においてユーザが公開したデータセット
アプリケーション	App.LinkData においてユーザが公開したアプリケーション
アイデア	Idea.LinkData においてユーザが公開したアイデア
ユーザ	LinkData.org にてアカウント登録を行ったユーザ

表2: LinkData.org におけるエンティティ間のリンクの定義

ID	リンク名	定義
<i>Ldd</i>	再利用	既存のデータのスキーマを再利用して新しいデータが作成された時の、新しいデータと元のデータの関係
<i>Ldu</i>	帰属	データとそれを作成したユーザの関係
<i>Laa</i>	フォーク	既存のアプリケーションのソースコードをコピーして新しいアプリケーションが作成された時の、新しいアプリケーションと元のアプリケーションの関係
<i>Lad</i>	ロード	アプリケーションとその入力として指定されているデータの関係
<i>Lau</i>	帰属	アプリケーションとそれを作成したユーザの関係
<i>Lii</i>	引用	既存のアイデアを引用して新しいアイデアが作成された時の、新しいアイデアと元のアイデアの関係
<i>Lid</i>	引用	アイデアとそこで引用されているデータの関係
<i>Lia</i>	引用	アイデアとそこで引用されているアプリケーションの関係
<i>Liu</i>	帰属	アイデアとそれを作成したユーザの関係
<i>Luu</i>	フォロー	あるユーザとそれをフォローしたユーザの関係
<i>Lud</i>	評価	データとそれを評価したユーザの関係
<i>Lua</i>	評価	アプリケーションとそれを評価したユーザの関係
<i>Lui</i>	評価	アイデアとそれを評価したユーザの関係

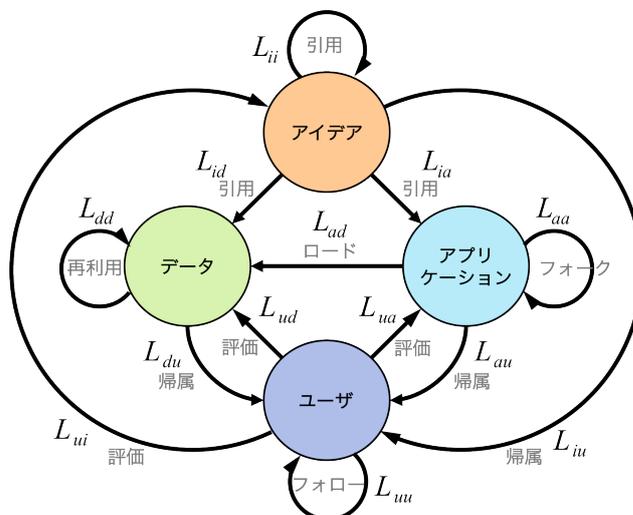


図2: データ・アプリケーション・アイデア・ユーザの対応関係を記録するための依存グラフ

3.4 データ・アプリケーション・アイデアの市町村毎の分類による地域のオープンデータ活動の促進

(1) 市町村毎分類と総合評価の表示

LinkData.org では、データ・アプリケーション・アイデアの公開時に関連する市町村名を登録することができる。この情報を元に、データ・アプリケーション・アイデアを市町村毎に分類して提供する CityData を開発した。

CityData では、前述の依存グラフ(図 2)に基づいて算出したデータ・アプリケーション・アイデアの評価指標を市町村毎に集計し、総合評価として提示している。また、この総合評価の数値の降順に市町村を一覧表示している。本機能は、市町村間での競争意識を高め、地域におけるオープンデータ活動を促進することを旨とするものである。

(2) 地図表示の自動化

CityData では、W3C RDF Interest Group による Geo vocabulary^{*10} のプロパティ(表 3)を用いて緯度・経度が記述されているデータが LinkData にアップロードされると、自動で地図表示画面を生成する機能を提供している。本機能により、プロパティ URI を使って緯度経度を記述することを推奨することで、RDF について詳細を知らないユーザが RDF のメリットを端的に理解できるようにする狙いがある。

表 3: LinkData で推奨する Geo vocabulary のプロパティ

URI	説明
http://www.w3.org/2003/01/geo/wgs84_pos#lat	緯度
http://www.w3.org/2003/01/geo/wgs84_pos#long	経度

3.5 ユーザ毎に束ねた検索結果表示による専門家探しと成果共有支援

(1) ユーザ検索による専門家探し支援

本研究において、データ・アプリケーション・アイデアの検索結果をユーザ毎に束ねて表示する検索機能を開発した。本機能では、特定のキーワードをタイトル・説明・タグのいずれかに含むデータ・アプリケーション・アイデアを検索することができる。また、検索結果を作成者ユーザ毎に束ねることで、特定のキーワードに関するデータ・アプリケーション・アイデアを多く公開しているユーザを特定することができる。こういったユーザはそのキーワードの分野に比較的精通していると見なすことができるため、ある分野の専門家を探す目的に本機能を利用できると考えられる。

(2) ユーザプロフィール表示による成果共有支援

本機能において提供するユーザのプロフィールページでは、これまでに当該ユーザが公開したデータ・アプリケーション・アイデアを一覧表示するとともに、連絡先情報や紹介文を追加することができる。ここでは研究者の業績評価で成功しているモデルを一般に展開することを目指している。すなわち、ハッカソンイベントやコンテスト等の成果をここに記録することで、個人の業績リストの役割を果たすものとなる。これにより積極的な成果発表の動機付けを行うことができ、従来イベントを越えて共有されにくかった成果の共有が可能となる。

*10 <http://www.w3.org/2003/01/geo/>

4. LinkData.org の活用事例

LinkData.org では 2013 年 3 月 9 日現在、1052 データセット(うち公開:670、限定公開:84、非公開:298)、428 アプリケーション(うち公開:255、限定公開:73、非公開:100)、155 アイデア(うち公開:137、限定公開:5、非公開:13)が登録されている。LinkData.org に公開されたデータセットについて、本基盤公開後の 2011 年から 2013 年にかけての推移を見ると、2012 年から 2013 年にかけて地方自治体によるオープンデータ、及び市民によるオープンデータが急増している(図 3)。この結果から、市民からデータ公開できるオープンデータの情報基盤が必要とされていることが分かる。

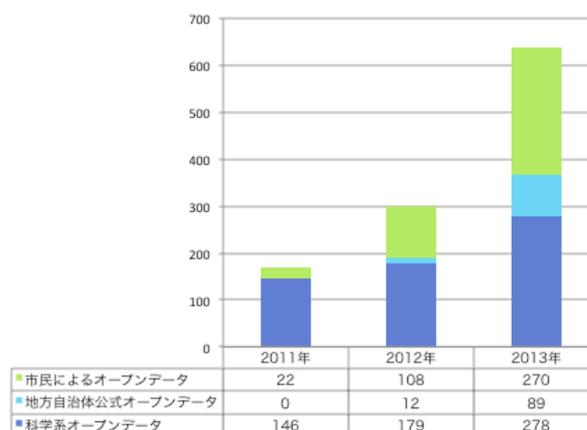


図 3: LinkData における公開データ数の推移

※科学系オープンデータのうち大規模なものは BioLOD^{*11} を使って公開

4.1 地方自治体職員によるオープンデータを活用したアプリケーションの公開

App.LinkData で公開された「スポット×混雑度マッシュアップ MAP^{*12}」は、地域の防災関連施設・設備を周辺の混雑度順に地図上にランキング表示できるアプリケーションであり、株式会社ゼンリンデータコムが提供する「いつも NAVI 開発キット API^{*13}」と、LinkData に公開されている避難施設等の位置情報の含まれたデータセットを順マッシュアップして作成された。

上記アプリケーション公開後に、秋田県横手市情報政策課のアカウントから「横手市×混雑度マッシュアップ MAP^{*14}」が公開された(図 4)。このアプリケーションでは入力データとして同アカウントから公開されている秋田県横手市のデータが指定されており、App.LinkData の「逆マッシュアップ」機能を使って横手市職員により作成されたアプリケーションであると特定できた。

さらに、静岡県裾野市のアカウントからも同様に逆マッシュアップにより作成された「裾野市×混雑度マッシュアップ MAP^{*15}」が公開された。

これらの事例から、プログラムを改変せずにデータを差し替えるだけで新たなアプリケーションが作成できる「逆マッシュアップ」はアプリケーション開発を支援する手法として有用であると考えられた。また、地域のためのアプリケーション開発においては、地域に適合するデータ作りが重要である。そのため、データ作成者が主体的にアプリケーションの再利用をできるようにすることが、地域を越えてアプリケーションを横展開させるのに有効であると考えられる。

*11 <http://biolod.org/>

*12 <http://app.linkdata.org/app/app1s420i>

*13 <http://lab.its-mo.com/service/densitymap.html>

*14 <http://app.linkdata.org/app/app1s454i>

*15 <http://app.linkdata.org/app/app1s505i>



図 4:逆マッシュアップを用いて公開されたアプリケーション

4.2 市民による自治体データの二次加工と公開

2013年に開催されたハッカソンイベント「第1回 Linked Open Data ハッカソン関西 in 大阪」では、大阪市より提供された施設情報・防災情報の CSV 等のデータを、イベント参加者が LinkData を使って RDF に変換した。その結果、9,238 件の施設情報^{*16}、4,694 件の防災情報^{*17}、18,780 地点分の大阪市内の標高情報^{*18}が市民の手により公開された。

さらにこれらのデータを用いてアプリを開発するハッカソンイベントも複数回に渡り開催され、App.LinkData を利用したアプリケーションも複数公開された。

こうした市民によるオープンデータ活動が広がりを見せていることから、政府や地方自治体によるデータカタログサイトだけではなく、市民が二次加工したデータやアプリケーション等の成果を登録できる基盤の需要は今後さらに大きくなると考えられる。

4.3 市民による地域組織の活動の可視化

Code for Japan^{*19}では、各地域において課題解決に取り組む単位である Code for Japan Brigade の活動を可視化するために、Idea.LinkData への組織紹介ページの公開を行っている(図 5)。ここではアイデアの引用の機能が活用されており、Code for Japan の組織紹介ページ^{*20}に Code for Japan Brigade が列挙される形で組織としてのまとまりが可視化されている。さらに、各 Code for Japan Brigade の登録時に市町村の情報が付加することで、CityData からもその地域において活動する組織が他のリソースと併せて可視化されるようになった。

Code for Japan Brigade のような地域における活動体は今後も増加すると考えられ、こういった活動を可視化し、地域毎に分類することは地域コミュニティの活性化と地域の課題解決への貢献につながると考えられる。

5. 今後の課題と展望

従来のオープンデータは、行政がデータを作成し、市民がアプリケーションを作成するというモデルで考えられていた。これに対し本研究における LinkData.org では、市民もデータ作成に参加し、行政もアプリケーションを公開することを可能にした。

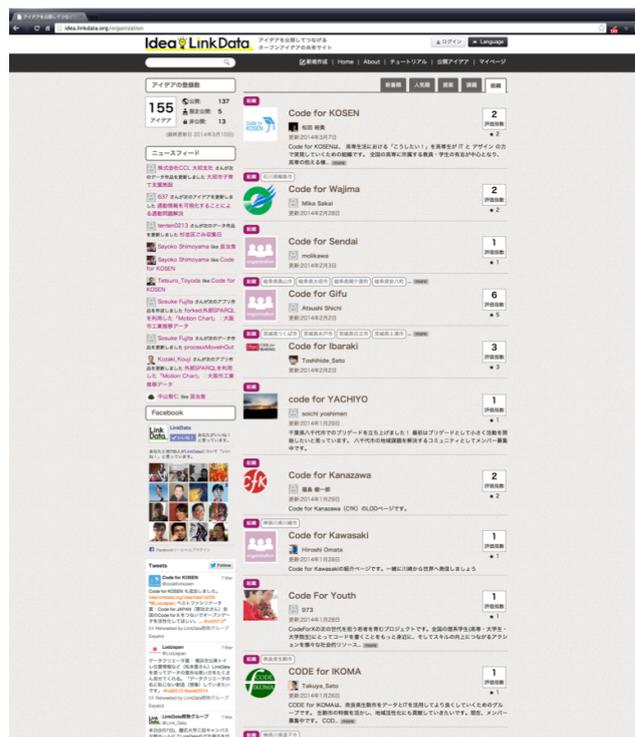


図 5: Idea.LinkData に公開された Code for JapanBrigade の組織紹介ページ

一方で、地方自治体ではオープンデータにかけられる予算や人的資源が限られており、データの加工や継続的なメンテナンスが困難となることが考えられる。そこで市民の協力を期待が寄せられるが、ハッカソン等のイベントだけでは長期的なデータのメンテナンスがなされないことや、データの品質を担保できないことが懸念される。

そこで、市民がオープンデータに係る作業を行った際に対価を得られるようにするための仕組みを作ることを提案したい。これは、行政が所有する一次データに対し、クレンジング、キュレーション、アノテーション等を施すことで、付加価値の高い二次加工データを創造する市民を「データクリエイター」と位置付けることにより、新たな雇用創出を目指すものである。職業としての「データクリエイター」が成立することで、在宅で働きたい人や、リタイア後の人材が活躍できる場を提供することが可能となる。

今後のオープンデータにおいては、市民による二次加工データを流通させるプラットフォームが必要である。

参考文献

[嘉村 2011] 嘉村哲郎, 加藤文彦, 松村冬子, 上田洋, 高橋 徹, 大向一輝, 武田英明: 芸術・文化情報の Linked Open Data 普及に向けた現状と課題—LODAC Museum を例に, 情報処理学会シンポジウムシリーズ, Vol.2011, No.8, pp.409-416, 2011

[下山 2012] 下山 紗代子, 西方 公郎, 吉田 有子, 豊田 哲郎: LinkData.org を使った RDF 教育とデータ公開化運動の推進, 人工知能学会全国大会 (第 26 回)論文集, No. 3C2-OS-13b-2, 2012

[Shimoyama 2013] Shimoyama S., Cox R. S., Gifford D., Toyoda T.: Organizing Scientific Competitions on the Semantic Web, Lecture Notes in Computer Science, Springer, Vol.8055, pp.311-318, 2013

*16 <http://linkdata.org/work/rdf1s933i>
 *17 <http://linkdata.org/work/rdf1s929i>, 他
 *18 <http://linkdata.org/work/rdf1s947i>
 *19 <http://code4japan.org/>
 *20 <http://idea.linkdata.org/idea/idea1s158i>