

海外における「Linked Data とオントロジー」コンテスト

"Linked Data and ontology" contest in foreign countries

乙守 信行*¹

Nobuyuki Otomori

*¹ 株式会社 MetaMoJi

MetaMoJi Corporation

The contest to advance the use of Linked Open Data by a lot of citizenry participation has been held in EU/United States. I introduce the Linked Open Data contest held in EU/United States and consider the problem when we hold the contest in Japan.

1. はじめに

Linked Open Data の運動は政府、メディア、出版、生命科学など広範な分野で進められ 2010 年 9 月現在、252 億の RDF トリプルという広大な Web of Data がネット上に存在している。[萩野他 11]この Web 上にできたグローバルなデータ空間をインフラとすることにより新しいサービスの創造とそれにより個人・組織の利益の享受を実践的に進めることを目的と EU、米国ではコンテストが開催されている。

本稿ではこれらのコンテストと EU と米国の事例を紹介するとともに、日本において同様のコンテストを開催する場合の課題について考察を行なう。

2. EU Open Data Challenge

2.1 主催者と開催期間

EU 第 7 次研究枠組計画(FP7)における「LOD2-Creating Knowledge out of Interlinked Data」プロジェクトの参加団体である Open Knowledge Foundation と Open Forum Academy が主催、Google, IBM, Microsoft などの大手 IT 企業がスポンサーとなり本コンテストは開催されている。

開催期間は 4/5 から 6/5 までの 60 日間である。[Open Data Challenge]

2.2 コンテストの募集内容

募集内容は表1の4部門、トータルで 20,000 ユーロの賞金が授与される。可視化部門では、デザイナーやアーティストも参加対象としていることは、特徴的である。

表 1 コンテストの募集部門概要

アイデア部門	公開された情報を利用したプロジェクトの提案	1位 2000 ユーロ
アプリケーション部門	公開された情報を利用したアプリケーションの提案	1位 5000 ユーロ
可視化部門	公開された情報に対して興味深く洞察に富む表現の提案	1位 3000 ユーロ

データセット部門	政府機関によるもの、政府データを変更した、または相互につなげた、いかなる形式でのオープンデータセットの提案	Talis Awarod 1000 ユーロ など
----------	---	--------------------------------

2.3 参加者を増やすための工夫

応募におけるルールは、EU に在住していることなどの基本的な9項目と補足的な8項目が設定されている。募集内容における可視化部門と同様に多くの参加を促進する以下の工夫が行われている。

- 応募するアイデアやアプリケーションは、コンテスト開始前に作成、または発表されたものでもよい
- エントリー数は自由である
- 年齢制限はない(但し 16 才以下は保護者の許可があるほうが望ましい)

応募方法も、Web から氏名、住所、アイデア、作品の URI などを入力するだけの簡易なものとなっている。

また、ライセンスについては、「4.国内での開催の課題」で後述するが Open Knowledge Foundation が策定を進めている Open Definition[Open Definition]に基づいたものを利用している。

3. アメリカ Challenge.gov

2009 年 1 月 21 日オバマ大統領の就任直後の大統領メモとして表明された以下のオープンガバメント3原則に沿って Data.gov サイトを開設し連邦政府各機関のデータのオープン化をアメリカは進めている。[佐藤他 11]

- 透明性(Transparency)
- 国民参加(Participation)
- 政府間および官民の連携・協業(collaboration)

EU の Open Data Challenge 同様に Linked Data を活用した Data.gov Hackason'09 などが米国でも開催されている。

そして、このオープンガバメント 3 原則を恒常的・効率的に行なうために Challenge.gov が開催されている。[Challenge.gov]

3.1 Challenge.gov とは

2010 年 3 月行政管理予算局 (Office of Management and Budget (OMB)) は各政府機関のコアとなるミッションを進めるためのポリシーと法的な枠組みを記述した覚書を発表した。

この覚書に基づき各課題を解決するための Web プラットフォームの開発を各政府機関が行ない、これらプラットフォームを **Challeng.gov** サイトとして米共通役務庁 **General Services Administration (GSA)** が管理することとなった。

3.2 Challenge.gov の応募内容

応募内容は防衛/経済/教育/エネルギー環境/健康/国際事務/ジョブ/科学技術/個人及び公共の安全の 9 カテゴリーに分類され、2011 年 5 月の時点では 32 の政府機関が 88 の課題の応募を行っている。一部を抜粋した表 2 により、必ずしも賞金がでるものだけではない。また、応募内容もアイデア提案から、ロゴの作成、ビデオ、デジタルゲームやモバイルアプリケーションまで非常に広範な項目となっている。

表 2 Challenge.gov から応募されている課題の一部

カテゴリー	政府機関	課題	応募期間と賞金
防衛	アメリカ空軍	正確なシューティングロケータの設計とシミュレーション	2011/03/02-5/02 \$50,000
経済	社会保障庁	“あなたの生活に違いをもたらした社会保障”をビデオコンテンツ	2011/06/29-08/27, なし
健康	アメリカ疾病管理予防センター	CDC のインフルエンザアプリケーションチャレンジ	2011/04/06-05/27 \$35,000
個人および公共の安全	保健福祉省	医療の質を可視化するコンシューマアプリケーション	2010/09/07-11/15 \$2,500

「CDC のインフルエンザアプリケーションチャレンジ」では **Data.Gov** 同様に以下のデータと API を提供している。

- インフルエンザ予防接種の見積もり(エクセル形式)
- ウィークリーインフルエンザ活動報告(XML 形式)
- データを統合した RSS Feed /JSON FEED

4. 国内でのコンテスト開催の課題

筆者が所属しているセマンティック Web 委員会では、**Linked Open Data** の運動を促進することを目的とした「**Linked Open Data チャレンジ**」の開催を計画している。本稿の最後として、本コンテスト開催における課題について考察する。

尚、本考察は筆者個人のものであり委員会としてのものではない。

4.1 海外と国内状況の違いについて

海外と国内の違いを明確にさせるためキャズムの段階と **Linked Open Data** 運動を対比させてみる。(表 3)

表 3 **Linked Open Data** とキャズム

イノベータ	Linked Open Data の提唱とアカデミックを中心とした Linked Open Data クラウドの形成とインフラとなる
-------	--

	技術開発
アーリーアダプタ	各分野における先進的な政府・企業の Linekd Open Data クラウドへのデータ提供とアプリケーション、サービスの開発
アーリーマジョリティ	アカデミック、一部の先進企業から国民全体が利用する Linked Open Data 基盤の構築とベストプラクティスの生成。これにより、

これまで紹介してきた海外事例はアーリーマジョリティ段階、これに対し国内は、理化学研究所サイネス、国立国会図書館の **NDLSH5.lod.ac** プロジェクト国立情報学研究所の **CiNii** などの実践的な成果はあるが[武田 11]企業参加がない現状からなど考えればイノベータ段階といえる。

4.2 今後の課題

日本における **Linked Data** の課題として[武田 11]では、情報公開・共有の文化、コミュニティの未成熟、中心的データの欠如、日本語のリソースの記述があげられている。国内でのコンテストを開催する場合の課題も多々あるが直近かつ直裁な問題は「情報公開・共有の文化」と「コミュニティの未成熟」である。

(1) データ公開ルールの整備

Eu で行われている **Open Data Challenge** のライセンスは **Open Definition** の 11 の定義に基いている。**Open Definition** は、アクセス、再分配、再利用の各項目において、技術的な制約がないこと(第 4 項)、ライセンスはパッケージに特化してはいけない(第 10 項)、ライセンスは他の作品に再分配に制限を与えてはならない(第 11 項)などオープン性が強い定義になっている。

しかし、現状の日本を考えた場合、本定義では、データ提供が進まない可能性があり、今後日本の現状と参加団体に合せたルールづくりを行なう必要がある。

(2) 個人・企業参加も含めたコミュニティの形成

アカデミック及び研究機関の先進的な実践とともに企業または個人として参加可能なコミュニティの形成が課題となる。**EU** もアメリカも政府主導でアーリーマジョリティ段階に進んでいるともいえるが、その前段階におけるアーリーアダプタ段階としては、より広範囲な個人・組織が参加したコミュニティ形成が必要である。

この二つは相互に強く関係しており、そのためコミュニティの形成の中でデータ公開ルールの整備など諸課題を解決していきたいと筆者は考えている。

参考文献

- [萩野他 11]リンクするデータ(Linked Data)-広がり始めたデータのクラウド- 情報処理学会誌 Vol.52 No3. Mar.2011
- [佐藤他 11]オープンガバメントとオープンデータ- 情報処理学会誌 Vol.52 No3. Mar.2011
- [Open Data Challenge] <http://opendatachallenge.org>
- [Open Definition] Defining the Open in Open Data, Open Content and Open Services <http://www.opendefinition.org>
- [Challenge.gov] <http://challenge.gov/>
- [武田 11]日本における **Linked Data** の現状と普及に向けた課題- 情報処理学会誌 Vol.52 No3. Mar.2011
- [乙守他 11]セマンティック Web コンファレンス 2011 <http://s-web.sfc.keio.ac.jp/conference2011/0405-otomori.pdf>