# あなたのスキルは社会に役立つ

2011年3月11日の東日本大震災発生の直後にHack For Japanは発足しました。今後発生しうる災害に対して過去の経験を活かすためにも、エンジニアがつながり続けるためのコミュニティとして継続しています。防災や減災、被災地の活性化や人材育成など、「エンジニアができる社会貢献」をテーマにした記事をお届けします。

## 第76回 選挙のオープンデータ化を目指す Code for 選挙

●Hack For Japanスタッフ　関 治之（せき はるゆき）　Twitter @hal_sk

去る2017年10月に、第48回衆議院議員総選挙が行われました。そのときに筆者が所属するCode for Japanと有志のメンバーで、候補者のオープンデータを作成する「Code for 選挙」プロジェクトが立ち上がりました。今回は誌面をお借りしてそのときの模様をお伝えします。

### 投票に役立つ客観的な情報を集めたい

そもそも、選挙のときに参考になる客観的な情報を集めたいと思ったことがこのプロジェクトを立ち上げるきっかけでした。2017年の衆院選では、希望の党が立ち上がり民進党がそれに合流するなど流動的で、各党の方向性もわかりにくい状況であったことなどから、党の名前ではなく信頼できそうな個人を選びたいと思ったのです。

「政治的な偏りを排除し、それぞれの政治家の事実情報のみを積み上げた客観的なデータベースを作ろう」とFacebookで仲間を呼びかけたところ、20名ほどのメンバーが集まりプロジェクトがスタートしました（**図1**）。プロジェクトメンバーは、最終的には30名ほどになっています。

### 実は、候補者のオープンデータは存在しない

2017年9月29日にプロジェクトをスタートさせてたくさんの仲間が集まったのですが、最初にびっくりしたのが、そもそも候補者の一覧というのが存在しないこと。総務省の公告などから引っ張ってくれば見られるのかと思ったら、一覧のPDFでさえも入手できなかったのです。

そこで、しかたなく各政党のHPなどからデータをクローリングして取ってくることにしました。

選挙区の区画データについては、幸いなことに、東京大学の西沢明先生が整備してくれていたデータ[注1]があり、利用することができました。

また、名前や政党などの情報だけでなく、Facebookやツイッターのアカウントも整理したかったので、それについては人海戦術で集めることにしました。そこで大活躍したのがGoogle Spreadsheet。

注1　東京大学 西沢明先生の選挙区画データ
http://www.csis.u-tokyo.ac.jp/~nishizawa/senkyoku/

▼図1　当時のFacebookへの書き込み

入力ボランティアに広く呼びかけて、必要なデータを埋めていきました(**図2**)。

そして、多くの人の手を借りて、なんとか公示日である10月10日にはデータを完成させ、Webサイトで配布を始めることができました。

## データを誰でも使えるように

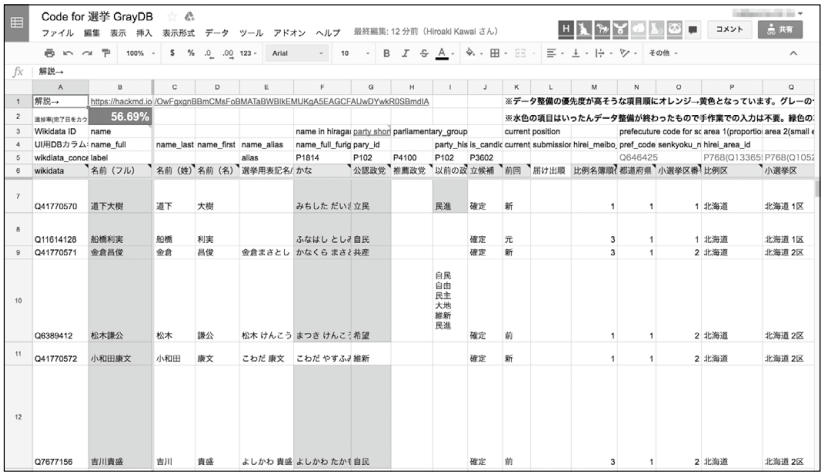
今回、Code for Japanを始めとする有志で、衆議院選挙候補者に関するオープンデータと、それを検索するフロントサイトを作成しました。https://kouhosha.info/にて公開しています(**図3**)。

このサイト自体は、郵便番号などから自分の住む地域の選挙区および候補者を探すことができるだけの単純なものですが、他のメディアのサイトと根本的に違うのは、我々は候補者のデータをオープンデータとして公開したという点です。同サイトで**図4**のように、候補者一覧および小選挙区のデータをCC0[注2]ライセンスで公開しています。

また、同じデータをWikidata[注3]という、オープンな構造化データベースにも登録しています。これによって、WikidataのビューアであるResonator[注4]を使ってビジュアライズできます。SPARQLクエリを投げることで、たとえば、候補者の男女比[注5]、各議員の職業のグラフ[注6]、出身校の割合[注7]などがグラフィカルに見られます。

注2　著作者が自身の著作物に対して、著作権法上の制約を一切課さないことを表明するための、クリエイティブ・コモンズが提供するツール。
https://creativecommons.jp/2015/05/01/cc0-jpver/
注3　https://www.wikidata.org/
注4　https://tools.wmflabs.org/reasonator/?q=Q661410&lang=ja
注5　http://tinyurl.com/ycnu7ltd
注6　http://tinyurl.com/y7r75csh
注7　http://tinyurl.com/y7sqypsk

▼図2　実際に使ったSpreadsheet




▼図3　kouhosha.info

▼図4　データのダウンロードページ

データ収集のプロジェクト自体オープンに組織されており、データの作成には、10名近くのボランティアエンジニア、10名以上のデータ入力ボランティアが協力してくれました。GitHubのリポジトリ注8で、ソースコードなども公開しています。

今回活用したツール類をまとめると、**図5**のようになります。

## Facebookの新機能にも使われた

このプロジェクトには英国のNPO、mySocietyが協力しており、彼ら経由で海外でも利用されています。

たとえば、Facebookが選挙直後に自分が住んでいる選挙区の政治家をフォローできる機能を公開しました(**図6**)。実はこの機能に使われているデータは、Code for 選挙プロジェクトで作ったものでした(この機能のデータ提供欄はmySocietyとなっていますが、彼らが使っているデータは、Code for 選挙プロジェクトで作ったデータです)。

これこそが、オープンデータの意義です。データと実装を切り離し、オープンなライセンスかつ世界標準のフォーマットで提供することで、他のサービスからでも利用可能になるのです。

また、VLED(一般社団法人オープン&ビッグデータ活用・地方創生推進機構)に、本取り組みを2017年度の優秀賞に選定いただきました。

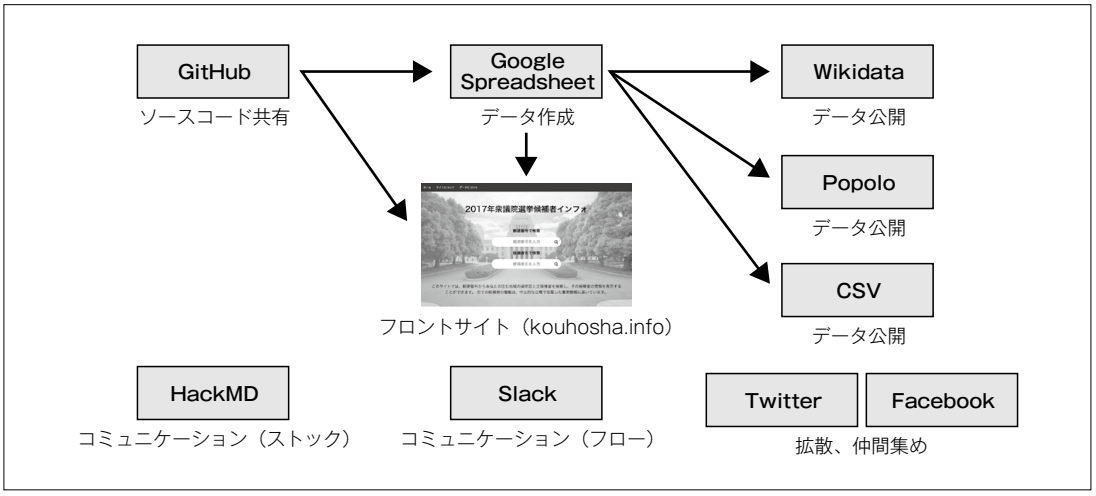
## 民主主義の根幹の情報をオープンデータに

今回データを作ってみて、ちゃんとしたデータを作るのはとにかくたいへんであることに気づきました。大手メディアは政党からExcel形式のデータを

注8 https://github.com/codeforjapan/codeforelection

▼図6 Facebookの機能

▼図5 活用システム概要



もらえるという話も聞きましたが、それにしてもデータの整形や確認作業などは発生しているはずです。それぞれのメディアが、事実情報という本来まったく差が出ない情報を収集するのに人手をかけているという無駄な状況には疑問を感じます。

国民の主権である選挙という重要な情報に対して、オープンなデータが存在しないというのもおかしな話です。選挙管理委員会にはぜひ候補者データの公開を検討してほしいものです。

また、選挙に関する情報だけではなく、議会議事録や予算といったデータも、生データの形でオープンになると良いと思っています。

今回プロジェクトを行うにあたって参考にしたのが、前述した英国のNPO、mySocietyが提供しているTheyWorkForYou[注9]というサイトです。このサイトでは、自分の住んでいるエリアで活動する議員が、いつどのような議会でどのような発言をして、どのような法案に賛成／反対をしているかを見ることができます(**図7**)。また、指定した候補者の活動をメールで知らせてくれるといった機能も備えています。

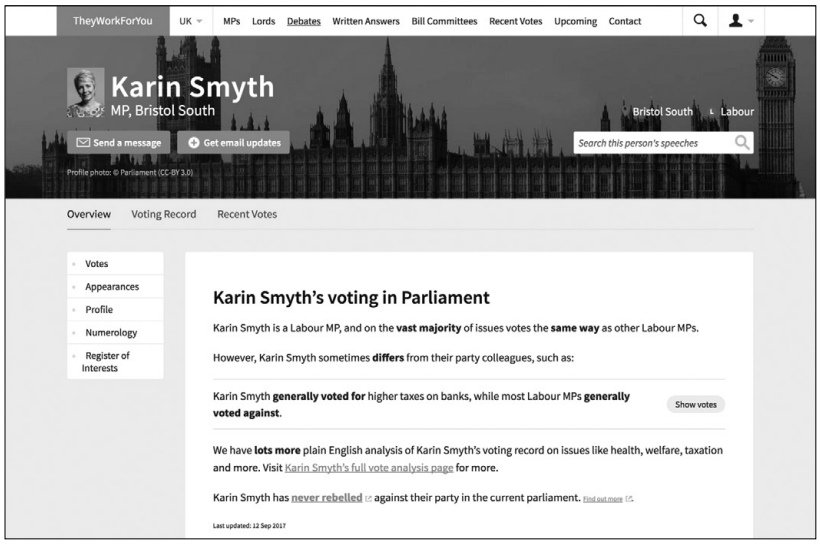
選挙のときには、マニフェストや演説など、未来のことを聞いたり見たりして投票を行う人が多いと思います。でも、その議員が今までどんな発言をし、どのような法案に賛成／反対をしてきたか、という客観的な情報はとても重要な判断材料になると思います(党議拘束があるので賛否情報に個人の意思が反映されにくい、という側面もあるので国会より地方議員のほうがおもしろそうですが)。

今後、Code for 選挙プロジェクトチームとしては、上記のようなEデモクラシー関連の活動を引き続き行っていく予定です。具体的には、選挙管理委員会との意見交換や、地方議員版でのデータ収集、議事録情報との紐付け、メディアとの連携などを通じて、データ公開の機運を高めていきたいと考えています。

最後になりますが、このたびデータ作りやシステム開発に協力いただいたメンバーのみなさん、本当にありがとうございました。みなさんの協力なしにはとてもこのような短時間でのデータ公開には至らなかったのではないかと思います。この場を借りて感謝を申し上げます。

Code for 選挙の活動に興味のある方は、Code for JapanのSlack[注10]に参加し、#codeforelectionチャンネルに参加いただければと思います。SD

注9　https://www.theyworkforyou.com/

注10　https://cfjslackin.herokuapp.com/

▼図7　候補者のページ