

「伝える」技術

LOD（ウェブ上でデータを公開して共有するための方法）

Linked Open Dataの略で、言葉通りに「公開されたデータ（Open Data）をつなぎ合わせたもの」です。今、各国の公的機関・企業・個人がデータをOpen Dataとして公開し始めています。そのOpen Dataをデータベースのように活用できる技術です。

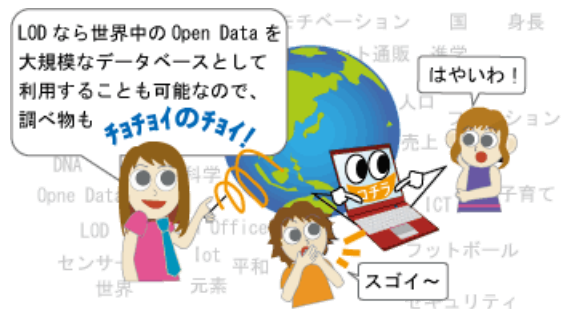
ご利用にあたっての注意

この講座は、2015年当時の情報です。予告なしに更新、あるいは掲載を終了することがあります。あらかじめご了承ください。

最終更新日 2015年8月5日

もくじ

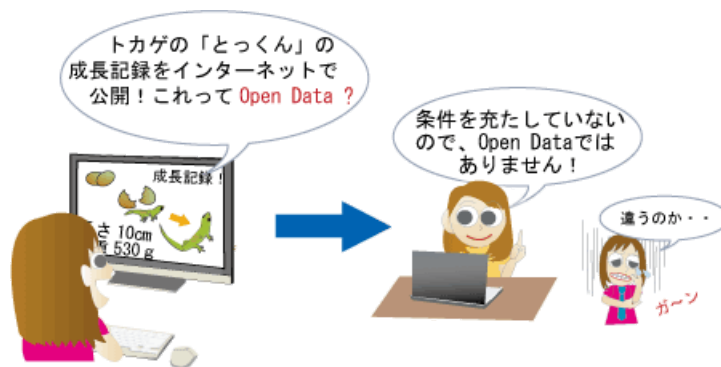
- ↓ Open Dataって？
- ↓ Open Dataにするメリット
- ↓ LODって？
- ↓ LODにする方法
- ↓ 富士通のLOD技術
- ↓ 小話
- ↓ 関連ページへのリンク



Open Dataについて学ぼう

webで公開すると、すべてOpen Dataなの？

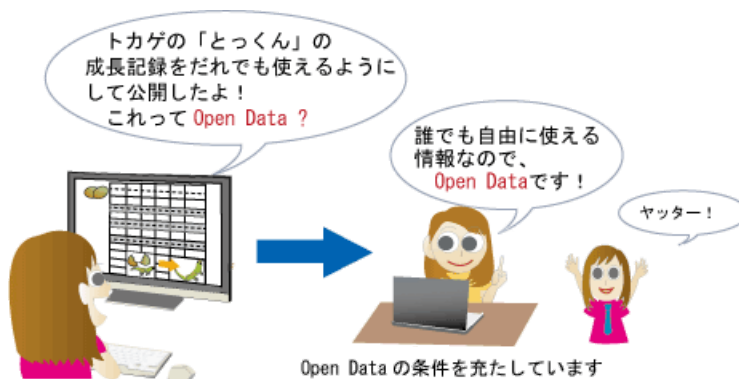
webで公開されれば、インターネットにつながる全世界のパソコンでそのデータを見ることができるので、公開データになります。その内、ある条件を満たすものが「Open Data」と呼ばれます。



Open Dataの条件

条件とは、「誰もがそのデータを自由に使えること、およびそのデータを使って得られた成果を自由に配布できること」です。ただし「作者の氏名や作品名を残す」、あるいは「同じ条件で配布する」程度の条件を付けることはできます。

現在webで公開されている形式には、いくつか種類があります。その中でOpen Dataの多くは、例えば**エクセル (xls)** や**ワード (doc)**、**テキスト (csv)** などのような形式で公開されています。



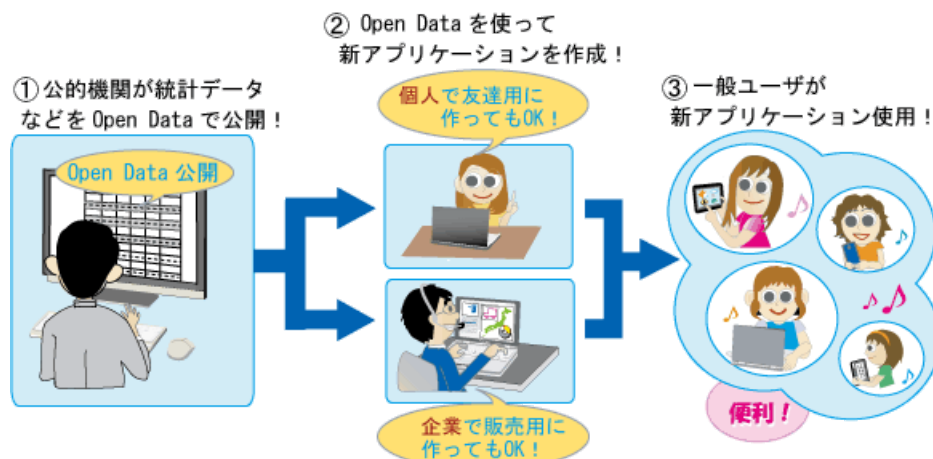
(Open Data のイメージ図)



Open Dataにするメリット

Open Dataの利用方法

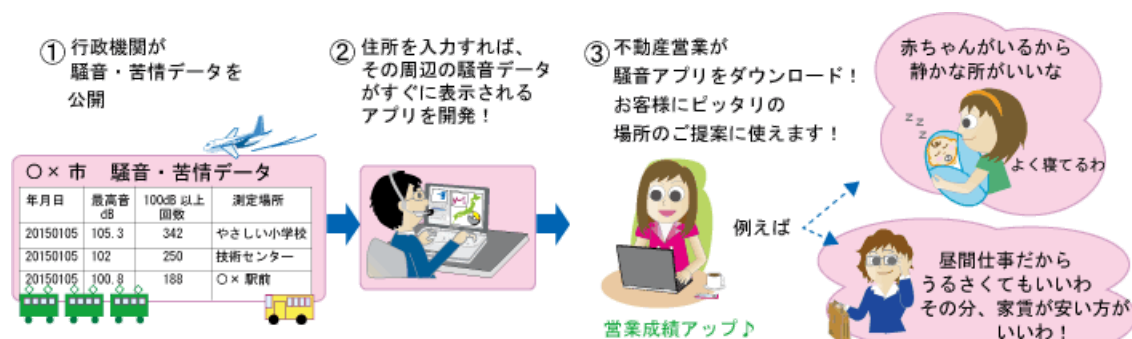
例えば、公的機関で持っているデータをOpen Dataとして公開したら、プログラム開発者（企業でも個人でもOK）がOpen Dataを使って、アプリケーションを開発することができます。そして一般ユーザは、その開発されたアプリケーションを利用することができます。



公的機関で公開しているOpen Dataってどんなデータがあるのかな

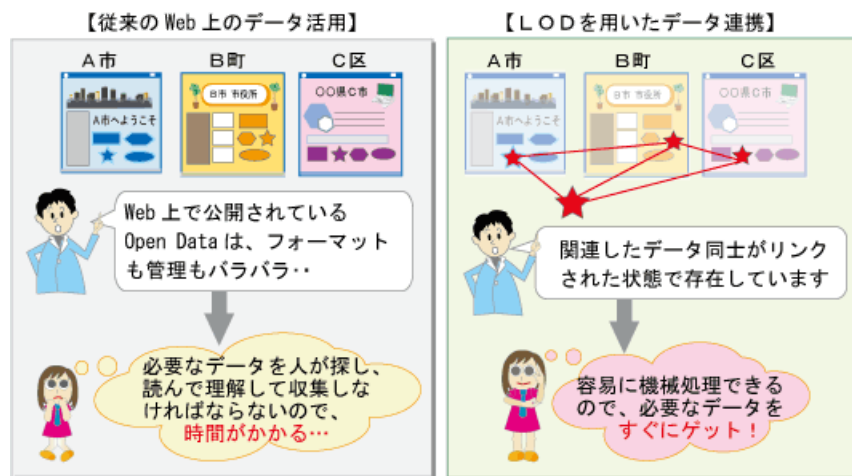
- ゴミの分別と出し方・地域別収集日（CSVファイル）
- 年齢別推計人口（Zip形式）
- 市内の駅別一日当たりの平均乗車人員数（CSVファイル）
- 航空機騒音・苦情統計（Zipファイル）
- 公共施設の所在地（図書館・こどもセンター・公園・プール・駐車場など）（CSVファイル）

（Open Dataとして公開されている「航空機騒音・苦情統計」を使った例）



LODって？

Linked Open Data（リンクト・オープン・データ）の略で、LOD（エルオーディー）と読み、ウェブ上でデータを公開して共有するための方法です。Open Dataとして公開されているデータ同士を結び付けて（リンクして）、誰でも自由に利用できるように公開されている（オープンライセンス）ものです。（データ同士を結び付けること＝データをリンクする、linked data）

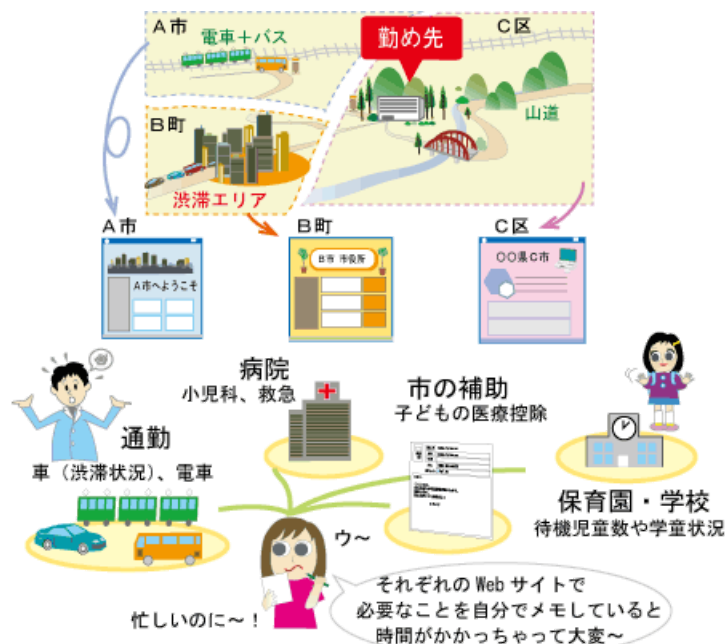


具体的にどんなことができるようになるの? ～引っ越し大作戦～

父親の転勤で引っ越すことになりました。インターネットを使って、家族の希望にあう場所を検索することにしました。

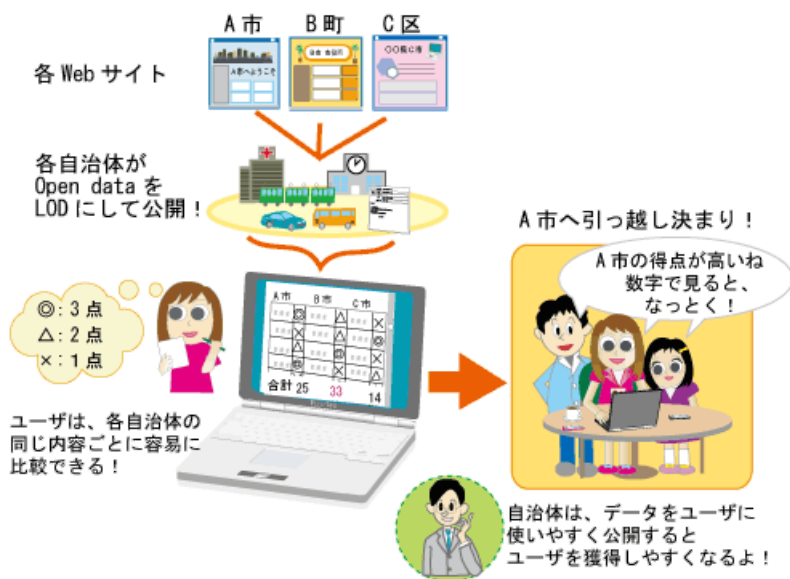
Web情報を使って情報収集した場合

父親の勤務地に近いA市、B町、C区について調べてみると、各自治体の特長がわかるOpen Dataを比較するには、各Webサイトから必要な情報を自分で整理しなければならず大変です。



LODを使って情報を収集した場合

各自治体がOpen DataをLODで公開した場合、ユーザはリンク先を参照することでデータを簡単に得ることができます。サイトをまたがっていても、同じ内容同士はリンクされているので、容易に比較することができます。また、自治体にとっても、ユーザに使いやすいデータを公開していることで、ユーザの信頼を得やすくなります。



LODにする方法

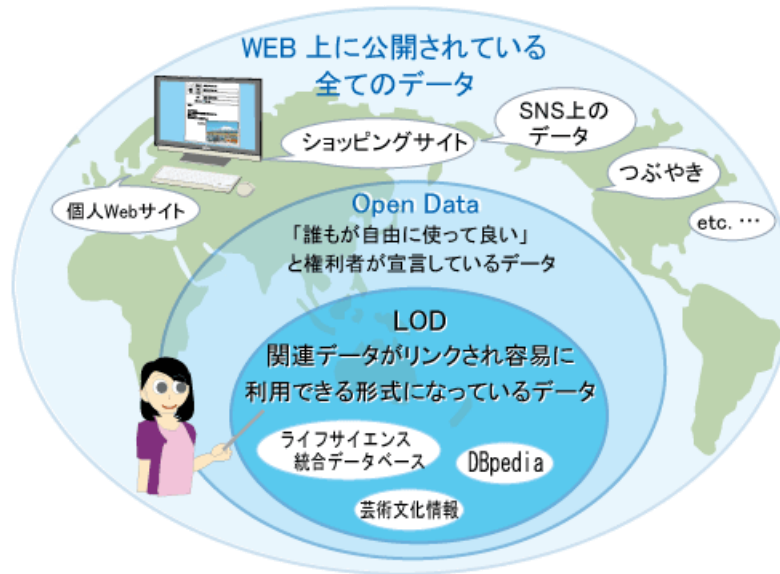
現在の公開方式「HTML」を使って、Open Data同士を結び付けたらLODになるかというところ、そうではありません。通常、私たちが目に見えているWeb上の多くは、作者がどのように「人」に見せたいか、見て欲しいかを考えて、「HTML」等で書いています。しかし、LODは「人」のためではなく、「機械同士」が理解できる形式で書く必要があり、4つの条件があります。

LODの条件

データがリンクされ、容易に利用できる形式になっているデータであればLODになります。

1. すべての事や物に名前（URI）をつける
2. 事や物をすべて参照できるように、HTTP URIを使用すること
3. URIを参照したときに、有用な情報を標準的なフォーマット（RDFなど）で提供できるようにすること
4. さらに多くの事や物を発見できるように、他のURIへのリンクを含むこと

(LODのイメージ図)



LODの条件をみたすのにはどうしたら良いの？

LODの条件をみたせる代表的な書き方（形式）として**RDF形式**があります。RDF形式は機械同士が理解できる形式で、リンクを付けることも可能なので、LODとして情報を公開するのに用いられます。（この講座でもRDF形式で書いたLODを紹介いたします）

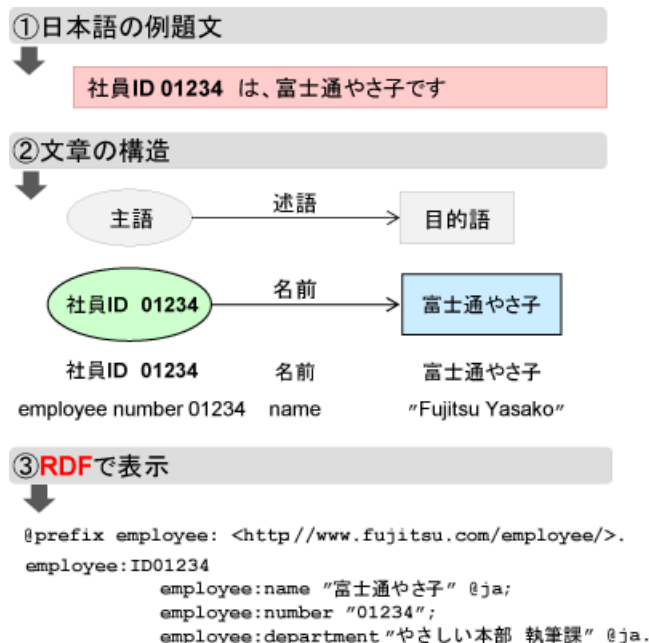
RDFは、Resource Description Framework（リソース・ディスクリプション・フレームワーク）の略です。

RDFはどうやって書くの？

文章を、主語（サブジェクト）、述語（プレディケート）、目的語（オブジェクト）の構造にわけます。

例えば「社員ID01234は富士通やさ子です」という文章があったとします。

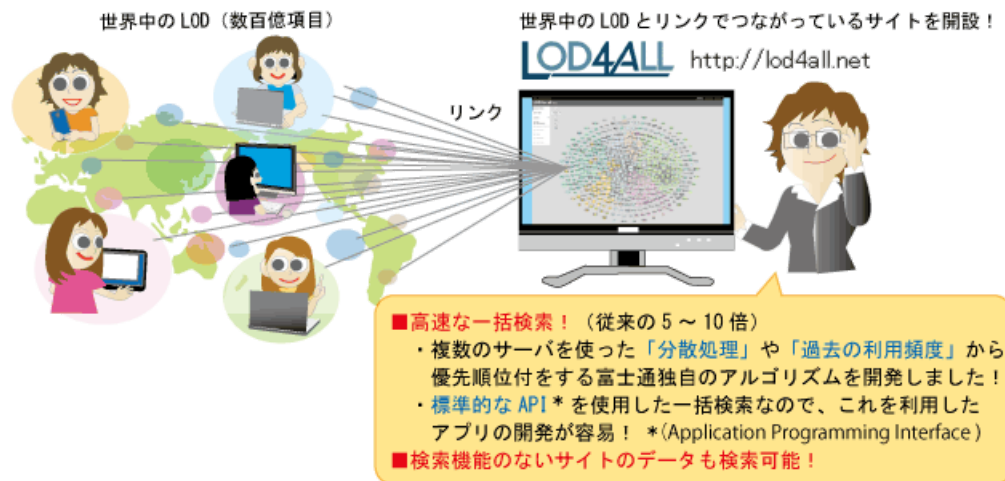
この文章の場合、主語は「ID01234」、述語は「名前」、目的語は「富士通やさ子」となります。



富士通のLOD技術

Open Dataの活用革新! LODの大規模データ格納・検索技術

世界中で公開されている数百億個のLODを集め、大規模なデータベースとして一括検索を可能にする技術を開発しました。この技術を盛り込んだサイト「LOD4ALL（エルオーディーフォーオール）」を無償公開しています。ユーザーはこのサイトを利用することにより、世界規模で必要なデータを即時に見つけることができます。

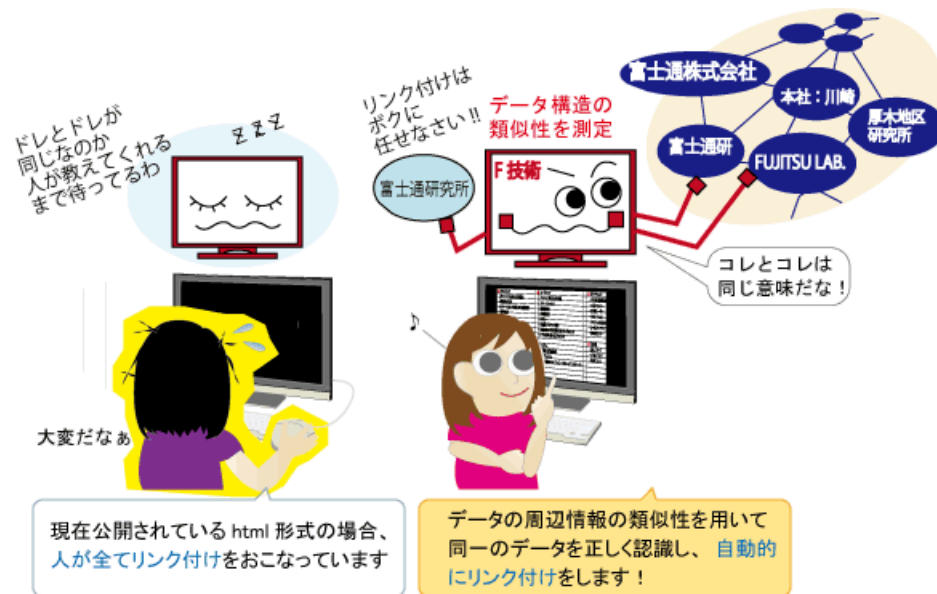


Open Dataへのリンクを自動的に付与する技術

これまでのLODは、同じ意味をもつ名前であっても異なる単語で登録されていると、機械は「同じである」と認識できないため、人間の目で確認してリンク付けする必要がありました。

富士通が開発した技術では、世界中で公開されているLODの中から、「同じ対象を表すデータ」を正確に、自動的にリンク付けすることができます。

データの周辺情報から「同じ対象を表すデータ」の類似性を測定し、機械学習手法を使って、同一性を判定します。



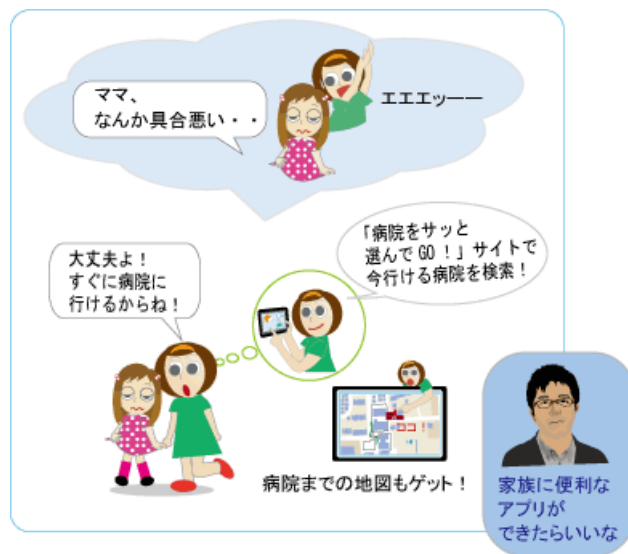
小話 ～こんなサイトがあったらいいな編～

案1「病院をサッと選んでGO！」サイト

家族が急な発熱や発疹など病気にかかってしまった時、各病院のWebサイトでその時の症状に合う診療科、病院の営業日、営業時間、予約制か否か、駐車場はあるかどうかなどを自分で調べてから病院に行きます。もし、LODサイトで「病院をサッと選んでGO！」なんていうものがあったら便利だな、と思います。

現在のインターネット検索でも、例えば「厚木市、内科」と入力すれば複数の病院がヒットしますね。その後は、ヒットした病院を自分で1件ずつ詳細を調べます。LODの場合は、更に一歩踏み込んでスマートデバイスに内蔵された曜日や時間を勝手に利用して、ユーザーがその時受診できる病院を表示し、各病院が提供する混雑状況（何分待ち等）も一覧で見ることができたら、便利ですよね（病院によっては、Webサイトで予約や現在の診察番号が表示される病院がありますので、それらをLODでつながれるといいかも）。

また、救急隊員が使えるような信頼度が高いサイトになったら、もっといいかも！現在、救急車の中に乗り込んでも隊員が受け入れ可能な病院を探すので、時間がかかりますね。それも解消されるといいな。



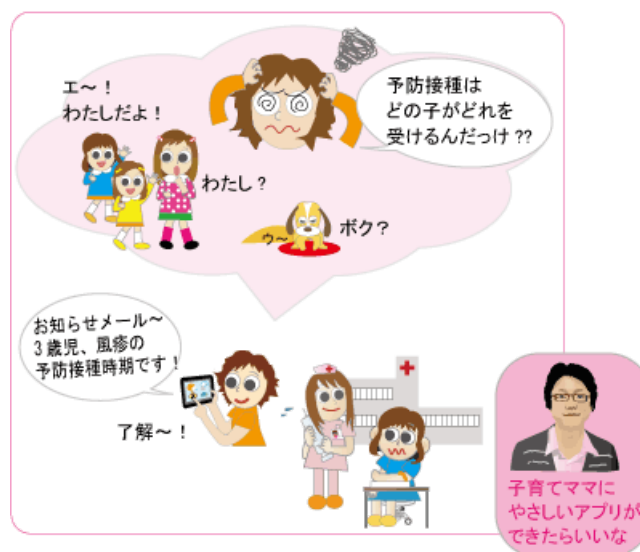
案2「予防注射お知らせメール！」サイト

子供は小学校に入るまでに色々な予防接種を受けています。現在、公的機関の保健委員がお母さんひとりずつと話し、どの予防接種をいつしたら良いか、それぞれで日程をたてます。もし子供が2人や3人いる場合、お母さんはどの予防接種をどの子が受けているか、受けていないか、混乱してしまうので母子手帳で確認しなければいけません。うっかり忘れていた予防接種もあるかもしれません。

もし、予防接種の間隔や順番などがOpen Dataになっていて、それを使ってスケジュールを立てて記録してくれるアプリがあると便利かもしれません。

さらに、自分の子どもの名前と生年月日を入力しておけば、その子供に適切な予防接種の時期にアプリからメールでお母さんに「お知らせ」が届くと、うっかりすることもなくなり安心ですね。

また現在は、以前住んでいた市の記録が引っ越し先の市に引き継がれていないこともあるので、お母さんが子どもの接種の有無を調べるしかありません。この場合も予防接種アプリでスケジュール管理していれば問題ありません。また、各自治体の病院に関するデータもLODでつなげて、予防接種と接種可能な病院名とその地図なども同時に表示されると大変助かりますね。



関連リンク

プレスリリース

- ＞ 川崎市で子育て支援アプリの実証実験を開始（2015年1月5日）
- ＞ Linked Open Data技術を適用した地域の特性を発見するツールを公開（2014年12月11日）
- ＞ 業界初！自治体の特性が見える化する評価ツールを開発（2014年2月28日）
- ＞ 世界中で公開されているオープンデータへのリンクを自動的に付与する技術を開発（2014年1月16日）
- ＞ オープンデータの活用革新! リンクが張られた公開データ（LOD: Linked Open Data）向け大規模データ格納・検索技術を開発（2013年4月3日）

その他

- ＞ 富士通の研究開発