

2020/10/31

# SPARQL ヒント&クエリ集

2020アーバンデータチャレンジ京都：  
ジャパンサーチ・タウン【参考資料】

国立国会図書館電子情報部電子情報流通課 横田志帆子

# 目次

1. SPARQLを使ってみよう！
2. クエリ集①：ある地域に関するコンテンツの種類や概数の調べ方
3. クエリ集②：各種コンテンツの探し方
4. 付録



SPARQLを  
使ってみよう！

# SPARQLを使ってみよう！

ジャパンサーチトップ > 開発者向け > EasySparql

<https://jpsearch.go.jp/rdf/sparql/easy/>

SPARQL query:

```
PREFIX jps: <https://jpsearch.go.jp/term/property#>  
PREFIX schema: <http://schema.org/>  
PREFIX type: <https://jpsearch.go.jp/term/type/>  
PREFIX cname: <https://jpsearch.go.jp/entity/cname/>
```

名前空間は指定済み  
(入力不要)

```
SELECT DISTINCT * WHERE {  
  ?s schema:creator cname:葛飾北斎 ;  
  rdfs:label ?label .  
  OPTIONAL { ?s schema:image ?image }  
}  
LIMIT 100
```

使い方①

ここにSELECT文を  
入力してクエリを  
実行する

使い方②

キーワードを入  
れてEasySPARQL  
を実行する

Results:  クエリ実行

だれ:  どこ:  いつ:  タイトル:  テキスト:

# SELECT文の基本的な構成

PREFIX — PREFIX句：接頭辞宣言

SELECT 変数1 変数2 ... 変数j

SELECT句：ここで列挙されている変数の順序組が出力される。

WHERE

{

RDFトリプル文.

～中略～

RDFトリプル文m.

WHERE句：ここで列挙されているRDFトリプル文がすべて真となるような、変数割り当てが検索される。

}

結果修飾語句 — 出力結果を制御する命令

出典：<http://www.aise.ics.saitama-u.ac.jp/~gotoh/IntroSPARQL.html>

# クエリ集①

ある地域に関するコンテンツの  
種類や概数の調べ方

## 場所情報として京都府が関連付けられたコンテンツの種類と数を調べる

```
SELECT ?type (count(?cho) as ?count)
WHERE {?cho a ?type; schema:spatial place:京都.}
GROUP by ?type
ORDER BY desc(?count)
```

type	count
<a href="#">type:図書</a>	51480
<a href="#">type:動物標本</a>	41897
<a href="#">type:雑誌</a>	32629
<a href="#">type:行政文書</a>	22763
<a href="#">type:植物標本</a>	19782
<a href="#">type:古書・古文書</a>	3239
<a href="#">type:地図資料</a>	2560
<a href="#">type:菌標本</a>	1227
<a href="#">type:電子資料</a>	713
<a href="#">type:建築</a>	610
<a href="#">type:考古</a>	588
<a href="#">type:建造物・建築</a>	557
<a href="#">type:芸術・美術</a>	550

- 「京都」を「大阪」や「奈良」に変更して検索することも可能。  
(「都」「道」「府」「県」は省略して正規化されている。)
- ?type、?cho、?count は任意の変数。

# 場所情報として京都府が関連付けられたコンテンツのプロバイダと数を調べる

```
SELECT ?provider (count(?cho) as ?count)
WHERE {?cho jps:sourceInfo/schema:provider ?provider ;
schema:spatial place:京都.}
GROUP BY ?provider
ORDER BY desc(?count)
```

provider	count
<a href="#">chname:NDLデジタルコレクション</a>	71743
<a href="#">chname:サイエンスミュージアムネット</a>	63061
<a href="#">chname:国立公文書館デジタルアーカイブ</a>	23438
<a href="#">chname:全国書誌</a>	17331
<a href="#">chname:文化遺産オンライン</a>	1947
<a href="#">chname:ColBase</a>	909
<a href="#">chname:ARC古典籍ポータル</a>	868
<a href="#">chname:ARC浮世絵ポータル</a>	775
<a href="#">chname:叡智の杜Web(宮城県図書館)</a>	643
<a href="#">chname:連歌・演能・雅楽データベース</a>	532
<a href="#">chname:みちしる</a>	436
<a href="#">chname:写真原板データベース</a>	400
<a href="#">chname:魚類写真資料データベース</a>	228
<a href="#">chname:アートコモンズ</a>	221



# 位置情報として京都府のGeohashコードを持つコンテンツのプロバイダと数を調べる

```
SELECT ?provider (count(?cho) as ?count)
WHERE {?cho jps:sourceInfo/schema:provider ?provider;
jps:spatial/jps:value/schema:geo <http://geohash.org/xn0x1pw0n> .}
GROUP BY ?provider
ORDER BY desc(?count)
```

provider (add en labels)	count
<a href="#">chname:NDLデジタルコレクション</a>	71756
<a href="#">chname:サイエンスミュージアムネット</a>	63061
<a href="#">chname:国立公文書館デジタルアーカイブ</a>	23438
<a href="#">chname:全国書誌</a>	18407
<a href="#">chname:文化遺産オンライン</a>	1947
<a href="#">chname:ARC浮世絵ポータル</a>	1050
<a href="#">chname:ColBase</a>	909
<a href="#">chname:ARC古典籍ポータル</a>	883
<a href="#">chname:叡智の杜Web(宮城県図書館)</a>	643
<a href="#">chname:連歌・演能・雅楽データベース</a>	532
<a href="#">chname:みちしる</a>	436
<a href="#">chname:写真原板データベース</a>	400
<a href="#">chname:魚類写真資料データベース</a>	228
<a href="#">chname:アートコモンズ</a>	221

# サムネイル画像を持つコンテンツのプロバイダと数を調べる

```
SELECT ?provider (count(?thumbnail) as ?count)
WHERE {?cho jps:sourceInfo/schema:provider ?provider ;
        rdfs:label ?label .
        OPTIONAL {?cho schema:image ?thumbnail}}
GROUP BY ?provider
ORDER BY desc(?count)
```

コンテンツのライセンスは別途確認が必要。

provider (add en labels)	count
<a href="#">chname:NDLデジタルコレクション</a>	480386
<a href="#">chname:魚類写真資料データベース</a>	119197
<a href="#">chname:ARC浮世絵ポータル</a>	116615
<a href="#">chname:東京大学学術資産アーカイブズ</a>	63169
<a href="#">chname:ADEAC</a>	38290
<a href="#">chname:データカタログサイト</a>	27291
<a href="#">chname:ARC古典籍ポータル</a>	21627
<a href="#">chname:ColBase</a>	16461
<a href="#">chname:国立美術館所蔵作品総合目録</a>	10894
<a href="#">chname:写真原板データベース</a>	7756
<a href="#">chname:慶應デジタルコレクション</a>	4600

# 場所情報として京都府が関連付けられ、サムネイル画像を持つコンテンツのプロバイダと数を調べる

```
SELECT ?provider (count(?thumbnail) as ?count)
WHERE {?cho jps:sourceInfo/schema:provider ?provider ;
        rdfs:label ?label ;
        schema:spatial place:京都.
        OPTIONAL {?cho schema:image ?thumbnail}}
GROUP BY ?provider
ORDER BY desc(?count)
```

コンテンツのライセンスは別途確認が必要。

provider (add en labels)	count
<a href="#">chname:NDLデジタルコレクション</a>	14194
<a href="#">chname:ARC浮世絵ポータル</a>	774
<a href="#">chname:ARC古典籍ポータル</a>	594
<a href="#">chname:みちしる</a>	436
<a href="#">chname:写真原板データベース</a>	326
<a href="#">chname:魚類写真資料データベース</a>	228
<a href="#">chname:ColBase</a>	203
<a href="#">chname:データカタログサイト</a>	16
<a href="#">chname:古地図コレクション</a>	8
<a href="#">chname:群馬県立図書館デジタルライブラリー</a>	8



# クエリ集②

各種コンテンツの探し方

# 京都府に関する陶磁/京都府で出土された陶磁

```
SELECT ?s ?label
WHERE {?s schema:spatial place:京都 ;
        rdfs:label ?label ;
        a type:陶磁 .}
```

s	label
<a href="https://ipsearch.go.jp/data/cobas-79738">https://ipsearch.go.jp/data/cobas-79738</a>	壺
<a href="https://ipsearch.go.jp/data/cobas-79290">https://ipsearch.go.jp/data/cobas-79290</a>	骨壺
<a href="https://ipsearch.go.jp/data/cobas-79445">https://ipsearch.go.jp/data/cobas-79445</a>	埴埴

class一覧：<https://ipsearch.go.jp/static/developer/introduction/jps-classes202008.xlsx>

```
SELECT ?s ?label
WHERE {?s jps:spatial [jps:value place:京都 ;
                    jps:relationType role:発見.出土] ;
        rdfs:label ?label ;
        a type:陶磁.}
```

s	label
<a href="https://ipsearch.go.jp/data/cobas-79738">https://ipsearch.go.jp/data/cobas-79738</a>	壺

role一覧：[https://ipsearch.go.jp/static/developer/property\\_structured/roles-definition.xlsx](https://ipsearch.go.jp/static/developer/property_structured/roles-definition.xlsx)

# 場所情報として京都府が関連付けられ、時間情報が12世紀のコンテンツ

```
SELECT DISTINCT ?s ?label ?when ?type
WHERE {?s schema:spatial place:京都 ;
        schema:temporal ?when ;
        rdfs:label ?label ;
        a ?type .
        ?when jps:start ?start; jps:end ?end .
        FILTER (?start >= 1101 && ?end < 1200) }
```

s	label	when	type
<a href="https://jpsearch.go.jp/data/bunka-147101">https://jpsearch.go.jp/data/bunka-147101</a>	醍醐寺薬師堂	<a href="#">time:1121</a>	<a href="#">type:建築</a>
<a href="https://jpsearch.go.jp/data/bunka-149681">https://jpsearch.go.jp/data/bunka-149681</a>	白描絵料紙墨書金 光明経〈巻第三ノ〉	<a href="#">time:1192</a>	<a href="#">type:芸 術・美術</a>
<a href="https://jpsearch.go.jp/data/bunka-185896">https://jpsearch.go.jp/data/bunka-185896</a>	蓮華王院本堂（三 十三間堂）	<a href="#">time:1112</a>	<a href="#">type:建築</a>
<a href="https://jpsearch.go.jp/data/bunka-156375">https://jpsearch.go.jp/data/bunka-156375</a>	教王護国寺蓮花門	<a href="#">time:1185</a>	<a href="#">type:建築</a>
<a href="https://jpsearch.go.jp/data/bunka-156455">https://jpsearch.go.jp/data/bunka-156455</a>	法界寺阿弥陀堂	<a href="#">time:1185</a>	<a href="#">type:建築</a>
<a href="https://jpsearch.go.jp/data/bunka-186096">https://jpsearch.go.jp/data/bunka-186096</a>	宇治上神社拝殿	<a href="#">time:1185</a>	<a href="#">type:建築</a>
<a href="https://jpsearch.go.jp/data/bunka-200779">https://jpsearch.go.jp/data/bunka-200779</a>	高山寺石水院（五 所堂）	<a href="#">time:1185</a>	<a href="#">type:建築</a>

## 場所情報として京都府が関連付けられ、 時間情報として平安時代が関連付けられたコンテンツ

```
SELECT DISTINCT ?s ?label ?type
WHERE {?s schema:spatial place:京都 ;
        jps:temporal [jps:era time:平安時代] ;
        rdfs:label ?label ;
        a ?type}
```

s	label	type
<a href="https://jpsearch.go.jp/data/cobas-72272">https://jpsearch.go.jp/data/cobas-72272</a>	木製経筒蓋	<a href="#">type:金工</a>
<a href="https://jpsearch.go.jp/data/cobas-72273">https://jpsearch.go.jp/data/cobas-72273</a>	木製経筒蓋	<a href="#">type:金工</a>
<a href="https://jpsearch.go.jp/data/cobas-72274">https://jpsearch.go.jp/data/cobas-72274</a>	火舎	<a href="#">type:金工</a>
<a href="https://jpsearch.go.jp/data/cobas-72276">https://jpsearch.go.jp/data/cobas-72276</a>	六器残缺	<a href="#">type:金工</a>
<a href="https://jpsearch.go.jp/data/cobas-71760">https://jpsearch.go.jp/data/cobas-71760</a>	皿	<a href="#">type:金工</a>
<a href="https://jpsearch.go.jp/data/cobas-71761">https://jpsearch.go.jp/data/cobas-71761</a>	小鉢	<a href="#">type:金工</a>
<a href="https://jpsearch.go.jp/data/cobas-71766">https://jpsearch.go.jp/data/cobas-71766</a>	銀塊	<a href="#">type:金工</a>
<a href="https://jpsearch.go.jp/data/cobas-66231">https://jpsearch.go.jp/data/cobas-66231</a>	聖宋元宝	<a href="#">type:金工</a>
<a href="https://jpsearch.go.jp/data/cobas-71758">https://jpsearch.go.jp/data/cobas-71758</a>	白磁皿	<a href="#">type:金工</a>
<a href="https://jpsearch.go.jp/data/cobas-71759">https://jpsearch.go.jp/data/cobas-71759</a>	白磁小盃	<a href="#">type:金工</a>
<a href="https://jpsearch.go.jp/data/cobas-72268">https://jpsearch.go.jp/data/cobas-72268</a>	菊花双鳥鏡	<a href="#">type:金工</a>

# 場所情報として京都府が関連付けられ、 人物情報として歌川豊国が関連付けられたコンテンツ

```
SELECT ?s ?label ?type
WHERE { ?s schema:spatial place:京都 ;
         jps:agential [jps:value/owl:sameAs* chn
         ame:歌川豊国] ;
         rdfs:label ?label ;
         a ?type . }
```

s	label	type
<a href="https://jpsearch.go.jp/data/arc_nishikie-MFA_11_16392">https://jpsearch.go.jp/data/arc_nishikie-MFA_11_16392</a>	「祇園御輿はらい ねり物姿」 「矢矧の長者牛若丸」 「うち屋君の」 「付添 みつ」	<a href="#">type:版画</a>
<a href="https://jpsearch.go.jp/data/dignl-2533806">https://jpsearch.go.jp/data/dignl-2533806</a>	[玉屋景物]	<a href="#">type:古書・古文書</a>



# 場所情報として京都府が関連付けられ、サムネイル画像を持つコンテンツ

```
SELECT ?s ?label ?thumbnail
WHERE {?s schema:spatial place:京都 ;
        schema:image ?thumbnail ;
        rdfs:label ?label ;
        jps:sourceInfo/schema:provider cname:NDLデジタルコレクション.}
LIMIT 200
```

11枚目のスライドにある他の  
プロバイダでも抽出が可能。

s	label	thumbnail
<a href="https://jpssearch.go.jp/data/dignl-11446228">https://jpssearch.go.jp/data/dignl-11446228</a>	頭書増補訓蒙図彙 21巻	
<a href="https://jpssearch.go.jp/data/dignl-1138123">https://jpssearch.go.jp/data/dignl-1138123</a>	日月日記	
<a href="https://jpssearch.go.jp/data/dignl-1138139">https://jpssearch.go.jp/data/dignl-1138139</a>	東北日記	
<a href="https://jpssearch.go.jp/data/dignl-2574945">https://jpssearch.go.jp/data/dignl-2574945</a>	義経記 8巻	
<a href="https://jpssearch.go.jp/data/dignl-2544762">https://jpssearch.go.jp/data/dignl-2544762</a>	源平盛衰記 48巻目 録1巻	



# 付録

# 各種記号

記号	使い方
?○○○	変数と組み合わせて未知の要素を表す（\$でも良い）
<b>a</b>	「rdf:type」の代わり
.	トリプルパターンの区切り
;	2回目以降の主語の省略
<>	接頭辞を用いずにURIを記述する際に使う
"""	文字列を記述する際に使う
/	プロパティ階層を連結して簡潔に記述する
+	プロパティが1回以上反復する構文を表す
*	プロパティが0回以上連続する構文を表す
?	0回もしくは1回を意味する

# 頻出語句

語句	使い方
<b>PREFIX</b>	名前空間の接頭辞を宣言する
<b>SELECT</b>	問い合わせ対象のRDFストアから条件にマッチするデータを取得する
<b>WHERE {～}</b>	条件句を表す
<b>DISTINCT</b>	<b>SELECT</b> の後ろに置き、検索結果から重複している結果を取り除く
<b>UNION</b>	複数のクエリ結果を結合する
<b>OPTIONAL {～}</b>	必須ではないパターンを指定する
<b>SERVICE</b>	異なるエンドポイントから得られるデータを統合して得る
<b>FILTER (～)</b>	特定の条件に合致したデータだけに絞り込む
<b>BIND～AS～</b>	<b>AS</b> の前の変数や式の演算結果を <b>AS</b> の後の変数と紐づける
<b>GROUP BY～</b>	結果修飾語句として使用。どのデータベースからどれだけのコンテンツ情報が提供されているかを一覧表示する
<b>ORDER BY～</b>	結果修飾語句として使用。指定した変数の値に応じて整列させる
<b>LIMIT</b>	結果修飾語句として使用。取得結果の最大数を指定する
<b>(COUNT(○○) AS ○○)</b>	集約グループごとに変数に含まれる値の数を数えて返す

# EasySPARQLのパラメータ

パラメータ	値	備考
<b>when</b>	YYYY[,YYYY]   時代名	YYYYは西暦年（の範囲）
<b>where</b>	都道府県名   nnn,mmm	nnn,mmmは緯度経度の数字
<b>who</b>	名前	先頭に~を加えると別名も含め検索
<b>what</b>	クラス名	先頭に~を加えると上位クラスも含め検索
<b>keyword</b>	主題／キーワード	先頭に=を加えると完全一致
<b>title</b>	タイトル（に含まれる文字列）	先頭に=を加えると完全一致
<b>text</b>	任意の文字列	すべての文字列フィールド対象
<b>format</b>	json   xml	返り値のフォーマット（略すとAcceptヘッダによる）
<b>limit</b>	最大取得件数	略すとシステム設定値