

SPARQL演習会

トルヴェェ アントワン / 後藤 孝行

2015年2月5日

SRPビル2階

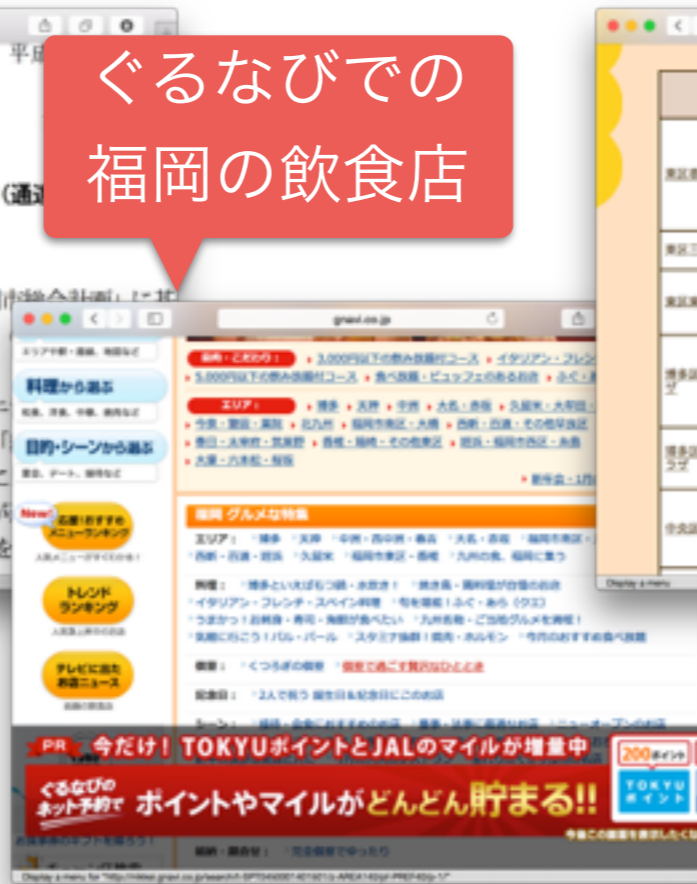
背景セマンティックウエツブ について

既存のウェブの課題：検索

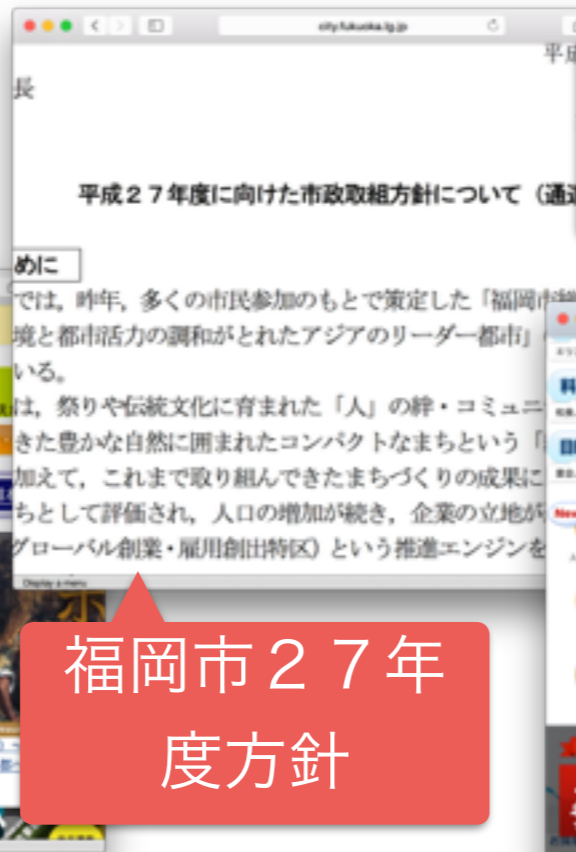
福岡市ホームページ



ぐるなびでの福岡の飲食店



福岡市27年度方針



福岡市のこどもプラザ一覧

施設名	住所	電話・FAX番号	休館日	アクセス	特徴	地図
東区東区子どもプラザ	福岡市東区東区2-1-1	092-980-3003	毎週月曜、毎月第3日曜	西鉄東区線「東区駅」徒歩10分	西鉄東区線「東区駅」徒歩10分	
東区三宮子どもプラザ	東区三宮5-1-40	092-980-4207	毎週水曜、毎月第2日曜	西鉄東区線「三宮駅」徒歩5分	西鉄東区線「三宮駅」徒歩5分	
東区東区子どもプラザ	東区東区1-1-1	092-290-0300	毎週水曜、毎月第3日曜	西鉄東区線「東区駅」徒歩10分	西鉄東区線「東区駅」徒歩10分	
東区山王子どもプラザ	東区山王5-1-1	092-672-0006	毎週日曜、毎月第1日曜、第3日曜、第5日曜	西鉄東区線「山王駅」徒歩5分	西鉄東区線「山王駅」徒歩5分	
東区東区子どもプラザ	東区東区1-1-1	092-980-0711	毎週水曜、毎月第4日曜	西鉄東区線「東区駅」徒歩10分	西鉄東区線「東区駅」徒歩10分	
中央区子どもプラザ	中央区東区2-2-4	092-741-9964	毎週月曜（祝日の場合は休館日）、毎月第1日曜、第3日曜、第5日曜	地下鉄東区線「東区駅」徒歩10分	地下鉄東区線「東区駅」徒歩10分	

- インターネットは膨大の情報が閲覧可能
- が、正規化・構造化されていないため、**検索は困難**



セマンティックウェブウェブによる 解決策

ページやページの中にあるエレメントに
属性情報を付加する

属性情報等は検索可能とする

標準技術を定義する

ページ属性情報

about: 福岡市
 type: ホームページ
 lastupdate: 2015-2-1



about: 福岡市
 type: 飲食店一覧
 lastupdate: 2015-2-5



about: 福岡市
 type: 市営方針資料
 fiscal year: 2015



about: 福岡市
 type: こどもプラザ一覧
 type: 子育て支援情報

施設名	住所	電話・FAX 番号	休館日	アクセス	備考	地図
東区東区子どもプラザ	福岡市東区 東区東区2-5-1 セビオプラザ 西館2階	092-663-3263	毎週月曜、毎月 第3日曜	西鉄東横線「西区東駅」、 下車徒歩5分。西鉄バス 「西区東駅」下車徒歩5分。 西鉄バス「東区東駅」下車 徒歩5分		
東区三宮子どもプラザ	東区三宮5-1-40 1階	092-662-6267	毎週水曜、毎月 第2日曜	西鉄東横線「三宮駅」、下 車徒歩5分		
東区東区子どもプラザ	東区東区1-1-1 1F 中の プラザ2階	092-292-6320	毎週水曜、毎月 第3日曜	西鉄バス「東区東駅」、下 車徒歩1分		
博多区山王子どもプラザ	博多区山王 5-13-10 博多区センター併設	092-472-6008	毎週月曜、毎月 第1土曜、毎月 第4土曜	西鉄バス「山王台駅」、 徒歩5分。西鉄バス「山王 台駅」下車徒歩5分		
博多区博多子どもプラザ	博多区博多 1-4-11 1 階	092-662-6711	毎週金曜、毎月 第4土曜	西鉄バス「博多駅」、下 車徒歩5分。西鉄バス「 博多駅」下車徒歩5分		
中央区子どもプラザ	中央区東区 2-2-4	092-741-3664	毎週月曜（祝日 の場合は翌日）、 毎月第1日（日 曜の場合は翌日）、 12月28日	地下鉄有明線「東区」、下 車徒歩5分。西鉄バス「 東区2丁目」、下車徒歩5 分		

ページエッセメント属性情報

about: 福岡市
 type: ホームページ
 lastupdate: 2015-2-1



about: 福岡市
 type: 飲食店一覧
 lastupdate: 2015-2-5



about: 福岡市
 type: 市営方針資料
 fiscal year: 2015



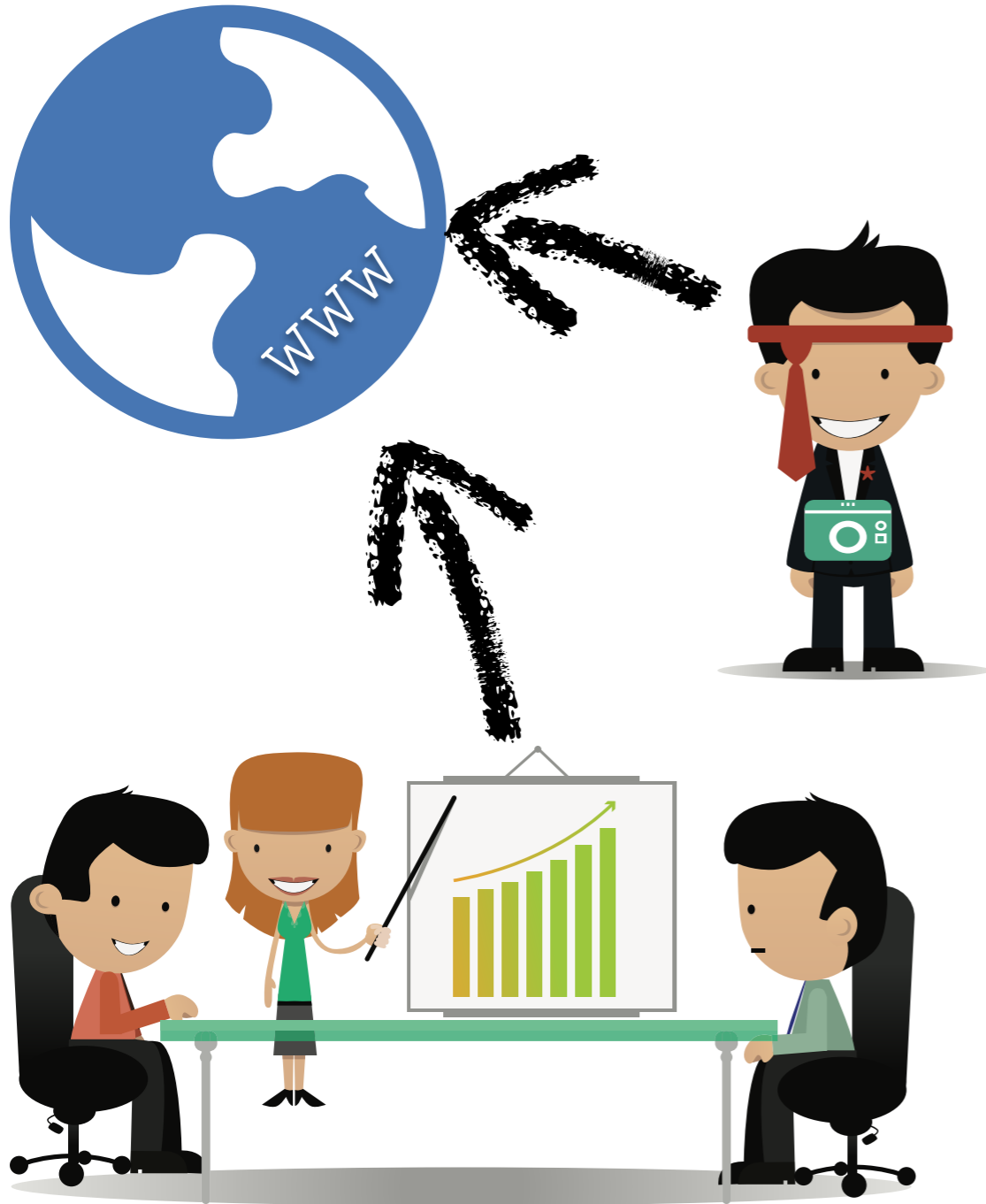
施設名	住所	電話・FAX 番号	休館日	アクセス	画像	地図
東区香椎子どもプラザ	福岡市東区香椎駅前2-5-2-1 セザアテラス西鉄香椎2階	092-663-3263	毎週月曜、毎月第3日曜	西鉄貝塚線「西鉄香椎」下車徒歩すぐ、西鉄バス「西鉄香椎駅前」下車徒歩すぐ、西鉄バス「香椎」下車徒歩3-5分、JR鹿児島本線・香椎線「香椎駅」下車徒歩3分		
東区三苦子どもプラザ	東区三苦5-1-40 1階	092-692-6267	毎週水曜、毎月第2日曜	西鉄貝塚線「三苦駅」下車徒歩5分		
東区東浜子どもプラザ	東区東浜1-1-1 ゆめタウン博多2階	092-292-5320	毎週水曜、毎月第3金曜	西鉄バス「東出三丁目」下車徒歩1分		
東区山王子どもプラザ	博多区山王1-1-1 ゆめタウン博多センター併設	092-472-6006	毎週日曜、毎月末日(土・日曜、祝日のときは、直後の土・日曜、祝日以外の日)、12月28日	西鉄バス「山王公園前」「山王一丁目」下車徒歩5分		
博多区博多南子どもプラザ	博多区竹丘1-4-11 1階	092-592-9711	毎週金曜、毎月第4土曜	西鉄バス「JR南福岡駅」下車徒歩6分、JR鹿児島本線「南福岡駅」下車徒歩6分		
中央区子どもプラザ	中央区長浜2-2-4	092-741-3564	毎週月曜(祝日の場合は翌日)、毎月末日(日曜の場合は閉館)、12月28日	地下鉄空港線「赤坂」下車徒歩6分、西鉄バス「長浜二丁目」下車徒歩すぐ		

about: 福岡市
 type: こどもプラザ一覧
 type: 子育て支援情報

東区三苦子どもプラザ	東区三苦5-1-40 1階	092-692-6267	毎週水曜、毎月第2日曜	西鉄貝塚線「三苦駅」下車徒歩5分		
------------	---------------	--------------	-------------	------------------	--	--

about: 東区東浜子どもプラザ
 type: こどもプラザ
 type: 子育て支援情報
 address: 東区東浜1-1-1ゆめタウン博多2階
 telephone: 092-292-5320

更に！



団体（自治体等）、各個人が生成・蓄積する膨大のデータをウェブで公開・構造化したら、**より便利な社会へ！**

しかし

データの利用条件（ライセンス）を明確する必要がある
オープンデータ

データ間の関係性を表さないと探索が困難
リンクオープンデータ（LOD）

オープンデータ級



二次利用可能
ライセンス

- ★オープンライセンス
- ★★構造化データ
- ★★★オープン形式
- ★★★★RDFを利用
- ★★★★★リンクRDFを利用

エクセルよりも、
CSV等

本題 1 RDFについて

セマンティックウェブの データモデル

Resource Description Framework

リソース (=物)

記述

枠組み

W3Cはインターネットで使われている技術 (HTML、Javascript、css等) を標準化する団体。1994年に結成されました。



- **W3C**の標準技術
- 複数の標準技術の集合 (抽象的なセマンティックス、テキスト表現等)
- 現在のバージョンは**1.1**です

RDF1.1データモデル(1)

トリプル=物3つ

ちなみに、カープル=物2つ/シングルトン=物1つ

全てデータは**トリプル**で表現する

主語

例：福岡

術後

=プロパティ名

例：人口

目的語

=プロパティ値

例：1521881人

RDF1.1データモデル(2)

主語

型：IRI

術後

型：IRI

目的語

型：IRI又は
リテラル

- IRIとは？
 - International (国際) Resource (リソース) Identifier (識別子)
 - 人物を指すユニークな識別子
 - URLはIRIの例です (ウェブサイトを指す)
- リテラルとは？
 - 単なる文字列
 - RDFではタイプや元号を付けることもできる

RDFモデルのサンプル

主語

術後

目的語

型・言語

city.fukuoka.lg.jp

rdfs:type

schema:website

city.fukuoka.lg.jp

schema:about

Fukuoka City

ja

city.fukuoka.lg.jp

schema:dateVisited

2015-2-5

xsd:date



さっきの福岡市ホームページ

RDFのグラフ形式

主語

術後

目的語

型・言語

city.fukuoka.lg.jp

rdfs:type

schema:website

city.fukuoka.lg.jp

schema:about

Fukuoka City

ja

city.fukuoka.lg.jp

schema:dateVisited

2015-2-5

xsd:date



city.fukuoka.lg.jp

rdfs:type

schema:website

schema:about

“Fukuoka City”@ja

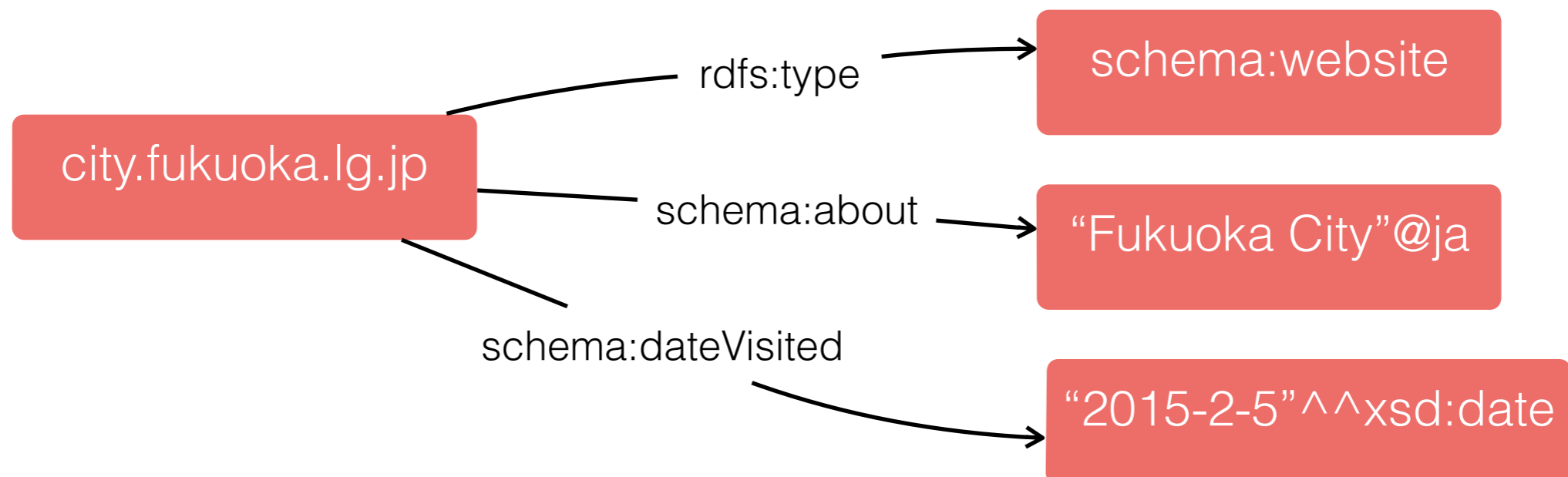
schema:dateVisited

“2015-2-5”^^xsd:date

RDFのグラフ形式

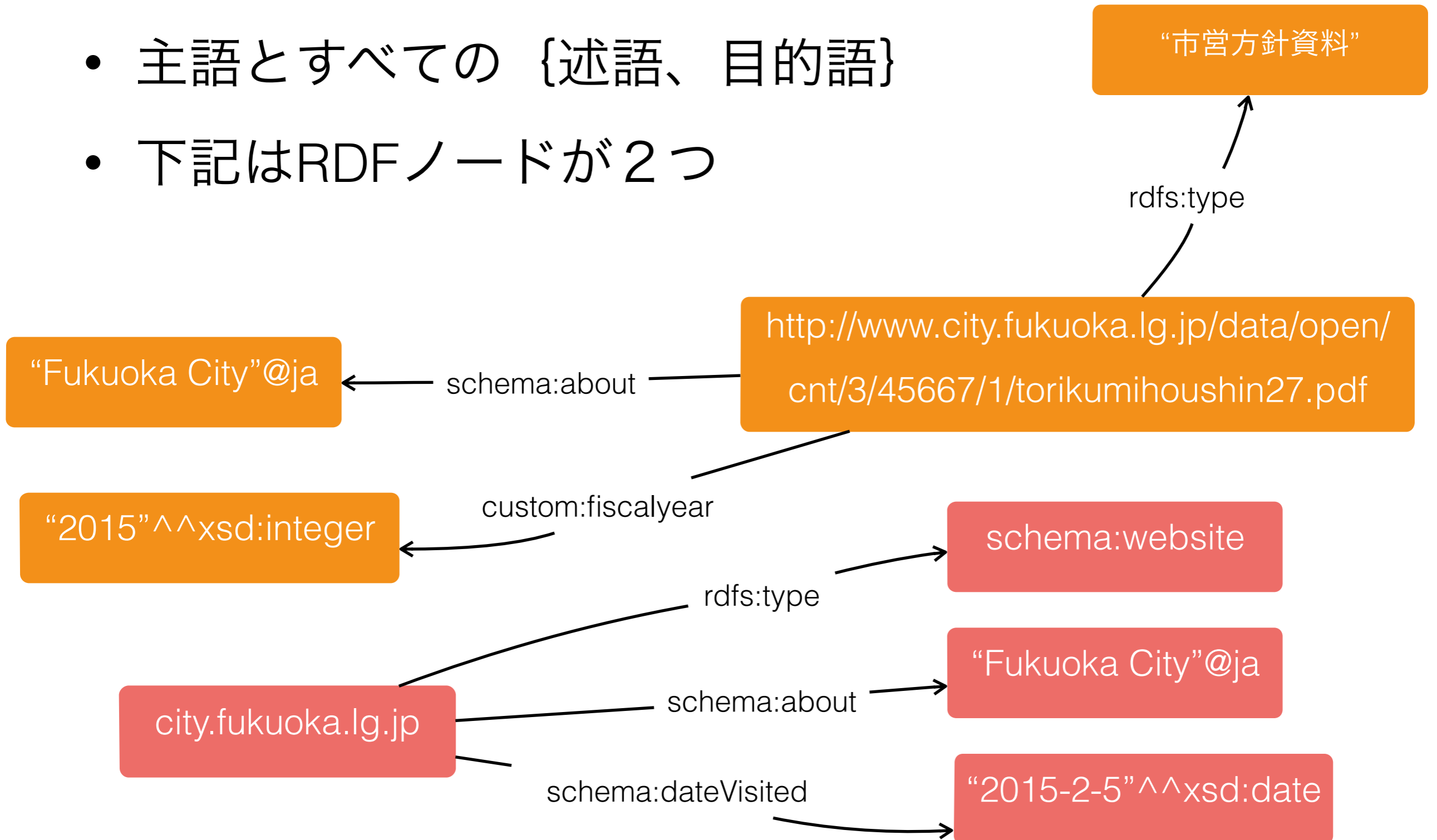
- 主語と目的語を矢印で繋ぐ
- 矢印は述語を表す（ラベル付き）
- 目的語の型や言語は下記のシンタックスで表す
 - 言語：“XXX”@lang
 - 型：“XXX”^^type

RDF/Turtle
シンタックス
(後で説明する)



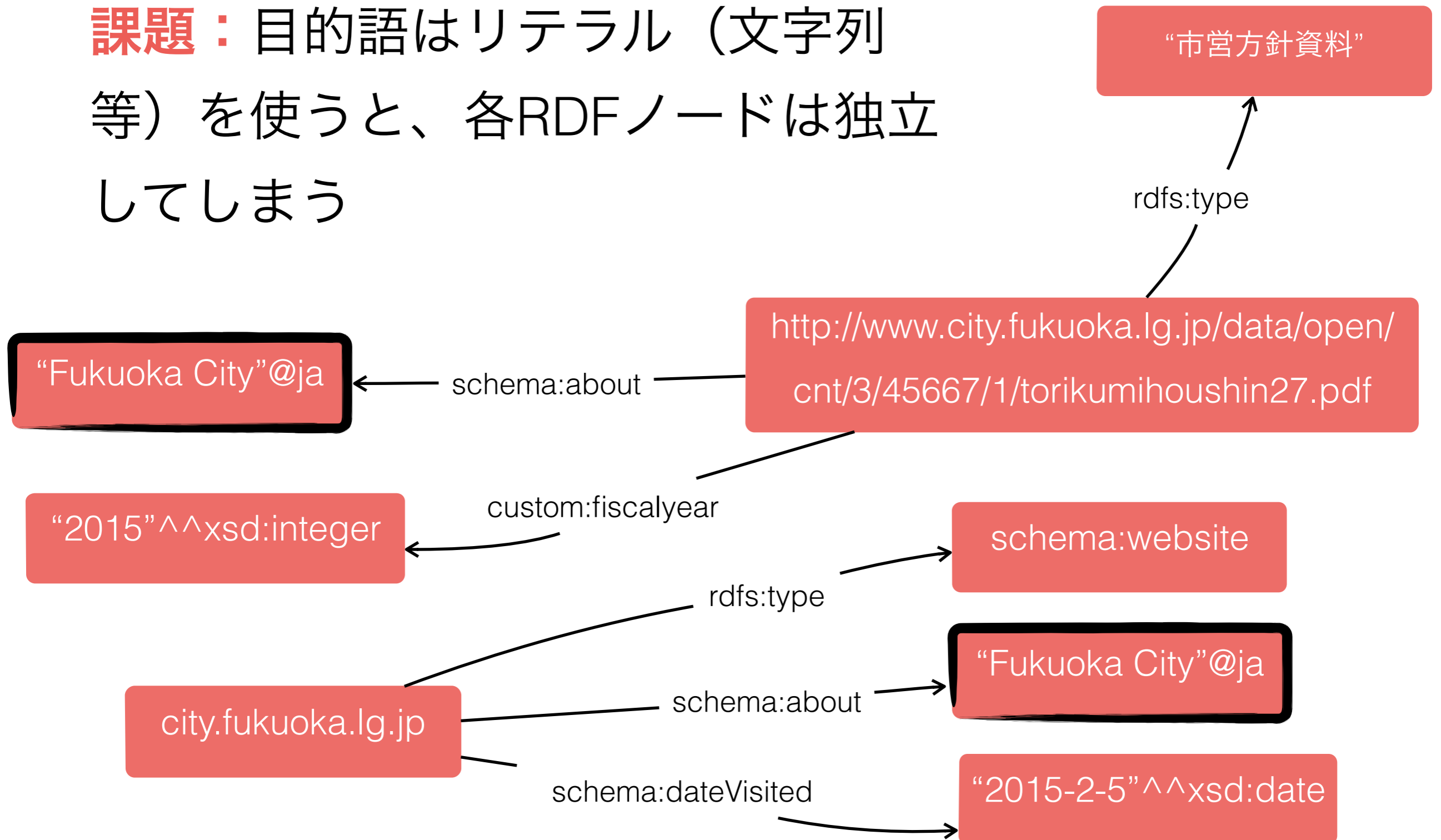
RDFノードとは

- 主語とすべての {述語、目的語}
- 下記はRDFノードが2つ



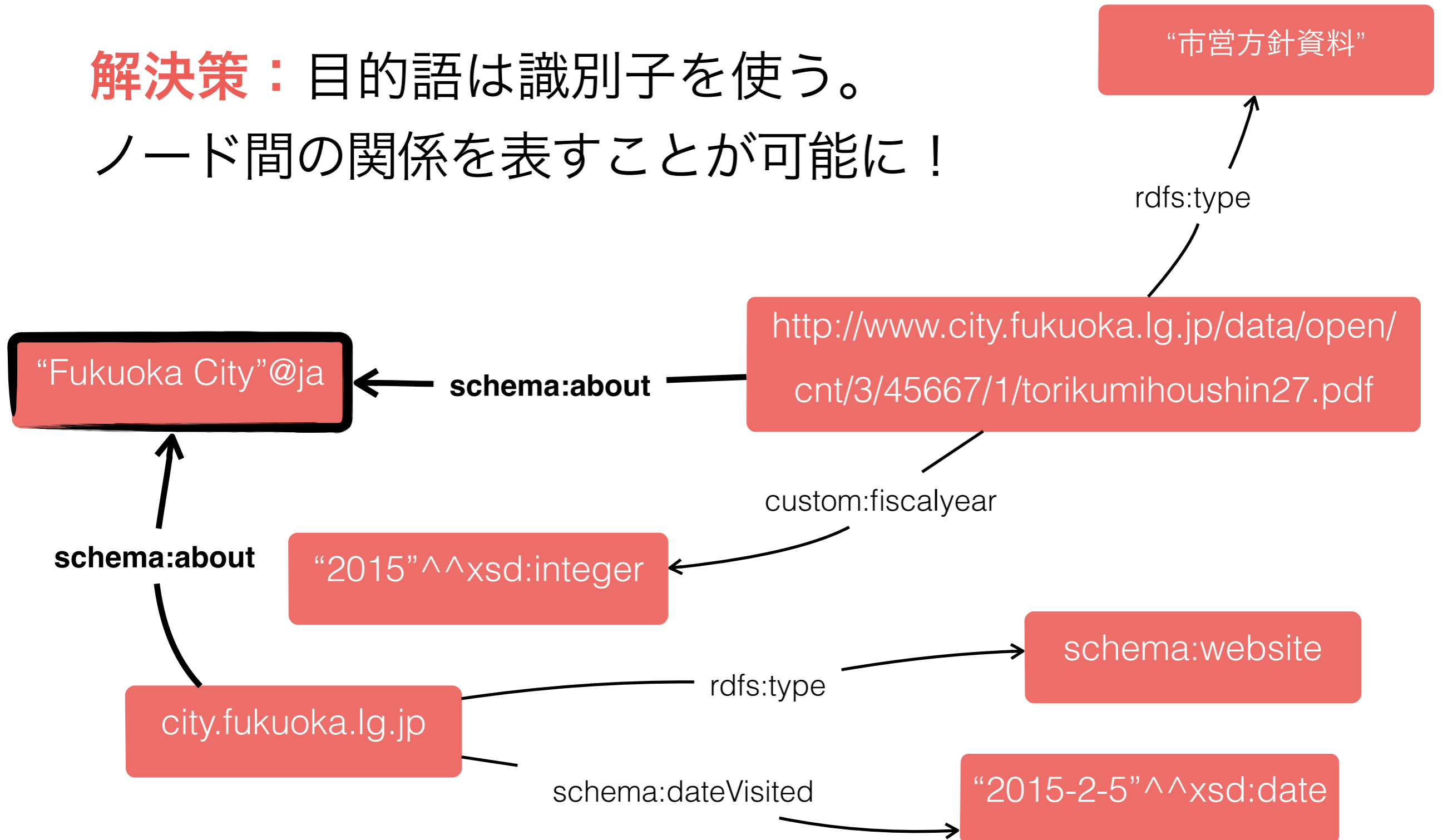
リンクRDF(1)

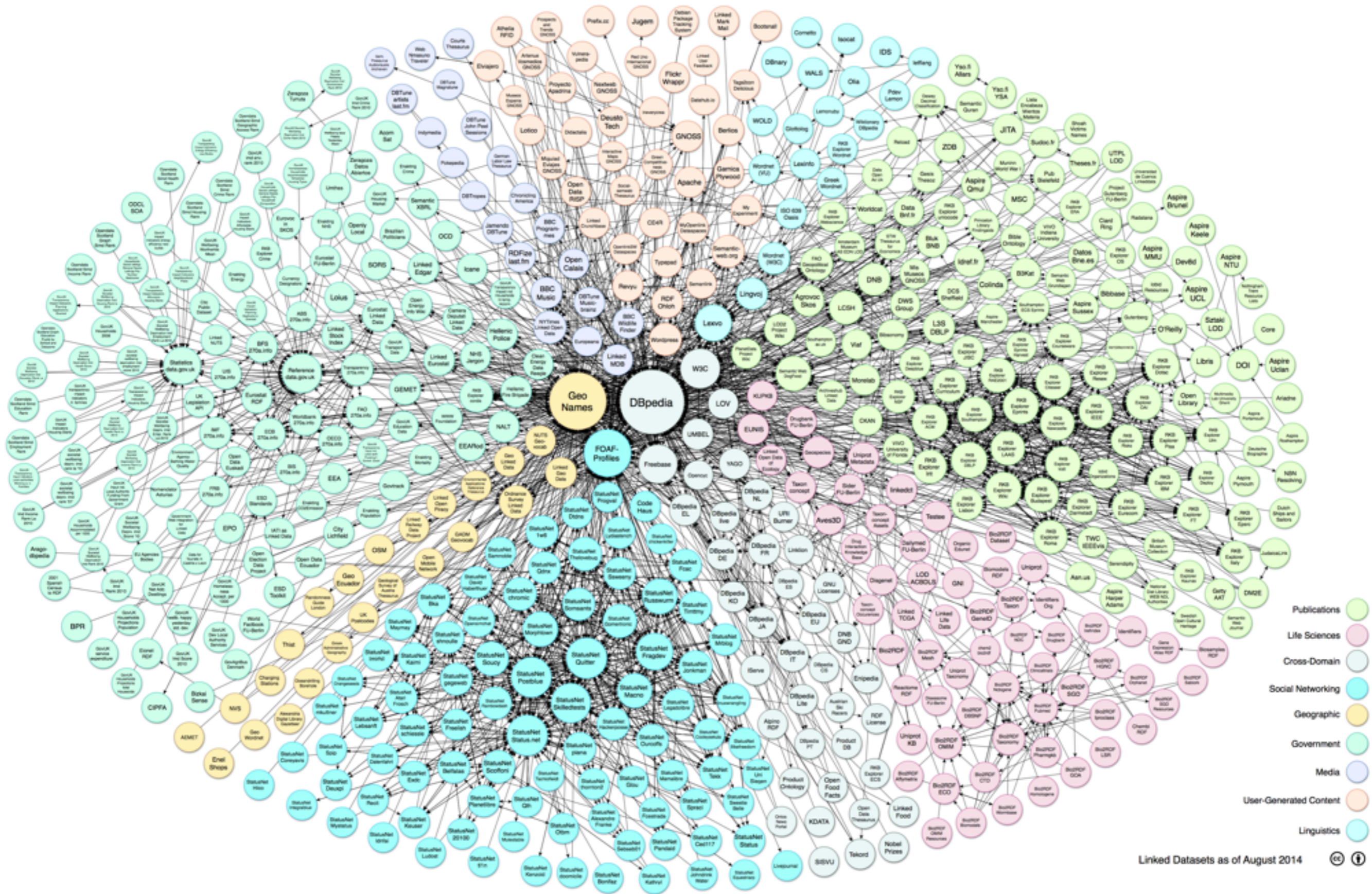
課題：目的語はリテラル（文字列等）を使うと、各RDFノードは独立してしまう



リンクRDF(2)

解決策：目的語は識別子を使う。
ノード間の関係を表すことが可能に！





<http://lod-cloud.net> より

語彙について

- RDFデータの使っている述語等は出来るだけ統一する必要がある
 - データベース間にリンクを貼るために
 - データベースを探索ために
- よく使われている語彙セットがあります
 - W3Cの標準語彙（少ない！）：RDF語彙集合、RDFスキーマ（RDFS）語彙集合
 - その他：Friend of a friend、Schema.org、Geo
- 人気語彙セットのランキング：<http://www.prefix.cc>
 - **名前空間**（接頭辞）で整理する

サンプルは次のスライド・・・



自分のパソコンでサイトを見
ても大丈夫ですよ！

語彙集合サンプル：RDFS

popular

「RDFS」語彙セット

- RDFスキーマ
- W3C標準語彙
- RDFデータベースの構造を記述するための語彙

RDFデータベースのスキーマはRDFで記述する

1. yago
2. rdf
3. foaf
4. dbp
5. dc
6. owl
7. rdfs
8. ont
9. dbo
10. onto
11. skos
12. geo
13. rss
14. gldp
15. sioc
16. sc
17. fb
18. geonames
19. xsd
20. gr
21. dcterms
22. dct
23. dbpedia
24. akt
25. org
26. commerce

prefix.ccによると人気
ランキングは7位
(2015/2/2の時点)

rdfs

 <http://www.w3.org/2000/01/rdf-schema#>  +1
-1
[Add alternative URI](#)

[ttl](#) [xml](#) [rdfa](#) [sparql](#) [txt](#) [json](#) [jsonld](#) [vann](#) | [lov](#) | [prefix.cc](#)

```
w3.org

@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .

<http://www.w3.org/2000/01/rdf-schema#> a owl:Ontology ;
  dc:title "The RDF Schema vocabulary (RDFS)" .

rdfs:Resource a rdfs:Class ;
  rdfs:isDefinedBy <http://www.w3.org/2000/01/rdf-schema#> ;
  rdfs:label "Resource" ;
  rdfs:comment "The class resource, everything." .

rdfs:Class a rdfs:Class ;
  rdfs:isDefinedBy <http://www.w3.org/2000/01/rdf-schema#> ;
  rdfs:label "Class" ;
  rdfs:comment "The class of classes." ;
  rdfs:subClassOf rdfs:Resource .

rdfs:subClassOf a rdf:Property ;
  rdfs:isDefinedBy <http://www.w3.org/2000/01/rdf-schema#> ;
  rdfs:label "subClassOf" ;
  rdfs:comment "The subject is a subclass of a class." ;
  rdfs:range rdfs:Class ;
  rdfs:domain rdfs:Class .

rdfs:subPropertyOf a rdf:Property ;
  rdfs:isDefinedBy <http://www.w3.org/2000/01/rdf-schema#> ;
  rdfs:label "subPropertyOf" ;
  rdfs:comment "The subject is a subproperty of a property." ;
  rdfs:range rdf:Property ;
  rdfs:domain rdf:Property .

rdfs:comment a rdf:Property ;
  rdfs:isDefinedBy <http://www.w3.org/2000/01/rdf-schema#> ;
  rdfs:label "comment" ;
  rdfs:comment "A description of the subject resource." ;
  rdfs:domain rdfs:Resource ;
  rdfs:range rdfs:Literal .

rdfs:label a rdf:Property ;
  rdfs:isDefinedBy <http://www.w3.org/2000/01/rdf-schema#> ;
  rdfs:label "label" ;
  rdfs:comment "A human-readable name for the subject." ;
  rdfs:domain rdfs:Resource ;
  rdfs:range rdfs:Literal .
```



このファイルは
なに？

RDFテキスト形式

- この発表の今まではRDFを抽象的にしか扱ってない
- が、RDFデータはテキスト式で格納する必要もある
- 標準フォーマットはRDF1.1スペックに述べられている：
 - RDF/Json
 - RDF/XML
 - RDF/Turtle (拡張し : ttl)

さっきのファイル
の形式

Turtleシンタックス(1)

SPARQLで使います
ので、覚えましょう！

- トリプルを効率よく表現できる
 - 特にRDF/XMLよりも文字数がすくなく、読みやすい
- トリプルをそのまま書きます (順序: 主語、述語、目的語)
 - ペリオッドで区切る「.」
 - リテラルはクォートで区切る「"”」 / IRIはアングルブラケット「<>」で区切る
 - 言語やリテラル型は「@」と「^^」を利用する (RDFグラフと同様)

```
<http://bodic.org/subject1> <http://www.w3.org/2000/01/rdf-schema#label> "名前"@ja.  
<http://bodic.org/subject1> <http://www.w3.org/2000/01/rdf-schema#label> "name"@en.  
<http://bodic.org/subject1>  
  <http://www.w3.org/2000/01/rdf-schema#type> <http://bodic.org/people>.
```

- 「PREFIX」キーワードで名前空間を定義し、IRIを短く書ける

```
PREFIX bodic: <http://bodic.org/>  
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>  
bodic:subject1 rdfs:label "名前"@ja.  
bodic:subject1 rdfs:label "name"@en.  
bodic:subject1 rdfs:type bodic:people.
```


Turtleシンタックス(2)

- 主語又は述語を合同し、更に短く書ける！
 - 主語の合同：セミコロンので {述語、目的語} を区切る
 - 述語の合同：コンマで目的語を区切る

```
PREFIX bodic: <http://bodic.org/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
bodic:subject1 rdfs:label "名前"@ja.
bodic:subject1 rdfs:label "name"@en.
bodic:subject1 rdfs:type bodic:people.
```



```
PREFIX bodic: <http://bodic.org/>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
bodic:subject1
  rdfs:label "名前"@ja,"name"@en;
  rdfs:type bodic:people.
```

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix owl: <http://www.w3.org/2002/07/owl#> .
@prefix dc: <http://purl.org/dc/elements/1.1/> .

<http://www.w3.org/2000/01/rdf-schema#> a owl:Ontology ;
    dc:title "The RDF Schema vocabulary (RDFS)" .

rdfs:Resource a rdfs:Class ;
    rdfs:isDefinedBy <http://www.w3.org/2000/01/rdf-schema#> ;
    rdfs:label "Resource" ;
    rdfs:comment "The class resource, everything." .

rdfs:Class a rdfs:Class ;
    rdfs:isDefinedBy <http://www.w3.org/2000/01/rdf-schema#> ;
    rdfs:label "Class" ;
    rdfs:comment "The class of classes." ;
    rdfs:subClassOf rdfs:Resource .

rdfs:subClassOf a rdf:Property ;
    rdfs:isDefinedBy <http://www.w3.org/2000/01/rdf-schema#> ;
    rdfs:label "subClassOf" ;
    rdfs:comment "The subject is a subclass of a class." ;
    rdfs:range rdfs:Class ;
    rdfs:domain rdfs:Class .

rdfs:subPropertyOf a rdf:Property ;
    rdfs:isDefinedBy <http://www.w3.org/2000/01/rdf-schema#> ;
    rdfs:label "subPropertyOf" ;
    rdfs:comment "The subject is a subproperty of a property." ;
    rdfs:range rdf:Property ;
    rdfs:domain rdf:Property .

rdfs:comment a rdf:Property ;
    rdfs:isDefinedBy <http://www.w3.org/2000/01/rdf-schema#> ;
    rdfs:label "comment" ;
    rdfs:comment "A description of the subject resource." ;
    rdfs:domain rdfs:Resource ;
    rdfs:range rdfs:Literal .

rdfs:label a rdf:Property ;
    rdfs:isDefinedBy <http://www.w3.org/2000/01/rdf-schema#> ;
    rdfs:label "label" ;
    rdfs:comment "A human-readable name for the subject." ;
    rdfs:domain rdfs:Resource ;
    rdfs:range rdfs:Literal .
```

例：rdfs:label の定義

rdfs:labelという述語を
定義するRDFノード


```
rdfs:label a rdf:Property ;
    rdfs:isDefinedBy <http://www.w3.org/2000/01/rdf-schema#> ;
    rdfs:label "label" ;
    rdfs:comment "A human-readable name for the subject." ;
    rdfs:domain rdfs:Resource ;
    rdfs:range rdfs:Literal .
```

rdfs:label ([<http://www.w3.org/2000/01/rdf-schema#label>](http://www.w3.org/2000/01/rdf-schema#label))
名前を指定するような標準的な述語。

RDFノード型について

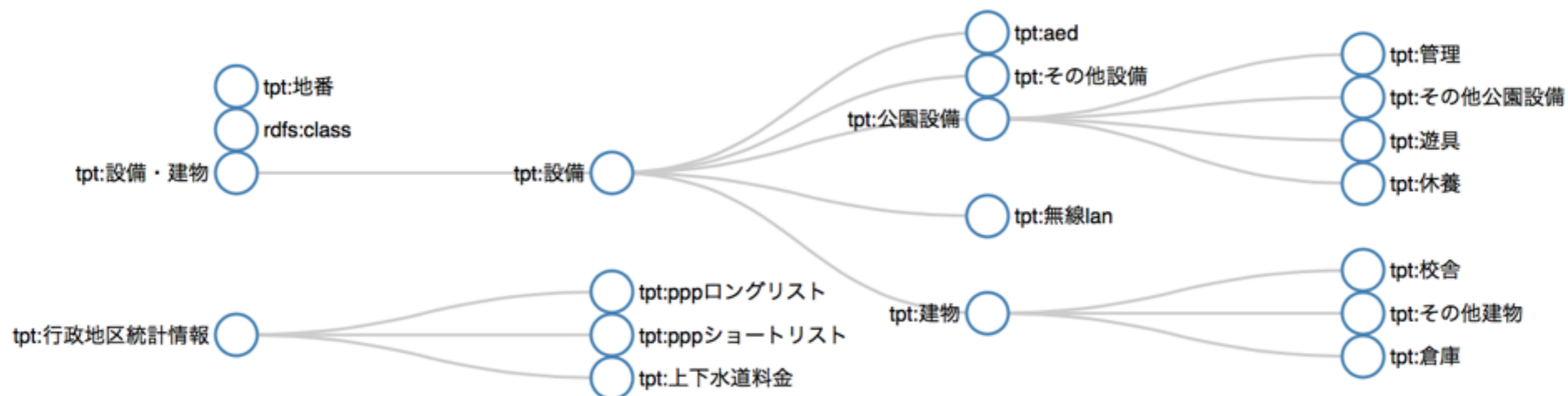
「a」キーワード又は「rdfs:type」でノード型を指定

- Turtle 「a」キーワードは「rdfs:type」述語と同義
- RDFスキーマ語彙を使うと、ノード型および利用可能述語を定義可能
- RDFで、データと同じデータベースの中にスキーマを定義するとのことです
- スキーマを守るのはデータ登録者の責任である（トリプルすたは確認してくれません）
- rdfs:subclassof述語で、更にノード型の親子関係を定義出来ます（オブジェクト指向言語に似てる）



```
rdfs:label a rdf:Property ;  
  rdfs:isDefinedBy <http://www.w3.org/2000/01/rdf-schema#> ;  
  rdfs:label "label" ;  
  rdfs:comment "A human-readable name for the subject." ;  
  rdfs:domain rdfs:Resource ;  
  rdfs:range rdfs:Literal .
```

RDFノード型ツリーのサンプル




http://teapot.bodic.org/voc_doc.htmlより

後でこのツリーを一緒に生成します・・・

その他の名前空間

popular

1. yago
2. rdf
3. foaf
4. dbp
5. dc
6. owl
7. rdfs
8. ont
9. dbo
10. onto
11. skos
12. geo
13. rss
14. gldp
15. sioc
16. sc
17. fb
18. geonames
19. xsd
20. gr
21. dcterms
22. dct
23. dbpedia
24. akt
25.  deri.org
26. commerce

yago. Wikipedia等から作られた
RDF知識データベース

rdf. W3DのRDF1.1スペーシフィ
ケーションの語彙集合

foaf. 人間関係を表す述語
(foaf=**F**riend **O**f **A** **F**riend)

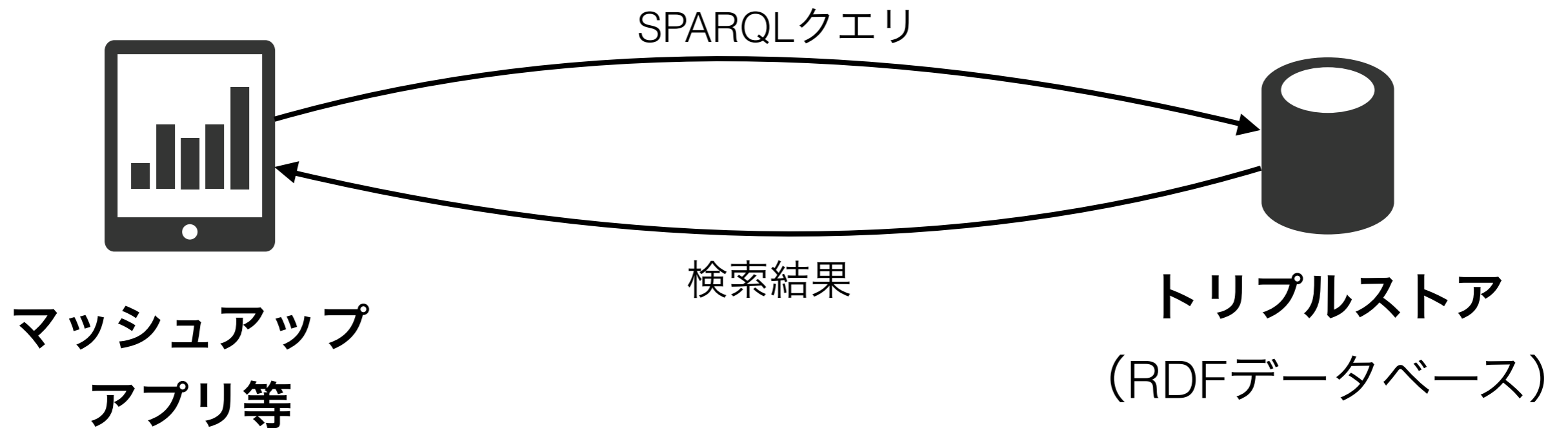
dppb. Wikipediaから作られた知
識データベース (DBペディア)

• • •

本題 2 SPARQLクエリの 基本

SPARQLとは

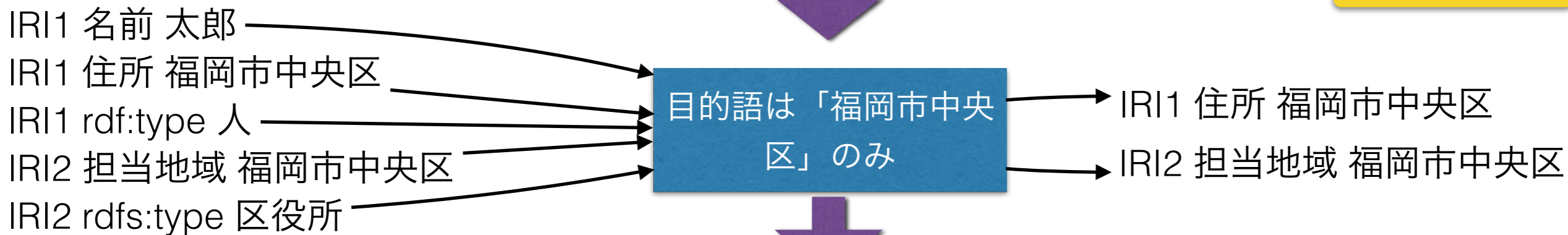
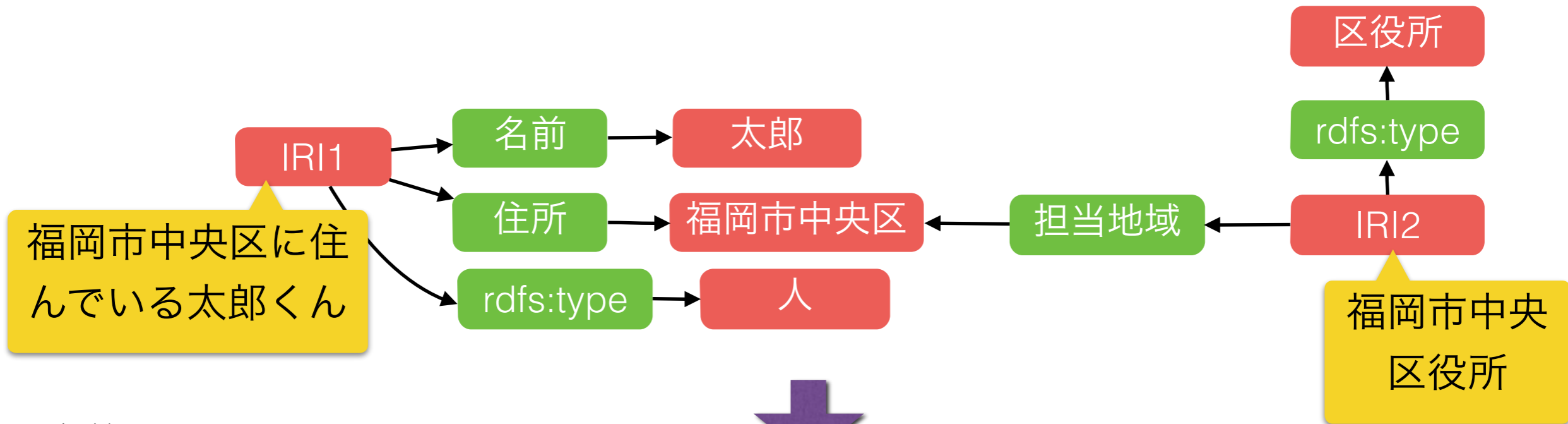
- RDF1.1のクエリ言語。W3Cの標準技術である
- クエリ言語とは
 - データの検索や操作を行うための言語
 - データモデルやデータベースシステムに合わせて設計



SPARQLの概念

- RDFトリプルを検索することで、**RDFグラフをマッチする**
- RDFトリプルを検索して、RDFトリプルを返す
- 検索条件はRDFトリプルで表現する
- RDF/Turtle Syntaxを利用

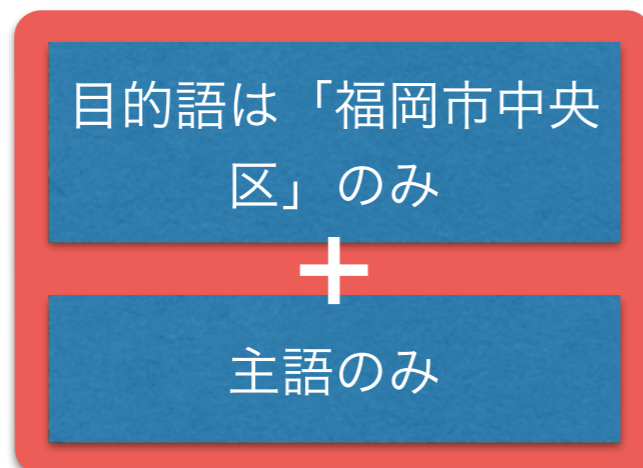
サンプル：次のスライド・・・



SPARQLシンタックス

- SPARQLクエリは主に2つに分けれる：
 - **SELECT部分**：返して欲しい変数を指定する（例：主語のみ）
 - **条件部分**：検索条件を指定する（例：目的語は「福岡市中央区のみ」）
- RDF/Turtleシンタックスを利用
 - 特にRDF/Turtleで書かれたRDFトリプルで検索条件を表現する
- RDFトリプルの主語・述語・目的語は変数または固定値である
 - **固定値**を使うと、条件とする
 - **変数**を使うと、選択する
- 変数名は必ず「？」で始まる

リテラル（文字列等）：クォートを利用（"..."）
IRI：ブラケットを利用（<...>）



SELECT部分

```
SELECT ?s  
WHERE { ?s ?p "福岡市中央区" }
```

条件部分

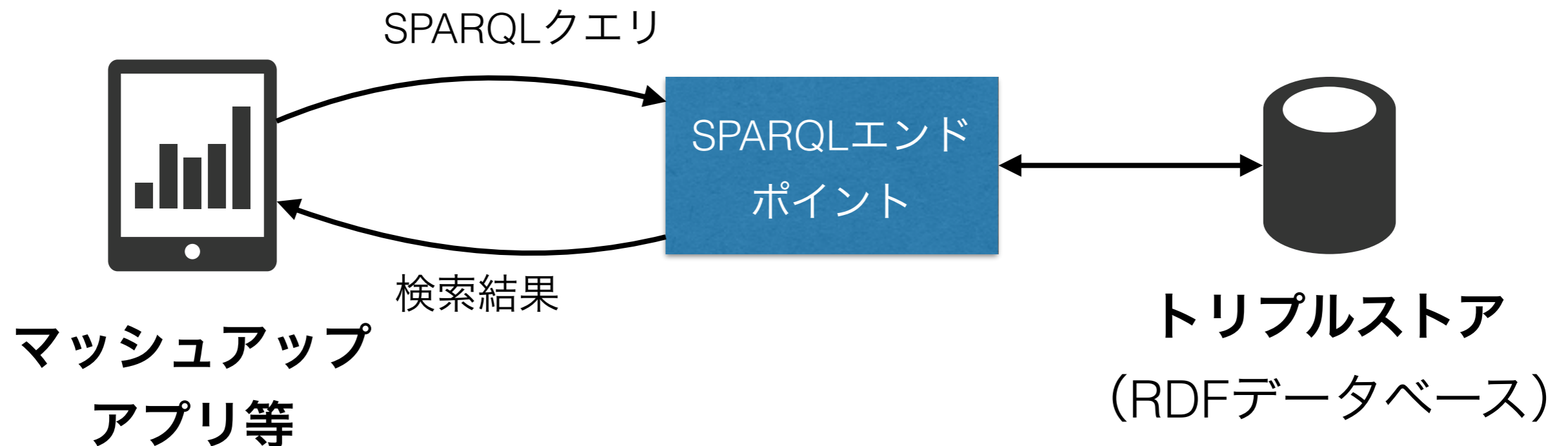
関係データベースとの比較

	関係データベース (MySQL等)	RDF
クエリ言語	SQL	SPARQL
データベース	関係データベース	トリプルストア (グラフストア)
データモデル	関係モデル	RDF1.1モデル



SPARQLエンドポイント

- SPARQLクエリを受ける場所
 - 普段はHTTP API経由
 - ウェブページ上のフォームを提供するエンドポイントも多い（裏はHTTP APIを利用）
- マッシュアップアプリ等がクライアント
 - サーバ側（PHP、Ruby等）
 - クライアント側（AJAX）



演習 1 SPARQLを使って

みよう



エンドポイント1

東日本大震災アーカイブ

<http://fukushima.archive-disasters.jp/>



The screenshot shows a web browser window with the address bar displaying "fukushima.archive-disasters.jp". The browser tabs include "SPARQLの活用 - 第4回AIツール入門講座", "Virtuoso SPARQL Query Editor", "sparql エンドポイント - Google Search", and "SPARQL Endpoint | 東日本大震災アーカイブ". The page content features the logo for "東日本大震災アーカイブ Fukushima" and the title "東日本大震災アーカイブ Fukushima - SPARQLエンドポイント(テスト版)". Three main sections are highlighted with red underlines: "SPARQLでさがす", "キーワードでさがす", and "データの構造". Each section contains a brief description of its functionality. At the bottom, there is a list of PREFIXes and a section for sample SPARQL queries.

東日本大震災アーカイブ Fukushima - SPARQLエンドポイント(テスト版)

[SPARQLでさがす](#) [キーワードでさがす](#) [データの構造](#)

このSPARQLエンドポイントには、[東日本大震災アーカイブ Fukushimaメタデータ](#)をRDFに変換したメタデータを蓄積しており、SPARQLを利用して検索することができます。

RDFで表現されている東日本大震災アーカイブ Fukushimaメタデータに対して全文テキスト検索を行うことができます。検索結果は、メタデータ1件単位で表示されます。

RDFに変換した東日本大震災アーカイブ Fukushimaメタデータの構造、SPARQL例、APIの利用方法に関する説明をします。

PREFIX(名前空間接頭辞)追加:
Dublin Core DCMi Metadata Terms NDL Metadata Terms Data Catalog Vocabulary Geo vocabulary GeoNames Ontology JM P2.0 The RDF Vocabulary (RDF) The RDF Schema vocabulary (RDFS) Creative Commons Friend of a Friend(FOAF) The Ontology for Media Resources NDL東日本大震災アーカイブメタデータ PREMIS vCard 「東日本大震災アーカイブ」基盤構築事業 福島プロジェクトメタデータ

サンプルSPARQLクエリ:
全検索 全データのタイトルを取得 URIと緯度経度情報を取得

 **SPARQLでさがす**

このSPARQLエンドポイントには、東日本大震災アーカイブ FukushimaメタデータをRDFに変換したメタデータを蓄積しており、SPARQLを利用して検索することがで

PREFIX(名前空間接頭辞)追加:

[Dublin Core](#) [DCMI Metadata Terms](#) [NDL Metadata Terms](#) [Data Catalog Vocabulary](#) [Geo vocabulary](#) [GeoNames Ontology](#) [JM P2.0](#) [The RDF Vocabulary \(RDF\)](#) [The RDF Schema vocabulary \(RDFS\)](#) [Creative Commons](#) [Friend of a Friend\(FOAF\)](#) [The Ontology for Media Resources](#) [NDL東日本大震災アーカイブメタデータ](#) [PREMIS](#) [vCard](#) 「東日本大震災アーカイブ」基盤構築事業 福島プロジェクトメタデータ

サンプルSPARQLクエリ:

[全検索](#) [全データのタイトルを取得](#) [URIと緯度経度情報を取得](#)

SPARQLクエリ

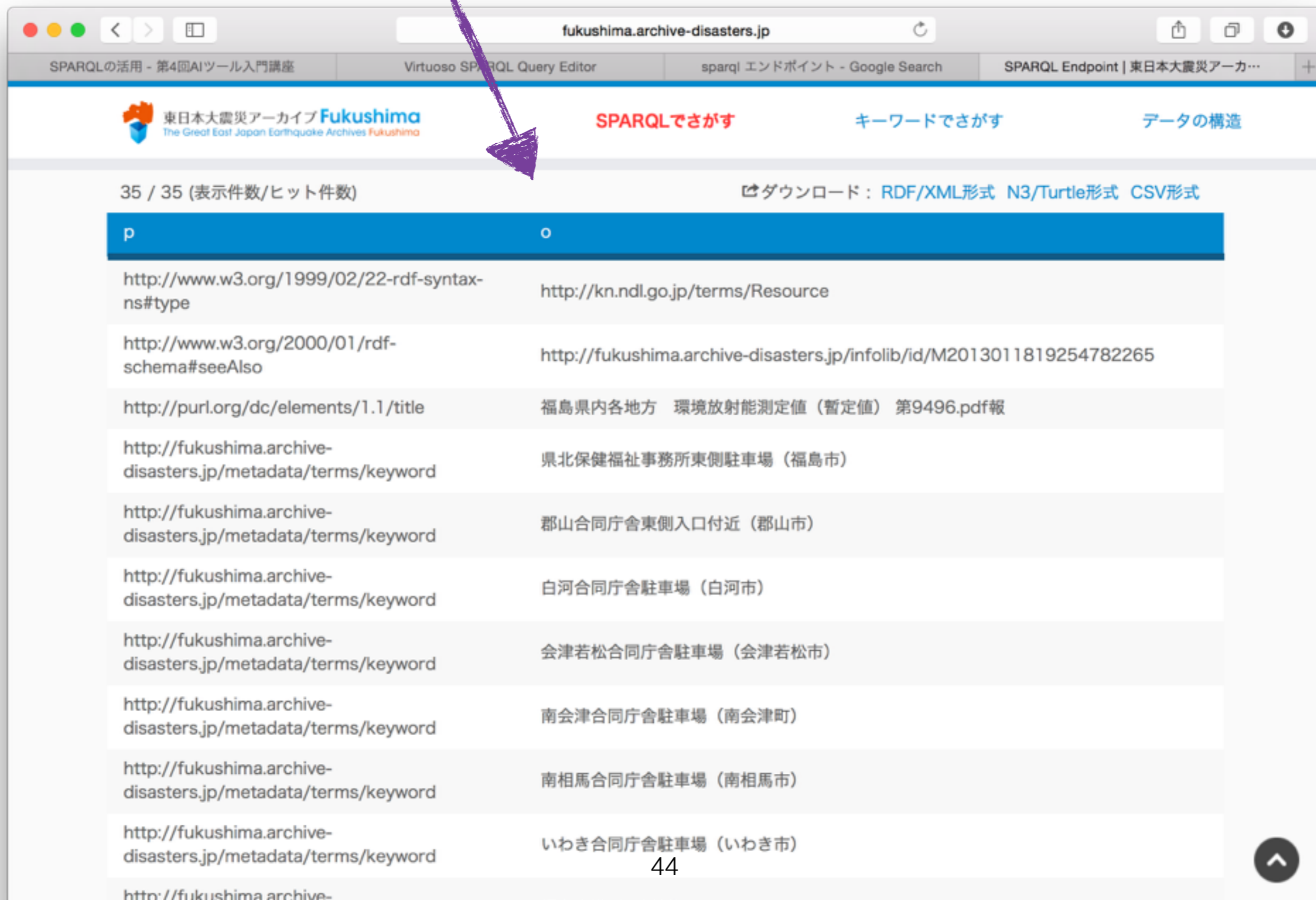
```
1 select ?p ?o [ <http://fukushima.archive-disasters.jp/id/resource/M2013011819254782265> ?p ?o ]
```

②SPARQLクエリを入力

SPARQL検索

③SPARQLクエリを実行する

④ 検索結果を確認する



The screenshot shows a web browser window with the URL `fukushima.archive-disasters.jp`. The browser tabs include "SPARQLの活用 - 第4回AIツール入門講座", "Virtuoso SPARQL Query Editor", "sparql エンドポイント - Google Search", and "SPARQL Endpoint | 東日本大震災アーカ...". The page header features the logo for "東日本大震災アーカイブ Fukushima" and navigation links for "SPARQLでさがす", "キーワードでさがす", and "データの構造". Below the header, it displays "35 / 35 (表示件数/ヒット件数)" and download options: "ダウンロード: RDF/XML形式 N3/Turtle形式 CSV形式". The main content area shows a table with two columns: "p" and "o".

p	o
http://www.w3.org/1999/02/22-rdf-syntax-ns#type	http://kn.ndl.go.jp/terms/Resource
http://www.w3.org/2000/01/rdf-schema#seeAlso	http://fukushima.archive-disasters.jp/infolib/id/M2013011819254782265
http://purl.org/dc/elements/1.1/title	福島県内各地方 環境放射能測定値 (暫定値) 第9496.pdf報
http://fukushima.archive-disasters.jp/metadata/terms/keyword	県北保健福祉事務所東側駐車場 (福島市)
http://fukushima.archive-disasters.jp/metadata/terms/keyword	郡山合同庁舎東側入口付近 (郡山市)
http://fukushima.archive-disasters.jp/metadata/terms/keyword	白河合同庁舎駐車場 (白河市)
http://fukushima.archive-disasters.jp/metadata/terms/keyword	会津若松合同庁舎駐車場 (会津若松市)
http://fukushima.archive-disasters.jp/metadata/terms/keyword	南会津合同庁舎駐車場 (南会津町)
http://fukushima.archive-disasters.jp/metadata/terms/keyword	南相馬合同庁舎駐車場 (南相馬市)
http://fukushima.archive-disasters.jp/metadata/terms/keyword	いわき合同庁舎駐車場 (いわき市)
http://fukushima.archive-disasters.jp/metadata/terms/keyword	

サンプルクエリを利用

① 「すべてのタイトルを取得」を選択

サンプルSPARQLクエリ:

全検索 全データのタイトルを取得 URIと緯度経度情報を取得

SPARQLクエリ

```
1 PREFIX dc: <http://purl.org/dc/elements/1.1/>
2 select ?s ?o from <http://fukushima.archive-disasters.jp/rdf/resource>
3 where {?s dc:title ?o}
4
```

②クエリが自動的に導入される ↓

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
select ?s ?o from <http://fukushima.archive-disasters.jp/rdf/resource>
where {?s dc:title ?o}
```

名前空間「dc」を定義

省略可能ですが、術後名などに「<http://purl.org/dc/elements/1.1/>」の代わりに「dc」を使える様になるため、便利。

クエリ対象ノード識別子の名前空間

指定する名前空間の中しかノードを検索しない（省略可能）

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
```

```
select ?s ?o from <http://fukushima.archive-disasters.jp/rdf/resource>
```

```
where {?s dc:title ?o}
```

クエリの条件

述語が「<http://purl.org/dc/elements/1.1/title>」のすべてのトリプルを探す

返して欲しい変数

クエリは他の変数を使ってもいいですが、そちらに書いてある変数のみ返される。

サンプルクエリの結果

```
select ?s ?o where { ?s dc:title ?o }
```

s	o
http://fukushima.archive-disasters.jp/id/resource/M2013011818013170719	お知らせ版No2
http://fukushima.archive-disasters.jp/id/resource/M2013011818013170718	お知らせ版No1
http://fukushima.archive-disasters.jp/id/resource/M2013011818013270720	お知らせ版No3
http://fukushima.archive-disasters.jp/id/resource/M2013011818013270721	お知らせ版No4
http://fukushima.archive-disasters.jp/id/resource/M2013011818013370722	お知らせ版No5
http://fukushima.archive-disasters.jp/id/resource/M2013011818013370723	お知らせ版No6
http://fukushima.archive-disasters.jp/id/resource/M2013011818013470724	お知らせ版No7
http://fukushima.archive-disasters.jp/id/resource/M2013011818013470725	お知らせ版No8
http://fukushima.archive-disasters.jp/id/resource/M2013011818013570727	お知らせ版No10
http://fukushima.archive-disasters.jp/id/resource/M2013011818013570726	お知らせ版No9
http://fukushima.archive-disasters.jp/id/resource/M2013011818013670729	お知らせ版No12



問題 1

タイトルが「お知らせ版No2」になっている
ノードについてすべての情報（述語、目的語）
を検索してください

答 1

案 1 : 前の検索結果を活かせる

前の結果では「dc:title」が「お知らせ版No2」となっているノードの識別子がありましたので、そのまま使う

案 1 :
select ?p ?o
where {<http://fukushima.archive-disasters.jp/id/resource/M2013011818013170719> ?p ?o}

案 2 :
PREFIX dc: <http://purl.org/dc/elements/1.1/>
select ?p ?o
where {?s dc:title "お知らせ版No2" . ?s ?p ?o }

案 2 : タイトルの文字列から検索

「?s dc:title "お知らせ版No2"」を検索条件として、すべての述語と目的語を検索

各クエーリを説明しましょう・・・

案1 クエリの説明

変数「p」と「o」を返す



```
select ?p ?o
```

```
where {<http://fukushima.archive-disasters.jp/id/resource/  
M2013011818013170719> ?p ?o}
```



主語を指定する

検索条件のトリプルは主語を指定するので、主語が指定したもののトリプルのみ検索する。述語と目的語は変数を指定するので、絞り込み条件に使わない。

案2クエリの説明

変数「p」と「o」を返す

検索トリプルで使う「s」は返さない。

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>  
select ?p ?o  
where {?s dc:title "お知らせ版No2" . ?s ?p ?o }
```

条件トリプル2：情報取得

条件トリプル1で選択した「s」に対して、すべての述語と目的語を選択し、それぞれ「p」と「o」に格納。

条件トリプル1：絞り込み条件

述語 = 「dc:title」と目的語 = 「お知らせ版No2」のトリプルのすべてを選択して、それらの主語は変数「s」に格納する。

で?



検索結果を見ると・・・

第4回 AIツール入門講座 | Virtuoso SPARQL Query Editor | sparql エンドポイント - Google... | SPARQL Endpoint | 東日本大... | rdf turtle - Google Search

東日本大震災アーカイブ Fukushima
The Great East Japan Earthquake Archives Fukushima

SPARQLでさがす | キーワードでさがす | データの構造

http://fukushima.archive-disasters.jp/metadata/terms/keyword	農畜産物損害賠償手続き説明会
http://fukushima.archive-disasters.jp/metadata/terms/keyword	国民健康保険への加入
http://purl.org/dc/terms/language	"jpn"^^<http://purl.org/dc/terms/ISO639-2>
http://purl.org/dc/terms/spatial	nodeID://b764270
http://purl.org/dc/terms/audience	一般
http://fukushima.archive-disasters.jp/metadata/terms/accessFree	無料
http://purl.org/dc/elements/1.1/creator	nodeID://b764267
http://purl.org/dc/elements/1.1/contributor	nodeID://b764268
http://purl.org/dc/terms/abstract	おしらせ版
http://xmlns.com/foaf/0.1/page	http://fukushima.archive-disasters.jp/infolib/id-c/M2013011818013170719
http://purl.org/dc/elements/1.1/source	http://fukushima.archive-disasters.jp/
http://fukushima.archive-disasters.jp/metadata/terms/accessPeriod	なし
http://ndl.go.jp/dcndl/terms/alternative	53 oshirase_no.02.pdf

URLっぽい目的語で、述語は「foaf:page」

URLをアクセスすると・・・

第4回 AIツール入門講座 Virtuoso SPARQL Query... sparql エンドポイント -... SPARQL Endpoint | 東... rdf turtle - Google Search お知らせ版No2 - 東日本...

東日本大震災アーカイブ Fukushima
The Great East Japan Earthquake Archives Fukushima

メタデータを見る コンテンツを見る Japanese English 検索画面へ

お知らせ版No2 コンテンツが表示されない場合はこちら

発行 浪江町災害対策本部
〒964-0904
福島県二本松市郭内一丁目196-1
(福島県男女共生センター内)
TEL 0243-62-0123(代)
FAX 0243-22-4261
<http://www.town.namie.fukushima.jp>

浪江町を離れ、避難生活を余儀なくされている町民の皆さまへ、
各種情報をお届けします。
※7月7日現在の情報を掲載しています。今後、内容等が変更されること
もありますので、あらかじめご了承ください。

平成23年6月
浪江町議会定例会
町長から
行政報告をし

3月11日午後2時46分
●地震発生
●大津波が沿岸部を襲う

夜
●福島第一原子力発電所から半径10km圏内に屋内退避指示

3月12日 54
早朝 ●半径10km圏内の住民に避難指示

〈表1〉避難の経緯

3月11日
午後2時46分 ●地震発生
●大津波が沿岸部を襲う

夜 ●福島第一原子力発電所から半径10km圏内に屋内退避指示

3月12日 54
早朝 ●半径10km圏内の住民に避難指示

県内、県外に避難を余儀なくされている中、町民の皆さまへ、避難生活を余儀なくされている町民の皆さまへ、各種情報をお届けします。今後、内容等が変更されることもありますので、あらかじめご了承ください。

RDFストアの探索について

悩み：初めてSPARQLエンドポイントを見る時に、使っている語彙（述語、ノード型等）が分からない



解決策：とりあえず、語彙の閲覧を出来る様なクエリを投げる！

データベースの中に使われているノード型をリストアップしましょう

下記のクエリーを使うと全てのノードのノード型を検索できる

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
select ?type
where {?s a ?type . }
```

- しかし、重複している。重複をなくすため、「**DISTINCT**」キーワードを利用できる

```
PREFIX dc: <http://purl.org/dc/elements/1.1/>
select distinct ?type
where {?s a ?type . }
```



問題 2

東日本大震災アーカイブから利用されている
すべての述語をリストアップして下さい

エンドポイント 2

公共施設等情報のオープンデータ実証

<http://teapot-test.bodic.org>

ニックネームは「Teapot」
(ティーポット)

総務省「情報流通連携基盤の公共施設等情報における実証」

公共施設等情報のオープンデータ実証 開発者サイト

総務省「情報流通連携基盤の公共施設等情報における実証」



本サイトは、平成26年度の総務省による公共施設等情報を取り扱うオープンデータ実証「情報流通連携基盤の公共施設等情報における実証」での、オープンデータ開発者サイトです。本実証は、地方自治体が保有する公共施設等情報や統計情報などを情報流通連携基盤共通APIを通し公開することで、公共施設に関する様々なアプリケーション開発が促進されることを目的としています。福岡市、福岡県、糸島市の協力を得て実施しています。データの利用に関しましては[利用規約](#)をよくお読みいただき、利用規約に同意の上、ご利用ください。

まず、スキーマを探索する

- **問題4**

- 利用されているノード型および述語を閲覧しましょう

「a」キーワードは使えない場合は「rdfs:type」述語を直接つかいましょう！

- ノード型が沢山あります！





問題 3

まず、データベースの構成を探索しましょう！

利用されているノード型および述語を閲覧しましょう

「a」キーワードは使えない場合は「rdfs:type」述語を直接つかいましょう！

・・・ノード型が沢山あります！

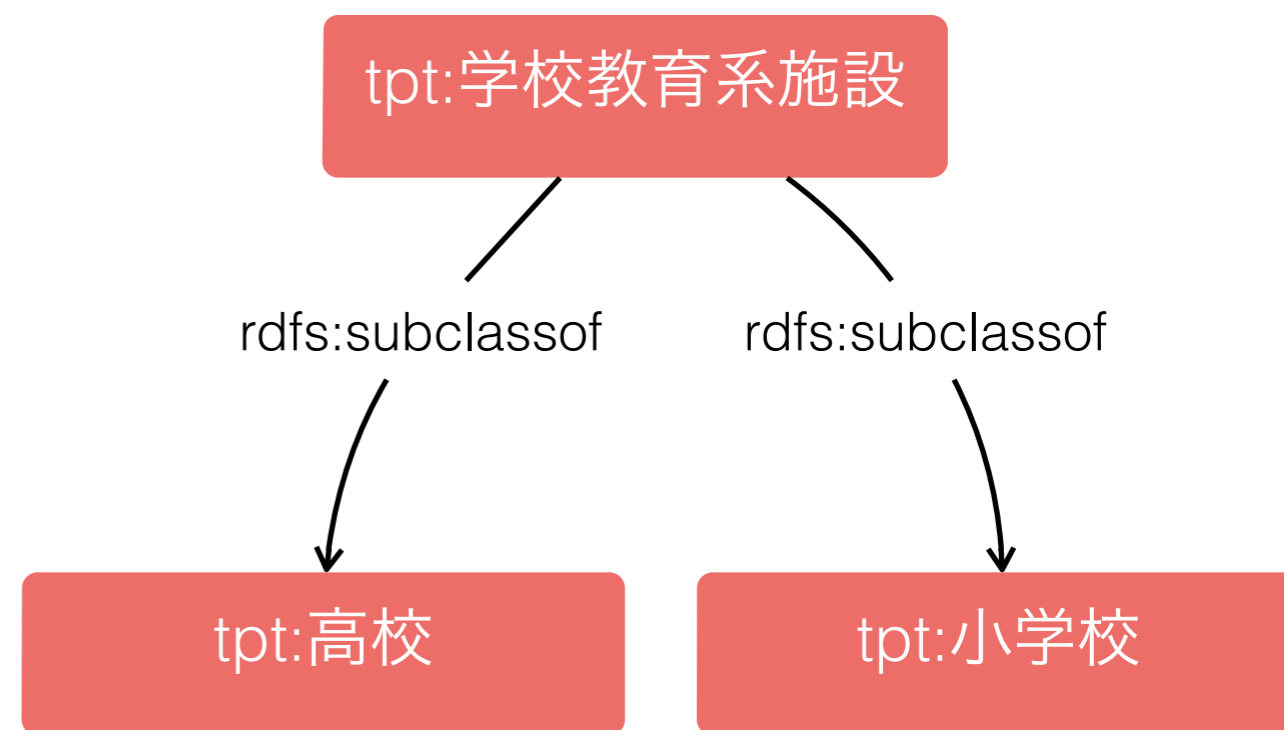


階層的ノード型の探索

TeapotはRDFS語彙を利用して、ノード型の親子関係を表す（述語：「rdfs:subclassof」）

```
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
SELECT ?t1 ?t2 { ?t1 rdfs:subclassof ?t2. }
```

```
{
  "t1": {
    "type": "uri",
    "value": "http://teapot.org/type/小学校"
  },
  "t2": {
    "type": "uri",
    "value": "http://teapot.org/type/学校教育系施設"
  }
},
{
  "t1": {
    "type": "uri",
    "value": "http://teapot.org/type/高校"
  },
  "t2": {
    "type": "uri",
    "value": "http://teapot.org/type/学校教育系施設"
  }
}
```



なんで？





問題 4

構造化ノード型を利用して、より正確なクエリを書きましょう！

全ての小学校の名前を検索しましょう

施設名は「`rdfs:label`」
述語で格納している



問題 5

全ての[学校教育系施設](#)の名前を[検索](#)しましょう

SPARQLでRDFグラフを辿る

SPARQLの「*」や「+」オペレータを使う
と、グラフを辿る

```
prefix rdfs: <...>
select ?s{ ノード 2 rdfs:subclassof ?s. }
```

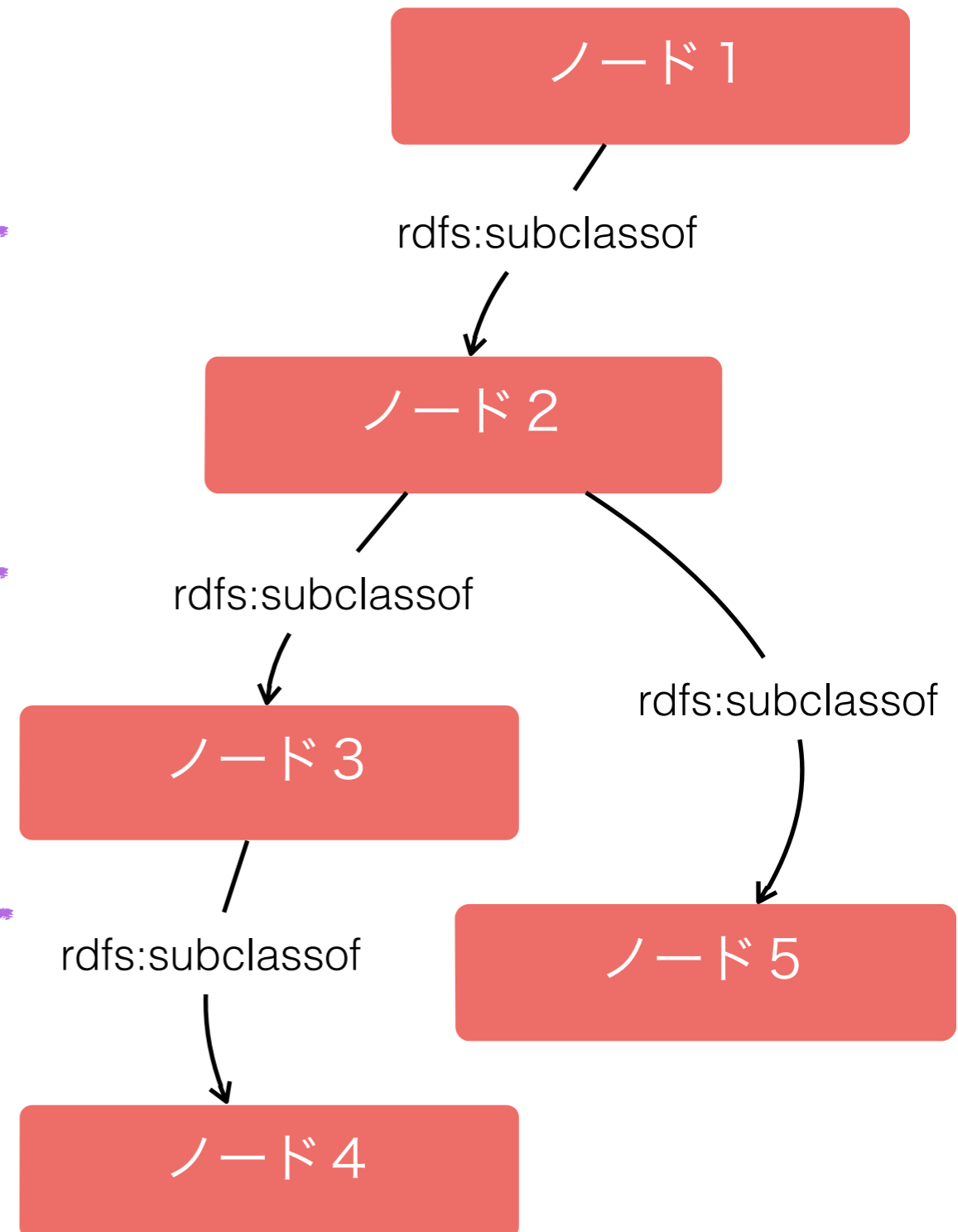
結果：ノード 3、 5

```
prefix rdfs: <...>
select ?s{ ノード 2 rdfs:subclassof+ ?s. }
```

結果：ノード 3、 4、 5

```
prefix rdfs: <...>
select ?s{ ノード 2 rdfs:subclassof* ?s. }
```

結果：ノード 2、 3、 4、 5





問題 6

施設の全てのサブタイプを検索しましょう！

SPARQLで文字列をフィルタをかける

SPARQLで「**FILTER**」キーワードを使うと諸々はフィルターをかけられる

例 1 (比較演算) : select ?s { ?s foaf:age ?age .

FILTER (age < 10 && age > 5) }

年齢 ∈]5,10[

例 2 (正規表現) : select ?s { ?s foaf:name ?name .

FILTER (regexp(?name, "taro")) }

"kentaro", "kantaro", "taro" ...

例 3 (正規表現) : select ?s { ?s foaf:name ?name .

FILTER (regexp(?name, "^山")) }

○山田、山口

×中山



問題7

中央区にある全ての施設を検索してください。

住所の文字列は「<http://teapot.org/predicate/所在地> (ぶん)」述語で格納している

その他の便利なSPARQL(1)

OPTIONAL キーワード

```
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>  
prefix tpp: <http://teapot.org/predicate/>  
select ?name ?address  
{ ?s rdfs:label ?name ; OPTIONAL { tpp:所在地 (文) ?address } }
```

「tpp:所在地 (文)」述語がなくても、
「rdfs:name」さえあれば合致する

その他の便利なSPARQL(2)

UNION キーワード

```
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
prefix tpt: <http://teapot.org/type/>
select ?name
{
  ?s rdfs:label ?name.
  { ?s rdfs:type tpt:小学校 } UNION { ?s rdfs:type tpt:高校 }
}
```

高校と小学校の名前を検索する

その他の便利なSPARQL(3)

COUNT キーワード

```
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
select COUNT(?name)
{
  ?s rdfs:label ?name.
}
```

rdf:labelのあるノード
の数を返す

その他の便利なSPARQL(4)

GROUP BY キーワード

```
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
select COUNT(?name)
{
  ?s rdfs:label ?name ; rdfs:type ?type
} group by ?type
```

各タイプ毎にrdfs:labelのある
ノードの数を返す

その他の便利なSPARQL(5)

LIMIT / OFFSET キーワード

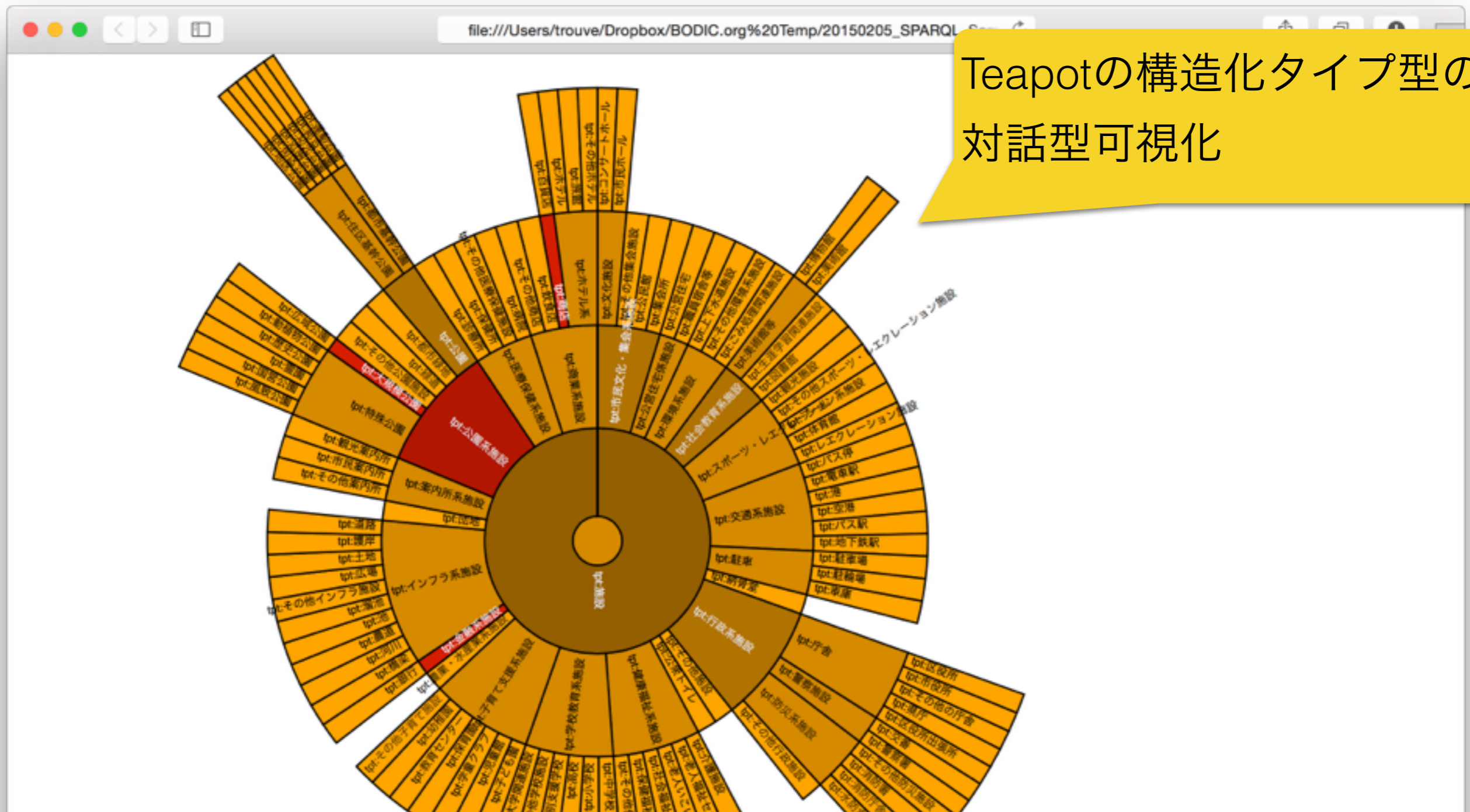
```
prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#>
select COUNT(?name)
{
  ?s rdfs:label ?name ; rdfs:type ?type
} group by ?type offset 50 limit 10
```

合致するデータの中から50個目から10個を返す（59個目まで）

演習 2 初めてのウェブ アプリケーション



この演習で出来上がるもの



その前に . . .

オープンデータで
みんなを
ハッピーに!

OPENDATA CONTEST 2014-2015 FUKUOKA

オープンデータコンテスト 2014-2015 福岡

応募受付開始日	2014年 12月22日(月)
応募締切日	2015年 2月23日(月)
表彰式	2015年 3月6日(金)
募集部門	アプリケーション部門 アイデア部門

[応募フォームはこちら >](#)

ホーム 開催概要 実証概要 応募方法 応募規約 審査・表彰 受賞作品 お問い合わせ

お知らせ

2014.12.22

[「オープンデータ・コンテスト2014-2015 Fukuoka」サイトオープン](#)



開催概要

オープンデータコンテスト2014-2015 Fukuokaの開催概要です。

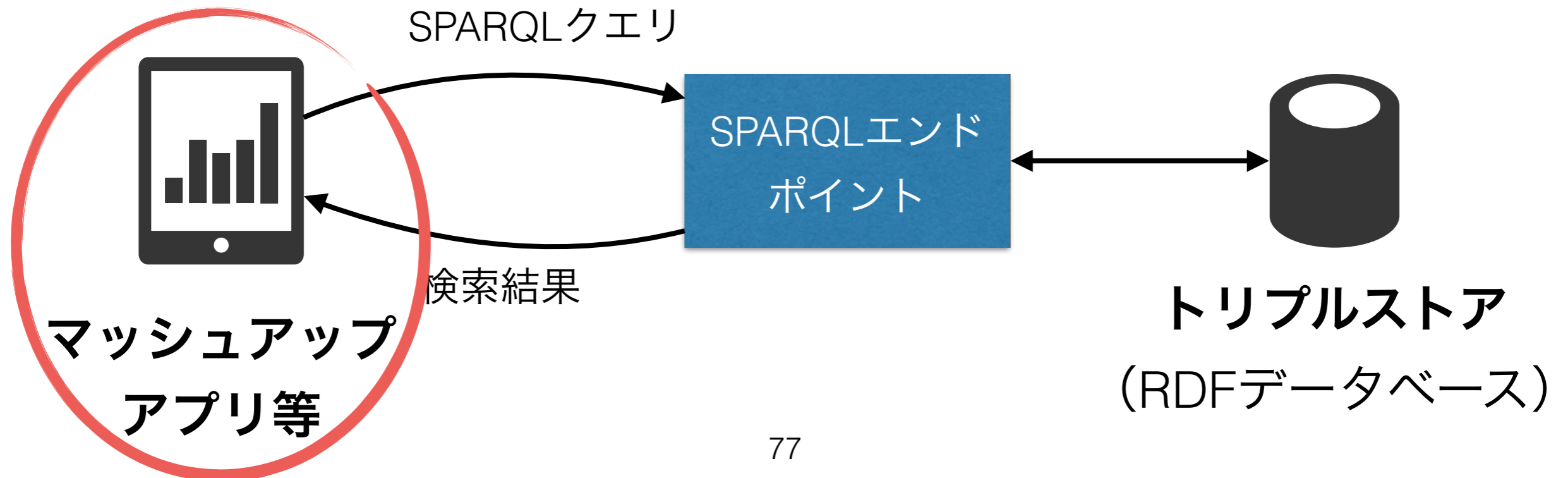


実証概要

今年度の実証実験の概要をご案内いたします。

ウェブアプリケーション (Web Application) とは

- Web Browserで利用できるアプリケーション
- クライアント側で利用する技術
 - 言語はHTML、CSS、Javascript
 - その他：AJAX (次のスライドで説明する・・・)
- サーバ側は今日しません



AJAXとは

Aynchronous **J**avascript **A**nd **X**ML

- クライアント側から非同期的に遠隔内容をアクセスするような技術
- 本演習はAJAX経由でクエリを投げて、動的にウェブページ内容を更新する

JQueryについて

- ウェブブラウザがJavascript経由で提供している機能を簡単に利用できるようなライブラリ
- 本演習では下記の機能を活躍する：
 - DOMエレメントへの選択（セレクトタ）
 - AJAX機能（POST関数）
 - DOMエレメントの内容を変更
- JQuery機能は「\$」オブジェクトから利用できる
 - DOMエレメント選択：\$(セレクトタ)
 - AJAX POST：\$.post(...)
 - DOMエレメント内容変更：\$(...).html()



問題 8

教科書の指示に基づいて、Teapotサーバへクエリを投げて、得られた結果はページに表示しましょう

クエリー結果を一部ずつロード

- **課題**：データが大量に合致した場合は、結果は数メガを超えます
 - サーバ側はGzip圧縮使うことも多いですが、ブラウザによってサポートしていない
- **解決策**：データを一気にダウンロードするのではなく、一部ずつ行う
- **実装方法**：SPARQLの「limit」と「offset」キーワードを利用する
- **例**：

```
select . . . where { . . . }  
offset 1000 limit 500
```

1000個目から1500個目
まで返す



問題 9

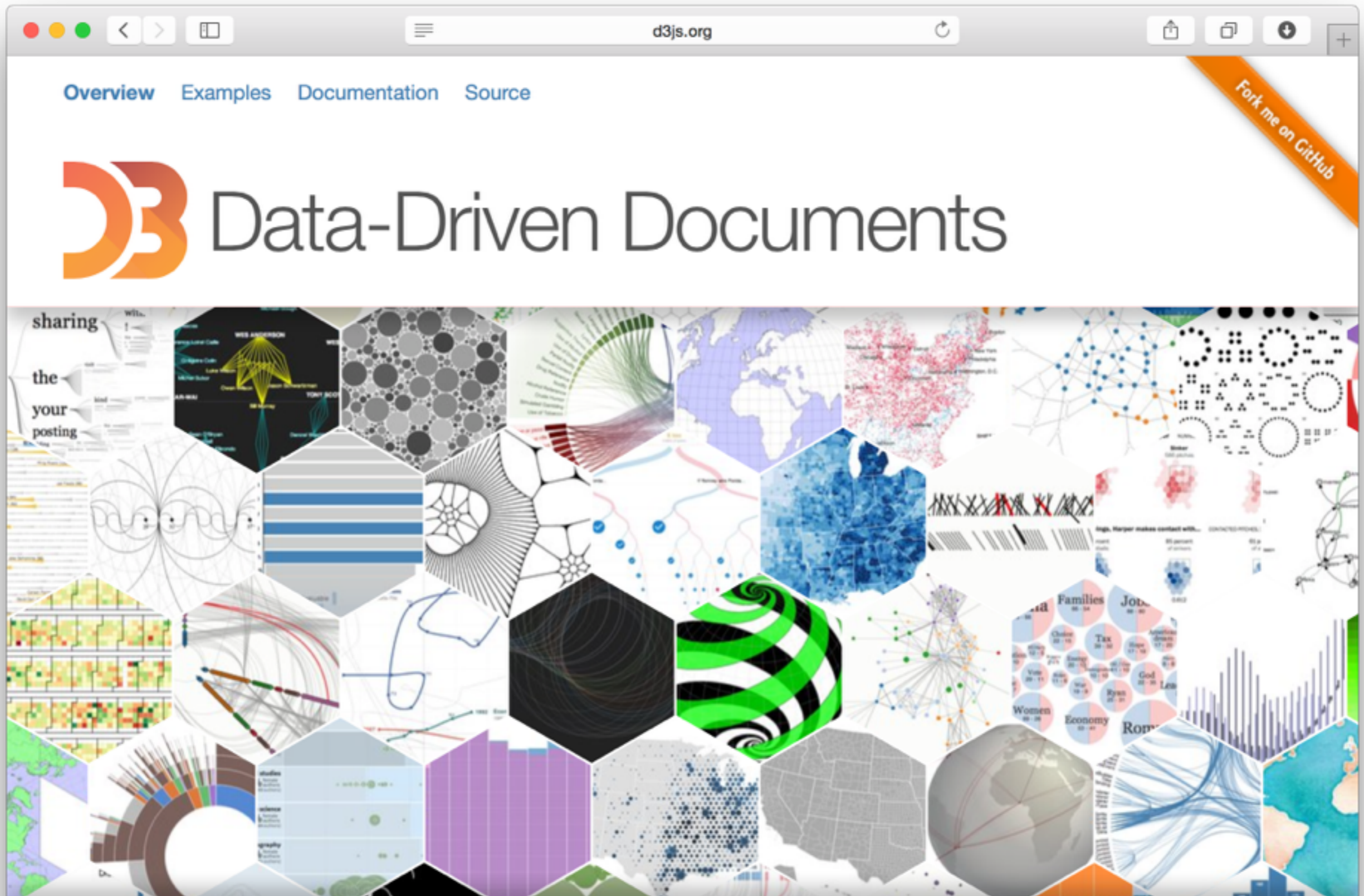
教科書の指示に基づいて、問題 8 のページを 10 個ずつクエリ結果をロードするように書き換えてください。



問題 10

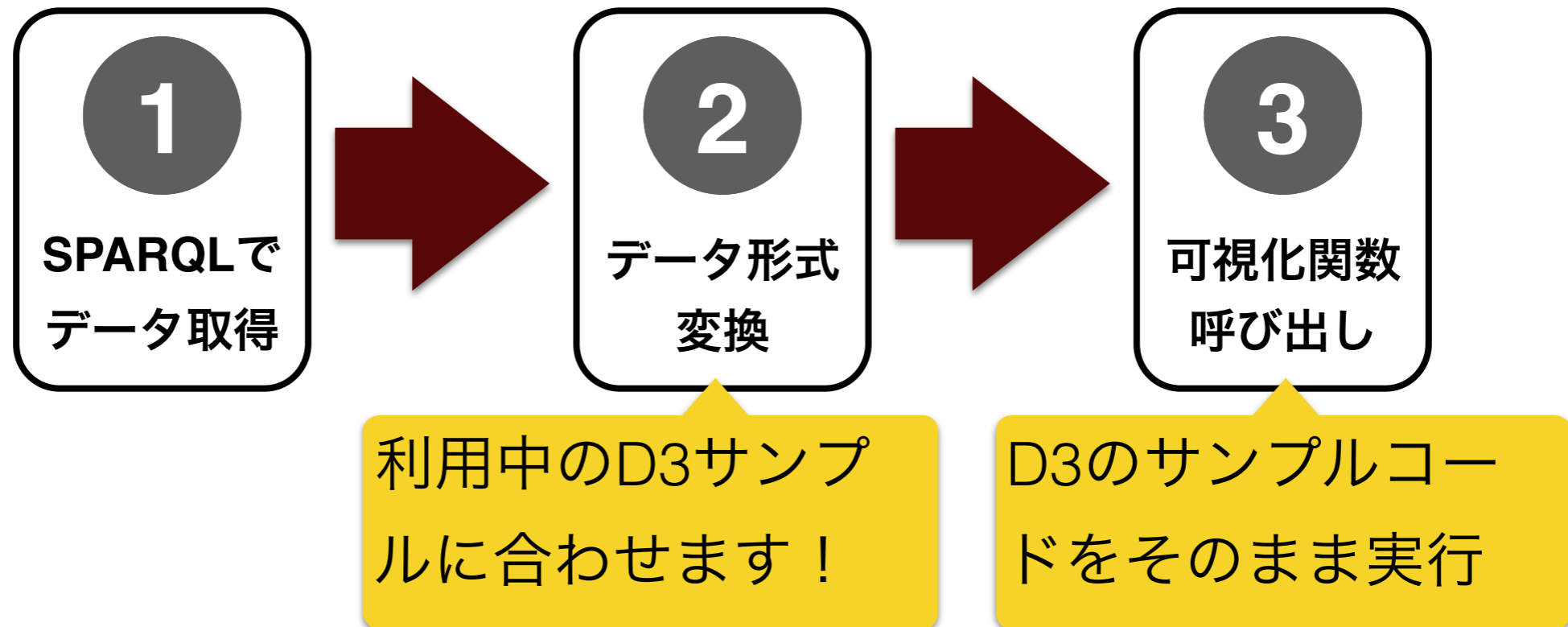
教科書の指示に基づいて、Teapotデータベースで利用されているノード型名を表示してください（生JSONではなく）。10個ずつロードしてください。

D3について

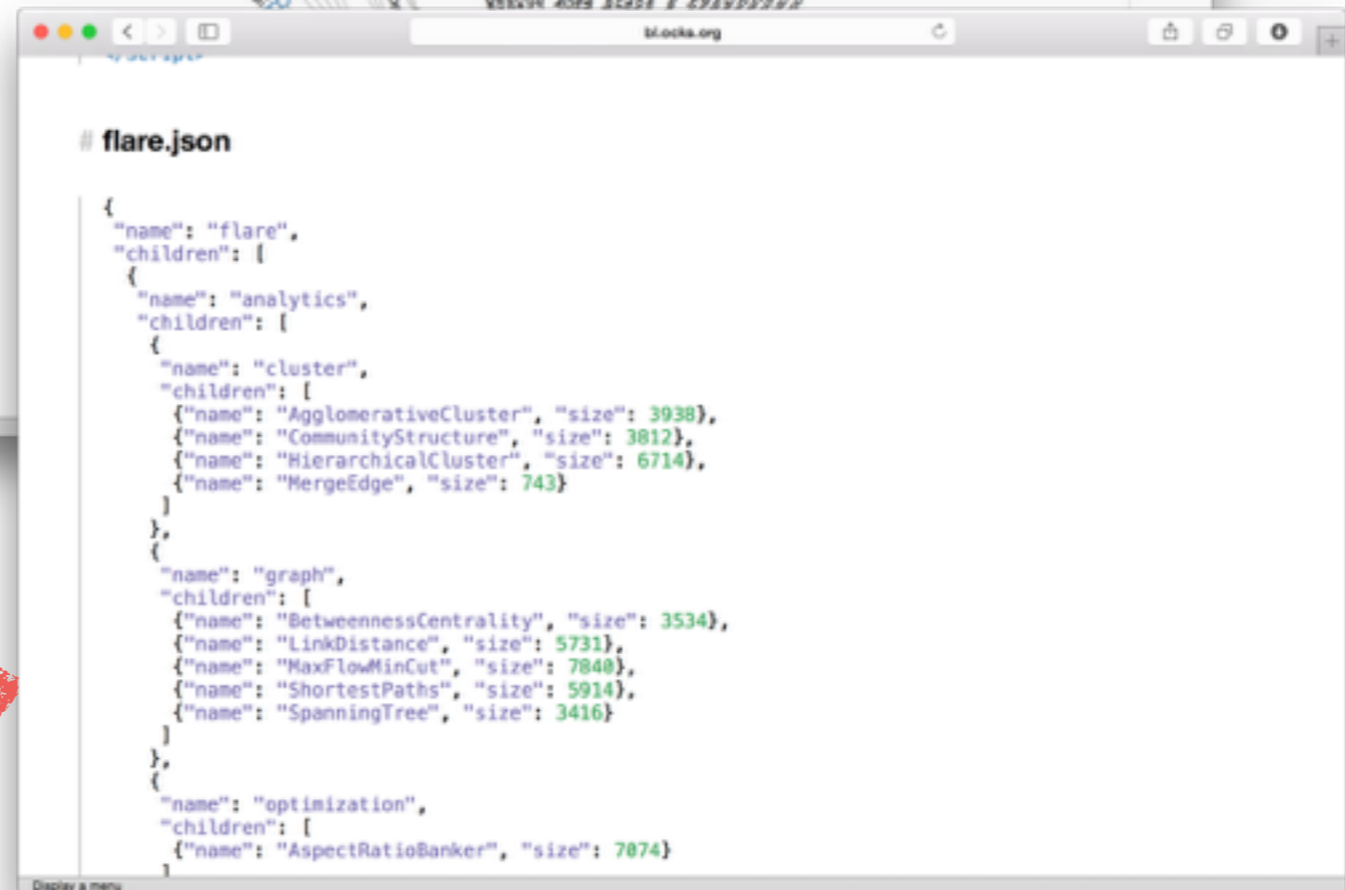
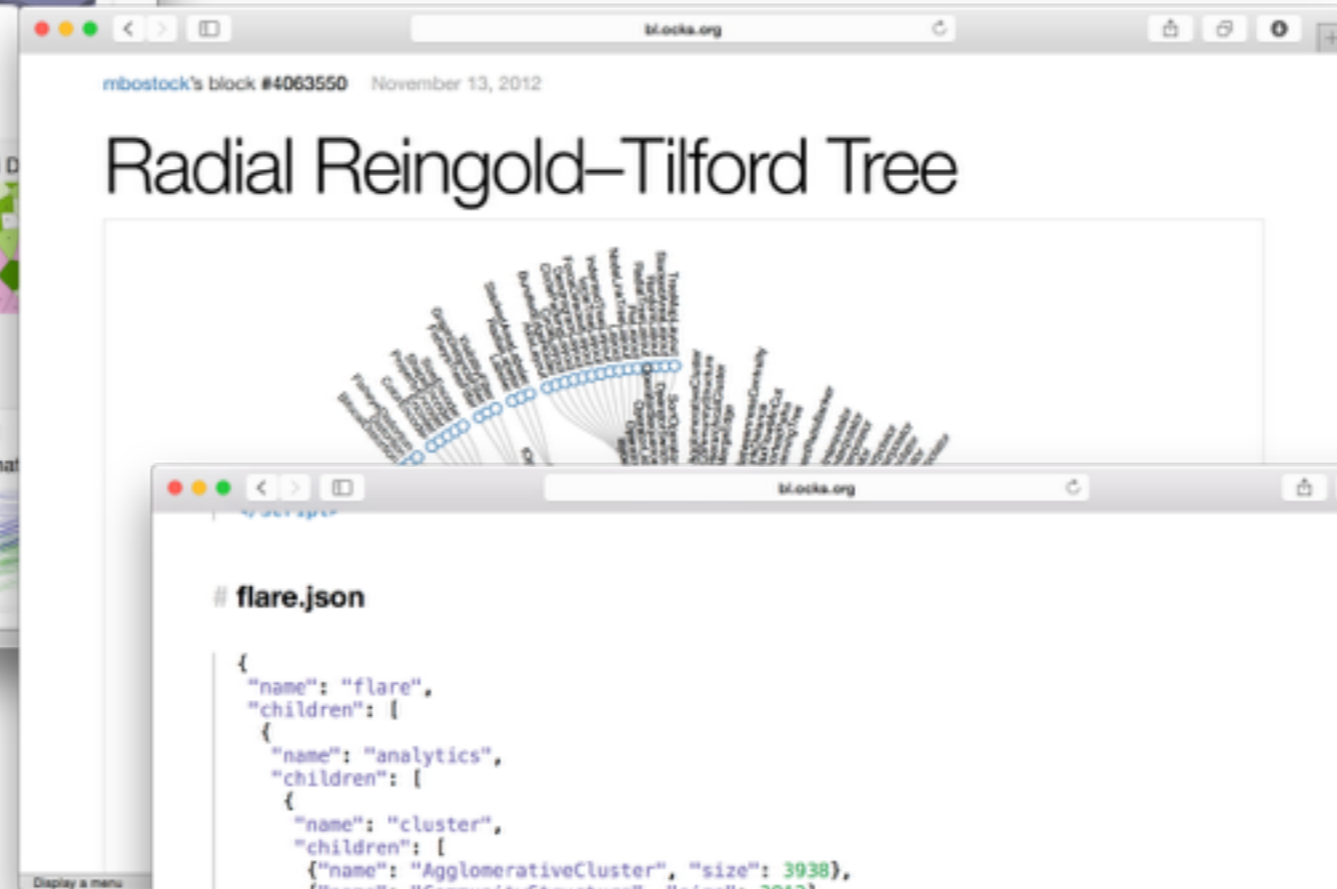
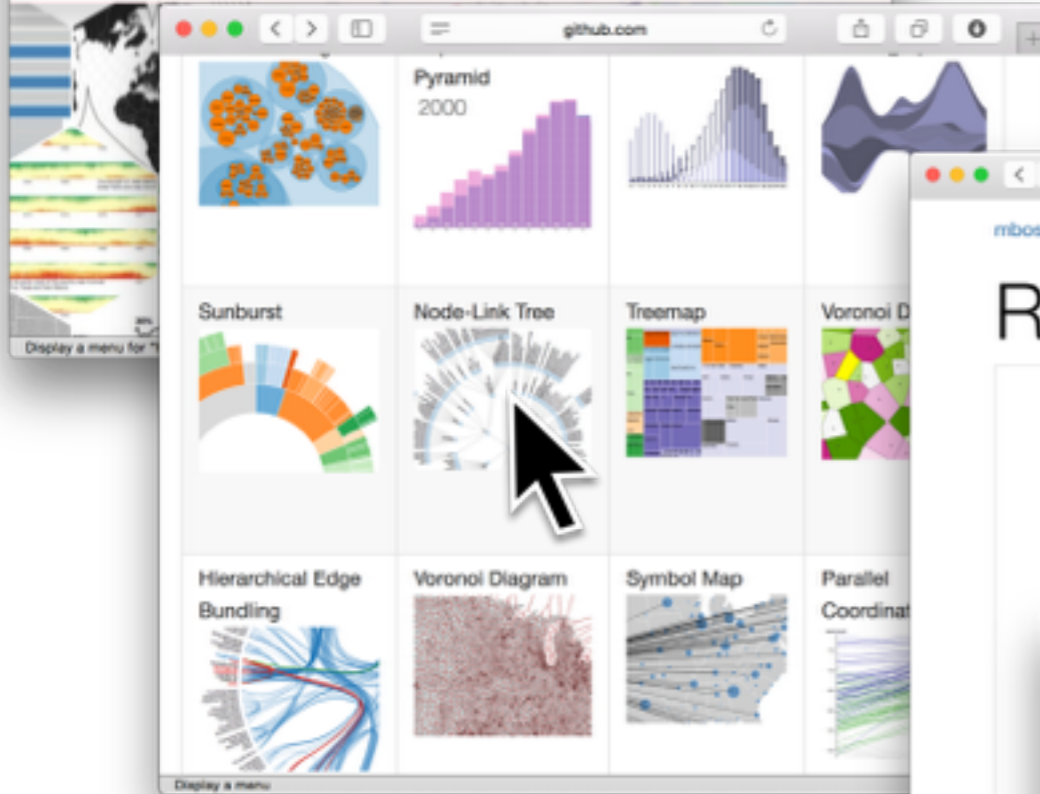


D3の使い方について

- D3のサイトに諸々なサンプルが置いてあります。そのままコピペしてもいいです！



サンプル



データ形式のサンプル

すこしわかり難い場合
もあります・・・



問題 10

教科書の指示に基づいて、Teapotのノード型を可視化しましょう！



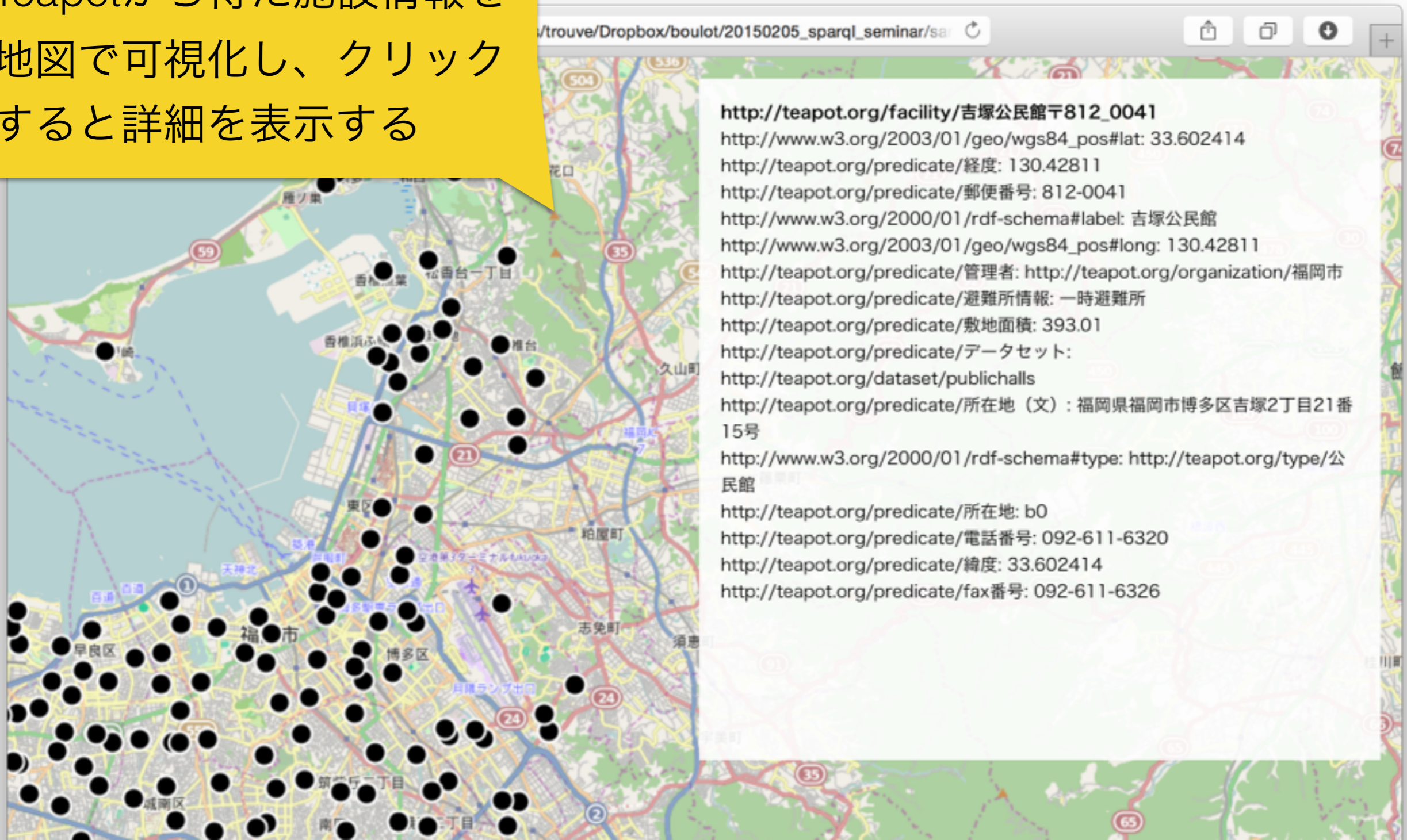
演習 3 地図アプリを作り

ましょう



この演習で出来上がるもの

Teapotから得た施設情報を
地図で可視化し、クリック
すると詳細を表示する



地图表示：OpenLayers3



- layer
- feature



問題X

教科書の指示に基づいて、地図を作成し、福岡周辺を表示してください。