

VLED 第 2 回技術委員会

議事録

1. 日 時 平成 27 年 2 月 10 日 (火) 15:30～17:30
2. 場 所 三菱総合研究所 4F 大会議室 D
3. 出席者 主 査：越塚登（東京大学大学院情報学環 教授）
副主査：武田英明（国立情報学研究所 教授）
委 員：中尾彰宏（東京大学大学院情報学環 教授）、深見嘉明（慶應義
塾大学 SFC 研究所 上席所員(訪問)）
社 員：KDDI、日本アイ・ビー・エム、日本電気、日本電信電話、日本マイクロソ
フト、日立製作所、富士通、三菱総合研究所
オブザーバー：総務省、内閣官房 IT 総合戦略室、国立国会図書館
自治体会員：千葉市
事務局：三菱総合研究所、YRP ユビキタスネットワークング研究所
4. 配布資料 資料 2-1 「オープンデータガイド」第 2 版案
資料 2-2 平成 27 年度活動テーマ案
資料 2-3 VLED 第 1 回技術委員会議事録
5. 議題
 1. 「オープンデータガイド」第 2 版検討
 2. 平成 27 年度活動テーマ検討
 3. その他
6. 議事録
 1. 「オープンデータガイド」第 2 版検討
 - 資料 2-1 を用いて、「オープンデータガイド」第 1 版及び第 2 版案を説明。
 - IT 総合戦略室が策定したオープンデータ導入ガイドラインを説明。
 - VLED のガイドラインと IT 総合戦略室のガイドラインの位置づけを考える必要がある。
 - オープンデータと情報公開制度は、別の取組として扱っている。
 - VLED のガイドラインは、今後も継続的に修正作業を行っていく予定。第 2 回委員会後に、第 2 版案に対するご意見を受け、それを反映させて修正したい。その上で 6 月くらいには CFC をし、正式反映版としては 7 月くらいかと思う。
 - 世の中いろいろと動いているので、なるべく急ぐように。

- ガイドラインを用いてオープンデータにするときは、どれほど理想に近い形で公開することが求められているのか。例えば RDF とかだと自治体には敷居が高い。
 - 機械判読できることを推奨している。5★Open Data で★3つほどのレベルを想定。最初から RDF、LOD 形式を目指すのは難しいため、公開を最初の目標とすべき。
- 具体的な手順は、IT 総合戦略室版よりも VLED のガイドラインを参照するといいい。
- 文章として公開するだけでなく、チュートリアルや公開されたオープンデータを教えてくれるポータルなどがあると良い。手順の提示や自動処理をするポータルがあれば、そのポータルを利用してオープンデータ化した事例を集計することが可能になり、ポータル自身の認知度も上げやすいと思われる。
- オープンデータ化を実現した人に資格などを与え、“この人に頼めばオープンデータとして公開してくれる”人などと認識させることが可能と考える。
- 他の活動で講習会を実施している。次年度の検討にも絡むが、今後の取り組みとして前向きに考えても良いと思う。
- データサイエンティスト分科会設立に向けた勉強会を開いた際には、データを活用できる人材が不足しているという声が多く出た。自治体の縦割り化が進行しているため、データを横断して活用し、その可能性や魅力を主張する人材が必要となっている。
- ビッグデータ関連のツールを保有する社員企業では、各社でそのツールの講習会をやっていると思われる。その講習会に協力をお願いしたり、VLED で後援をしたりで、各社のリソースをうまく活用できると良い。
- オープンデータはビッグデータよりも新たな取り組みであり、興味を持っている組織は多いと思われる。そのため、VLED からアプローチを取り講習を展開していくことが考えられる。VLED の講師として、社員が所属する企業の担当者向けに講習会を行ったり、さらにその人達が講師となって広めていくような活動があっても良い。
 - VLED からその講師の方へ資格などを与えて、その肩書で地方へいくなどもあって良い。まだみんな方向性がわかっていないので、最初は VLED として引っ張っていくというのはあるかもしれない。方向がわかれば、各社それぞれでやっていける。
- オープンデータは義務感で実施するのではなく、オープンデータを経済活性化に結びつけ、必要性から実施する取り組みに変えていく必要がある。そのためには方向性を設計し、必要なコンテンツを盛り込んでいった上で、講習なり資格なりを考える必要がある。まずは社員の意見を聞きながら内容の設計を行っていきたい。

2. 平成 27 年度活動テーマ検討

- 資料 2-2 を用いて、27 年度の活動テーマ検討について説明。
- 利活用の拡大が現在のオープンデータの課題であることは、誰しもの共通認識。データカタログを導入したとしても、CC-BY が付いている点は評価できるが、以前から公開していたデータをまとめるのみ

に止まっている場合が多い。今後は、利活用への展開が期待できるような、ニーズに合ったデータを出すことが必要となる。

- 一方で、官公庁としては公開時点からデータの用途が把握できていないと、公開に踏み切れなという現状あり。この点はオープンイノベーションに対する阻害要因である。
- 今後はデータを出す方から、使う方に注力していく必要があり、そこに技術委員会としてできることは何か、という議論になる。テーマ1はデータの利用者側、テーマ2はデータの公開側からの取り組みである。
- VLEDとして利活用の方向性を打ち出すことはデータガバナンス委員会でも議論になっている。活用するためのワンストップサイトが必要であり、それに加えて人材育成や講習会等も含めて活動できると良い。
- テーマ1だけでは、利活用の促進に不足していると思える。自治体のへの取組の話が多いが、国や政府のところでもエッジの効いた取組を促すべきではないか。テーマ2のマッシュアップは魅力的な取り組みと思われる。
- 事例を示すだけでなく、一步踏み込んだ利活用促進を目指すべきであり、技術委員会としては、ツールの配布が考えられる。大規模・小規模な取り組み、独創性のある取り組み、他の利活用の基盤となる取り組みなど、多種多様な活用に適応できるよう、技術的にもっと踏み込めないかと思う。
- ゼロからサービスを設計することは難しい。そのため、利活用の普及を考えたとき、既存のサービスにオープンデータを適応するという視点が鍵になると考えられる。その適応した結果、どれほど効果があったかを事例で示せると良い。
- ツールの利用者が不明確なため（専門家／非専門家、サービス事業者・データ管理者など）、ツールの開発に必要な要素が明らかになっていない。
 - オープンデータの活用方向として、地域の課題解決を考える人、自分のビジネス利用を考える人、コンサル、学生など、さまざまに1つにはならない。そういう意味でもエッジの効いたものが必要である。
- 特許業界から見れば、生み出されたテクノロジーに対してその活用方法は想定できない。ビジネスになる活用は、他者が想定しなかった利用をすることで生まれるものである。そうした観点から考えれば、利用の仕方を想定してデータの公開可否を決めるのは適当でない。自治体のデータも手間をかけて公開するのではなく、日々の業務がシームレスにデータ公開に結びつくようなデフォルトで公開する仕組みが重要である。
- 自治体がデータを出すモチベーションをどう上げていくかという視点が必要。ガイドラインを見て、導入した利用者がどのような効果を得たか示すべき。ガイド内容と結果の繋がりが確認できる事例紹介をすべきと考える。
- ガイドラインを読む、具体的な対象者が想定しにくいと感じた。また、利活用の視点で技術委員会はどのような貢献をできるのか検討する必要がある。
 - 利活用においても技術的障害はある。自前で全てやるのは困難で、ツールへの要望が出てきたりする。オープンデータでは、それらツールを誰が開発しているかが認知されておらず、VLEDと

してツールを扱いまとめることは有意義。他分野だと、簡単に集まる要素（専門書や講習会、各種ツールなど）がまだ社会に存在していないため、いち早くその要素を用意するという意義がある。

- ツールパッケージについては、ベンダの立場としては最終的には自社のツールを使ってほしいのが実情。社員企業のパッケージを複数並べた場合には、どのように社員企業の保有ツールを利用してもらうのか工夫の余地がある。
- IT 室の地方公共団体のガイドでは、どういうデータがあるのか。データの品質を★で表わす指標はあるが、データのフレッシュネスを示す指標がない。仮に三ツ星でも、古いデータとなると、クオリティ的に利用することはできない。利用者側が一つ一つ更新日を確認する手間を省くため、データ自体に対する認定制度などがあると、利用する側にも役に立つと思う。また、利用側から、こういったデータで公開することを望むのか、見本を示しても良いと思える。
- 似て非なるデータというものがある。オープンデータを提供する仕組みとしての環境整備から一步踏み込んで、データの中身の標準化も出来ると良い（例えば、年齢の区切りは5歳おきにするなど）。ただ一方で、標準化に従うことは自治体の負荷にもなる。従って、標準化された形式に変換するツールなどが必要になってくる。
 - オープンソースの利用方法が近年変化してきている。1つ1つ必要なモジュールを集め、それらを構築して利用するのは、手間と考えるのが一般的になってきた。現在は、クリック一つで全てのパッケージを自動利用できるものが増えており、敷居を下げる仕組みを考える必要がある。公開前にデータを標準フォーマットにする際にも、ワンクリックで実現するポータルサイトが求められている。簡単に実現できるツールが出てくると、活用するユーザーが広がる。
 - オープンソースを利用したビジネスは、入り口が狭いことが原因で活性化が妨げられているように思える。アカデミックの範囲でも良いから、ツールの利用をフリーにすることなどを考慮すべき。
- 技術委員会としてどこにフォーカスを置くか、検討する必要がある。利活用を進めるためには、エンジニアが楽しく感じる仕組みを作ることが一番の近道と考える。自治体は公開したデータの利用を促すだけでなく、楽しくハックすることを促進すべき。そのためには、こういったツールが必要とってくるか。社員企業が展開するツールと競合したら意味が無い。イノベーションを起こすのは、ハッカソンなどには参加しないエンジニアであり、その人たちを具体的に想定してツール開発を検討すべきと考える。ハッカー的にはフレームワークが使えないものを作ると楽しく、そういうところが盛り上がると技術が楽しいということになる。
- 一社単独では出来ないようなところを、オープンソースとしてやっていく。例えば、コミュニティが手弁当でやっているような CKAN の日本語化や日本独自のモジュール開発、社員企業が持つシステムを連結させるツールの開発、各社の技術担当をオープンソース開発コミュニティに入れる手助けをすること等が考えられる。技術的委員会だからこそ貢献できることは多数ある。

3. その他

- VLED 技術委員会では、何が能力的にできるのかが不明。事務局が手を動かし、委員会で確認をしているというのが現状。委員会・事務局・WG・コミュニティなど、それぞれの単位や組織で何を担当するかを明らかにし、やるべきことを割り振り対応していくべき。
- 委員会の位置付けとしては、オープンデータをどういう分野で使うのか技術の視点で議論をする。統計データならどういう形でフォーカスするか、スキーマとか、日本のオープンデータを技術的な視点からストラテジを決めてやっていくべきである。一方、サイト作りの話しは、企業の方にも入ってもらって一緒になってやるのがいい。今日の概要の議論なども非技術者向けのもの。公開と利活用は技術レベルとすれば差がなく、切り口としては技術者向けかどうか、コードを書くか／書かないかに分けて考えると良い。
- 技術視点からの講習は本委員会で計画を作る必要がある。実際の講師は外部の人間でも良いが、交通整理はここで行う。
 - 内容面とか講習会とかいろいろなアクションアイテムが出たが、どのような体制で活動をしていくか、受け皿がないと、絵に描いた餅になってしまう。委員会は何をやり、プロジェクトは何処を受け皿にするのか、外部の資金を利用してプロジェクトを起こすのか、外部コミュニティとの連携など考えるべきことは多い。今年度はあと2回あるので、実施体制も含めて、事務局を通じて運営委員会とも議論をし、外部との関係も考えながら、次回以降の検討を協力して進めたい。
 - 技術委員会から運営委員会に対して提案していくのも良いと思う。
- 技術委員会としては、三ツ星データの RDF 化やボキャブラリの付加、Sparql 適用データを突き詰めていくのか、地方のオープンデータを生成するための仕組み・モデルを考えるのか、方向性を決めていく必要はある。
- 例えば人口統計なら、このくらいまでの粒度、精度で出せば全国で使えるように成るとか、技術的視点で切って、日本のオープンデータを良く出来ればそれでいい。それは地方創生に反するものではないと思う。

以上