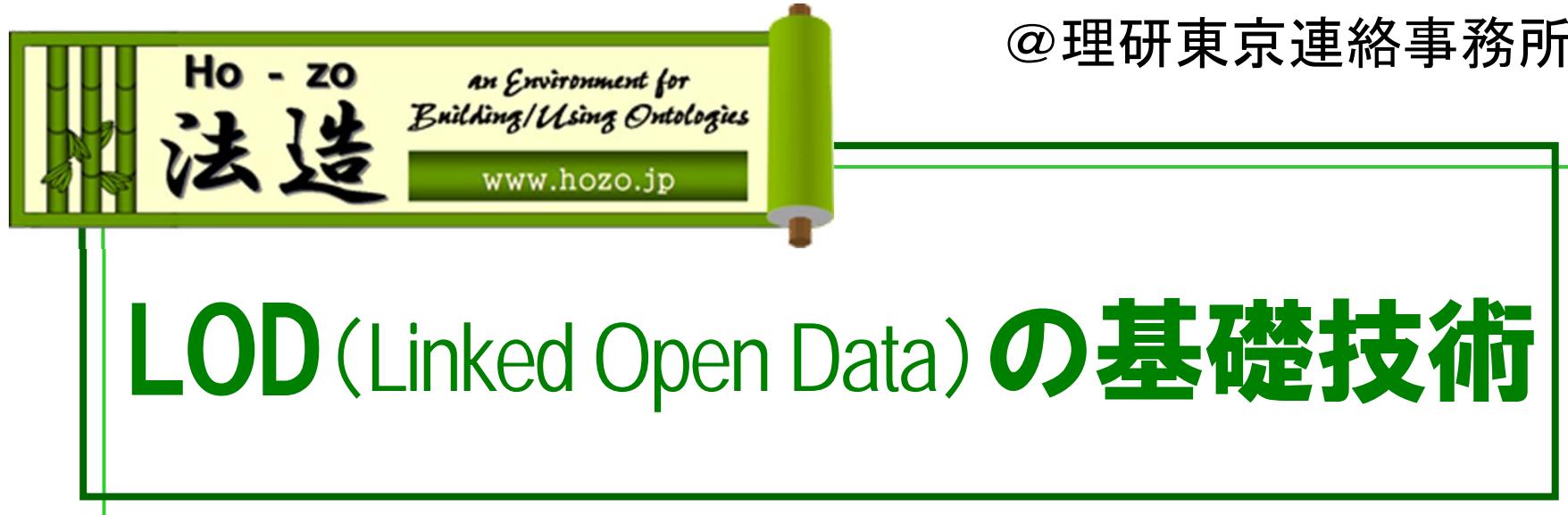


# 第1回Linked Open Data (LOD) 活用ワークショップ

2017年8月3日(木)

@理研東京連絡事務所



大阪大学 産業科学研究所

古崎晃司

# 自己紹介



- 古崎(こざき)晃司  
@koujikozaki
  - 本職: 大阪大学・准教授
  - 専門: オントロジー工学(情報科学・人工知能)
- 人工知能学会  
セマンティックウェブとオントロジー(SWO)研究会 主査
- コミュニティ活動



LOD普及を目的としたコンテスト  
<http://lodc.jp>



関西を中心とした  
LODの普及活動



研究成果として  
公開中のソフト



大阪をITの力でよくしようと活動している団体です。  
<http://code4.osaka/about/>

# **Linked Open Dataとは**

## **Linked Open Data (LOD)**

= Linked Data + Open Data(オープンデータ)

= Linked Dataとして公開されたOpen Data

## **Linked Data**

= Webの仕組みを用いて相互に“リンクされた”データ

# オープンデータ



## ■ オープンデータ

- 誰でも自由に使える形で公開されているデータ.
- 世界各国でオープンデータの公開・利活用が推進されている.
- 官民データ活用基本法(2016/12/7)

## ■ オープンデータの目的・意義

- 国民参加・官民協働の推進
- 透明性・信頼性の向上
- 経済の活性化・行政の効率化

「オープンデータ基本方針」  
(IT戦略本部, 2017/5/20) より



日本政府の  
データカタログサイト  
(<http://www.data.go.jp>)



・19,422のデータセット  
・476のデータベース  
が掲載されている。  
(2017/6/19調べ)

# LOD (Linked Open Data)

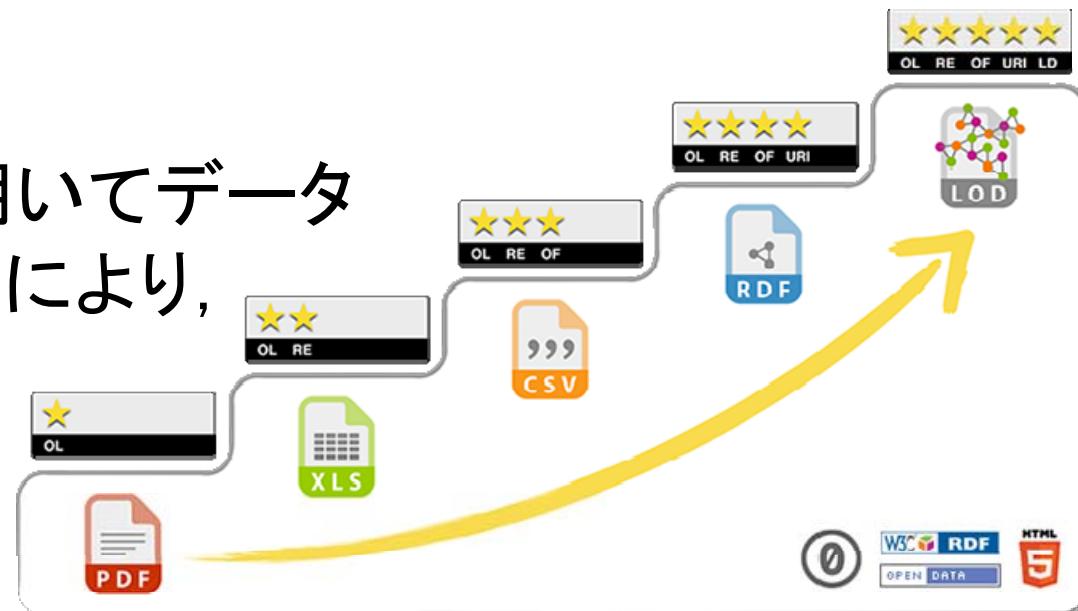


## ■ LOD(Linked Open Data)

- Webの発明者であるティム・バーナーズ=リーが提案している「5つ星スキーム」において、5★とされる「**Webの仕組みを用いたオープンデータ公開のベストプラクティス**」

## ■ Linked Data

- Webの仕組みを用いてデータを“リンク”することにより、「**データの新たな価値の創出**」をめざした技術



<http://5stardata.info/ja/>

# Linked Data



- **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/>

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.

**“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。**

Login to edit  
(authorised users only at present)

Username: \*

Password: \*

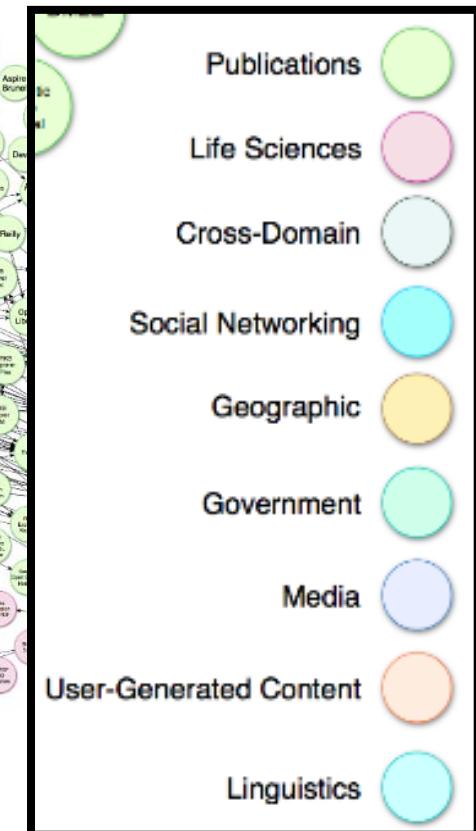
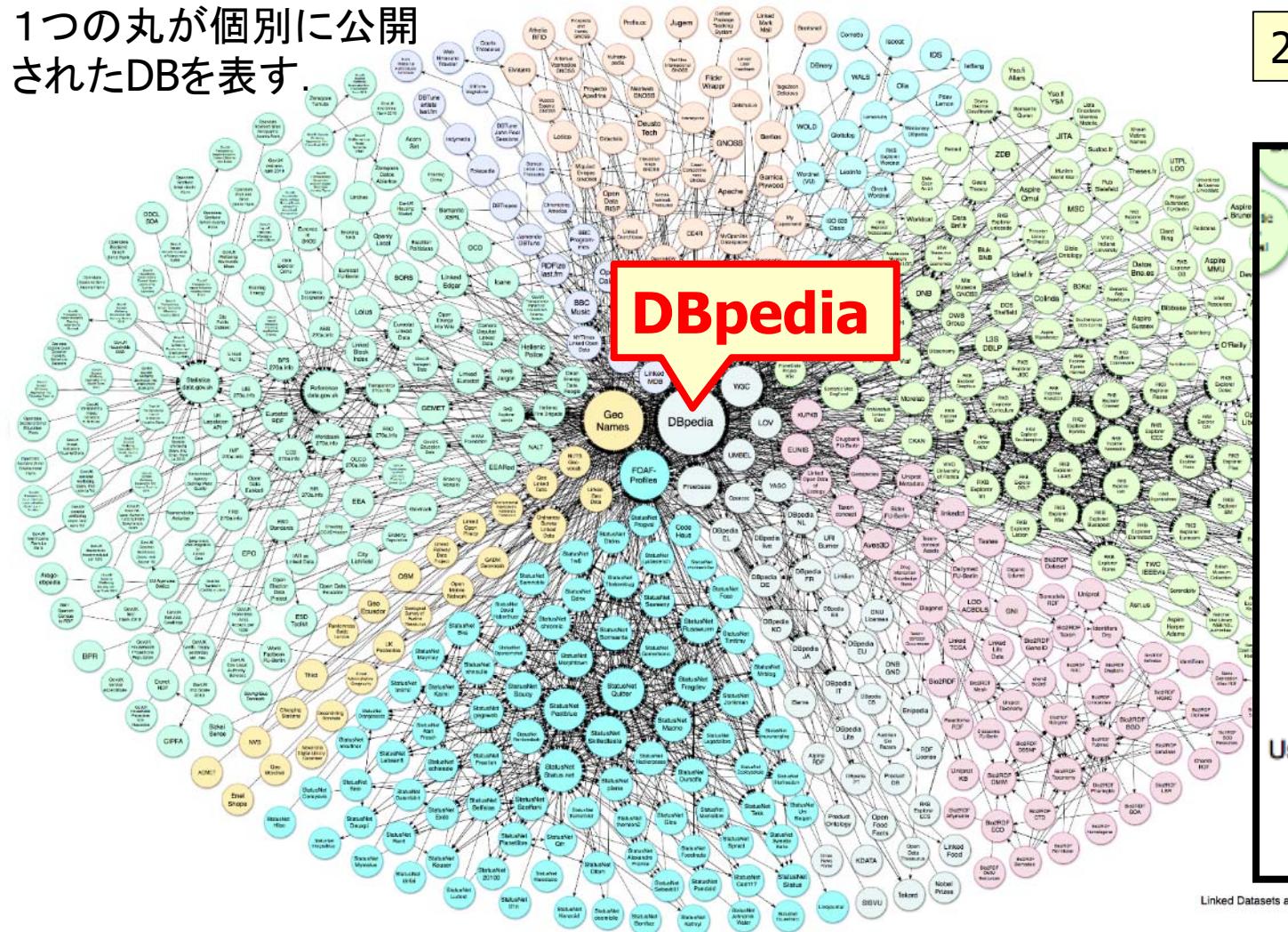
Log in using OpenID  Request new password

# LOD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/>

# 既に公開・リンクされているLOD ～LODクラウド～



Domains	# of dataset	%
Government	183	18.05
Publications	96	9.47
Life sciences	83	8.19
User-generated content	48	4.73
Cross-domain	41	4.04
Media	22	2.17
Geographic	21	2.07
Social web	520	51.28
<b>Total</b>	<b>1014</b>	

# LODクラウドの最新版



## Legeru

Cross Domain

Geography

Government

Life Sciences

Linguistics

Media

Publications

Social Networking

User Generated

Incoming Links

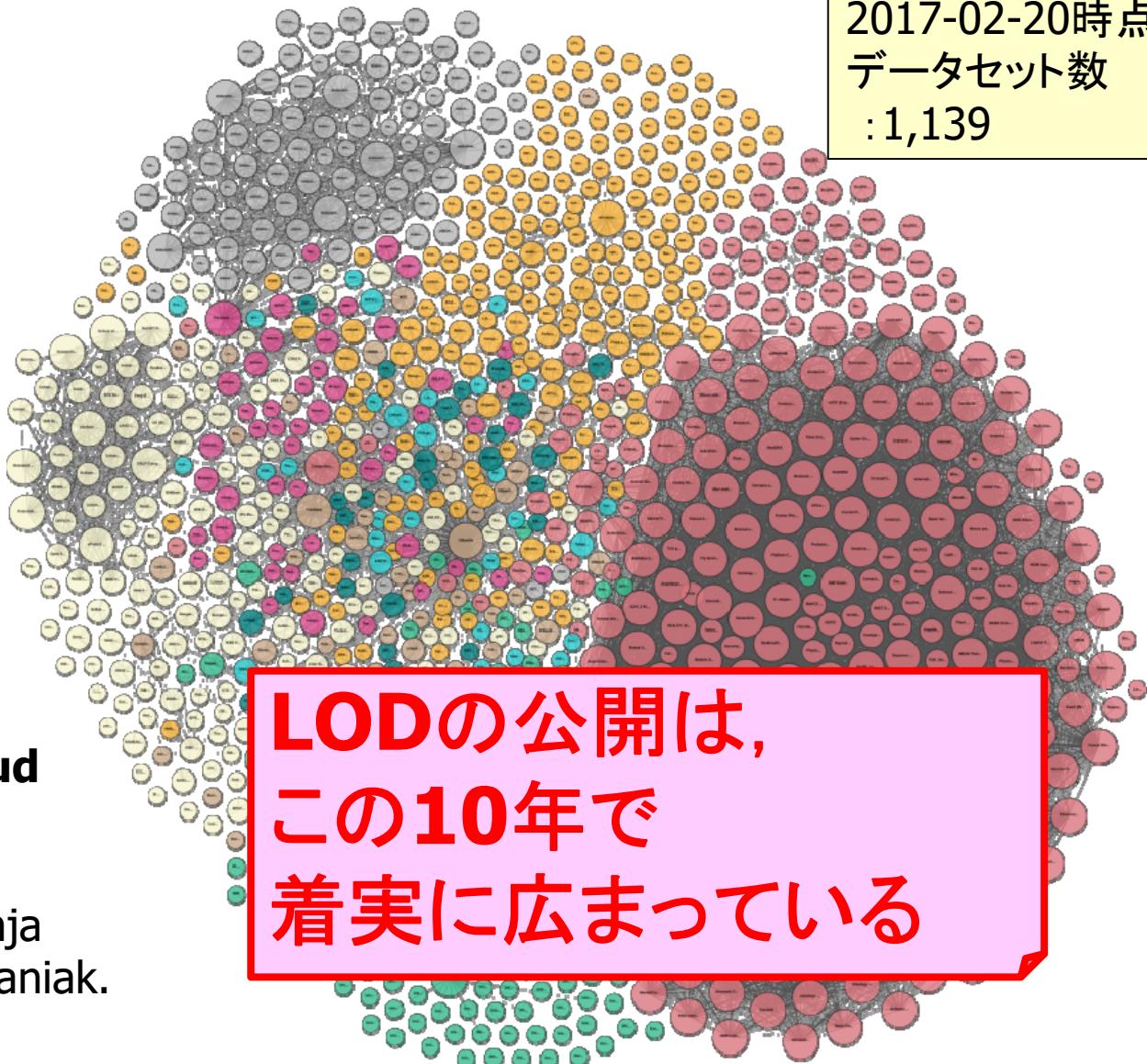
Outgoing Links

**Linking Open Data cloud  
diagram 2017,**

by Andrejs Abele, John P.  
McCrae, Paul Buitelaar, Anja  
Jentzsch and Richard Cyganiak.

<http://lod-cloud.net/>

2017-02-20時点  
データセット数  
: 1,139



# Linked Dataの基礎技術

## Linked Data

=Webの仕組みを用いて相互に“リンクされた”データ

# 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.hozo.jp>

オントロジー構築・利用環境「法造」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

(例) <http://www.hozo.jp>

**URI・IRI**

ロジ一構築・利用環: **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)

# Linked Dataの要素技術



## ■ **URI(Uniform Resource Identifier)**

- 一定の記述方式によって、リソース（モノやコト）を示すための識別子。リソースの場所を表す**URL**を拡張したもの。

## ■ **IRI(Internationalized Resource Identifier)**

- URIを拡張し、漢字、日本語文字など多言語に対応したもの。

## ■ **RDF(Resource Description Framework)**

- リソースを記述するための統一的枠組み。
- 主語(subject)、述語(predicate)、目的語(object)の3つ組み(トリプル)の組み合わせで、リソースに関する情報を表現する。

## ■ **オントロジー(Ontology)**

- データを記述するための語彙やスキーマを定義したもの

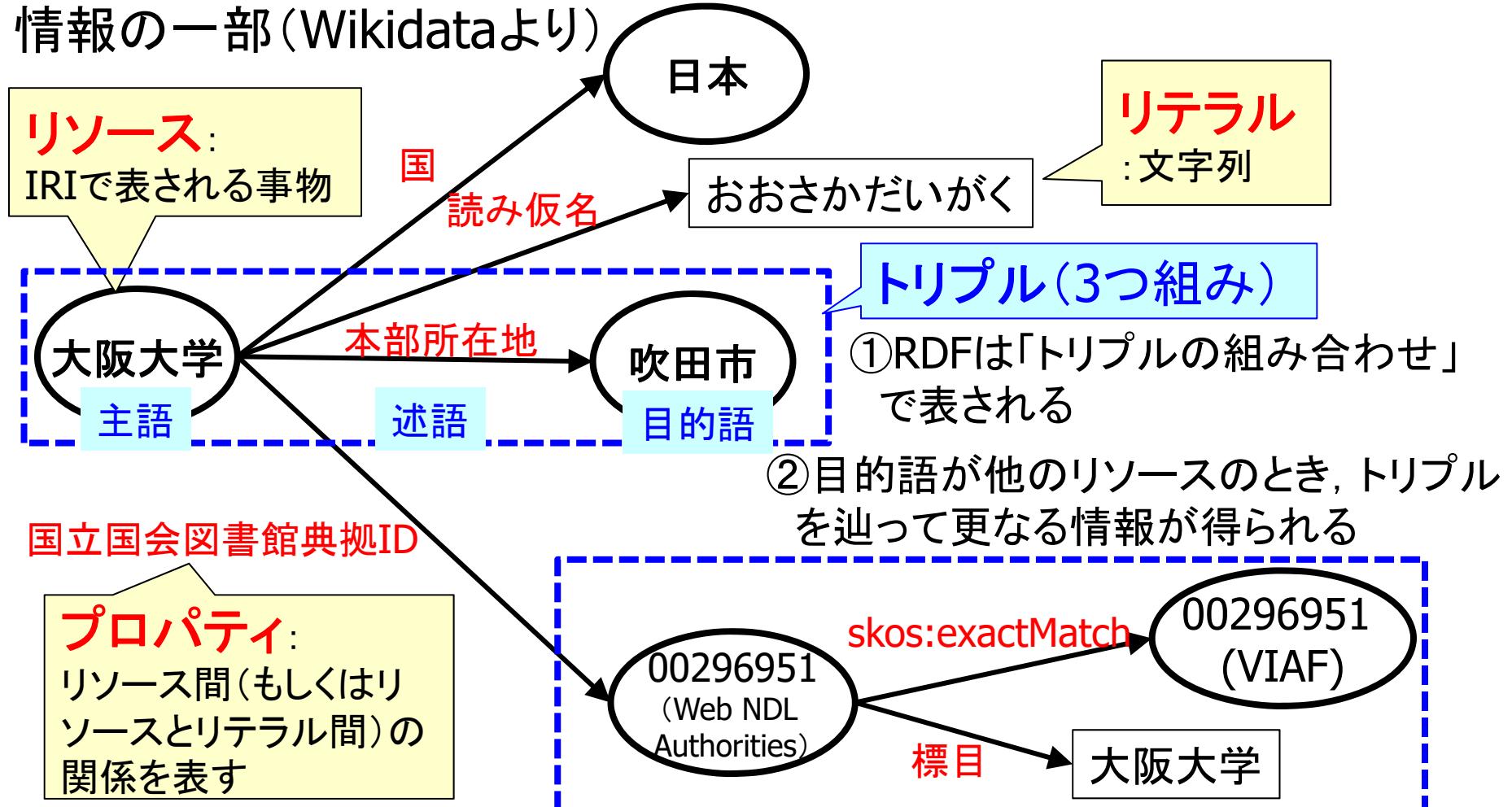
## ■ **SPARQL**

- RDFで記述されたデータを検索するためのクエリ言語

# Linked Data (RDF) の表現例



<http://www.wikidata.org/entity/Q651233> というIRIから得られる情報の一部(Wikidataより)



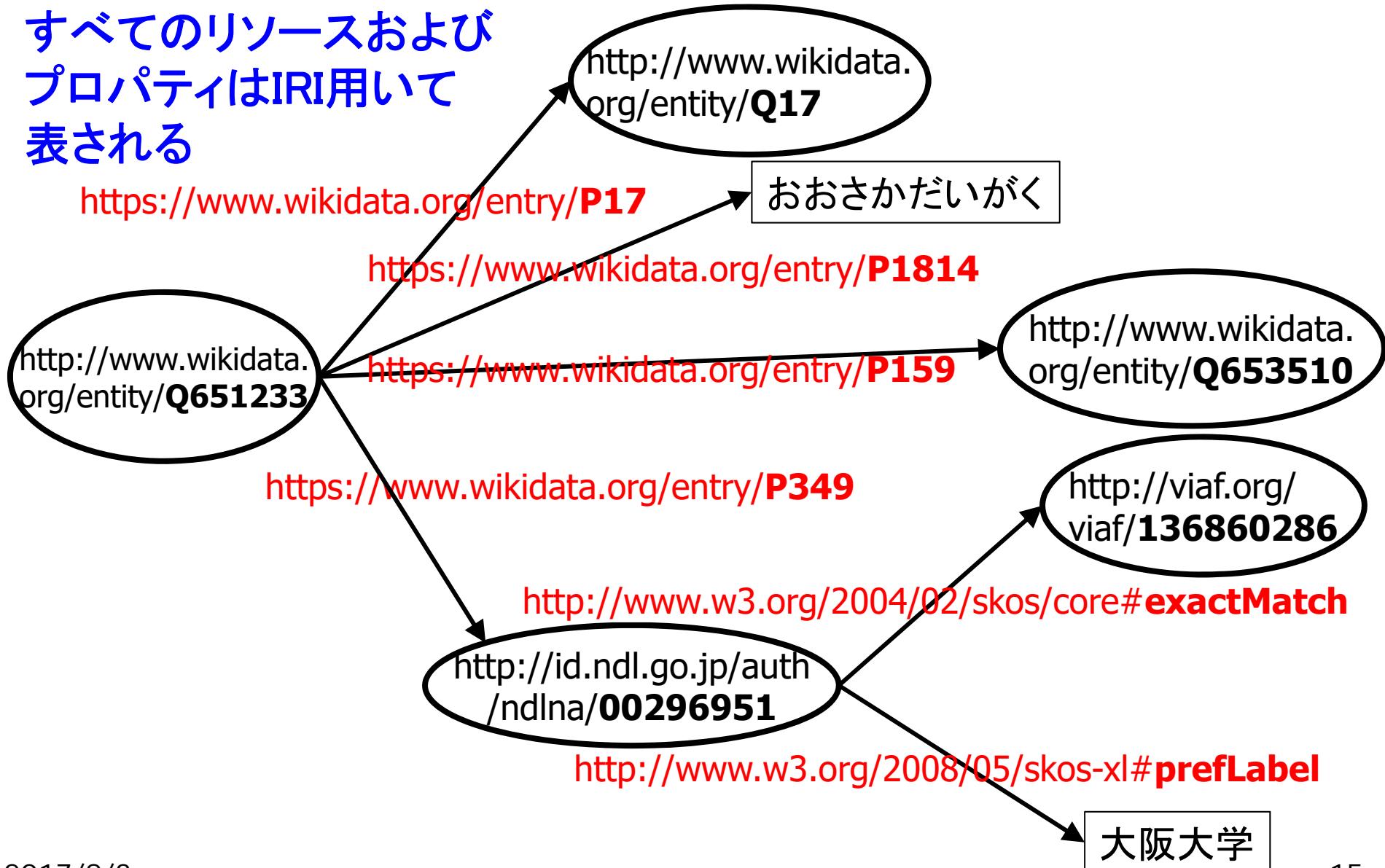
注: この図中においてリソースとプロパティは名前を示しているが、正確にはすべてのリソースおよびプロパティがIRI用いて表されている。

2017/8/3

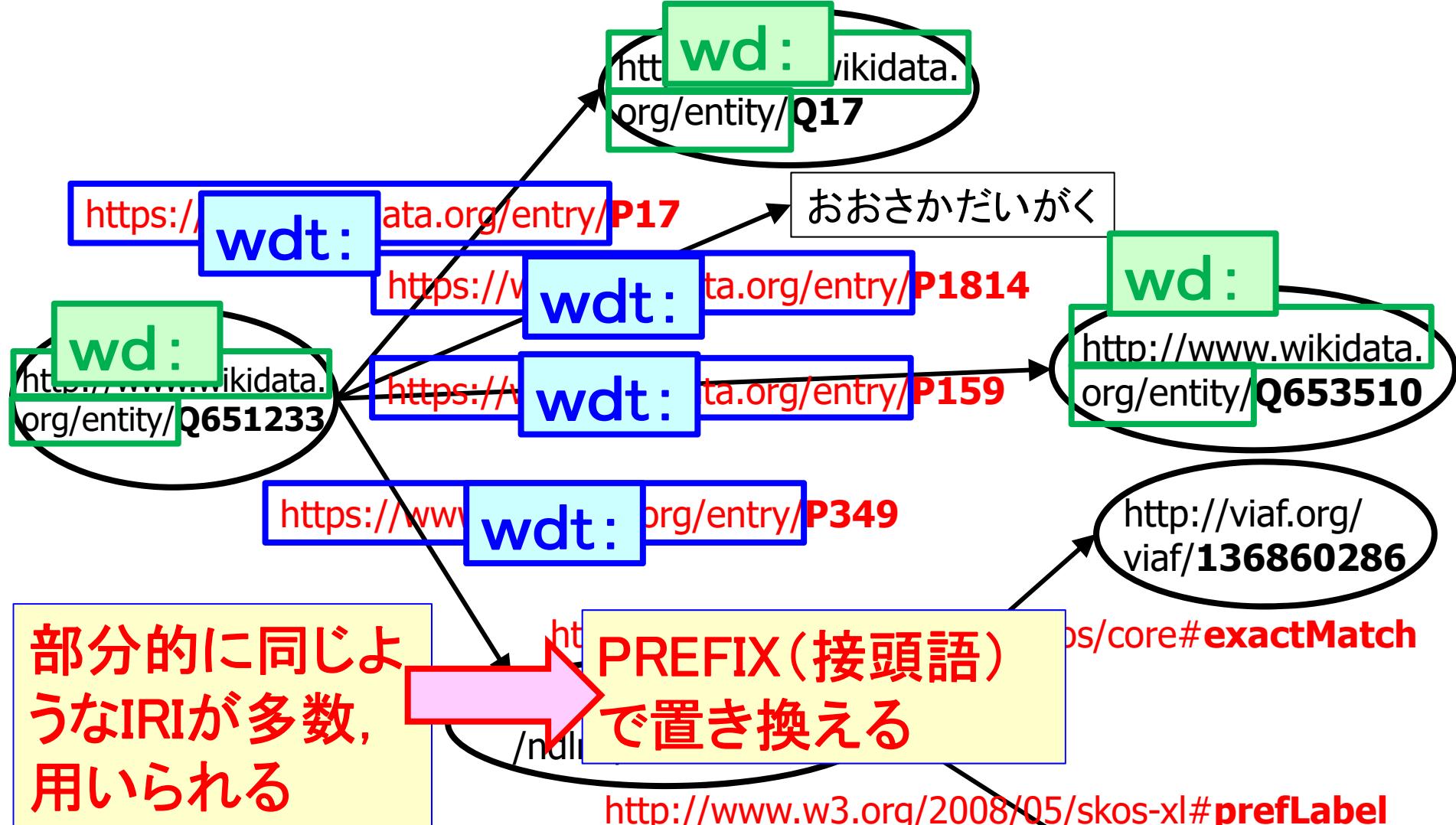
# Linked Data (RDF) の表現例



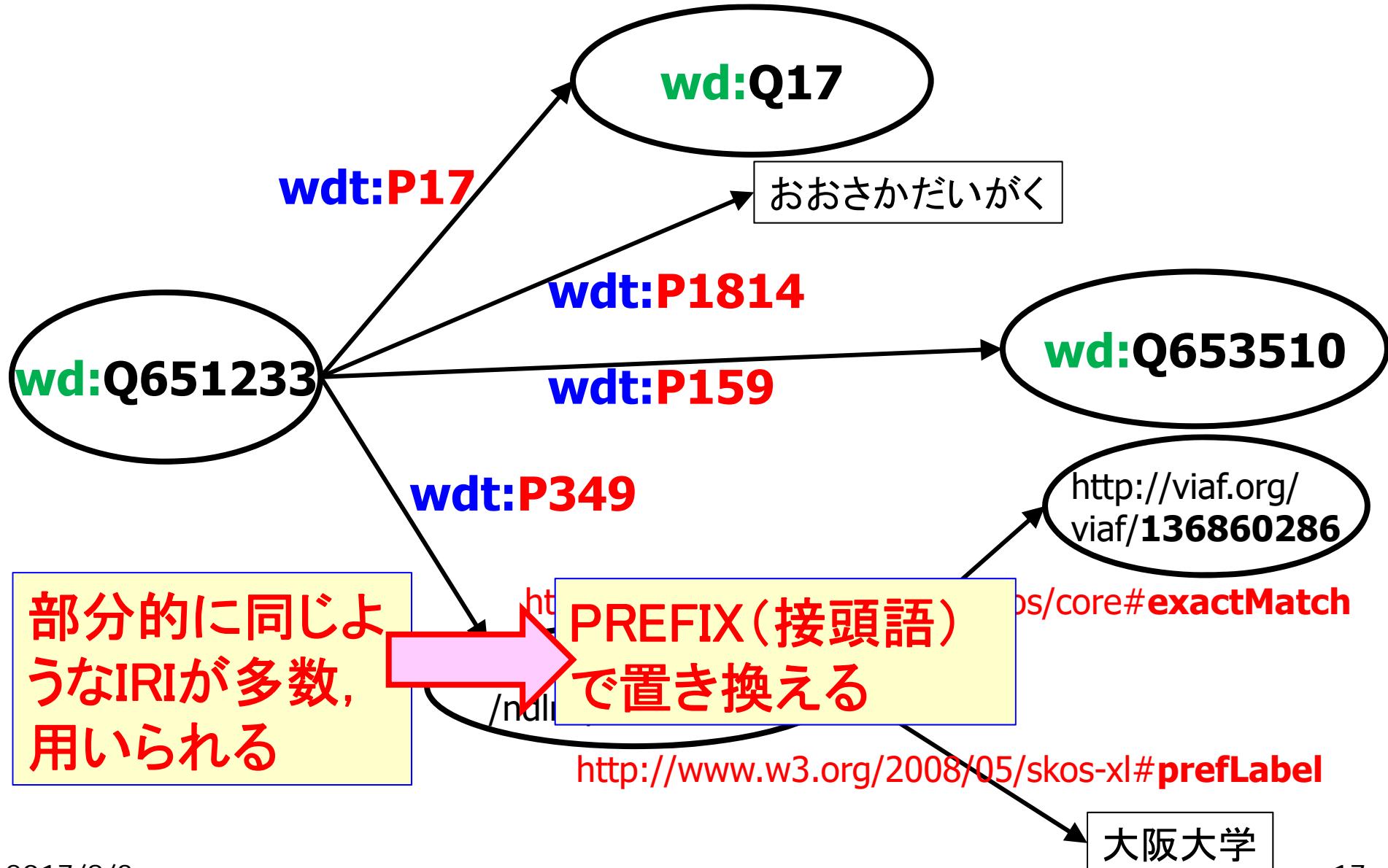
すべてのリソースおよび  
プロパティはIRI用いて  
表される



# PREFIX : IRIの省略表現



# PREFIX : IRIの省略表現



# Linked Data/LOD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の公開例

Wikidata(ウィキデータ)を例として、  
公開されているLOD眺めてみる。

1. IRIによるアクセス
2. SPARQLエンドポイントの利用

# Wikidata (<http://wikidata.org/>)



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

日本語 Koujikozaki ログアウト

メインページ 議論 閲覧 編集 履歴表示 Wikidata内を検索

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  
コモンズ  
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

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・

ようこそ!

Wikidataは、人間とコンピューターの双方が同時に参照・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。

Wikimediaの姉妹プロジェクト群（WikidataやWikivoyage、Wikisources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



W 大阪大学 - Wikipedia × 保護された通信 | https://ja.wikipedia.org/wiki/大阪大学

ページ ノート

## 大阪大学

大阪大学（おおさかだいがく、英語: Osaka University）は、大阪府吹田市山田丘1番1号に本部を置く日本の国立大学である。1931年に設置された。大学の略称は阪大（はんだい）。

文部科学省が実施しているスーパーグローバル大学事業のトップ型指定校である。

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  
ウィキデータ項目  
このページを引用

4.1.3.5 学部附属施設  
4.1.3.6 大学院研究科附属施設  
4.1.3.7 その他  
4.2 研究  
4.2.1 21世紀COEプログラム

閲覧 編集 履歴表示 Wikipedia内を検索

ログインしていません トーク 投稿記録 アカウント作成 ログイン

座標: 北緯34度49分7秒 東経135度31分26秒

大阪大学

大学設置 1931年  
創立 1838年  
学校種別 国立  
設置者 国立大学法人大阪大学  
本部所在地 大阪府吹田市山田丘1番1号  
北緯34度49分7秒 東経135度31分26秒  
キャンパス 吹田（大阪府吹田市）  
豊中（大阪府豊中市）  
箕面（大阪府箕面市）  
学部 人間科学部  
文理学部

Wikipediaの各記事から  
対応する  
**Wikidata項目へのリンク**

2017/8/3

# 参照解決可能IRIによる公開



- 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  
=システムによる処理(リンク解析等)が可能
- コンテンツネゴシエーションに対応していると、アクセスするツールに応じた形式での取得が可能
  - 例) Webブラウザでアクセス → HTML  
Linked Dataブラウザでアクセス → RDF/XML など
  - 拡張子を変えることで、取得形式を変えることも可能
    - 例) .rdf → RDF/XML, .ttl → Turtle, .nt → N-triple  
.json → JSON/LD など
  - LODブラウザの例
    - (Yet Another) Linked Data Browser  
<http://www.kanzaki.com/works/2014/pub/ld-browser>

# IRIによるLODへのアクセス例



さまままな言語での  
「ラベル」、「概要説明」、「別名」

Osaka University (Q651233)

higher education institution in Osaka Prefecture, Japan

In more languages [Configure](#)

Language	Label	Description	Also known as
English	Osaka University	higher education institution in Osaka Prefecture, Japan	
Japanese	大阪大学	No description defined	千里門 豊中キャンパス 大阪医学校 大阪帝国大学 大阪帝大 大阪帝國大学 大阪大学豊中キャンパス ISIR 府立大阪医科大学 吹田キャンパス 阪大 大阪帝國大學 産業科学研究所 大阪大学吹田キャンパス
?ucināguci	No label defined	No description defined	
Korean	오사카 대학	No description defined	오사카 대학교 오사카대 오사카 제국대학 오사카대학

All entered languages

Statements

instance of

Wikidataにおいて  
「大阪大学(Q651233)」にアクセスした例  
<http://www.wikidata.org/entity/Q651233>

# IRIによるLODへのアクセス例



Osaka University - Wiki... 保護された通信 https://www.wikidata.org/wiki/Q651233 Kozaki

Statements

述語(プロパティ) (Red Box)

- instance of
- image
- inception
- name in kana
- country
- located in the administrative territorial

目的語(オブジェクト) (Blue Box)

- university
  - 0 references
  - Osaka University Faculty of Medicine.jpg ↗
    - 1 reference
- 1931
  - 1 reference
- おおさかだいがく
  - 0 references
- Japan
  - 0 references
- Osaka I...

他のリソースへのリンク (Pink Box)

2017/8/3 24

# IRIによるLODへのアクセス例



The screenshot shows a web browser window with the URL <https://www.wikidata.org/wiki/Special:EntityData/Q651233.ttl>. The page displays RDF triples in Turtle format. The triples describe the entity wd:Q651233, which is a schema:Dataset. The triples include information about the dataset's schema:about, cc:license, schema:softwareVersion, schema:version, schema:dateModified, wikibase:statements, wikibase:identifiers, and wikibase:sitelinks. Below this, another set of triples describes the entity wd:Q651233 as a wikibase:Item, providing its rdfs:label, skos:prefLabel, schema:name, and language variants in zh-hans, zh-hant, and zh-hk.

```
@prefix wdno: <http://www.wikidata.org/prop/novalue/> .
@prefix skos: <http://www.w3.org/2004/02/skos/core#> .
@prefix schema: <http://schema.org/> .
@prefix cc: <http://creativecommons.org/ns#> .
@prefix geo: <http://www.opengis.net/ont/geosparql#> .
@prefix prov: <http://www.w3.org/ns/prov#> .

wdata:Q651233 a schema:Dataset ;
    schema:about wd:Q651233 ;
    cc:license <http://creativecommons.org/publicdomain/zero/1.0/> ;
    schema:softwareVersion "0.1.0" ;
    schema:version "504812667"^^xsd:integer ;
    schema:dateModified "2017-06-22T16:56:13Z"^^xsd:dateTime ;
    wikibase:statements "27"^^xsd:integer ;
    wikibase:identifiers "15"^^xsd:integer ;
    wikibase:sitelinks "32"^^xsd:integer .

wd:Q651233 a wikibase:Item ;
    rdfs:label "大阪大学"@zh-hans ;
    skos:prefLabel "大阪大学"@zh-hans ;
    schema:name "大阪大学"@zh-hans ;
    rdfs:label "大阪大學"@zh-hant ;
    skos:prefLabel "大阪大學"@zh-hant ;
    schema:name "大阪大學"@zh-hant ;
    rdfs:label "大阪大學"@zh-hk ;
    skos:prefLabel "大阪大學"@zh-hk ;
    schema:name "大阪大學"@zh-hk ;
```

<http://www.wikidata.org/entity/Q651233.ttl>

とすると、**Turtle**形式でデータ取得が可  
(Wikidataは、ほかに、.rdf, .nt, .json等に対応)

# Linked Dataブラウザによる LODへのアクセス例



https://www.wikidata.org ×

www.kanzaki.com/works/2014/pub/ld-browser?u=https%3A%2F%2Fwww.wikidata.org%2Fwiki%2FSpecial%3AEntityData%2FQ651233&n=2

https://www.wikidata.org/wiki/Special:EntityData/Q651233

wdata:Q651233の説明記述

rdf:type	schema:Dataset
schema:about	wd:Q651233 [+] ( 大阪大学)
cc:license	<http://creativecommons.org/publicdomain/zero/1.0/> [+] +
schema:softwareVersion	"0.1.0"
schema:version	504812667
schema:dateModified	"2017-06-22T16:56:13Z" ^^^xsd:dateTime
wikibase:statements	27
wikibase:identifiers	15
wikibase:sitelinks	32

グラフ描画 (9トリブル) ● PNG ○ SVG ( 選択描画データにこのリソースを追加)

大阪大学

wd:Q651233 [+] の説明記述

rdf:type	wikibase:Item
rdfs:label (41 more)	"大阪大学" @zh-hans
skos:prefLabel (41 more)	"大阪大学" @zh-hans
schema:name (41 more)	"大阪大学" @zh-hans
schema:description (6)	"higher education institution in Osaka Prefecture, Japan" @en

## (Yet Another) Linked Data Browser

<http://www.kanzaki.com/works/2014/pub/ld-browser>

を利用した閲覧例

→URIの欄に閲覧したいデータ  
のIRI(URI)を入力する

# WikidataのSPARQLエンドポイント



ここに、クエリを入れる

```
1 select distinct * where {  
2   <http://www.wikidata.org/entity/Q651233> ?p ?o . }
```

- ・ブラウザからのアクセス  
<https://query.wikidata.org/>
- ・プログラムからのアクセス  
<https://query.wikidata.org/sparql>

クエリ結果の表示

p	o
wdt:P3876	<a href="#">wd:Q6376024</a>
wdt:P910	<a href="#">wd:Q8706109</a>
wdt:P3500	11948
p:P268	<a href="#">wds:Q651233-944A4822-8B0C-48FF-92C6-C35BBE3AE634</a>

# まとめ:Linked Open Data

- Linked Dataは、Web上で公開されたデータを“つなぐ”仕組み
- Web上に公開された様々なデータを“つなげて”，1つの大きなデータベースへとする
- LOD(**Linked Open Data**)  
= Linked Dataとして公開されたオープンデータ
- データを“つなげる”ことで新たな価値が生まれることが期待される
- いろんなデータをつなぐことで、「おもしろい」組み合わせが見つかるかも！？