

# 合同会社 緑 I T 事務所

Midori IT Office, LLC

## LODとSPARQL入門(1)

この記事は1年以上前に書かれました。  
内容が古くなっている可能性がありますのでご注意ください。

[ビジュアライゼーション\(1\)](#)で使用したオープンデータはExcel形式のファイルだったため、CSV形式に変換する必要がありました。  
CSVファイルはプログラムによる処理に便利ですが、それ以上にデータの利活用を向上させる技術として期待されているのがリンクト・オープン・データ (LOD : Linked Open Data) です。

ExcelやCSVでは、データは表形式で表現されます。一方LODでは、RDF (Resource Description Framework) グラフと呼ばれる、主語 (subject) と目的語 (object) を述語 (predicate) でリンクしたトリプル (triple) で表現します。

```
夏目漱石 → 作品 → 吾輩は猫である
```

「夏目漱石」が主語、「作品」が述語、「吾輩は猫である」が目的語です。  
作品の発行年のデータでは、

```
吾輩は猫である → 発行 → 1905
```

というようなグラフができます。  
1つめのグラフでは、作品名は目的語でしたが、2つめのグラフでは主語になっています。そこで、

```
夏目漱石 → 作品 → 吾輩は猫である → 発行 → 1905
```

このように、RDFグラフをリンクさせることができます。

LODのデータ取得には、SPARQL (スパークル) というクエリ言語を使用します。  
一見するとSQLに似たような構文ですが、SPARQLの対象はグラフモデルですので、関係モデルを対象とするSQLとは似て非なるものと考えた方が良さそうです。

SPARQLクエリを受け付けるURIを、SPARQLエンドポイントと呼びます。  
SPARQLエンドポイントによっては、SPARQLクエリを入力するフォームが用意されていることがあります。Wikipediaから抽出した情報をLODとして公開しているDBpedia Japaneseでは、<http://ja.dbpedia.org/sparql>の入力フォームからSPARQLクエリを発行することができます。

それでは、DBpediaのフォームを使ってSPARQLクエリを試してみましょう。  
以下のクエリ

```
SELECT DISTINCT *  
WHERE {  
  ?s ?p ?o .  
}  
LIMIT 10
```

は、すべてのトリプルから、データを10件だけ取得します。

```
SELECT DISTINCT *
```

は、重複しないすべての結果を返します。次の

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

WHERE節には、データを取得する際の条件を記述します。「?s」「?p」「?o」は、それぞれ主語、述語、目的語を表す変数です。主語、述語、目的語のすべてが変数になっていますので、すべてのトリプルがヒットします。

そこで、

```
LIMIT 10
```

で取得するデータを10件に制限しています（そうしないと、大変なことになってしまいます）。

```
SELECT DISTINCT ?s
WHERE {
  ?s ?p ?o .
}
LIMIT 10
```

とすると、主語だけを10件取得することができます。

次に、主語が「フクジュソウ」のすべてのトリプルを取得してみましょう。

WikipediaのリソースのURIは、DBpediaでは

```
<http://ja.dbpedia.org/resource/リソース名>
```

になりますので、フクジュソウは

```
<http://ja.dbpedia.org/resource/フクジュソウ>
```

です。

最初のクエリで変数「?s」になっていた箇所をこのURIに変えて、

```
SELECT DISTINCT *
WHERE {
  <http://ja.dbpedia.org/resource/フクジュソウ> ?p ?o .
}
```

とします。今回は、件数を制限しなくても大丈夫です。

p	o
<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a>	<a href="http://www.w3.org/2002/07/owl#Thing">http://www.w3.org/2002/07/owl#Thing</a>
<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a>	<a href="http://dbpedia.org/ontology/Species">http://dbpedia.org/ontology/Species</a>
<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a>	<a href="http://dbpedia.org/ontology/Plant">http://dbpedia.org/ontology/Plant</a>
<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a>	<a href="http://dbpedia.org/ontology/FloweringPlant">http://dbpedia.org/ontology/FloweringPlant</a>
<a href="http://www.w3.org/1999/02/22-rdf-syntax-ns#type">http://www.w3.org/1999/02/22-rdf-syntax-ns#type</a>	<a href="http://dbpedia.org/ontology/Eukaryote">http://dbpedia.org/ontology/Eukaryote</a>
<a href="http://www.w3.org/2000/01/rdf-schema#label">http://www.w3.org/2000/01/rdf-schema#label</a>	“フクジュソウ”@ja
<a href="http://www.w3.org/2000/01/rdf-schema#comment">http://www.w3.org/2000/01/rdf-schema#comment</a>	“フクジュソウ (福寿草、学名:Adonis ramosa)は、キンポウゲ科の多年草。別名、ガンジツソウ (元日草)。毒草である。1月1日の誕生花。”@ja
<a href="http://www.w3.org/2002/07/...">http://www.w3.org/2002/07/...</a>	<a href="http://ja.dbpedia.org/resource/Amur-Adonisröschen">http://ja.dbpedia.org/resource/Amur-Adonisröschen</a>

フクジュソウを主語とするトリプルの、重複しないすべての述語と目的語のペアを取得できました。

上の画像では表示されていませんが、

```
http://dbpedia.org/ontology/family
```

```
http://ja.dbpedia.org/resource/キンポウゲ科
```

というデータがあり、フクジュソウがキンポウゲ科の植物であることが分かります。

そこで、この述語と目的語を使用して、Wikipediaに掲載されているキンポウゲ科のすべての植物を調べてみることにします。

今回は、主語だけが変数となります。

```
SELECT DISTINCT ?s
WHERE {
```

```
?s <http://dbpedia.org/ontology/family> <http://ja.dbpedia.org/resource/キンポウゲ科> .
}
```

クエリを実行すると、以下のような結果が得られます。

s
<a href="http://ja.dbpedia.org/resource/トリカブト">http://ja.dbpedia.org/resource/トリカブト</a>
<a href="http://ja.dbpedia.org/resource/ミシマバイカモ">http://ja.dbpedia.org/resource/ミシマバイカモ</a>
<a href="http://ja.dbpedia.org/resource/キンポウゲ科">http://ja.dbpedia.org/resource/キンポウゲ科</a>
<a href="http://ja.dbpedia.org/resource/アネモネ">http://ja.dbpedia.org/resource/アネモネ</a>
<a href="http://ja.dbpedia.org/resource/エゾノリュウキンカ">http://ja.dbpedia.org/resource/エゾノリュウキンカ</a>
<a href="http://ja.dbpedia.org/resource/ミヤマオダマキ">http://ja.dbpedia.org/resource/ミヤマオダマキ</a>
<a href="http://ja.dbpedia.org/resource/シナノキンバイ">http://ja.dbpedia.org/resource/シナノキンバイ</a>
<a href="http://ja.dbpedia.org/resource/フクジュソウ">http://ja.dbpedia.org/resource/フクジュソウ</a>
<a href="http://ja.dbpedia.org/resource/シラネアオイ">http://ja.dbpedia.org/resource/シラネアオイ</a>
<a href="http://ja.dbpedia.org/resource/シュウメイギク">http://ja.dbpedia.org/resource/シュウメイギク</a>
<a href="http://ja.dbpedia.org/resource/ヘレボルス">http://ja.dbpedia.org/resource/ヘレボルス</a>
<a href="http://ja.dbpedia.org/resource/クレマチス">http://ja.dbpedia.org/resource/クレマチス</a>
<a href="http://ja.dbpedia.org/resource/セツブンソウ">http://ja.dbpedia.org/resource/セツブンソウ</a>
<a href="http://ja.dbpedia.org/resource/オダマキ属">http://ja.dbpedia.org/resource/オダマキ属</a>
<a href="http://ja.dbpedia.org/resource/ハナキンポウゲ">http://ja.dbpedia.org/resource/ハナキンポウゲ</a>
<a href="http://ja.dbpedia.org/resource/ツルシロカネソウ">http://ja.dbpedia.org/resource/ツルシロカネソウ</a>

少し残念なことに、植物名だけではなく「キンポウゲ科」や「オダマキ属」などの分類名まで取得してしまいました。

先ほどの

```
SELECT DISTINCT *
WHERE {
  <http://ja.dbpedia.org/resource/フクジュソウ> ?p ?o .
}
```

の結果には、

<http://ja.dbpedia.org/property/種>

"フクジュソウ A. ramosa"@ja

という述語と目的語のペアがあります。そこで、

```
http://ja.dbpedia.org/property/種
```

を述語とするトリプルを持っていることを条件に加えれば良さそうです。

```
SELECT DISTINCT ?s
WHERE {
  ?s <http://dbpedia.org/ontology/family> <http://ja.dbpedia.org/resource/キンポウゲ科> .
  ?s <http://ja.dbpedia.org/property/種> ?o .
}
```

クエリを実行すると、

s
<a href="http://ja.dbpedia.org/resource/ミシマバイカモ">http://ja.dbpedia.org/resource/ミシマバイカモ</a>
<a href="http://ja.dbpedia.org/resource/アネモネ">http://ja.dbpedia.org/resource/アネモネ</a>
<a href="http://ja.dbpedia.org/resource/ミヤマオダマキ">http://ja.dbpedia.org/resource/ミヤマオダマキ</a>
<a href="http://ja.dbpedia.org/resource/フクジュソウ">http://ja.dbpedia.org/resource/フクジュソウ</a>
<a href="http://ja.dbpedia.org/resource/シユウメイギク">http://ja.dbpedia.org/resource/シユウメイギク</a>
<a href="http://ja.dbpedia.org/resource/ハナキンボウゲ">http://ja.dbpedia.org/resource/ハナキンボウゲ</a>
<a href="http://ja.dbpedia.org/resource/パツルシロカネソウ">http://ja.dbpedia.org/resource/パツルシロカネソウ</a>
<a href="http://ja.dbpedia.org/resource/ミヤマキンボウゲ">http://ja.dbpedia.org/resource/ミヤマキンボウゲ</a>
<a href="http://ja.dbpedia.org/resource/ゼンニンソウ">http://ja.dbpedia.org/resource/ゼンニンソウ</a>
<a href="http://ja.dbpedia.org/resource/キタダケソウ">http://ja.dbpedia.org/resource/キタダケソウ</a>
<a href="http://ja.dbpedia.org/resource/ウマノアンガタ">http://ja.dbpedia.org/resource/ウマノアンガタ</a>
<a href="http://ja.dbpedia.org/resource/アズマイチゲ">http://ja.dbpedia.org/resource/アズマイチゲ</a>
<a href="http://ja.dbpedia.org/resource/タガラシ">http://ja.dbpedia.org/resource/タガラシ</a>
<a href="http://ja.dbpedia.org/resource/キクザキイチゲ">http://ja.dbpedia.org/resource/キクザキイチゲ</a>
<a href="http://ja.dbpedia.org/resource/キツネノボタン">http://ja.dbpedia.org/resource/キツネノボタン</a>
<a href="http://ja.dbpedia.org/resource/リュウキュウヒキノカサ">http://ja.dbpedia.org/resource/リュウキュウヒキノカサ</a>
<a href="http://ja.dbpedia.org/resource/ミツバオウレン">http://ja.dbpedia.org/resource/ミツバオウレン</a>
<a href="http://ja.dbpedia.org/resource/ミヤマハンショウヅル">http://ja.dbpedia.org/resource/ミヤマハンショウヅル</a>
<a href="http://ja.dbpedia.org/resource/アキカラマツ">http://ja.dbpedia.org/resource/アキカラマツ</a>
<a href="http://ja.dbpedia.org/resource/オウレン">http://ja.dbpedia.org/resource/オウレン</a>
<a href="http://ja.dbpedia.org/resource/レンゲショウマ">http://ja.dbpedia.org/resource/レンゲショウマ</a>
<a href="http://ja.dbpedia.org/resource/サラシナショウマ">http://ja.dbpedia.org/resource/サラシナショウマ</a>
<a href="http://ja.dbpedia.org/resource/ケキツネノボタン">http://ja.dbpedia.org/resource/ケキツネノボタン</a>
<a href="http://ja.dbpedia.org/resource/ミスミソウ">http://ja.dbpedia.org/resource/ミスミソウ</a>
<a href="http://ja.dbpedia.org/resource/オキナグサ">http://ja.dbpedia.org/resource/オキナグサ</a>
<a href="http://ja.dbpedia.org/resource/ボタンヅル">http://ja.dbpedia.org/resource/ボタンヅル</a>
<a href="http://ja.dbpedia.org/resource/オオヤマオダマキ">http://ja.dbpedia.org/resource/オオヤマオダマキ</a>
<a href="http://ja.dbpedia.org/resource/ヒメイチゲ">http://ja.dbpedia.org/resource/ヒメイチゲ</a>
<a href="http://ja.dbpedia.org/resource/ヒメウス">http://ja.dbpedia.org/resource/ヒメウス</a>
<a href="http://ja.dbpedia.org/resource/イチリンソウ">http://ja.dbpedia.org/resource/イチリンソウ</a>
<a href="http://ja.dbpedia.org/resource/サンリンソウ">http://ja.dbpedia.org/resource/サンリンソウ</a>
<a href="http://ja.dbpedia.org/resource/バイカモ">http://ja.dbpedia.org/resource/バイカモ</a>
<a href="http://ja.dbpedia.org/resource/オオバショウマ">http://ja.dbpedia.org/resource/オオバショウマ</a>
<a href="http://ja.dbpedia.org/resource/ヒキノカサ">http://ja.dbpedia.org/resource/ヒキノカサ</a>
<a href="http://ja.dbpedia.org/resource/バイカオウレン">http://ja.dbpedia.org/resource/バイカオウレン</a>
<a href="http://ja.dbpedia.org/resource/クサボタン">http://ja.dbpedia.org/resource/クサボタン</a>
<a href="http://ja.dbpedia.org/resource/チトセバイカモ">http://ja.dbpedia.org/resource/チトセバイカモ</a>
<a href="http://ja.dbpedia.org/resource/ミツバノバイカオウレン">http://ja.dbpedia.org/resource/ミツバノバイカオウレン</a>
<a href="http://ja.dbpedia.org/resource/オオバイカモ">http://ja.dbpedia.org/resource/オオバイカモ</a>

今度は、キンボウゲ科の植物名のみを取得することができました。

PREFIX句を用いると、URIの接頭辞をあらかじめ宣言しておくことにより、クエリの本体を簡潔に記述することができます。

```

PREFIX dbp:<http://ja.dbpedia.org/resource/>
PREFIX dbp-owl:<http://dbpedia.org/ontology/>
PREFIX prop:<http://ja.dbpedia.org/property/>
SELECT DISTINCT ?s
WHERE {
  ?s dbp-owl:family dbp:キンボウゲ科 .
  ?s prop:種 ?o .
}

```

WHERE節で同じ主語が続くときは、行末をピリオドではなくセミicolonにすることで、次の行の主語を省略することができます。

```

PREFIX dbp:<http://ja.dbpedia.org/resource/>
PREFIX dbp-owl:<http://dbpedia.org/ontology/>
PREFIX prop:<http://ja.dbpedia.org/property/>
SELECT DISTINCT ?s
WHERE {
  ?s dbp-owl:family dbp:キンボウゲ科 ;
  prop:種 ?o .
}

```

WHERE節が読みやすくなったと思います。

