

8/27(火) Nagatacho GRiD

人工知能学会SWO研究会ワークショップ

「ナレッジグラフ推論チャレンジ2019技術勉強会」

# ナレッジグラフ入門

大阪電気通信大学

/セマンティックウェブとオントロジー(SWO)研究会

古崎 晃司

## ▶ 講演のねらい

- ナレッジグラフ(知識グラフ)の基礎技術について、Linked Data / Linked Open Data(LOD)を中心に解説する。
- LODの利用方法の基礎をハンズオンを通して学ぶ。

## ▶ 講演内容

### ➤ ナレッジグラフの基礎知識

- ✓ ナレッジグラフ(Knowledge Graph / 知識グラフ)
- ✓ Linked Data / Linked Open Data(LOD) / RDF
- ✓ ナレッジグラフ(RDF)を利用するための基礎技術

### ➤ SPARQLクエリによる検索演習【ハンズオン】 ～Wikidataを例として～

# ナレッジグラフ (Knowledge Graph)

- ・ さまざまな「知識」の関係(つながり)をグラフ構造で表したもの.
- ・ 知的システム開発の基盤となるデータベース(知識ベース)として用いられる.

# Google Knowledge Graph



- Knowledge Graph (知識グラフ)と呼ばれる、知識の“つながり”を活用した様々なサービスの開発が、Google, Yahoo!, Facebookなど多くのネット企業で進められている。  
→Google Knowledge Graphは2012年サービス開始

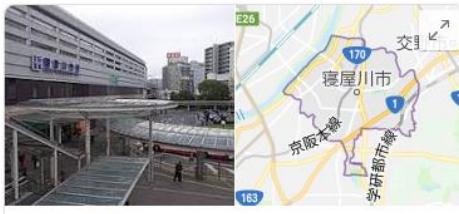
<https://www.youtube.com/watch?v=mmQI6VGvX-c>

# Google Knowledge Graphの検索例

分類

属性

関連事項



寝屋川市

わやがわし  
日本の市

寝屋川市は、大阪府東部大阪地域に位置する人口約23万人の市。市名は、御伽草子の「鉢かづき」に登場する主人公・初瀬姫の父である藤原実高の別荘が、高野街道を行く旅人に宿を提供して「寝屋」と呼ばれるようになったことに由来する。2019年4月1日に施行時特例市から中核市に移行した。ウィキペディア

現在の天気： 温度: 22°C、風向: 北東、風速: 0 m/s、湿度: 92%

総人口： 231,407人；（推計人口、2019年4月1日）

人口密度： 9,369人/km<sup>2</sup>

市町村長： 広瀬慶輔 (2019年)

人口： 23.75万 (2015年) 国際連合

大学



筑南大学



大阪電気通信  
信大

他の人はこちらも検索



茨木市



枚方市



熊本市



福井市

旅行の計画を立てる、概要

フィードバック

Google検索結果  
のナレッジパネル  
の表示例

君の名は。 / キャスト



悠木碧  
名取早耶香



上白石萌音  
Mitsuha Miyamizu



花澤香菜  
Teacher



神木 隆之介  
立花謙



成田凌  
黒使河原克彦



長澤まさみ  
奥寺ミキ



石川界人  
高木真太



島崎信長  
藤井司

>

谷花音

神木 隆之介

[https://ja.wikipedia.org/wiki/%E7%A5%BD%E6%9C%8B\\_%E9%80%9A](https://ja.wikipedia.org/wiki/%E7%A5%BD%E6%9C%8B_%E9%80%9A)  
部分ブロックに登場する人物  
『君の名は。』の監督  
事務所：アミューズ  
職業：俳優、声優  
ハブルの動く城



その他の画像

トップニュー



「君の名は。リフを神木隆之介と白石萌音本人が歌った『入れ替わる』

YouTubeで予告編を再生

2016年・ドラマ/ファンタジー・1時間 52分

97% のユーザーがこの映画を高評価しました

Google ユーザー

『君の名は。』は、2016年に公開された新海誠監督による日本の長編アニメーション映画である。

ウィキペディア

公開日： 2016年8月26日 (日本)

監督： 新海誠

製作国： 日本

挿入歌： 前前前世

興行収入： 3.58億アメリカ合衆国ドル

キャスト



悠木碧  
名取早耶香



上白石萌音  
Mitsuha Miyamizu



花澤香菜  
Teacher



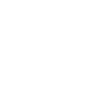
神木 隆之介  
立花謙



長澤まさみ  
奥寺ミキ



石川界人  
高木真太



島崎信長  
藤井司

<

神木 隆之介

かみいりゅうのすけ  
俳優

神木 隆之介は、日本の男性俳優、声優。埼玉県出身。  
アミューズ所属。 ウィキペディア

生年月日： 1993年5月19日 (年齢 26歳)

生まれ： 埼玉県

身長： 167 cm

学歴： 堀越高等学校 (2009年 - 2012年)

血液型： B型

TV/映画



SPECへ  
視察公安部  
六小路

## ■ 知識獲得

- どのようにして「知識」を集めるか？
  - Webを知識源とした知識収集  
例) Webマイニング, セマンティックWeb

## ■ 知識表現

- あつめた「知識」をどのようにコンピュータ上で表現するか？
  - グラフ構造 = 意味ネットワーク

## ■ ナレッジグラフの公開

- Web-API
- **Linked Data, Linked Open Data (LOD)**

Webの仕組みを用いた  
ナレッジグラフの公開方法

# Google Knowledge Graphの場合



## ■ 情報源

- Freebase, Wikipediaなどのオープンな情報源
- Googleのサービスから得られる様々なデータ

## ■ データの規模(2012時点)

- 5億(500 million)以上のobjects
- 35億(3.5 billion)以上のfacts/relationships

## ■ 想定している利用目的

- 1. Find the right thing
- 2. Get the best summary
- 3. Go deeper and broader

## Introducing the Knowledge Graph: things, not strings

May 16, 2012, <https://googleblog.blogspot.jp/2012/05/introducing-knowledge-graph-things-not.html> より

# **Linked Data (LD)**

## **Linked Data**

=Webの仕組みを用いて相互に“リンクされた”データ

# Linked Data (2007頃-)

- **Linked Data**: Web上のデータを, つなぐ(linkする)ことで, 新しい価値を生み出そうとする取り組み. Webの創始者Tim Berners-Lee氏が提唱
- ※ **Linked Open Data(LOD)**: オープンな形で公開されたLinked Data

Linked Data - Connect Distributed Data across the Web <http://linkeddata.org/>

Home  
Guides and Tutorials  
Frequently Asked Questions  
Glossary  
Images and Posters  
Presentations  
Data Sets  
Tools  
Events  
Calls for Papers  
Research

News and Blogs  
See All

Linked Data

Linked Data is about using the Web to connect related data that wasn't previously linked, or using the Web to lower the barriers to linking data currently linked using other methods. More specifically, Wikipedia defines Linked Data as "a term used to describe a recommended best practice for exposing, sharing, and connecting pieces of [data](#), [information](#), and [knowledge](#) on the Semantic Web using [URIs](#) and [RDF](#)."

This site exists to provide a home for, or pointers to, resources from across the Linked Data community.



Login to edit  
(authorised users only at present)

Username: \*

Password: \*

Log in  
 Log in using OpenID  
 Request new password

Navigation  
Search  
Sitemap  
Feedback  
Help  
About  
Syndicate  
[RSS](#)

**“Webの仕組み”に基づいてデータを公開することで、  
Web上に公開された膨大なデータを  
統合した1つの知識ベースとして利用可能にする。**

# Webの仕組み



- URLを指定することで、Webページにアクセス
  - 例) <http://www.osakac.ac.jp> 「大阪電気通信大学」のページ
  - URLは、世界中“すべて”的Webページの場所(ID)を一意に特定できる仕組み
- ハイパーリンクにより、Webページを“つなげる”
  - リンク先のURLを指定することで、好きなWebページと自由に“リンク”できる
    - リンクを辿って、様々な情報にたどり着ける
    - リンクを解析による様々なビジネス
      - 例) Googleなどの検索エンジン

# Webの仕組み→Linked Data



- URLを指定することで, Webページにアクセス

**URI·IRI**

//www.osakac.ac.jp 「大阪電気通信大学」の

データ (モノ・コト)

- URLは, 世界中“すべて”的 Webページの場所(ID)を一意に特定できる仕組み
- ハイパーリンクにより, Webページを“つなげる”
  - リンク先のURLを指定することで, 好きな Webページと自由に“リンク”できる

**Linked Data**

Webと同じ仕組みでデータを“公開”し,  
相互に“つなぐ”(リンクする)

**重要**

# データを「つなげる」仕組み

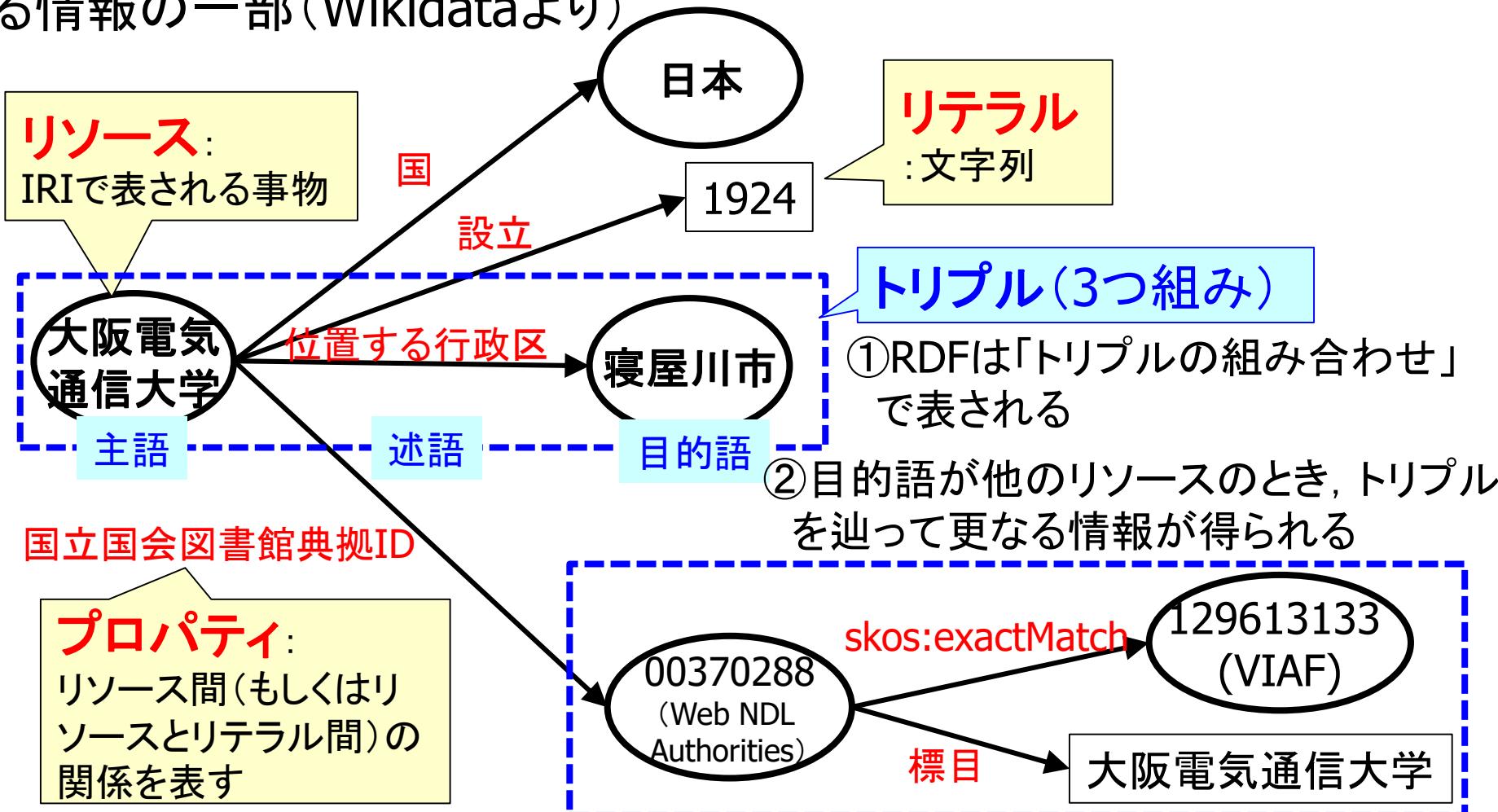


- 「3つ組(トリプル)」により様々な構造のデータの“つながりを柔軟に表現”できる
  - Webサイトのリンクを辿ると同様に, プロパティ(リンク)を辿ることで関連するデータの情報を辿ることが出来る.
  - RDFのプロパティは, “関係の意味”を定義できる.
    - cf. Webのハイパーリンクは単に“つながり”を表すのみ
- データ(リソース)をIRIで表すことで, “外部のデータとつなげる”ことができる.
  - URIは, グローバルに一意のIDを表す.
    - WebサイトのURLと同じ仕組み.
    - cf. 単なる数字をIDとすると, 異なるDBが同じIDを使っている可能性がある

- Linked Data = 外部のデータとつながったデータ  
≠ RDFフォーマットのデータ

# Linked Data (RDF) の表現例

<http://www.wikidata.org/entity/Q7105556> というIRIから得られる情報の一部(Wikidataより)



# Linked Data (RDF) の表現例

すべてのリソースおよび  
プロパティはIRI用いて  
表される

<https://www.wikidata.org/entry/P17>

1924

<https://www.wikidata.org/entry/P571>

<http://www.wikidata.org/entity/Q7105556>

<https://www.wikidata.org/entry/P131>

<http://www.wikidata.org/entity/Q389633>

<https://www.wikidata.org/entry/P349>

<http://viaf.org/viaf/129613133>

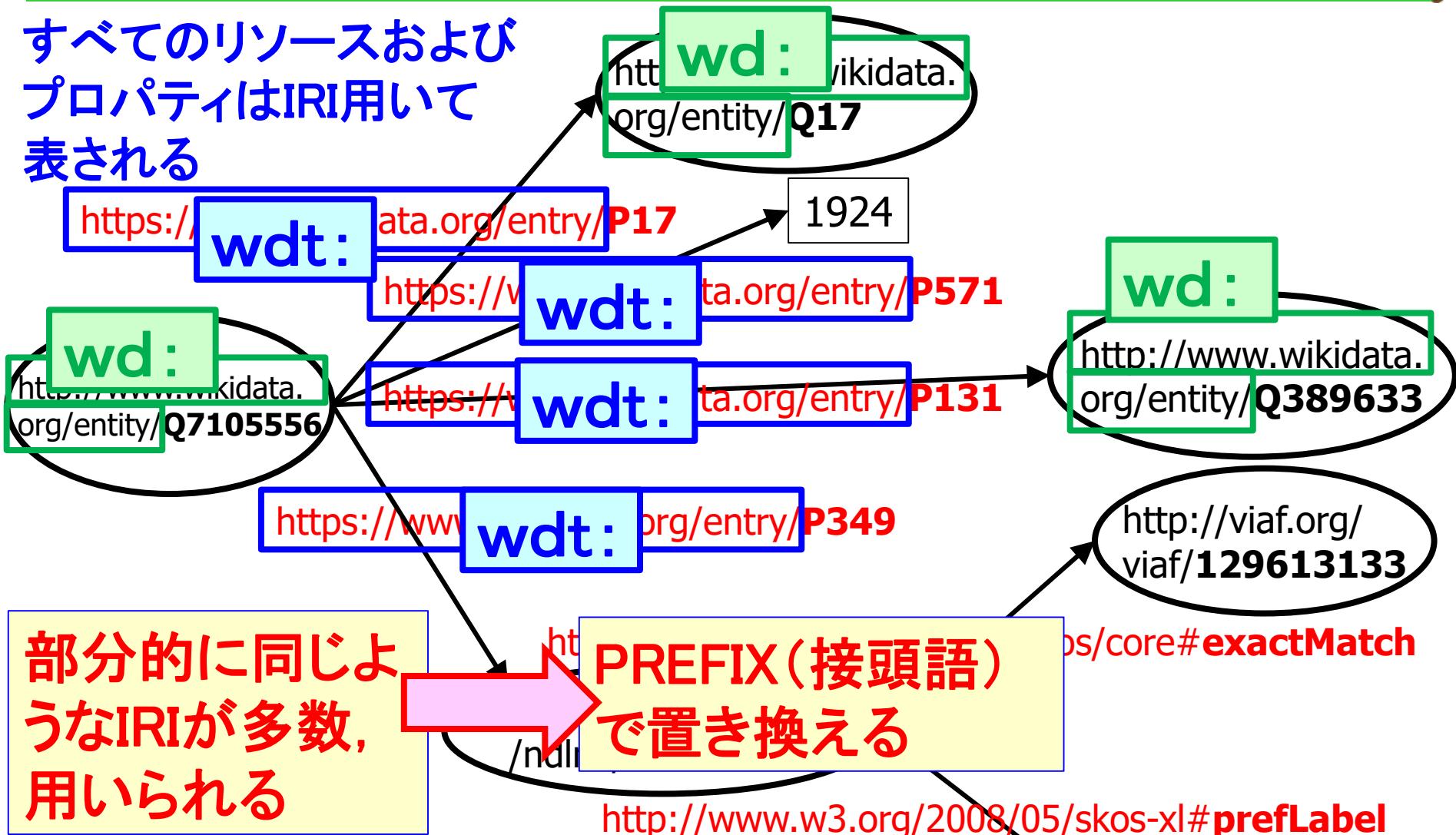
<http://id.ndl.go.jp/auth/ndlina/00296951>

<http://www.w3.org/2008/05/skos-xl#prefLabel>

大阪電気通信大学

# Linked Data (RDF) の表現例

すべてのリソースおよび  
プロパティはIRI用いて  
表される

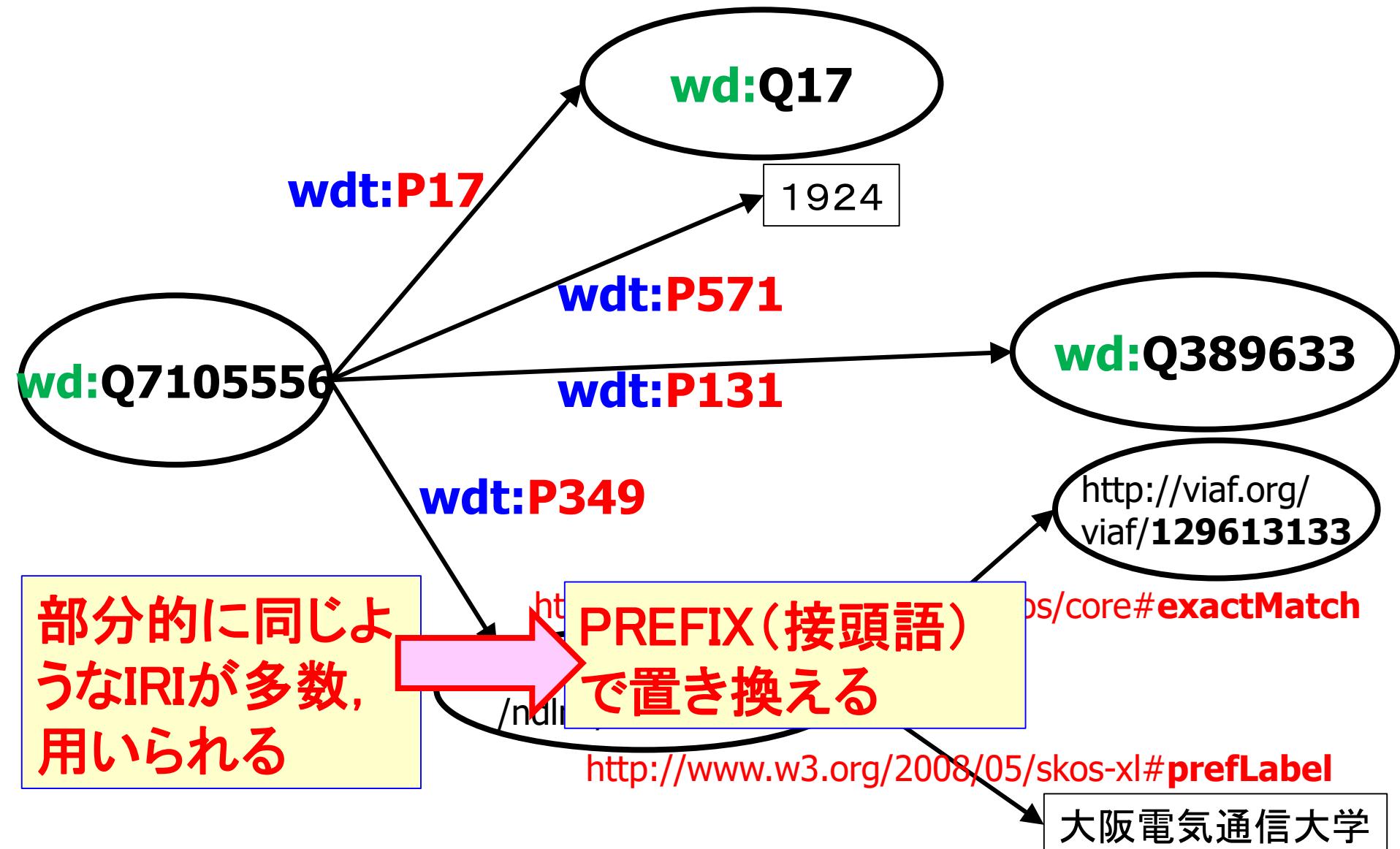


部分的に同じようなIRIが多数、用いられる

## PREFIX(接頭語) で置き換える

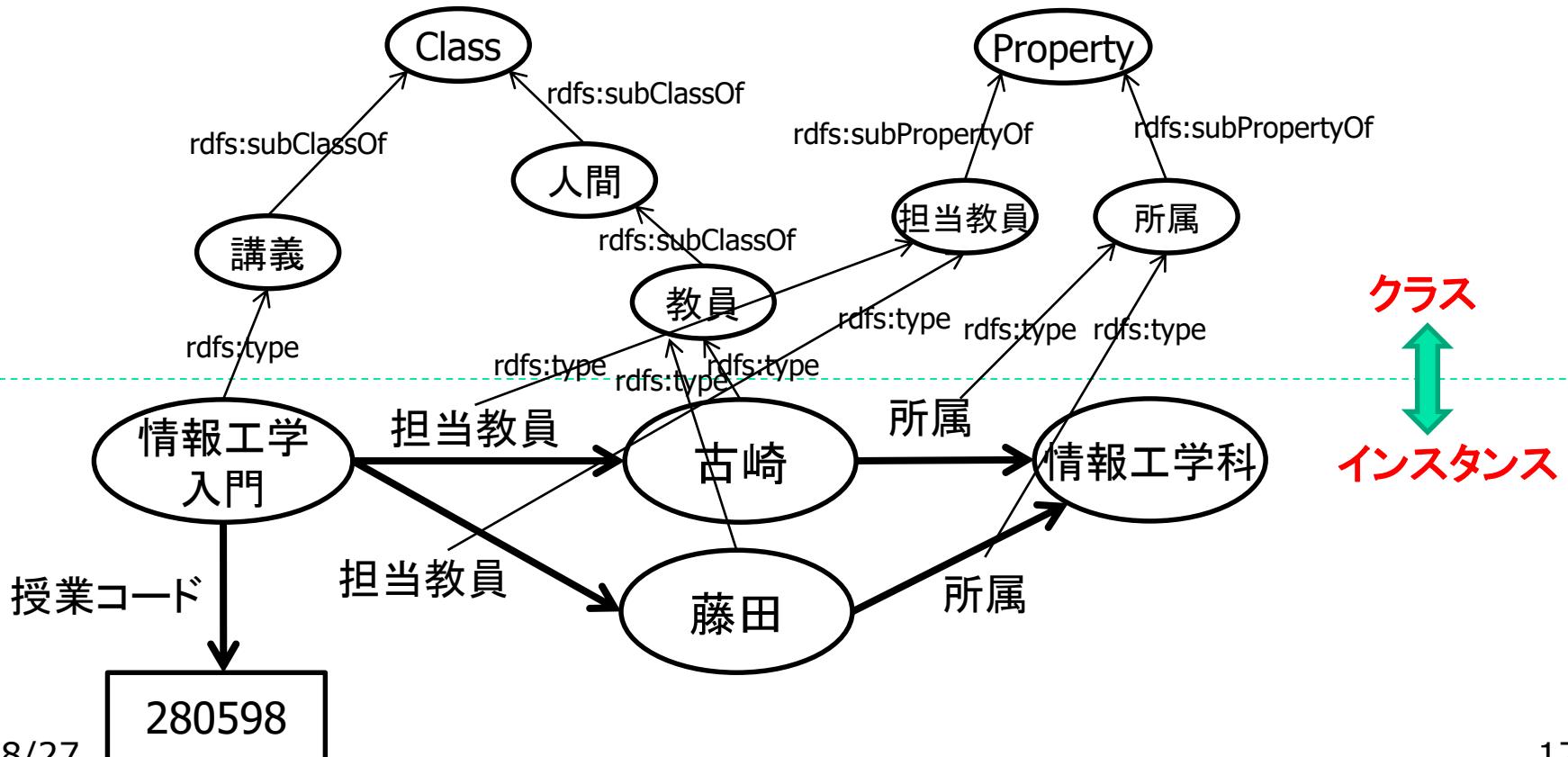
<http://www.w3.org/2008/05/skos-xl#prefLabel>

# PREFIX : IRIの省略表現



# Linked Dataとオントロジー

- RDFで表現されたデータの意味を明確にし、Web上で共有するには、RDF記述に用いる「語彙」を統一する必要がある。  
→用いる「語彙」をオントロジーとして定義する。
- Linked Data**におけるオントロジーの役割
  - データ記述に用いる「語彙」(=リソースとプロパティの種類)を規定する。



# RDFSとOWL

- セマンティックウェブにおいてオントロジー記述に用いられる言語
- **RDFS**(RDF Schema)
  - クラスおよびプロパティの**is-a**階層
  - プロパティの定義域(domain)/値域(range) : 主語/目的語となれるリソースの種類
- **OWL**(Web Ontology Language)
  - あるクラスの持つプロパティの値域に関する制約  
→述語論理における全称記号(▽), 存在記号(∃)に相当する表現
  - あるクラスが持つプロパティの数に関する制約
  - プロパティの推移律(例:「子孫の子孫」は「子孫」)
  - . . . など, より詳細なオントロジーの記述が可能

# Linked Dataの公開方法



- 参照解決可能なhttp IRIs(URL,URI)を用いた公開
  - IRIでデータにアクセスが可能
    - 通常のWebページと同様に、データのURIを用いて「つながり」を辿ることが出来る  
=システムによる処理(リンク解析等)が可能
- SPARQLエンドポイントの公開
  - RDF用のクエリ言語SPARQLにより検索可能なAPIを公開
    - クエリによるデータ検索・抽出が可能
- RDFファイルのダンプの公開
  - 全データをダウンロードできる形で公開
    - ダウロードしたファイルをRDFパーサー, RDF-DBなどのツールを用いて処理可能

# **Linked Open Data (LOD)**

## **Linked Open Data**

=**Linked Data + Open Data(オープンデータ)**

=**Linked Dataとして公開されたOpen Data**

# オープンデータとは



## ■ オープンデータとは

- 誰でも自由に使える形で公開されているデータ

## ■ オープンデータの定義 (Open Definition)

- "*Open data and content can be **freely used, modified, and shared** by **anyone** for **any purpose**"*  
(<http://opendefinition.org/>)

## ■ オープンデータでない例

- 改変や再配布が禁止されている
- 利用者を限定 例) 学術機関のみ, 個人利用不可
- 利用目的を限定 例) 商用利用不可, コンテスト応募目的のみ

# 5★オープンデータ

<http://5stardata.info/> より

オープン  
ライセンス  
(形式問わず)

機械可読な  
フォーマット

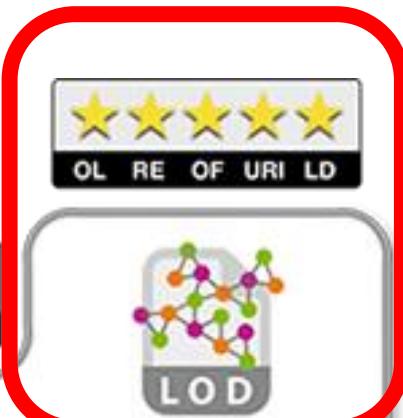
オープンな  
フォーマット



XLS



RDF



# 5★オープンデータにおけるLOD



- ★★★★ (RDF) <http://5stardata.info/ja/> より引用  
物事を示すのに**URL(IRI)**を使いましょう、そうすることで  
他の人々があなたのデータにリンクすることができます
- ★★★★★ (LOD)  
あなたのデータのコンテキストを提供するために  
他のデータへリンクしましょう

→**LOD(Linked Open Data)**  
=Web上で相互に**リンクされたOpen Data**

※ リンクする際には「**Webの仕組み**」を利用する

- データ(物事)を示すのに**URL**(正確にはIRI)を用いる
- データ間を(名前付き)**Hyper-link**でリンクする

# TED Talk by Tim Berners Lee (1) 2009/2 ～Raw Data Now!～



[http://www.ted.com/talks/tim\\_berners\\_lee\\_on\\_the\\_next\\_web.html](http://www.ted.com/talks/tim_berners_lee_on_the_next_web.html)



Speakers	TED Conferences	TED Conversations	About TED
Playlists <small>NEW</small>	TEDx Events	TED Community	TED Blog
Translations	TED Prize	TED-Ed <small>NEW</small>	TED Initiatives

Search

New TED Talks are released every weekday. Be the first to know!

Enter your email for TED updates

Daily  Weekly

Subscribe

## TALKS

### ティム・バーナーズ=リーが示す次のウェブ

FILMED FEB 2009 • POSTED MAR 2009 • TED2009



会場:「データを」

データを抱え込むのでは無く、誰もが使えるように(オープンに)することを呼びかける  
**=Raw Data Now !**  
(生データをすぐに！)

789,259 Views

94

20年前、ティム・バーナーズ=リーは、World Wide Webを発明しました。彼の次のプロジェクトは、ウェブが文字や画像、動画を対象に行ったように、私たちのデータを解放し、データの相互利用方法の再構成を意図とした、オープンでリンクするデータの為の新しいウェブの構築です。

Tim Berners-Lee invented the World Wide Web. He leads the World Wide Web Consortium, overseeing the Web's standards and development. [Full bio »](#)

Translated into Japanese by [Yuki Okada](#)

Reviewed by [Yasushi Aoki](#)

Comments? Please email the translators above.

[More talks translated into Japanese »](#)



# TED Talk by Tim Berners Lee (1) 2010/2 ～LODの活用事例～



Talks TED Conferences TED Conversations About TED  
Speakers TEDx Events TED Community TED Blog  
Playlists NEW! TED Prize TED-Ed NEW! TED Initiatives

[http://www.ted.com/talks/tim\\_berniers\\_lee\\_the\\_year\\_open\\_data\\_went\\_worldwide.html](http://www.ted.com/talks/tim_berniers_lee_the_year_open_data_went_worldwide.html)

New TED Talks are released every weekday. Be the first to know!

Enter your email for TED updates

Daily  Weekly Subscribe

Follow TED [f](#) [t](#) [r](#)

## TALKS | IN LESS THAN 6 MINUTES

### ティム・バーナーズ=リー「オープンデータとマッシュアップで変わる世界」

FILMED FEB 2010 • POSTED MAR 2010 • TED2010



494,797 Views

[Like](#) 233

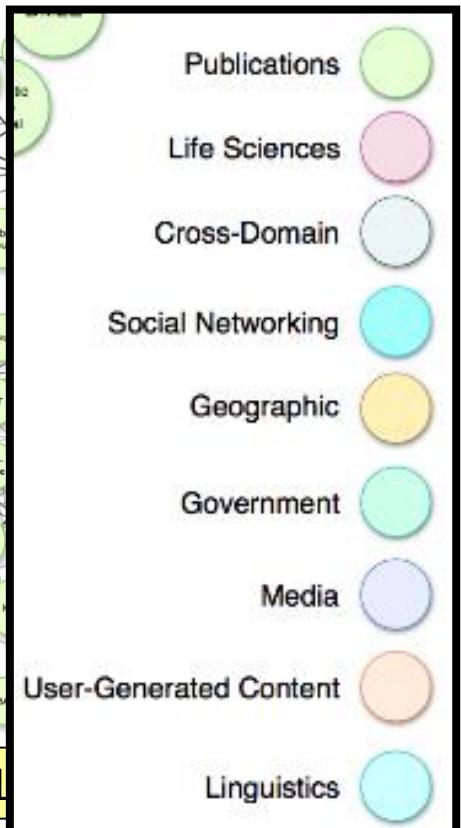
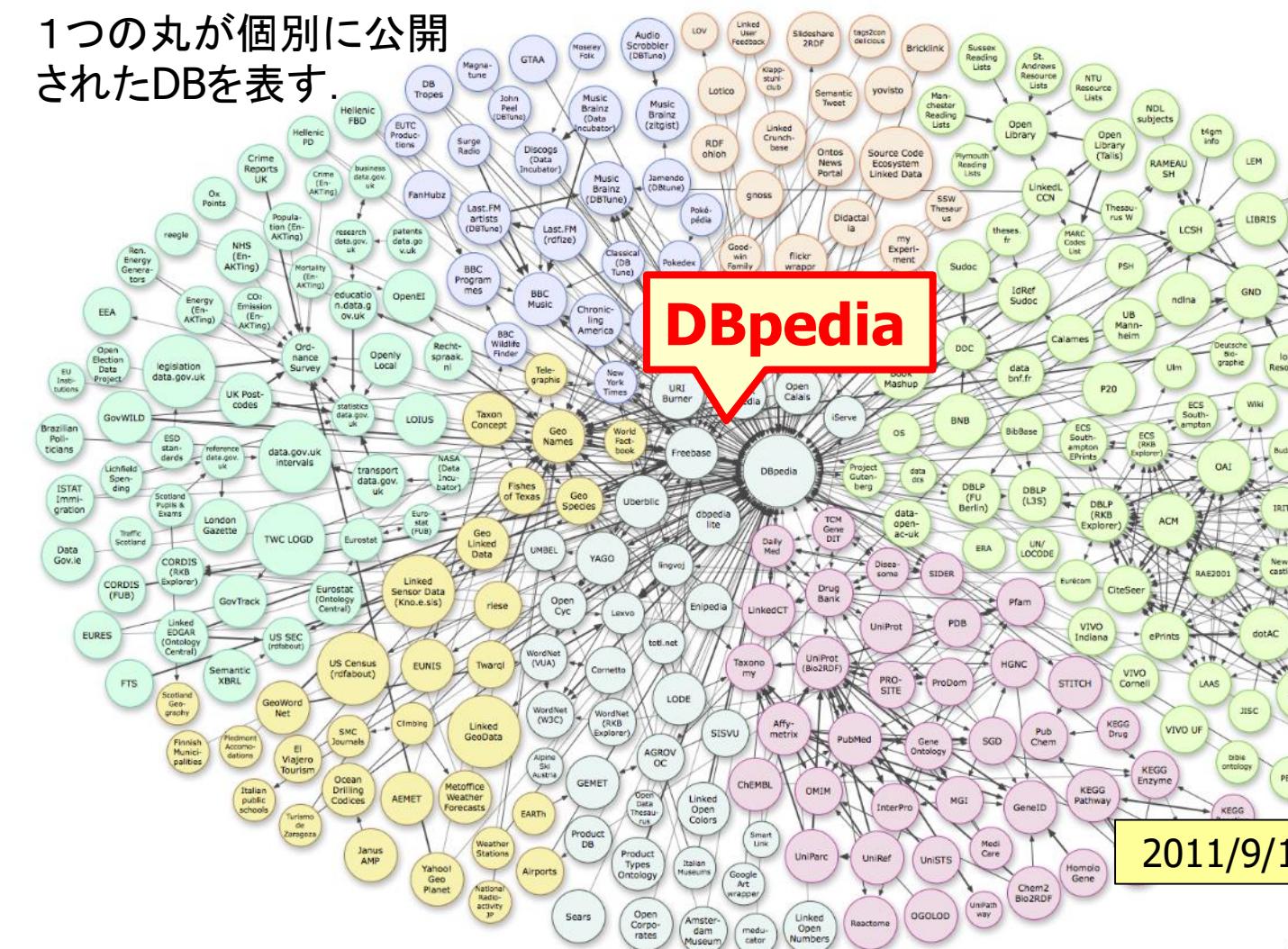
オープンデータの活用事例を紹介

- 白人/黒人の住む家と水道管の整備状況の相関を見える化→裁判で勝利へ
- Where does my money go?  
(税金はどこへ行った?)  
(英国) <http://wheredoesmymoneygo.org/>  
(横浜市) <http://spending.jp/>

# LOD公開の世界的なひろがり ～LODクラウド～

1つの丸が個別に公開されたDBを表す。

2014/08/30時点

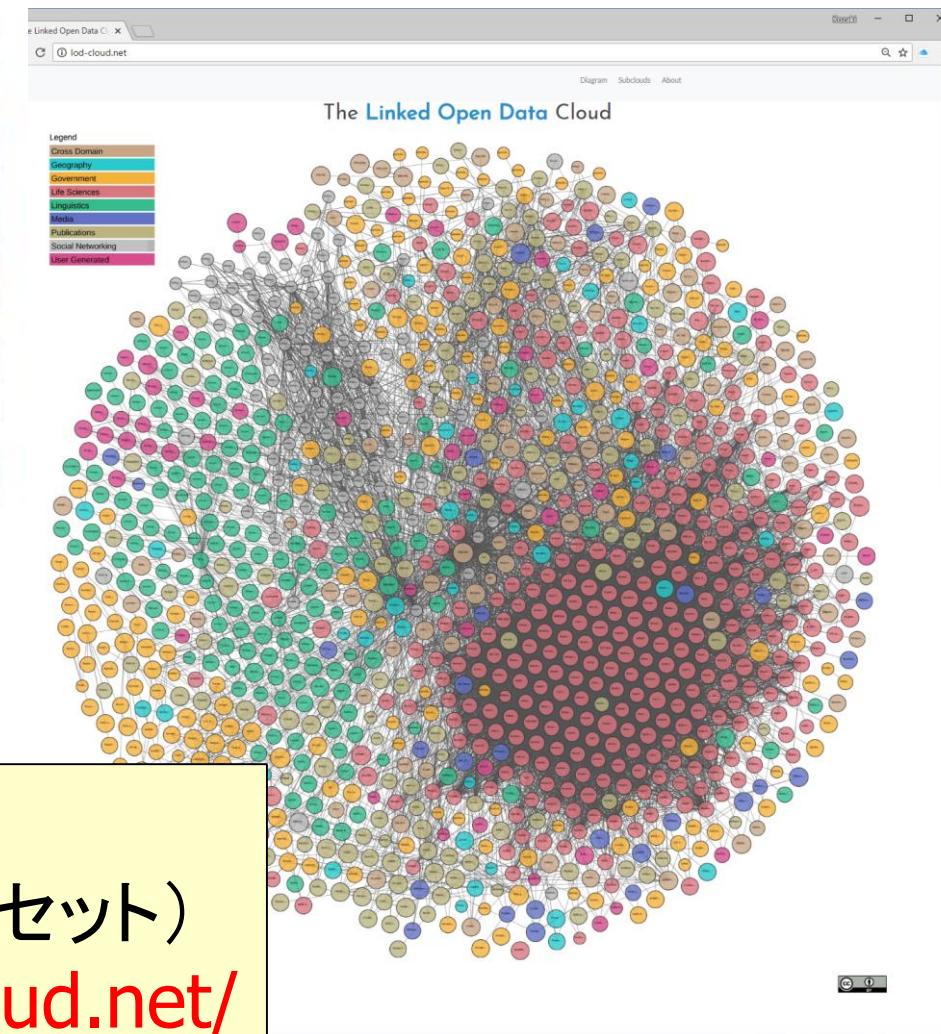


Linking Open Data cloud diagram 2014, by Max Schmachtenberg, Christian Bizer, Anja Jentzsch and Richard Cyganiak. <http://lod-cloud.net/>

# ★LOD公開の世界的なひろがり ～LODクラウド～

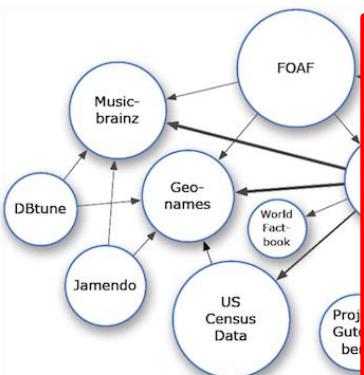


## Legend



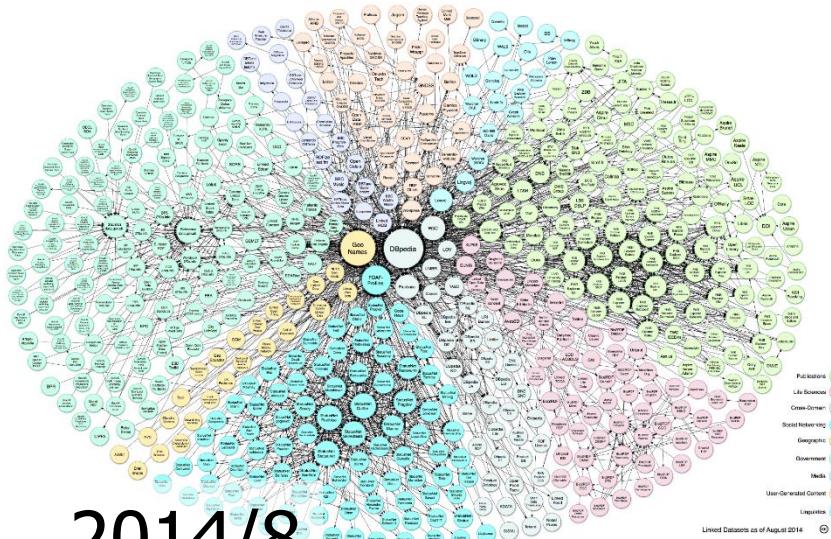
2018-05-30  
(1,186データセット)  
<http://lod-cloud.net/>

# LODの世界的なひろがり

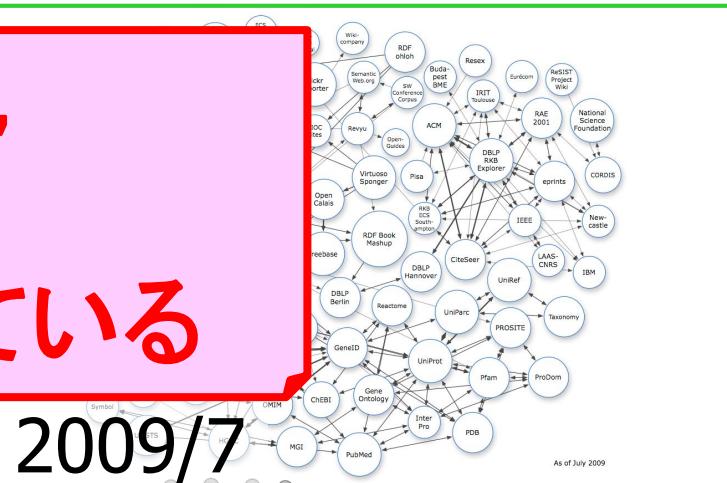


LODの公開は、  
この10年で  
急速に広まっている

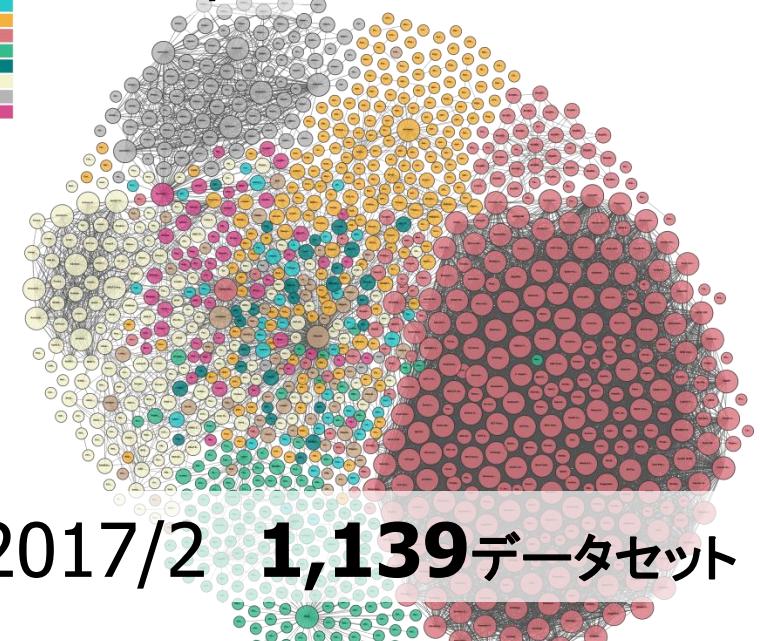
2007/5 12データセット



2014/8



2009/7



2017/2 1,139データセット

Linking Open Data cloud diagram, <http://lod-cloud.net/>

# 日本語で使えるLODの例

法人番号や法人名から  
企業等の登録情報が検索できます。

法人番号または法人名

当サイトの情報について

当サイトでは、法人が役所より受けた補助金や表記、交付可否の法人活動情報を掲載しています。  
ただし、政府の実施事業であっても、業務執行十件が改訂と異なる(都道府県、独立行政法人等)事業の情報は、現在、掲載していません。  
ご了承ください。

お知らせ

2017年03月03日  
中期経済分析システム(RESAS)について  
RESASアソシエイターズは事業を始め切らしました。  
3月26日(日)に新規審査が開始されました。  
詳しくは→こちら <https://open-data.reses-portal.go.jp/>

2017年03月03日  
API利用規約のバージョンを1.1上に更新しました。

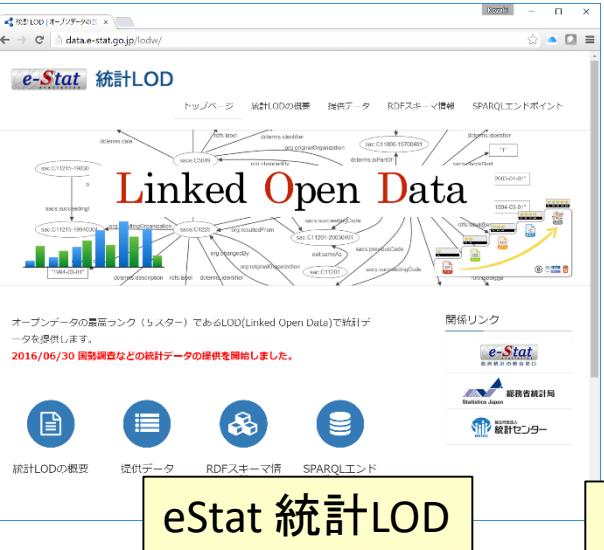
2017年01月24日  
地域経済分析システム(RESAS)について  
RESAS-AP機能を備えたアプリケーションを開発中です。法人インフォーションとの組み合わせによる「法人インフォメーション」を使用していますので、ぜひ、ご利用ください。  
詳しくは→こちら <https://open-data.reses-portal.go.jp/>

法人インフォについて

活用事例集

本サイトへのご意見・ご要望  
をお寄せ下さい。(外部サイト)

# 法人インフォ(経済産業省)



The screenshot shows the homepage of the Osaka Open Data Portal. At the top, there's a navigation bar with links for HOME, DataCatalog, Organization, Group, FAQ, and Help/FAQ. Below the navigation is a large banner featuring a night view of a bridge over a river with illuminated signs. The main title "OSAKA Open Data Portal" is centered above a search bar. Below the search bar are four large numerical values: 112, 51, 12, and 40, each with a corresponding label: データセット (Dataset), 表示 (Display), グループ (Group), and タグ (Tag). Further down, there are several cards with icons and text: a green card for "データカタログ DataCatalog", an orange card for "活用事例", a blue card for "API", a white card for "アプリ等を作成された方へ", a green card for "FAQ", and a purple card for "ご要望に沿ってデータを公開する方へ". A sidebar on the right is titled "新着データ" and lists recent datasets. At the bottom, there's a "お知らせ" (Announcement) section and a large banner at the very bottom with the text "大阪市オープンデータポータル".



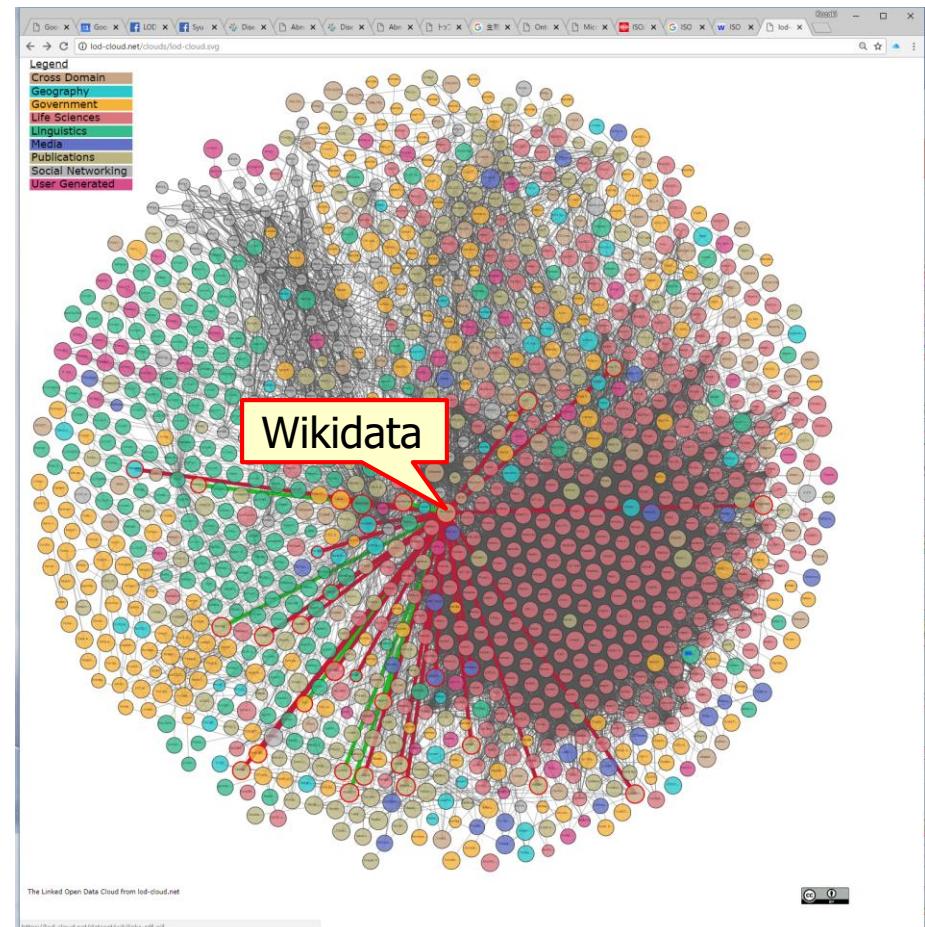
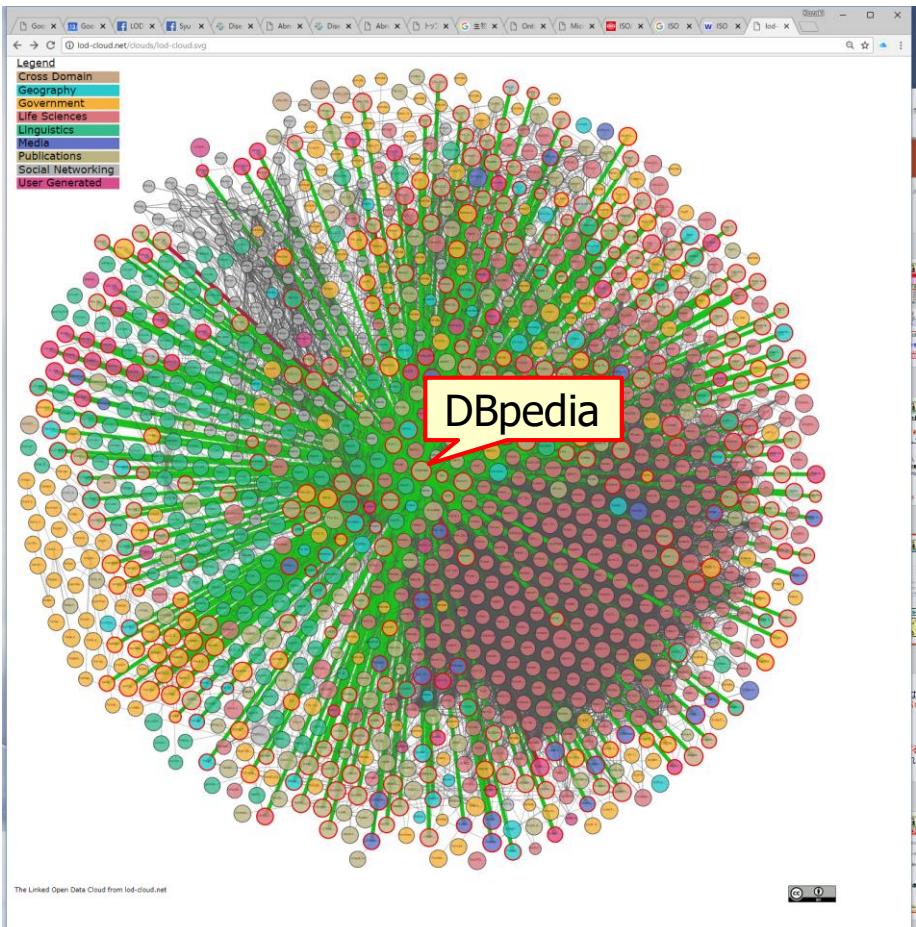
井戸端・コミュニティ・ポータル

はじめに • 井戸端 • コミュニティ • ポータル • ヘルプ

データについて

Wikidata

# DBpedia/Wikidata in LODクラウド



# DBpedia

<http://dbpedia.org/>

The screenshot shows the official DBpedia website at <http://dbpedia.org/>. The page features the DBpedia logo (a stylized tree of orange and yellow nodes) and the text "Towards a Public Data Infrastructure for a Large, Multilingual, Semantic Knowledge Graph". Below the logo, there are social media icons for Facebook, Google+, Twitter, YouTube, and RSS. A yellow box highlights the Japanese version URL: <http://jp.dbpedia.org/>.



大阪市旗	
国	日本
地方	近畿地方
都道府県	大阪府
団体コード	27100-4
面積	223.00km <sup>2</sup> (境界未定部分あり)
総人口	2,687,287人 (推計人口、2014年11月1日)
人口密度	12,050人/km <sup>2</sup>
隣接自治体	豊中市、吹田市、摂津市、門真市、大東市、東大阪市、守口市、八尾市、堺市、松原市、兵庫県尼崎市
市の木	サクラ
市の花	パンジー

インフォボックスの例

Wikipediaの各記事のインフォボックスの情報を抽出して自動生成されるLOD  
様々なデータをつなぐ**LODのハブ**的な存在となっている。  
※英語版と日本語版で、データ構造の違いがあるので注意

# DBpediaのデータ例（大阪市）

検索API

About: 大阪市

An Entity of Type : **settlement**, from Named Graph : <http://ja.dbpedia.org>, within Data Space : [ja.dbpedia.org](http://ja.dbpedia.org)

大阪市（おおさかし）は、日本の近畿地方（関西地方）、大阪府のほぼ中央に位置する市で、同府の府庁所在地である。政令指定都市に指定されている。

Raw Data

生データの取得

Property	Value
dbpedia-owl:abstract	大阪市（おおさかし）は、日本の近畿地方（関西地方）、大阪府のほぼ中央に位置する市で、同府の府庁所在地である。政令指定都市に指定されている。
dbpedia-owl:areaCode	27100-4
dbpedia-owl:location	dbpedia-ja:大阪府
dbpedia-owl:neighboringMunicipality	dbpedia-ja:兵庫県 dbpedia-ja:大東市 dbpedia-ja:豊中市 dbpedia-ja:堺市
dbpedia-owl:postalCode	530-8201

すべてのWikipediaの記事が  
<http://ja.dbpedia.org/resource/大阪市>  
のようなURL(URI)でデータ化されている

プログラムからの  
データ取得も可能

# DBpediaの利用例

- DBpediaを百科事典的な“知識”(事実情報)を取得するための汎用情報源として利用

Disease Compass

日本語 Find Find In Page View

循環器内科(疾患) 心臓弁狭窄症 借帽弁疾患 借帽弁狭窄症

心臓弁狭窄症 借帽弁狭窄症

★疾患：リウマチ熱 借帽弁の硬化・肥厚  
 借帽弁交連部の癒合 借帽弁の開閉制限  
 借帽弁の開閉不全  
 借帽弁腱索の癒合・強縮  
 (借帽弁狭窄をもつ先天的要因)  
 借帽弁輪部石灰化

DBpedia

借帽弁狭窄症 (そうばうべんきょうさくしきょう, mitral stenosis, MS) は、心疾患の一種であり、借帽弁口の狭窄によって左心房から左心室へ血液が流れにくくなる疾患である。

icd9	746.5
icd10	I05.0,I34.2,Q23.2
meshId	D008946

image

BodyParts 3D

[View on J View on J](#)

医療分野での利用例  
<http://lod.hozo.jp/DiseaseChainViewer/>

CLOSE

マンボウ

マンボウ (翻車魚, Mola mola) は、フグ目マンボウ科の1種である。巨体と独特の体型が特徴の海水魚である。ウォノタユウ (瀬戸内海)、ウキ、ウギギ、ウキキ (浮木) (以上東北地方)、マンザイラク (神奈川)、マンボウザメなど多数の地方名で呼ばれる。英語名はオーシャンサンフィッシュ (ocean sunfish) だが、ブルーギルなどを含むスズキ目サンフィッシュ科とは関係がない。



[View on J](#) [View on J](#)

CLOSE

多言語対応も可能！

Slender sunfish

The slender sunfish (Ranzania laevis) is a mola of the family Molidae, the only member of the genus Ranzania, found globally in tropical and temperate seas. Its length is up to 1 m (3.3 ft). Several stranding and mass stranding events have been discovered on beaches near Albany, Western Australia.



生物分野での利用例  
<http://biomimetics.hozo.jp/>

# DBpediaを利用したアプリ例

The screenshot shows a web browser window with the title 'バーチャル宇宙旅行'. The URL in the address bar is [museums-info.net/spacemachine/navi/](http://museums-info.net/spacemachine/navi/). The page content includes:

- A search form titled '条件からさがす' (Search by condition) with dropdowns for '出発' (Departure) set to 'アメリカ航空宇宙局' (NASA), '行先' (Destination) set to '月' (Moon), and 'ロケット' (Rocket) set to '指定しない' (Not specified). A button 'さがす' (Search) is present.
- A list titled '宇宙機一覧' (List of space vehicles) showing 3 items: 'エルクロス', 'GRAIL', and 'Nautilus-X'. Navigation buttons '<前の10件 次の10件>' are at the bottom of this list.
- A detailed description of the 'エルクロス' (LCROSS) mission, stating its purpose is to confirm the presence of water ice at the Moon's poles. It mentions the mission ended in operation and has a total weight of 2249 kg (satellite) and 621 kg (body). Four image links are provided.
- A small image of the LCROSS satellite in space, positioned next to the Moon.

**DBpediaの情報の“つながり”を辿ることで、バーチャルな宇宙旅行を！**  
※「第7回LODハッカソン関西 in IODD大阪(1日イベント)」の成果

<http://museums-info.net/spacemachine/navi/>

# Wikidata



保護された通信 | <https://www.wikidata.org/wiki/Wikidata:メインページ>

A あ 日本語 Koujikozaki ノート トーケ 個人設定 ベータ版 ウォッチリスト 投稿記録 ログアウト

メインページ 講論 閲覧 編集 履歴表示 Wikidata内を検索

<http://wikidata.org/>

WIKIDATA

メインページ  
コミュニティ・ポータル  
井戸端  
項目の新規作成  
ページ名から項目を探す  
最近の更新  
おまかせ表示  
Query Service  
付近  
ヘルプ  
寄付

印刷/書き出し  
ブックの新規作成  
PDF形式でダウンロード  
印刷用バージョン

他のプロジェクト  
commons  
MediaWiki  
メタウィキ  
Wikispecies

ツール  
リンク元  
関連ページの更新状況  
特別ページ  
この版への固定リンク  
ページ情報  
Wikidata項目

open  
multilingual  
free

collaborative

ウィキデータへようこそ  
25,265,213 件の項目を持つ、誰でも編集可能なフリーの知識ベース  
はじめに・井戸端・コミュニティ・ヘルプ

ようこそ!

ウィキデータは、人間とコンピューターの双方が同等に参照・編集できるフリーかつオープンな知識データベースです。

ウィキメディアの姉妹プロジェクト群（WikidataやWikivoyage、Wikisourceなど）の構造化データを集中的に保管・管理します。

Wikidataは、Wikimediaの各プロジェクトへサービス提供に留まらず、他のサイトやサービスも支援を行います。Wikidataのコンテンツは、フリーライセンスであり、標準化された形式で出力可能で、ウェブ上にある他のオープンデータベースと相互にリンクできるようになっています。

• ウィキメディア財団が運営する Wikipedia の「データ版」  
• Wikipediaと同じようにデータをコミュニティで編集、公開できる  
• Wikipediaの「多言語リソース」の相互リンクのために整備  
• SPARQLエンドポイントや各種検索ツールなども提供

項目: 地球 (Q2) プロパティ: 最も高い地点 (P610)

# WikipediaからWikidataへ



Wikidata Query Service | Wikidata Query Service | 大阪電気通信大学 - Wikipedia

https://ja.wikipedia.org/wiki/大阪電気通信大学

Koujikozaki 会話 下書き 個人設定 ベータ版 ウォッチリスト 投稿記録 ログアウト

ページ ノート 閲覧 編集 履歴表示 Wikipedia内を検索

大阪電気通信大学

関連ページの更新状況  
ファイルをアップロード  
特別ページ  
この版への固定リンク  
ページ情報  
Wikidata項目  
このページを引用

座標: 北緯34度45分37.2秒 東經135度37分37.1秒

大阪電気通信大学  
寝屋川キャンパス  
大学設置 1961年  
創立 1941年

Wikidata項目へのリンク

2019/8/27

# Wikidataのデータ例（1/2）



さまざまな言語での  
「ラベル」、「概要説明」、「別名」

Osaka Electro-Communication University (Q7105556)

higher education institution in Osaka Prefecture, Japan

In more languages Configure

Language	Label	Description	Also known as
English	Osaka Electro-Communication University	higher education institution in Osaka Prefecture, Japan	
Japanese	大阪電気通信大学	No description defined	OECU 大阪電通大
?ucināguci	No label defined	No description defined	
Korean	No label defined	No description defined	

Statements

instance of university edit

0 references

image

Wikidataにおいて  
「大阪電気通信大学(Q7105556)」にアクセスした例  
<https://www.wikidata.org/wiki/Q7105556>

# Wikidataのデータ例（2/2）

The screenshot shows a browser window with three tabs: 'Wikidata Query Service' (selected), 'Wikidata Query Service', and 'Osaka Electro-Communication U'. The main content area displays 'Statements' for item Q7105556. A red box highlights the 'instance of' statement, which points to 'university'. A blue box highlights the 'image' statement, showing a photograph of a large brick building labeled 'OECU.jpg'. A yellow box highlights the 'country' statement, pointing to 'Japan'. A pink box highlights the 'located in the administrative territorial entity' statement, pointing to 'Neyagawa'. A large yellow arrow on the left points right, containing the text '述語(プロパティ)'. A large blue arrow on the right points left, containing the text '目的語(オブジェクト)'. A pink box at the bottom contains the text '他のリソースへのリンク'.

Wikidata Query Service | Wikidata Query Service | Osaka Electro-Communication U

Statements

instance of

image

inception

country

located in the administrative territorial entity

university

0 references

OECU.jpg  
1,555 × 1,166; 273 KB  
1 reference

1924  
1 reference

Japan  
1 reference

Neyagawa

edit

edit

述語(プロパティ)

目的語(オブジェクト)

他のリソースへのリンク

# ウィキデータの可視化ツール例 : Reasonator



大坂城 - Reasonator

<https://tools.wmflabs.org/reasonator/?q=Q321242&lang=ja>



大坂城 (Q321242)

오사카 성 | 긴조 | Osaka-jo | Ōsaka-jō | Ōsaka-jō | Château d'Osaka | Château d'Ōsaka | Castello di Osaka | Kastil Osaka | Burg Osaka | Burg von Osaka | Schloss Osaka | Osakajo | 大阪城 | 浪華城 | 金城 | 錦城 | Lâu đài Osaka | Castelo Ōsaka | Osako | Замок Осака | Осакский замок | Замок Хидэёси в Осаке | Осака замок | Замок Хидэёси в Осаке | ปราสาทโอซาก้า | Osaka Castle | Castell d'Osaka | Осацкий замок | ოსაკის ციხე-სიმაგრე | ოსაკის ციხე | hrad Ōsaka | 天守閣 | قلعه اوساکا

日本の城

Location properties

分 日本の城  
類

Maps



Other Wikidata items within 15km | Geohack | TagInfo | Overpass | 34.6872222222 / 135.52583333333

位置

名前	説明
地球	太陽系第三惑星
アジア	大陸、大州
日本	東アジアに位置する国家
近畿地方	日本の地域
大阪府	日本の都道府県
大阪市	日本の大阪府にある政令指定都市。府厅
中央区	大阪市の24行政区のひとつ



外部サイト

公式ウェブサイト

外部情報源

BabelNet	03134837n
Commons maps	Maps of Osaka Castle
category	place/Osaka-Castle
Encyclopædia Britannica	
Freebase	/m/024p_g
GeoNames	6324544
NDL Auth	00953775
Structurae structure	20033111
VIAF	244717962

外部識別子

ウィキメディアプロジェクト

現在の言語のウィキペディア

ja	大阪城
de	Burg Ōsaka
en	Osaka Castle
es	Castillo de Osaka
fi	Osakan linna
fr	Château d'Osaka
hu	Oszakai várkastély

ウィキデータのID  
を指定するだけで、  
データの種類に応  
じた可視化が可能

- ・人物
  - ・場所
  - ・生物
  - ・カレンダー

...など

<https://tools.wmflabs.org/reasonator/>

表示例 <https://tools.wmflabs.org/reasonator/?q=Q7105556>

# ウィキデータの可視化ツール例 : Scholia



<https://tools.wmflabs.org/scholia/>

Denny Vrandečić (Q18618629)

Related: Egon Willighagen · Dietmar Schomburg · Patrick O. Brown · Christoph Steinbeck · Geoffrey Burnstock · Joseph Schlessinger · Aled Edwards · Tyler Cowen · Ion Idriess · Arne Skerra

<https://orcid.org/0000-0002-9593-2294>

List of publications (RSS)

Show 10 entries

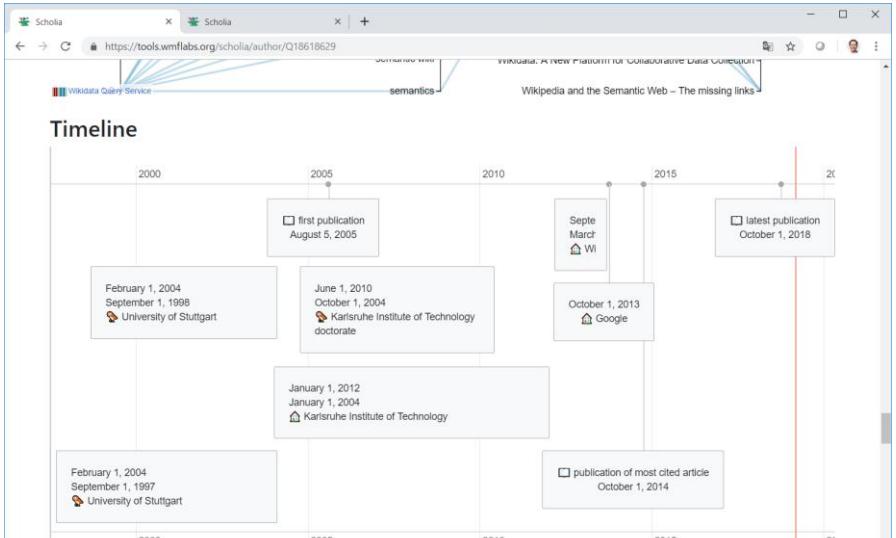
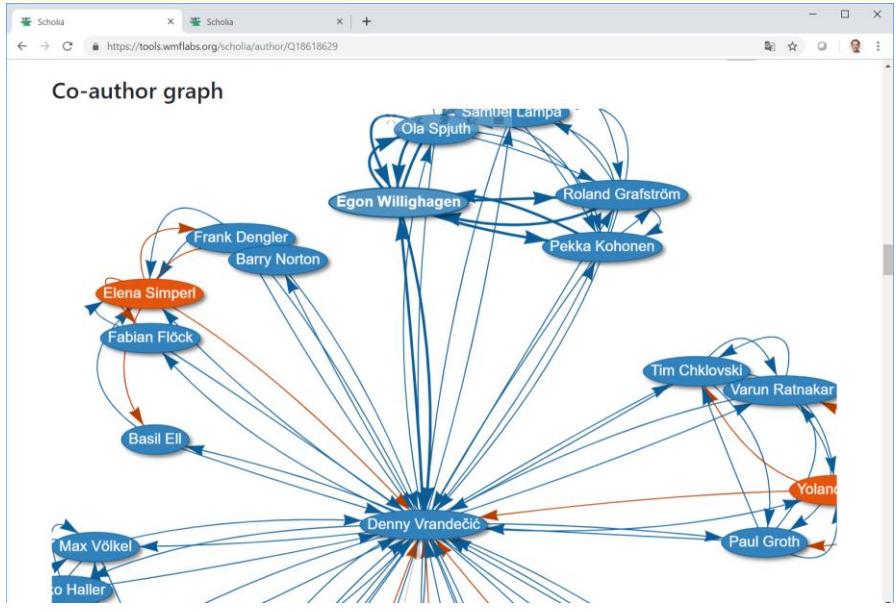
Date	Work	Type	Pages	Venue	Authors
2018-10-01	Capturing meaning:Toward an abstract Wikipedia	document	4		Denny Vrandečić
2017-09-04	RDFIO: extending Semantic MediaWiki for interoperable biomedical data management.	scholarly article	13	Journal of Biomedical Semantics	Egon Willighagen, Denny Vrandečić, Ola Sjöström, Roland Grafström, Pekka Kohonen, Samuel Lampa
2016-01-01	From Freebase to Wikidata: The Great Migration	scholarly article	10	Proceedings of the 25th International Conference on World Wide Web	Sebastian Schaffert, Lydia Pintscher, Thomas Steiner, Denny Vrandečić, Thomas Pellissier Tanon
2014-10-01	Wikidata: A Free Collaborative Knowledgebase	scholarly article	8	Communications of the ACM	Markus Krötzsch, Denny Vrandečić

Showing 1 to 10 of 29 entries

Missing publications? Search for unlinked works on Wikidata or add new ones with sourceme.

Number of publications per year

Year	First author	Last author	Middle author	Solo author	Unknown	Total
2005	0	0	0	0	1	1
2006	0	0	0	0	2	2
2007	0	0	0	0	1	1
2008	0	0	0	0	1	1
2009	1	0	0	0	0	1
2010	0	0	0	0	0	0
2011	0	2	0	0	0	2
2012	0	1	3	0	0	4
2013	0	0	0	2	0	2
2014	1	0	0	0	0	1
2015	0	0	0	0	0	0
2016	0	0	1	0	0	1
2017	0	0	0	1	0	1
2018	0	0	0	1	0	1



# LOD/Linked Dataの例 Japan Search

The screenshot shows the Japan Search BETA website interface. At the top, there's a navigation bar with links like 'HOME', 'SEARCH', 'ABOUT', 'CONTACT', and 'LOGOUT'. Below the header, there's a large 'JAPAN SEARCH BETA' logo with a colorful geometric icon. A 'GALLERY' section is visible with a red ribbon icon. The main content area displays several search results in a grid:

- ことごとくと箏曲**  
箏と琴、どちらも「こと」と読みますが、別の楽器です。現在私たちが「こと（お琴）」と呼んでいる十三弦の楽器は、一般的には「箏」を指します。  
ここでは事にまつわる歴史や文化を紹介します。
- 富士山**  
古代より人々に讃仰されてきた日本の最高峰
- あれもこれも和菓子**  
和菓子には意外な歴史的背景や伝統行事との関わり、有名作家の和菓子にまつわるエピソードなどが詰まっています。
- 茶**  
葉の加工品を飲料とするツバキ常緑樹。茶の交易は世界史上に大きな影響を与えた
- 懸女房染分手綱 PDM(パリック・ドメイン)  
国立国会図書館デジタルコレクション**

Below the grid, there's a yellow banner with the URL <https://jpsearch.go.jp/> and two bullet points:

- 国内デジタルアーカイブの統合検索サービス
- SPARQLエンドポイントあり

At the bottom, there are more search categories:

- 横断検索**  
全てのデータベースを横断してキーワード検索します。
- 日本食について調べる**  
日本食に関する資料を検索します。キーワードを入力してください。
- 葛屋重三郎が版元の作品や作家について調べる**  
葛屋重三郎が版元の作品や作家について調べます。「詳細」をクリックしてください。
- 北斎の作品を調べる**  
北斎の作品を調べます。作品名を入力してください。
- 水墨画について調べる**  
水墨画の作者名、題材や季節（竹、雪、松、秋など）を入力してください。

# JapanSearchの利用例



- ジャパンサーチが提供するLODとその活用例、
  - 中村 覚(東京大学), LODチャレンジ2019ミートアップ(2019/07/05)  
[https://2019.lodc.jp/archives/2019/img\\_meetup/lodc2019-0705-LT05.pdf](https://2019.lodc.jp/archives/2019/img_meetup/lodc2019-0705-LT05.pdf)
  - ジャパンサーチ非公式サポート ページ(神崎正英氏)
    - [https://www.kanzaki.com/works/ld/jp\\_search/](https://www.kanzaki.com/works/ld/jp_search/)
  - Japan Search Apps(中村覚氏)
    - <https://github.com/nakamura196/jpsearch>

# LOD/Linked Dataの例

## 統計LOD

統計LODへようこそ | 統計LOD | 保護されていない通信 | data.e-stat.go.jp/lodw/ | English | 検索 |

## e-Stat 統計LOD

統計LODの概要 | 統計LODの使い方 | 提供データ | RDFスキーマ情報 | SPARQLエンドポイント

Linked Open Data

オープンデータの最高ランク（5スター）であるLOD(Linked Open Data)で統計データを提供します。

お知らせ | 一覧

2019年4月16日 サイト更新情報 データ追加のお知らせ（2019年4月）

2018年4月13日 サイト更新情報 データ追加のお知らせ（2018年4月）

2017年12月8日 トピックス イベント開催のお知らせ

<http://data.e-stat.go.jp/lodw/>

- 日本政府の統計データのLOD
- SPARQLエンドポイントあり

## ■ LODチャレンジ2018 最優秀賞(林 正洋)

- <https://www.mirko.jp/townpower/>

市区町村・行政区

都道府県

東京都

埼玉県

東京23区

東京23区 VS 埼玉県

部門勝利数 3 対 3 で同点  
合計ポイント 76 対 79 で埼玉県の勝利！

指標	東京23区	埼玉県
人口力	18点	18点
女子力	勝 8点	6点
国際力	勝 14点	10点
家族力	6点	勝 11点
キッズ力	6点	勝 10点
シルバーカー	3点	3点
婚姻力	5点	勝 9点
労働力	勝 16点	12点
合計ポイント	76点	79点

人口力  
労働力  
婚姻力  
シルバーカー  
女子力  
国際力  
家族力  
キッズ力

東京23区 埼玉県

# 統計LODの使用例

## エリアベンチマークリング

: 統計LODで似ている町を探してみよう

LODチャレンジ2016  
Inspire the LOD賞



政府統計データが提供される統計LODより取得した市町村区の特徴データを主成分分析し、特徴の近い地域をデータから探すサービス。

<http://area-benchmarking.meta-dog.com/>  
<http://idea.linkdata.org/idea/idea1s2203i>

他市区町村への通勤者数 607728 0.24138

このサービスは政府統計データがODataで提供されます

# LOD/Linked Dataの例

## 法人インフォ



トップページ|法人インフォ × +

← → C 経済産業省 [JP] | https://hojin-info.go.jp/hojin/TopPage ☆ ○ ① ⋮ ? ヘルプ

### 法人インフォ

法人番号や法人名から  
企業等の活動情報が検索できます。

簡易検索 詳細検索 簡易地図検索 ダウンロード API利用方法

法人番号または法人名  

#### 当サイトの情報について

当サイトでは、法人が政府より受けた補助金や表彰、許認可等の法人活動情報を掲載しています。  
ただし、政府の実施事業であっても、業務執行主体が政府と異なる都道府県等の事業の情報は、現在、掲載していません。  
今後、システム更新時に法人活動情報のデータ構造を変更することがあります。  
変更する場合は、お知らせ情報でご連絡致します。ご了承ください。

お知らせ

2019年06月26日

<https://hojin-info.go.jp/>

・国内企業情報がSPARQLで取得可能

# LOD/Linked Dataの例 Web NDL Authorities



Web NDL Authorities (国) 大阪大学 - Linked Data B https://id.ndl.go.jp/auth/ 大阪大学 - Linked Data B Kozaki ー × 保護された通信 https://id.ndl.go.jp/auth/ndla/ ☆ ☁ ☒ ⋮ English ヘルプ

## Web NDL Authorities

国立国会図書館典拠データ検索・提供サービス

キーワード検索 分類記号検索

検索語を入力してください 検索

すべて  名称のみ  普通件名のみ

<https://id.ndl.go.jp/auth/ndla/>

- SPARQLエンドポイントあり
- 参照解決可能

.rdf .ttl .json

Web NDL Authoritiesは、国立国会図書館で作成し、維持管理する典拠データを一元的に検索・提供するサービスです。サービスの概要については、「[Web NDL Authoritiesについて](#)」をご覧ください。

お知らせ

- [NDC10版の適用を開始します](#) (2017年3月23日)
- [英語版を改修しました／English version renewed](#) (2016年12月15日)
- [英語版を公開しました／English version is now available.](#) (2016年3月23日)

国立国会図書館ホーム | 国立国会図書館サーチ Web NDL Authoritiesについて | お問い合わせ

Copyright © National Diet Library. All Rights Reserved

# LOD/Linked Dataの例 CiNii（メタデータAPI）



利用者のみなさまにご不便をおかけしておりますことをお詫び申し上げます。NII-ELSの終了にともない学協会との調整が必要な論文を除き、従前通りのサービス（ダウンロード機能を含む）を再開しました。詳細についてはこちらをご覧ください。

## CiNii Articles - 日本の論文

論文検索 著者検索 全文検索

フリーワード 検索

すべて 本文あり ▲ 閉じる

タイトル  
著者名  
刊行物名  
出版者

著者ID  
著者所属

<http://ci.nii.ac.jp/>  
[https://support.nii.ac.jp/ja/cia/api/a\\_rdf](https://support.nii.ac.jp/ja/cia/api/a_rdf)  
[https://support.nii.ac.jp/ja/cia/api/a\\_json](https://support.nii.ac.jp/ja/cia/api/a_json)

・参照解決可能 **rdf .json**

CiNiiについて ヘルプ  
CiNiiについて CiNii Articles - 日本の論文  
収録刊行物について CiNii Books - 大学図書  
利用規約 CiNii Dissertations - 日本の博士論文をさがす  
お問い合わせ Twitter NII-REO - 電子リソースリポジトリ  
国立情報学研究所 (NII) メタデータ・API

# LOD/Linked Dataの例 BioPortal



米国The National Center for Biomedical Ontology(NCBO)により運営されている  
ライフサイエンス分野のオントロジーレポジトリ

Welcome to BioPortal, the world's most comprehensive repository of biomedical ontologies

Search for a class

Enter a class, e.g. Melanoma

Advanced Search

Find an ontology

Start entering ontology name, e.g. Cancer, then choose from list

Browse Ontologies

Ontology Visits (April 2017)

Ontology	Visits (April 2017)
CPT	~75,000
MEDDRA	~20,000
SNOMEDCT	~18,000
RXNORM	~15,000
NDDF	~10,000

More

BioPortal Statistics

Ontologies	559
Classes	8,134,819
Resources Indexed	48
Indexed Records	39,537,360
Direct Annotations	95,468,433,792
Direct Plus Expanded Annotations	144,789,582,932

<https://bioportal.bioontology.org/>

PRODUCTS SUPPORT ABOUT CONNECT

国内のライフサイエンス関連のRDFデータ  
を集めたポータルサイト

The NBDC RDF Portal provides a collection of life science datasets in RDF (Resource Description Framework). The portal aims to accelerate integrative utilization of the heterogeneous datasets deposited by various research institutions and groups. In this portal, each dataset comes with a summary, downloadable files and a SPARQL endpoint.



<https://integbio.jp/rdf/>

- Linked Dataは、Web上で公開されたデータを「つなぐ」仕組み
  - URL(IRI)を用いたグローバルに一意なデータの識別
  - データ間の“リンク”
  - Webと同じ仕組みを用いたデータのアクセス
    - IRIによる直接アクセス
    - SPARQLエンドポイント(API)によるアクセス
- これらの仕組みが標準化されているため、  
**データの共有**が容易に行える。
  - オープンに公開されているLinked Data(**LOD**)を利用することで、
    - データ構築コストの削減、
    - データへのセマンティクスの付与 が可能に！

# ナレッジグラフ利用技術の紹介



## → RDFを扱うための技術・ツール

➤ RDF用検索言語: **SPARQL**

- ✓ SPARQLエンドポイント(検索用API)からWeb経由の検索が可能

## ➤ RDF用のライブラリ

- ✓ <https://github.com/KnowledgeGraphJapan/sparql-library-examples> にプログラム言語でのサンプルあり

- ✓ Javaを使うなら, **Apache Jena**がおススメ

## ➤ OWL形式のファイルを開くには

- ✓ **protégé**などのオントロジエディタを使用(RDFファイルも開ける)  
<https://protege.stanford.edu/>

## ➤ RDFファイルをDBに格納して使用するには

- ✓ **Fuseki**や**Virtuoso**などのRDF-DBを使用

## → チュートリアル資料

## ➤ Jena, Fusekiの簡単な使用方法, SPARQLクエリの書き方.etc

<https://github.com/KnowledgeGraphJapan/LODws2nd>

# Wikidataを例とした SPARQLによる検索演習

- SPARQLとは？
- SPARQLによる検索例

# SPARQLによるRDFの検索

## ■ SPARQL

- RDFデータに対するクエリ言語
- 「指定したグラフ構造」に一致するトリプルを検索する

## ■ 最も基本的な検索

```
select ?s ?p ?o
where {
  ?s ?p ?o .
}
LIMIT 100
```

↑取得する数の制限

←返す要素

?x(x:任意の文字列)は変数を表す

←検索するグラフのパターン

「.」(ピリオド)  
を忘れない

このパターンを変  
えることで、欲しい  
データを取得する

# SPARQLによる DBpedia Japaneseの検索例



## ■ 「各都道府県で生まれた政治家の数」を調べる

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

PREFIX dbpedia-owl: <http://dbpedia.org/ontology/>

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

PREFIX dbpedia-ja: <http://ja.dbpedia.org/resource/>

PREFIX category-ja: <http://ja.dbpedia.org/resource/Category:>

```
select distinct ?pref (count(?s) AS ?c) where {
  ?pref rdf:type dbpedia-owl:Place.
  ?pref dbpedia-owl:wikiPageWikiLink category-ja:日本の都道府県.
  ?s rdf:type dbpedia-owl:Politician;
    dbpedia-owl:birthPlace ?pref.
} GROUP BY ?pref
ORDER BY ?c
```

SPARQLエンドポイント  
(<http://ja.dbpedia.org/sparql>)

(解説) **Qiita**: DBpediaを使った都道府県別ランキング  
<http://qiita.com/koujikozaki/items/439fa7ce3e28b738fe10>

実行結果

# SPARQLによる DBpedia Japaneseの検索例



## ■ 「各都道府県で生まれた芸人の数」を調べる

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

PREFIX dbpedia-owl: <http://dbpedia.org/ontology/>

PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>

PREFIX dbpedia-ja: <http://ja.dbpedia.org/resource/>

PREFIX category-ja: <http://ja.dbpedia.org/resource/Category:>

```
select distinct ?pref (count(?s) AS ?c) where {
  ?pref rdf:type dbpedia-owl:Place.
  ?pref dbpedia-owl:wikiPageWikiLink category-ja:日本の都道府県.
  ?s rdf:type dbpedia-owl:Comedian;
    dbpedia-owl:birthPlace ?pref.
}GROUP BY ?pref
ORDER BY ?c
```

SPARQLエンドポイント

(<http://ja.dbpedia.org/sparql>)

実行結果

ここを、  
**Politician**→**Comedian**  
に変えるだけ！

# SPARQLによる検索例

## Wikidataを用いたランキング



### ■ 「日本の政治家の出身大学」ランキング

→ 実行結果

```
select ?univ ?univl (count(?s) As ?c) where{
?univ wdt:P31/wdt:P279* wd:Q3918. #大学の一覧取得
?s wdt:P27 wd:Q17; #国籍=日本
      wdt:P106 wd:Q82955; #職業=政治家
      wdt:P69 ?univ. #出身大学

OPTIONAL{
  ?s rdfs:label ?name. FILTER(lang(?name)="ja")
  ?univ rdfs:label ?univl . FILTER (lang(?univl) = "ja") .
}

}GROUP BY ?univ ?univl
ORDER BY DESC(?c)
LIMIT 100
```

政治家の出身大学ランキング  
<https://qiita.com/koujikozaki/items/a049e2ac1051e0e43be6>

# SPARQLによる検索例

## Wikidataを用いたランキング



### ■ 「日本の総理大臣の出身大学」ランキング

→ 実行結果

```
select ?univ ?univl (count(?s) As ?c) where{
```

```
?univ wdt:P31/wdt:P279* wd:Q3918. #大学の一覧取得
```

```
?s wdt:P27 wd:Q17; #国籍=日本  
      wdt:P106 wd:Q82955; #職業=政治家  
      wdt:P39 wd:Q274948; #公職=内閣総理大臣  
      wdt:P69 ?univ. #出身大学
```

この1行を  
追加する

```
OPTIONAL{
```

```
?s rdfs:label ?name. FILTER(lang(?name)="ja")
```

```
?univ rdfs:label ?univl . FILTER (lang(?univl) = "ja") .
```

```
}
```

```
}GROUP BY ?univ ?univl
```

```
ORDER BY DESC(?c)
```

```
LIMIT 100
```

政治家の出身大学ランキング  
<https://qiita.com/koujikozaki/items/a049e2ac1051e0e43be6>

# SPARQLクエリのハンズオン

- WikidataのSPARQLエンドポイント(検索用API)  
<https://query.wikidata.org/>  
を使った、SPARQLクエリの演習.
- クエリ例は、すべて、  
[https://github.com/koujikozaki/SPARQLexamples/  
blob/master/WikkidataEx1.md](https://github.com/koujikozaki/SPARQLexamples/blob/master/WikkidataEx1.md)  
から試せる.

## ■ 例題の解説

- スライドを使って例題のクエリを解説する
- 解説後、例題のクエリを実行してみる

## ■ 演習課題

- 例題のクエリの「**一部を変更**」し、演習課題で指定したクエリを作成する
- 作成したクエリは、テキストファイルなどにメモしておくとよい。

# 検索例1：

## 主語と述語を指定して「目的語」を取得

### “<主語>の<述語>は何？”

検索するグラフパターン



SPARQLでの記述

<大阪電気通信大学> <位置する行政区> ?o

主語

述語

目的語

# 検索例1：主語と述語を指定

- 例1)「大阪電気通信大学」(主語)の「位置する行政区」(述語)となる目的語(?o)を取得する

```
select ?o
where {
    wd:Q7105556 wdt:P131 ?o .
```

}

大阪電気  
通信大学  
(主語)

位置する  
行政区  
(述語)

目的語  
(変数)

# 検索例1：主語と述語を指定

- 例1-1) 「大阪電気通信大学」(主語)の「設立」(述語)となる目的語(?o)を取得する

```
select ?o
where {
    wd:Q7105556 wdt:P571 ?o .
```

}

大阪電気  
通信大学  
(主語)

設立  
(述語)

目的語  
(変数)

※述語を変えるといろんな  
目的語が取得できる

# 検索例1：主語と述語を指定

- 例1-2) 「大阪電気通信大学」(主語)の「位置する行政区」(述語)となる**目的語(?o)**を取得する  
※ 検索結果がデータのIDとなる場合、下記の記述を  
追加することで「ラベル」をあわせて取得可能

PREFIX wd: <<http://www.wikidata.org/entity/>>

PREFIX wdt: <<http://www.wikidata.org/prop/direct/>>

select ?o ?oLabel

元の変数名 + Label

where {

wd:Q651233 wdt:P159 ?o .

SERVICE wikibase:label { bd:serviceParam  
wikibase:language "[AUTO\_LANGUAGE],ja". }  
}

Wikidata  
のみの  
仕様

英語にしたいならen

# 例1-3:複数の述語を指定して、目的語を取得する



- ※複数行ならべると、まとめて目的語を取得できる。  
(変数名は変える)
- 「大阪電気通信大学」(主語)の「位置する行政区」(述語)となる目的語(?o), および「設立」(述語)となる目的語(?o2)を取得

```
select ?o ?oLabel ?o2
```

```
where {
```

```
wd:Q7105556 wdt:P131 ?o .
```

```
wd:Q7105556 wdt:P571 ?o2 .
```

```
SERVICE wikibase:label { bd:serviceParam  
wikibase:language "[AUTO_LANGUAGE],ja". }  
}
```

位置する行政区

設立

## 【演習1】主語と述語を指定して「目的語」 を取得する

## ■ 演習1-a:「述語」を変えてみる

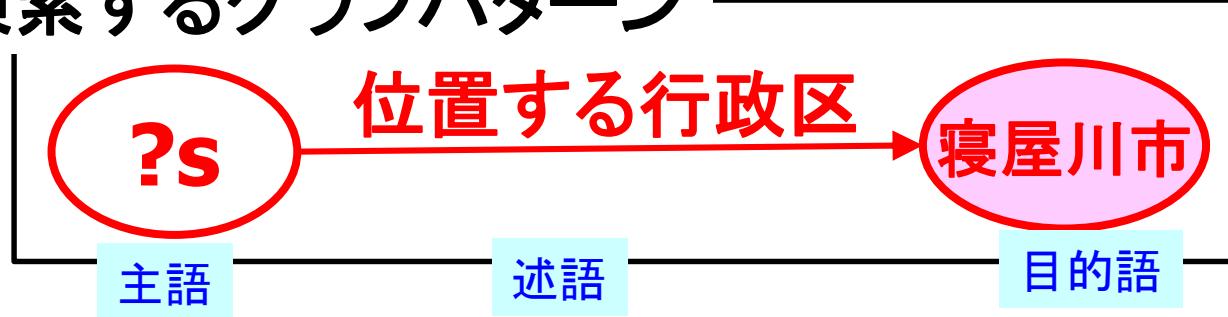
- 「検索例1-1, 2, 3」の「述語」を変えて、「大阪電気通信大学」(主語)のいろんな情報(目的語?o)を取得してみる
  - 述語のIDは「大阪電気通信大学」のWikidataのページを見て探す。  
→プロパティにマウスを持っていくと表示される「P〇〇〇」の番号を使えばよい。

## ■ 演習1-b:「主語」を変えてみる

## 検索例2：

述語と目的語を指定して「主語」を取得  
“<述語>が<目的語>となる<主語>は？”

検索するグラフパターン —————



SPARQLでの記述 —————

?s <位置する行政区> <寝屋川市>

## 検索例2：述語と目的語を指定



- 検索例2-1 「位置する行政区」(述語)が「寝屋川市」(目的語)となる「主語(?s)」の一覧を取得する
  - 「主語」の取得は検索結果が多数となる場合が多い

```
select ?s ?sLabel
```

```
where { ?s ?p ?o . }
```

```
?s wdt:P131 wd:Q389633 .
```

```
SERVICE wikibase:label { bd:serviceParam  
wikibase:language "[AUTO_LANGUAGE],ja". }  
}
```

```
LIMIT 100
```

位置する行政区(述語)

寝屋川市(目的語)

結果が多いので、件数の制限をかける

## 検索例2：述語と目的語を指定

### ■ 例) 2-2: 「分類が“大学”となる主語(?s)を取得

- 分類(instance-of)を使うと同じ種類のデータ一覧が取得できる

Wikidataのみ

他のLODでは、  
instance-ofではなく、**rdf:type**  
が用いられることが多い。

select ?s ?sLabel

where {

分類(instance-of)

?s wdt:P31 wd:Q3918 .

SERVICE wikibase:label { bd:serviceParam  
wikibase:language "[AUTO\_LANGUAGE],ja". }  
}

LIMIT 100

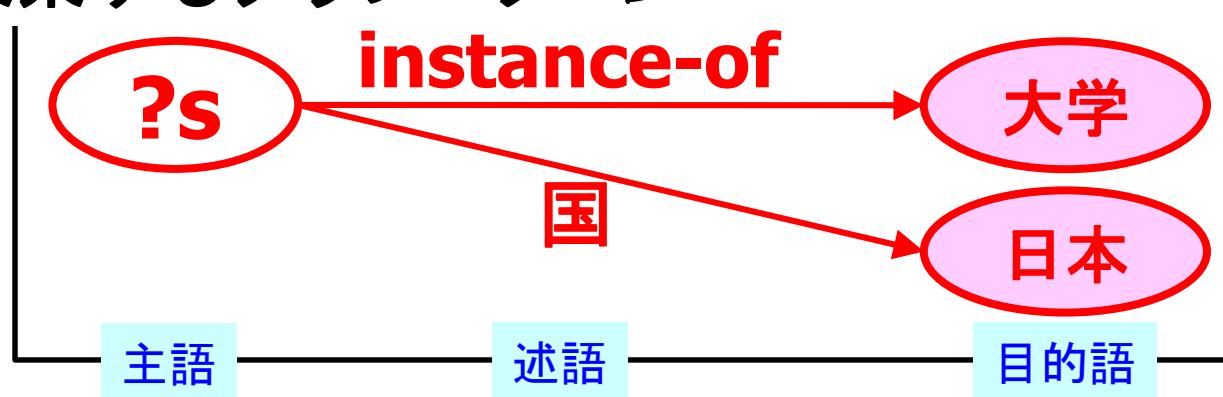
## 【演習2】述語と目的語を指定して 「主語」の一覧を取得する



- 演習2-a: 「目的語」を変えて、いろんな種類(クラス)のデータ一覧を取得してみる
  - 「検索例2-2」の「目的語」となるクラスを変える  
クラスのIDは適当なデータの「Wikidataのページ」で、  
instance of (P31)の目的語を調べると良い。
- 演習2-b: 「述語」と「目的語」の組み合わせを変えて、  
いろんなデータ一覧を取得してみる
  - 「検索例2-1」の「述語」や「目的語」を変える

## 検索例3: 「主語」の一覧の「絞り込み」 “「... という条件を満たす」となる<主語>は？”

検索するグラフパターン —————



SPARQLでの記述 —————

```
?s <instance-of> <大学>.  
?s <国> <日本>.
```

### 検索例3

## ：「主語」の一覧の「絞り込み」



- 検索例3-1 「大学の一覧(主語)」を「国(述語)」の「目的語(?country)」と共に取得する
  - 同じ主語のトリプルを続けることで、合わせて取得したい内容を指定する。※2行以上の指定も同様に可

```
select ?s ?sLabel ?country ?countryLabel
```

```
where {
```

```
?s wdt:P31 wd:Q3918 .
```

例2-2)「分類」が「大学」

の主語一覧

```
?s wdt:P17 ?country .
```

```
SERVICE wikibase:label/sd { bd:serviceParam
```

```
wikibase:language "[A-Z]{2} [!LANGUAGE],ja". }
```

```
}
```

```
LIMIT 100
```

取得した主語(?s)の「国(述語)」  
の「目的語(?country)」

### 検索例3

## ：「主語」の一覧の「絞り込み」



- 検索例3-2 「大学の一覧」を「国(述語)が日本(目的語)」のものに絞り込む
  - 同じ主語のトリプルを続けることで、主語が満たす条件を指定する。※2行以上の指定も同様に可

```
select ?s ?sLabel
where {
  ?s wdt:P31 wd:Q3918 .
  ?s wdt:P17 wd:Q17 .
  SERVICE wikibase:label SILENT
  wikibase:language ?lang
}
LIMIT 100
```

例2-2)「分類」が「大学」の主語一覧

取得した主語(?s)の「国(述語)」の「目的語(?country)」が「日本(wd:Q17)」

### 検索例3

#### ：「主語」の一覧の「絞り込み」



- 検索例3-2 「大学の一覧」を「国(述語)が日本(目的語)」のものに絞り込む
  - 「国(述語)」の「目的語」も合わせて取得し、「国＝日本」で絞り込まれていることを確認する。

```
select ?s ?sLabel ?country ?countryLabel
where {
  ?s wdt:P31 wd:Q3918 .
  ?s wdt:P17 ?country .
  ?s wdt:P17 wd:Q17 .
  SERVICE wikibase:label { bd:serviceParam
    wikibase:language "[AUTO_LANGUAGE],ja". }
}
LIMIT 100
```

### 検索例3

## ：「主語」の一覧の「絞り込み」



- 検索例3-3 「日本にある大学の一覧」を「設立日」と共に取得する
  - 「絞り込んだ主語(?s)」の一覧と、同じ主語のトリプルを続けることで、その主語が持つ情報を取得できる。

```
select ?s ?sLabel ?o  
where {
```

```
?s wdt:P31 wd:Q3918 . # ?Sの「分類」が「大学」  
?s wdt:P17 wd:Q17 . # ?sの「国」が「日本」
```

```
?s wdt:P571 ?o . # ?sの「設立」を?oとする
```

```
SERVICE wikibase:sparql Param  
wikibase:language "ja" E],ja". }  
}LIMIT 100
```

検索例3-2 「日本にある  
大学の一覧」の取得

合わせて取得したい内容

※ クエリ内の#以降は、「コメント」

### 検索例3

## ：「主語」の一覧の「絞り込み」



- 検索例3-3-a 「日本にある大学の一覧」を「設立日」と共に取得し、「設立日」で並び替え
  - **ORDER BY**: 並び替えの関数

```
select ?s ?sLabel ?o
where {
    ?s wdt:P31 wd:Q3918 . # ?Sの「分類」が「大学」
    ?s wdt:P17 wd:Q17 . # ?sの「国」が「日本」
    ?s wdt:P571 ?o .     # ?sの「設立」を?oとする
    SERVICE wikibase:label { bd:serviceParam
        wikibase:language "[AUTO_LANGUAGE],ja". }
} ORDER BY ?o
LIMIT 100
```

?o(設立日)で並び替え  
※ DESC(?o)とする「降順」に

## 【演習3】

# いろんなデータの一覧を取得してみる



### ■ 演習3-a

- :「目的語」を変えて、さまざまな「絞り込み」を試す
- 「検索例3-2」の「述語」と「目的語」の組み合わせを  
変え、様々な条件で絞り込んだ「大学の一覧」を取  
得する

### ■ 演習3-b

- :いろんなデータ一覧を取得する

- 「検索例3-1,2,3」を変更して、いろんなデータ一覧  
を取得してみる
- 演習2-aや演習2-bに、「主語の条件」を追加して、  
データ一覧を絞り込む

# SPARQL演習の補足説明

- データの集約
- Wikidata以外でSPARQLを使う際の注意

# 補足1:PREFIXの定義

- 例1)「大阪電気通信大学」(主語)の「位置する行政区」(述語)となる目的語(?o)を取得する

PREFIX wd: <<http://www.wikidata.org/entity/>>

PREFIX wdt: <<http://www.wikidata.org/prop/direct/>>

※wd:やwdt:といった省略表現を使うためには、  
本来はPREFIXの定義が必要

select ?o  
where {

wd:Q7105556 wdt:P131 ?o .

}

大阪電気  
通信大学  
(主語)

位置する  
行政区  
(述語)

目的語  
(変数)

## 補足1: PREFIXを利用しない表現

### ■ 例1-1)をPREFIX(接頭語)を用いず書いた場合

```
select ?o
where {
  <http://www.wikidata.org/entity/Q7105556>
  <http://www.wikidata.org/prop/direct/P131> ?o .
}LIMIT 100
```

PREFIXによる省略表現

PREFIX **wd:** <http://www.wikidata.org/entity/>  
 PREFIX **wdt:** <http://www.wikidata.org/prop/direct/>  
 select ?o
 where {
 **wd:Q7105556 wdt:P131 ?o .**
}LIMIT 100

#### PREFIXの定義

省略表現に用いる文字列は任意に設定できるが、できるだけ慣習的に利用されるものにあわせるとよい。

# 参考:よく用いられるPREFIX



## ■ RDF一般のもの

- **rdfs:** <<http://www.w3.org/2000/01/rdf-schema#>>  
RDFスキーマ(基本的な語彙定義)
- **schema:** <<http://schema.org/>>  
Webのメタデータに記述される語彙
- **skos:** <<http://www.w3.org/2004/02/skos/core#>>  
Web上でのシソーラス, 用語集などに用いられる語彙

## ■ Wikidataで使われるもの

- **wd:** <<http://www.wikidata.org/entity/>>  
エンティティ(もの, コト, データ)
- **wdt:** <<http://www.wikidata.org/prop/direct/>>  
プロパティ(述語/関係)※ 主にSPARQL検索用の直接関係

## ■ PREFIXの検索サービス

- <https://prefix.cc/>

WikidataのRDFでは、詳細情報を記述するため、同じ内容のプロパティが3種類記述されているが、今回はwdt:を使う。

## ラベルの取得方法【Wikidata】

- 例1-2) 「大阪電気通信大学」(主語)の「位置する行政区」(述語)となる目的語(?o)を取得する  
 ※検索結果がデータのIDとなる場合、下記の記述を追加することで「ラベル」をあわせて取得可能

select ?o ?oLabel

元の変数名 + Label

where {

**wd:Q7105556 wdt:P159 ?o .**

SERVICE wikibase:label { bd:serviceParam  
 wikibase:language "ja". }

}

**Wikidata  
のみの仕様**

- 英語にしたいなら**en**
- [AUTO\_LANGUAGE]とすると、**自動選択**
- “**ja, en**”のように併記することも可  
 →記述順が取得する言語の優先順位

# ラベルの取得方法【RDF一般】

- 補足例) 「大阪電気通信大学」のラベルとなる目的語(?o)を取得

PREFIX wd: <<http://www.wikidata.org/entity/>>

PREFIX rdfs: <<http://www.w3.org/2000/01/rdf-schema#>>

select distinct ?o

where {

wd:Q7105556 **rdfs:label** ?o .

}LIMIT 100

RDFで一般に  
「ラベル」を表すプロパティ(述語)

実行例 <https://w.wiki/646>

- 補足例) 「大阪電気通信大学」のラベルとなる目的語(?o)を取得
  - 「言語の種別」を合わせて取得

PREFIX wd: <<http://www.wikidata.org/entity/>>

PREFIX rdfs: <<http://www.w3.org/2000/01/rdf-schema#>>

「言語種別」を取得する関数

結果を別の変数に代入

```
select ?o (lang(?o) AS ?ln)
```

```
where {
```

```
  wd:Q7105556 rdfs:label ?o .
```

```
}
```

実行結果 <https://w.wiki/648>

## 補足2:

# 言語種別の絞り込み【RDF一般】



- 補足例) 「大阪電気通信大学」のラベルとなる目的語(?o)を取得
  - 「言語の種別=日本語(ja)」をのみ

```
PREFIX wd: <http://www.wikidata.org/entity/>
```

```
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
```

```
select ?o (lang(?o) AS ?ln)
```

```
where {
```

```
wd:Q7105556 rdfs:label ?o .
```

```
FILTER (lang(?o) = "ja") .
```

```
}
```

絞り込みの条件を記述(様々な条件記述できる)

実行結果 <https://w.wiki/64A>

## 補足3: カウントの利用

### ■ 補足例) 「大学」のインスタンスの数を取得する

PREFIX wdt: <<http://www.wikidata.org/prop/direct/>>  
PREFIX wd: <<http://www.wikidata.org/entity/>>

データ数をカウントする関数

```
select (count (?s) AS ?c) where {  
?s wdt:P31 wd:Q3918.  
}
```

分類(instance-of)

大学

実行例

<https://w.wiki/64B>

※本来は、**select (count (distinct ?s) AS ?c)**  
とした方が正確な数が得られる。  
(このクエリの場合、数は変わらず)

## 補足4: グループ化の利用

- 補足例) 「大学の一覧(主語)」を「国(述語)」の「目的語(?country)とそのラベル」と共に取得し、「国ごとのインスタンス数」を取得する

グループごとの数  
をカウント

```
select ?country ?countryLabel (count(?s) As ?c)
where {
  ?s wdt:P31 wd:Q3918 . 主語が大学のインスタンス
  ?s wdt:P17 ?country . 主語の「国」
  SERVICE wikibase:label { bd:serviceParam
    wikibase:language "ja". }
} GROUP BY ?country ?countryLabel
```

実行例  
<https://w.wiki/64D>

「目的語(?country)とそのラベル(?countryLabel )」  
でグループ化

## 補足4: グループ化の利用



- 補足例) 「大学の一覧(主語)」を「国(述語)」の「目的語(?country)とそのラベル」と共に取得し、「国ごとのインスタンス数」を取得し、**多い順にソート**。

```
select ?country ?countryLabel (count(?s) As ?c)
where {
  ?s wdt:P31 wd:Q3918 .
  ?s wdt:P17 ?country .
  SERVICE wikibase:label { bd:serviceParam
    wikibase:language "ja". }
} GROUP BY ?country ?countryLabel
ORDER BY DESC (?c)
```

多い順にソート

実行例  
<https://w.wiki/64E>

# グループ化の利用例（1/2）

## ■ 鉄道路線の総線長の取得

- 参考: Wikidataを使って鉄道会社ごとの総路線長をランキングしてみる  
<https://qiita.com/RK-miha/items/6d94f425871c4e9f5f73>

```
select ?s ?sLabel ?o
```

```
where {
```

```
?s wdt:P31 wd:Q728937.
```

分類=鉄道路線

```
?s wdt:P17 wd:Q17.
```

国=日本

```
?s wdt:P2043 ?o.
```

全長

```
SERVICE wikibase:label { bd:serviceParam  
wikibase:language "[AUTO_LANGUAGE],ja". }  
} order by desc(?o)
```

全長で降順にソート

**【実行結果】**

# グループ化の利用例（2/2）

## ■ 鉄道運営会社ごとの路線全長の合計

select **(SUM(?o) as ?total)** ?op ?opLabel

where {

全長の合計

?s wdt:P31 wd:Q728937 .

?s wdt:P17 wd:Q17.

?s wdt:P2043 ?o.

?s **wdt:P137** ?op.

運営者

SERVICE wikibase:label { bd:serviceParam  
wikibase:language "[AUTO\_LANGUAGE],ja". }

}**GROUP BY ?op ?opLabel**

運営者で集約

**order by desc(?total)**

【実行結果】

# まとめ



## ■ ナレッジグラフ

- “知識のつながり”をグラフ構造で表すことで、さまざまな知識システムの知識基盤を与える。
- **Linked Data**はナレッジグラフをWebで公開する技術

## ■ Linked Data / Linked Open Data (LOD)

- **Linked Data**は、Webの仕組みを用いて“データのつながり”を公開する仕組み。
- **LOD**は、Linked Dataとして公開されたオープンデータ。
- すでに、多くの**LOD**が公開されている！
- まずは、**Wikidata**や**DBpedia**から使ってみるとよい。

# 参考資料

- **SPARQL仕様(W3Cのドキュメント)**
  - **SPARQL 1.1 Query Language**  
<https://www.w3.org/TR/sparql11-query/>
- **SPARQLの解説本**
  - オープンデータ時代の標準Web API SPARQL  
<http://sparqlbook.jp/>
- **SPARQL入門スライド(by古崎)**
  - DBpedia Japaneseを例にした解説  
<https://www.slideshare.net/KoujiKozaki/4lod>
  - 大阪市のオープンデータを例にした解説  
<https://www.slideshare.net/KoujiKozaki/apisparql>
- **解説記事**
  - DBpediaを使った都道府県別ランキング  
<http://bit.ly/2oDPI0Q>
  - Wikidataを使った日本の政治家の出身大学ランキング  
<http://bit.ly/2PBt8fn>



<http://sparqlbook.jp/>より