

### イベント概要

- 13:00-13:30 環境設定について質問受付(希望者)
- 13:30-13:40 ナレッジグラフ推論チャレンジの紹介
- 13:40-15:00 ナレッジグラフ/LOD利用技術の入門
  - ・ナレッジグラフ(RDF)の基礎
  - ・ナレッジグラフ(RDF)用クエリ言語SPARQLの基礎
  - •RDFデータベース「GraphDB」の導入方法
  - Pythonライブラリ「RDFLib」によるRDF利用プログラミング
- 15:00-15:20 休憩/環境設定について質問受付(希望者)
- 15:20-18:00 推論チャレンジ応募に向けた開発事例の実演
  - ・利用可能なデータ, 追加する背景知識の紹介
  - チャレンジのタスクの推論処理の事例紹介



### 講演概要



### ねらい

- ナレッジグラフの公開に用いられるRDFデータ
- RDF用の検索言語SPARQLの基本的な使い方の基礎的な技術を紹介する.

### ■ 内容

- RDFの基礎
- SPARQL検索の基本~Wikidataを例として~
- SPARQLを用いた推論チャレンジのナレッジグラフの検索例

### RDFの基礎

■ RDFとは?

■ RDFの例(Wikidata)

# RDF (Resource Description Framework) とは?



### - RDFとは

- Web上でメタデータを記述するためのデータモデル.
- データの「意味」を記述することが想定されている。
- 近年は、ナレッジグラフ(知識グラフ)を記述するデータモデルとして広く利用されている

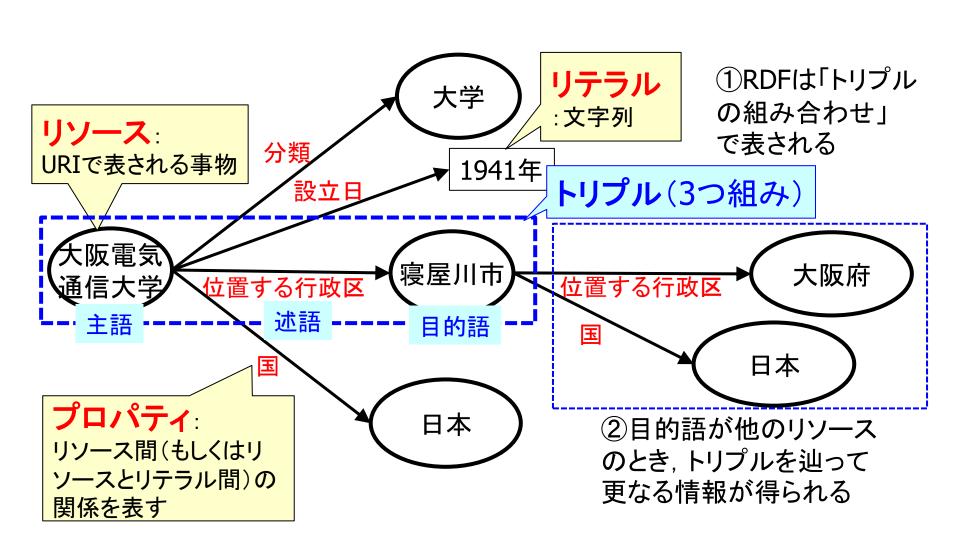
### ■ <u>RDFの基本構造</u>

- すべてのモノやコトをURIで表す.
- データの構造(つながり)を 「主語(Subject)-述語(Predicate)-目的語(Object)」 からなるトリプル(3つ組み)で表す。



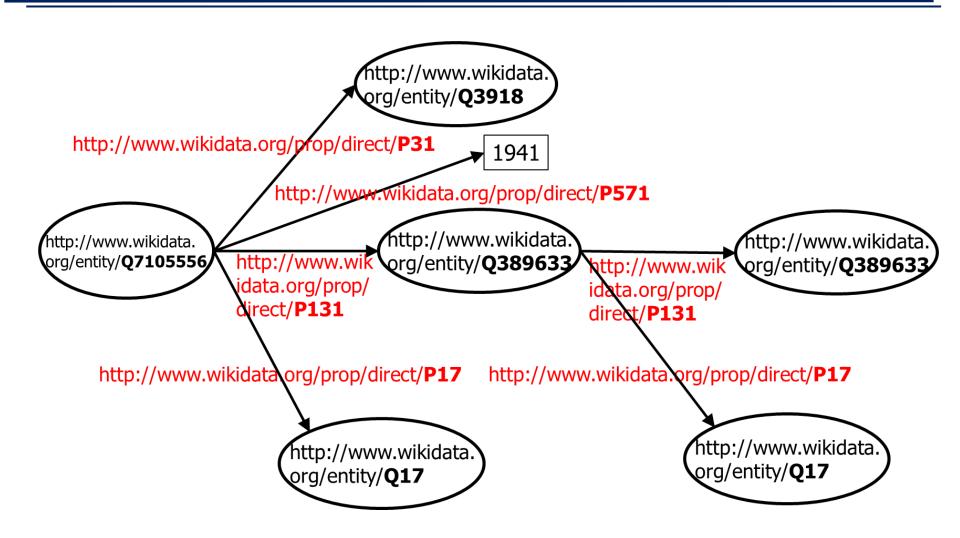
### WikidataにおけるRDFデータの例





### WikidataにおけるRDFの例





※実際のデータは、URIで表されている

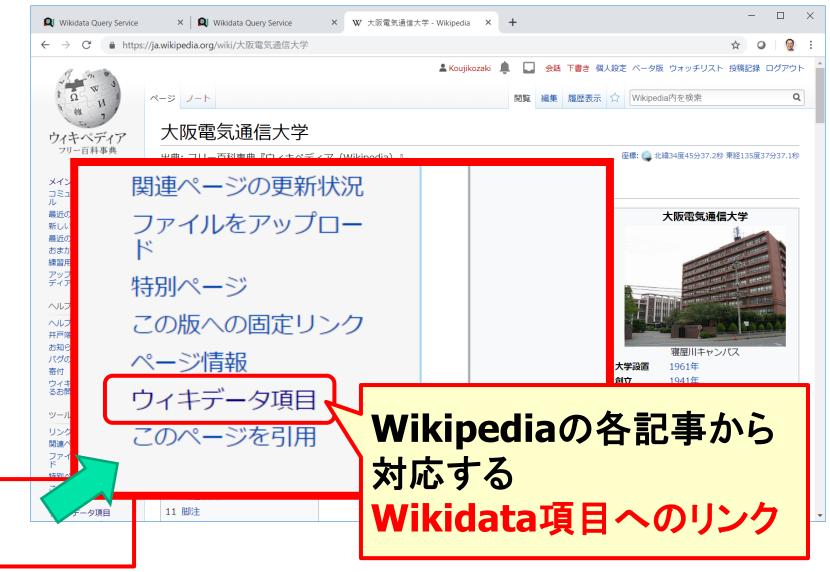
### ウィキデータ(wikidata)とは? そ プレッジグラフ推論チャレンジ





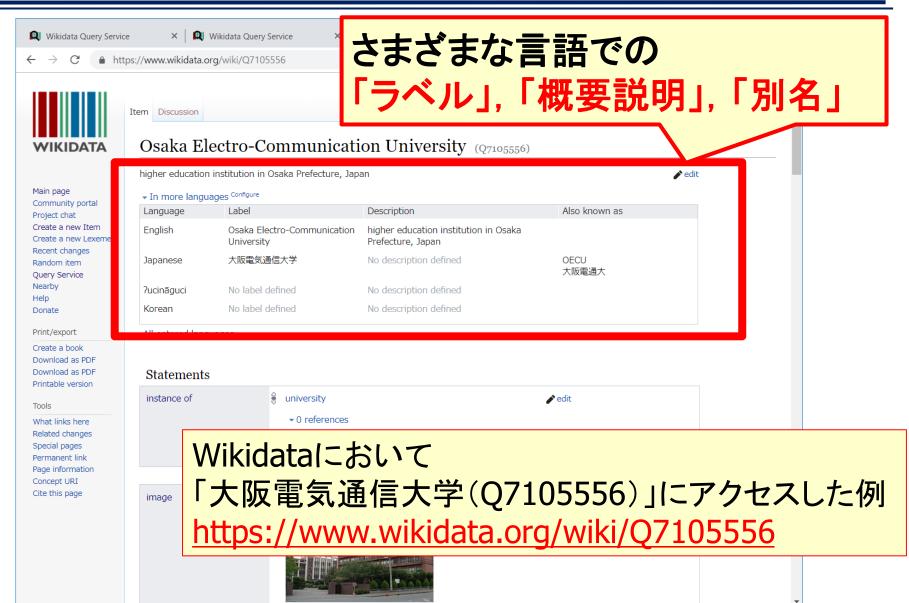
### WikipediaからWikidataへ





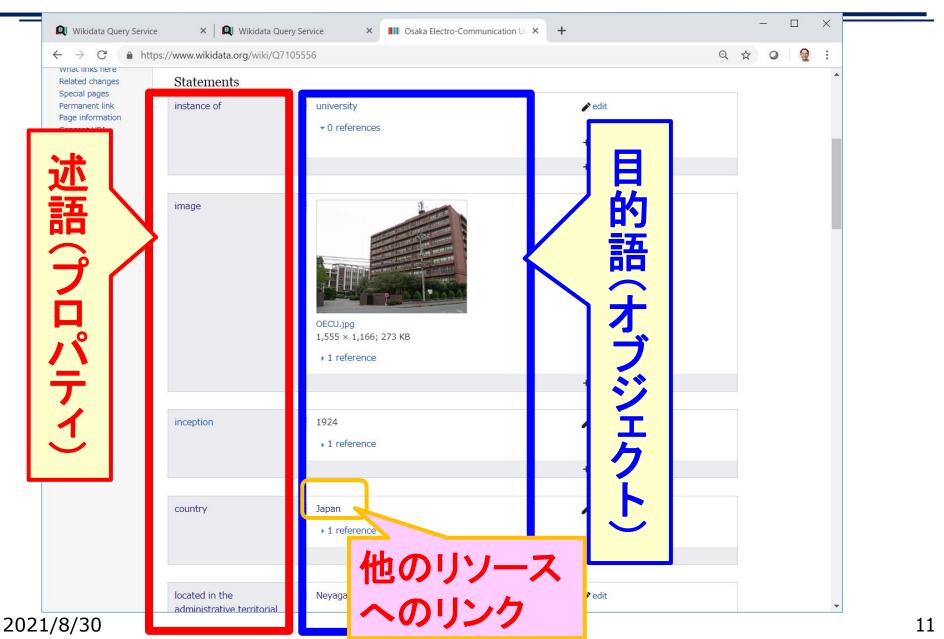
### Wikidataのデータ例(1/2)





### Wikidataのデータ例(2/2)





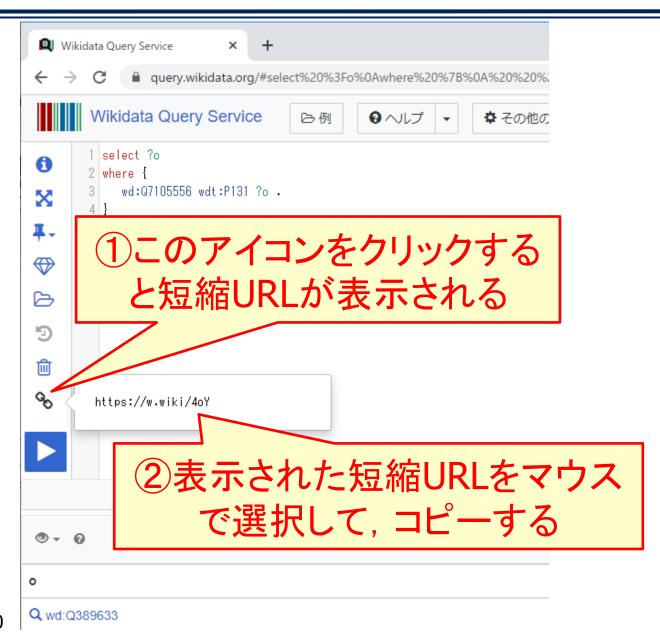
#### ウィキデータの検索サービス : SPARQLエンドポイント





### (参考) WikidataのSPARQL検索ページでの実行用「短縮URL」の取得方法





# Wikidataを例としたSPARQLクエリの例

WikidataのSPARQLエンドポイント(検索用API)

https://query.wikidata.org/

を使ったSPARQLクエリ例

・ 検索例は下記から実行可能

https://github.com/KnowledgeGraphJapan/KGRC-ws-2021/blob/main/Section2/SPARQL-sample.md

#### 紹介する検索例

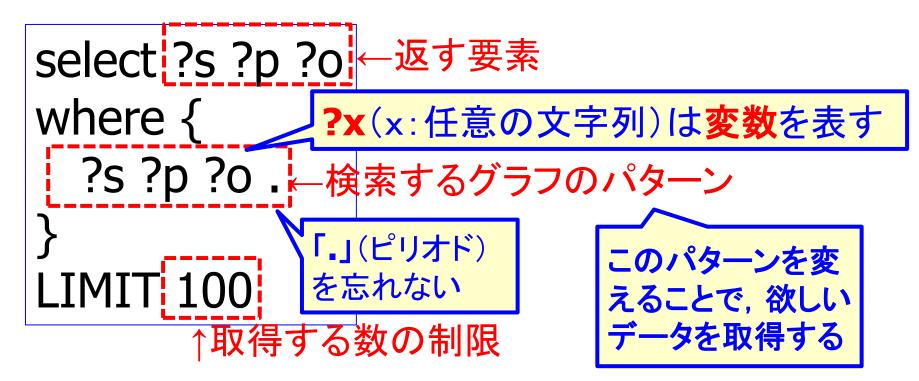
- ・ 検索例1:<主語>-<述語>を指定して<目的語>を取得する
- ・ 検索例2:<述語>-<目的語>を指定して<主語>を取得する
- ・ 検索例3:より複雑な検索例

### SPARQLによるRDFの検索

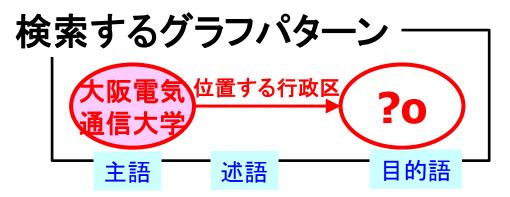


#### SPARQL

- 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



### 主語と述語を指定して「目的語」を取得 "〈主語〉の〈述語〉は何?"



 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

■ **例1-1**)「大阪電気通信大学(Q7105556)」(主語) の「位置する行政区(P131)」(述語)となる 目的語(?o)を取得する



クエリを実行

■ **例1-2)「大阪電気通信大学**(Q7105556)」(主語) の「設立(P571)」(述語)となる目的語(?o)を取得

クエリを実行

※述語を変えるといろんな 目的語が取得できる

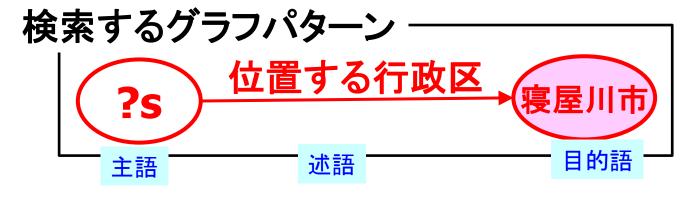
- **例1-1+)「大阪電気通信大学**(Q7105556)」(主語)の「位置する行政区(P131)」(述語)となる目的語(?o)を取得する. 【名称(ラベル)も合わせて取得】
  - ※検索結果がデータのIDとなる場合、下記の記述を追加することで「名称(ラベル)」をあわせて取得可能

クエリを実行

- 例1-3)「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)を同時に取得
  - 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)

### 検索例2:<述語>-<目的語>を指定して KGRC: Knowledge Graph Reasoning Challenge く主語>を取得する

### 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?"



SPARQLでの記述

?s <位置する行政区> <寝屋川市>

### 検索例2:<述語>-<目的語>を指定して、KGRC: Knowledge Graph Reasoning Challenge く主語>を取得する

- 例2-1)「位置する行政区(P131)」(述語)が 「寝屋川市(Q389633)」(目的語)となる 「主語(?s)」の一覧を取得する
  - ■「主語」の取得は<u>検索結果が多数</u>となる場合が多い

結果が多いので、件数の制限をかける

クエリを実行

### 検索例2:<述語>-<目的語>を指定して、KGRC: Knowledge Graph Reasoning Challenge く主語>を取得する

■ 例2-1+)「位置する行政区(P131)」(述語)が 「寝屋川市(Q389633)」(目的語)となる 「主語(?s)」の一覧を取得する【名称あり】

### 検索例 2:<述語>-<目的語>を指定して KGRC: Knowledge Graph Reasoning Challenge ・ 主語>を取得する

■ **例2-2**「分類(instance-of)(P31)」(述語)が 「大学(Q389633)」(目的語)となる 「主語(?s)」+名称の一覧を取得する

→<u>同じ種類のデーター覧</u>が取得できる

クエリを実

2021/8/3<mark>0</mark>

rdf:typeが用いられることが多い.

### 検索例2:<述語>-<目的語>を指定して、KGRC: Knowledge Graph Reasoning Challenge く主語>を取得する

■ **何2-2+**)「分類(instance-of)(P31)」(述語)が 「大学(Q389633)」(目的語)となる 「主語(?s)」+名称(※あれば…)の一覧を取得

クエリを写

2021/8/30

必須でない(OPTINAL)とした例.

### より複雑な検索例



#### **例3** <u>集約(グループ化)とカウントを利用したランキング</u>

- 分類(P31)が大学(Q3918)となる主語(?s)と、その主語(?s)の国(P17) の目的語(?country)を取得する.
- さらに、取得した結果を目的語(?country)が同一のもので集約(グループ化)し、
- それぞれのグループに含まれる主語(?s)の数をカウントする.
- その結果「国ごとの大学の数のランキング」を取得できる.

2021/8/30

結果を降順に並び変え

クエリを実行



### 考:PREFIXによる省略表現

PREFIX(接頭語)を用いるとURIの省略表記ができる

```
select ?o
where {
   <a href="http://www.wikidata.org/entity/Q7105556">http://www.wikidata.org/entity/Q7105556></a>
      <a href="http://www.wikidata.org/prop/direct/P131">http://www.wikidata.org/prop/direct/P131</a> ?o .
}LIMIT 100
                                             PREFIXによる省略表現
```

```
PREFIX wd: <a href="http://www.wikidata.org/entity/">http://www.wikidata.org/entity/>
PREFIX wdt: <a href="http://www.wikidata.org/prop/direct/">http://www.wikidata.org/prop/direct/</a>
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:** <a href="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.

### SPARQLを用いた推論チャレンジの ナレッジグラフの検索例

- 推論チャレンジのナレッジグラフ
- SPARQLによる検索例

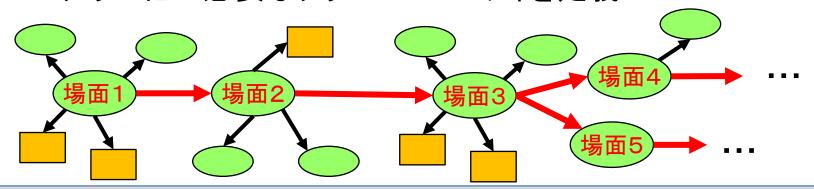
### ナレッジグラフ化の考え方

### → ナレッジグラフの要求仕様

- ▶ 犯人を推論(推理)するのに必要な知識を提供する
- ▶「推理小説」で描かれる様々な状況を、できるだけ統一的な 形式で計算機処理(検索・推論・etc.)可能にする

#### → ナレッジグラフ化の基本方針

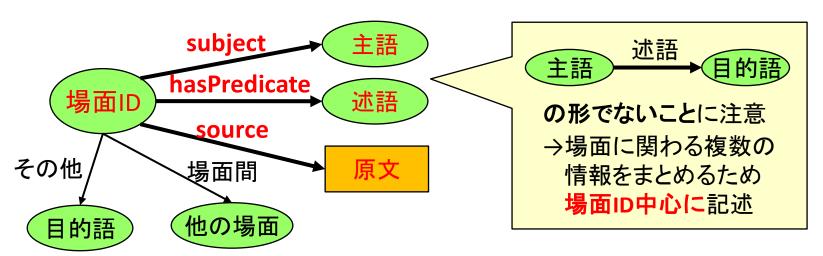
- ▶「推理小説」の内容を、最小単位の「場面(シーン)」に分割 →場面ごとにID(IRI)を付与
- ▶ 「各場面の記述内容」および「場面間の関係」をグラフ化 →グラフ化に必要なクラス・プロパティを定義



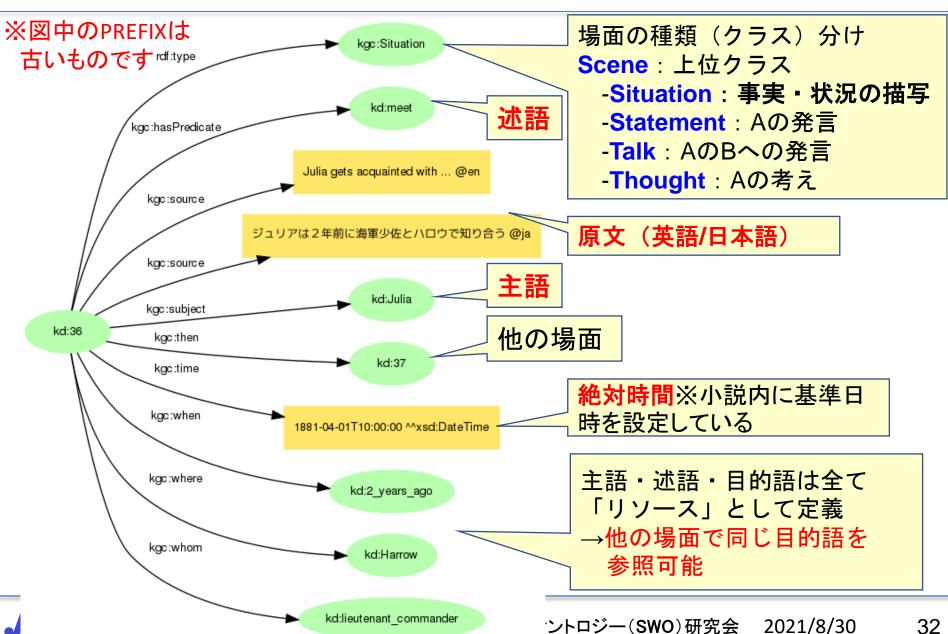
### 場面(シーン)のスキーマ

#### → 場面を表現するプロパティ

- ➤ subject:その場面の記述において主語となる人や物
- ➤ hasPredicate:その場面の内容を表す述語
- ▶ 場面の詳細を表す目的語:whom(だれに), where(どこで), when(いつ), what(何を), how(どのように), …etc.
- ➤ 場面間の関係:then, if, because, …etc.
- ➤ time: その場面が起こった絶対時間(xsd:DateTime)
- ➤ source: その場面の原文(英語/日本語のリテラル)



### 場面(シーン)スキーマ 記述例

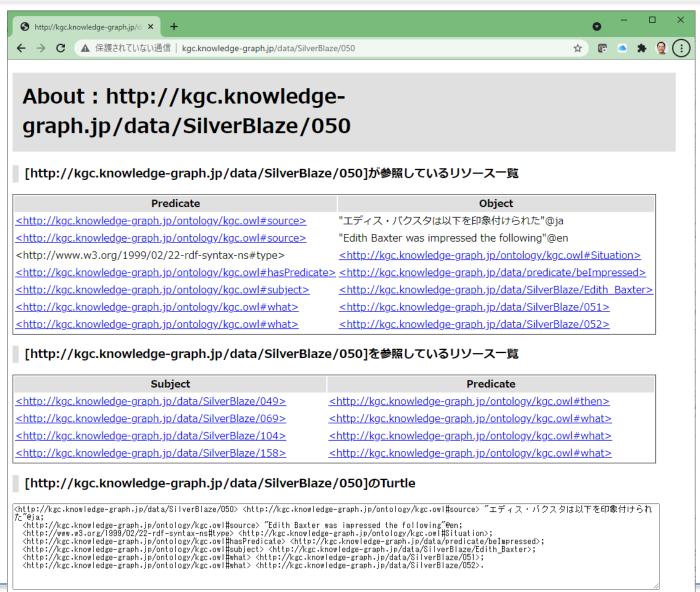


### ナレッジグラフの公開

- → ナレッジグラフの公開場所
  - ➤「推論チャレンジ」のサイト→ナレッジグラフの公開

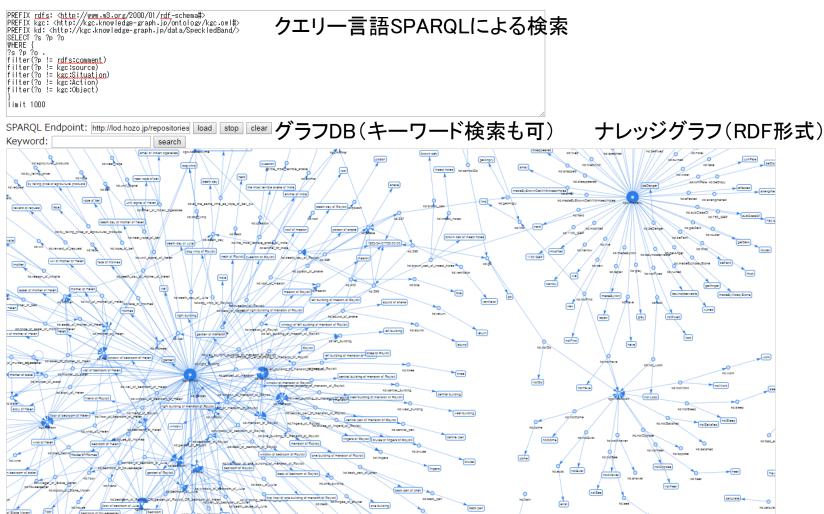
    https://github.com/KnowledgeGraphJapan/KGRC-RDF
  - ➤ 最新版は, 2020/01/19版 (ver.2)
- → 公開形式
  - ➤ 小説ごとのRDFファイル(Tutle形式)のダウンロード
  - ➤ SPARQLエンドポイント[ver.2(2021/01/19版)]
    (APIとして使う) <a href="http://kg.hozo.jp/fuseki/kgrc2020v2/sparql">http://kg.hozo.jp/fuseki/kgrc2020v2/sparql</a>
    (ブラウザから) <a href="http://knowledge-graph.jp/sparql2020v2.html">http://knowledge-graph.jp/sparql2020v2.html</a>
  - ➤ 簡易閲覧ツール <a href="http://kgc.knowledge-graph.jp/">http://kgc.knowledge-graph.jp/</a>
  - ➤ 可視化ツール http://knowledge-graph.jp/visualization/

### 簡易閲覧ツール



### 推理小説のナレッジグラフ

#### → SWO研究会・勉強会での予備的作業を経て、有志数名でナレッジグラフ化



http://knowledge-graph.jp/visualization/



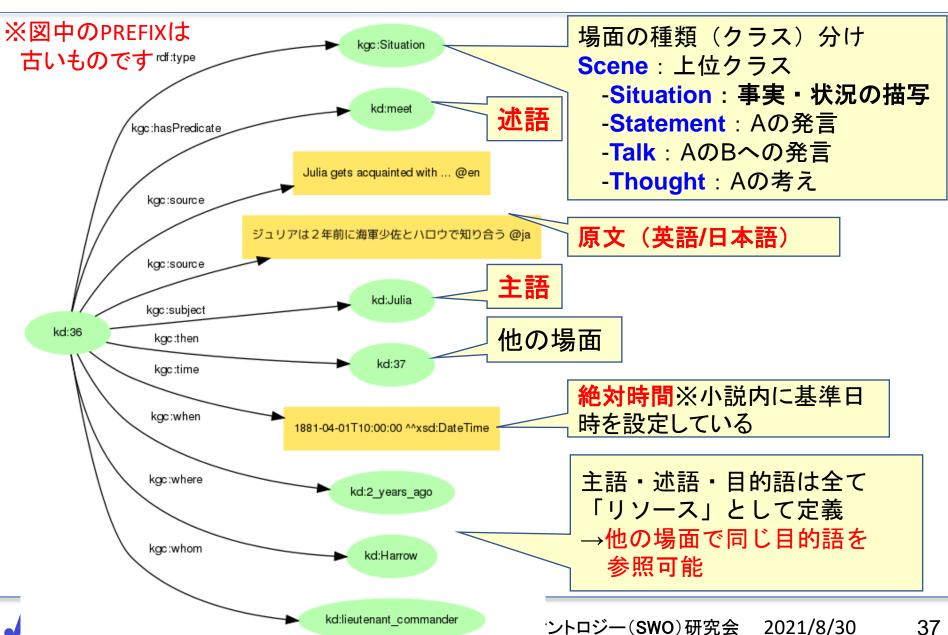
### ナレッジグラフの記述・検索例

→ 以下, 提供するナレッジグラフの記述例. および. そ れを踏まえたSPARQLクエリによる検索例を示す.

- → クエリのサンプルは、下記に公開している
  - ➤ https://github.com/KnowledgeGraphJapan/LOD-ws-2020/blob/master/kgrc2020ws/SPARQL-Sample-KGRC2020.md
  - > これをコピーして http://knowledge-graph.jp/sparql2020v2.html

のクエリ欄に入力することで実行できる

## 場面(シーン)スキーマ 記述例



→ 「まらだのひも」の場面36の全トリプル(述語・目的語) を取得する

→ 「まらだのひも」の場面36のhasPredicate(その場面の内容を表す述語)を取得する

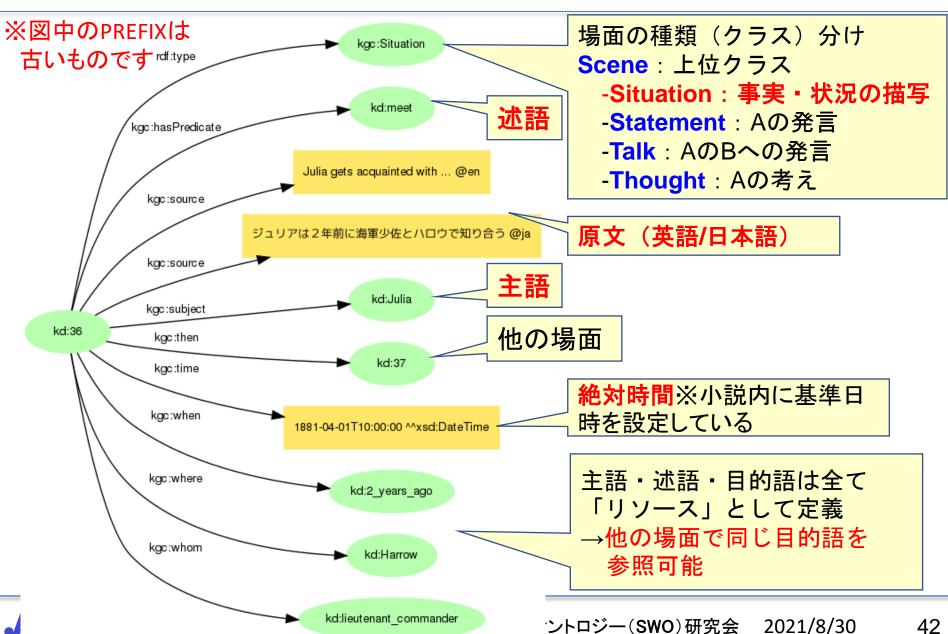
```
PREFIX kgc: <a href="http://kgc.knowledge-graph.jp/ontology/kgc.owl#">
2
3 SELECT DISTINCT *
4 FROM <a href="http://kgc.knowledge-graph.jp/data/SpeckledBand">http://kgc.knowledge-graph.jp/data/SpeckledBand</a>
5 ▼ WHERE {
6 <a href="http://kgc.knowledge-graph.jp/data/SpeckledBand/36">http://kgc.knowledge-graph.jp/data/SpeckledBand/36</a>> kgc:hasPredicate ?o .
7 }
```

→「まらだのひも」の場面36のsubject(その場面の主語) を取得する

→ 「まらだのひも」で用いられているhasPredicate(その場面の内容を表す述語)の一覧を取得する

```
1  PREFIX kgc: <http://kgc.knowledge-graph.jp/ontology/kgc.owl#>
2  
3  SELECT DISTINCT ?o
4  FROM <http://kgc.knowledge-graph.jp/data/SpeckledBand>
5  WHERE {
6     ?s kgc:hasPredicate ?o .
7  }
```

## 場面(シーン)スキーマ 記述例



→ 「まらだのひも」のSituation (事実・状況の描写)の
一覧を取得する

```
1  PREFIX kgc: <http://kgc.knowledge-graph.jp/ontology/kgc.owl#>
2  
3  SELECT DISTINCT ?o
4  FROM <http://kgc.knowledge-graph.jp/data/SpeckledBand>
5  WHERE {
6     ?s kgc:hasPredicate ?o .
7  }
```

## →「まらだのひも」の全場面の一覧を取得する

```
|PREFIX_rdfs: <http://www.w3.org/2000/01/rdf-schema#>
                                 PREFIX rdf: <a href="mailto://www.w3.org/1999/02/22-rdf-syntax-ns#">PREFIX rdf: <a href="mailto://www.w3.org/1999/02/22-rdf-syntax-ns#">http://www.w3.org/1999/02/22-rdf-syntax-ns#</a>>
                                 PREFIX kgc: <a href="http://kgc.knowledge-graph.ip/ontology/kgc.owl#>">http://kgc.knowledge-graph.ip/ontology/kgc.owl#>">http://kgc.knowledge-graph.ip/ontology/kgc.owl#>">http://kgc.knowledge-graph.ip/ontology/kgc.owl#>">http://kgc.knowledge-graph.ip/ontology/kgc.owl#>">http://kgc.knowledge-graph.ip/ontology/kgc.owl#>">http://kgc.knowledge-graph.ip/ontology/kgc.owl#>">http://kgc.knowledge-graph.ip/ontology/kgc.owl#>">http://kgc.knowledge-graph.ip/ontology/kgc.owl#>">http://kgc.knowledge-graph.ip/ontology/kgc.owl#>">http://kgc.knowledge-graph.ip/ontology/kgc.owl#>">http://kgc.knowledge-graph.ip/ontology/kgc.owl#>">http://kgc.knowledge-graph.ip/ontology/kgc.owl#>">http://kgc.knowledge-graph.ip/ontology/kgc.owl#>">http://kgc.knowledge-graph.ip/ontology/kgc.owl#>">http://kgc.knowledge-graph.ip/ontology/kgc.owl#>">http://kgc.knowledge-graph.ip/ontology/kgc.owl#>">http://kgc.knowledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontology/kgc.owledge-graph.ip/ontol
                                  SELECT DISTINCT *
                                  FROM <a href="http://kgc.knowledge-graph.jp/data/SpeckledBand">http://kgc.knowledge-graph.jp/data/SpeckledBand</a>
                               |WHERE |
                                                   ?s rdf:type/rdfs:subClassOf kgc:Scene .
                                                 ?s rdf:type ?o .
10.
```

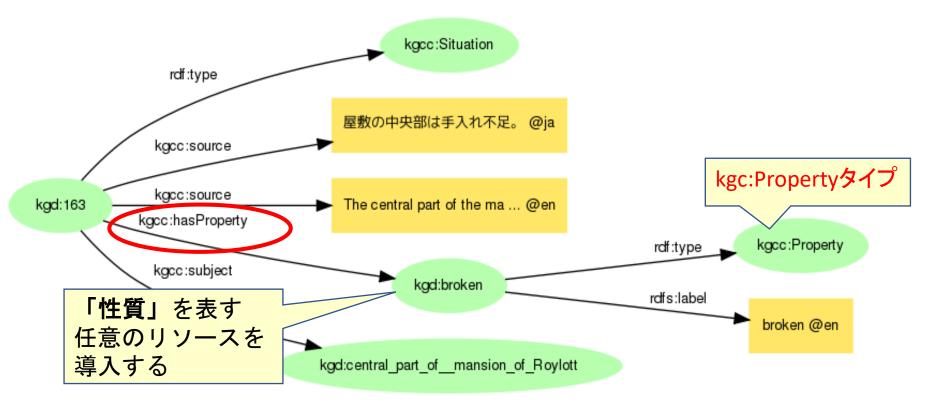
44

## → 「全小説」で述語がmeetの場面の一覧を取得する

```
SELECT DISTINCT *
                     FROM <a href="fromthtp://kgc.knowledge-graph.jp/data/SpeckledBand">http://kgc.knowledge-graph.jp/data/SpeckledBand</a>
                    FROM <a href="FROM">FROM</a> <a href="FROM">http://kgc.knowledge-graph.jp/data/DancingMen">FROM</a> <a href="FROM">FROM</a> <a
5.
                    FROM <a href="http://kgc.knowledge-graph.jp/data/ACaseOfIdentity">http://kgc.knowledge-graph.jp/data/ACaseOfIdentity</a>
                     FROM <a href="http://kgc.knowledge-graph.jp/data/DevilsFoot">http://kgc.knowledge-graph.jp/data/DevilsFoot</a>
                     FROM <a href="fromthtp://kgc.knowledge-graph.jp/data/CrookedMan">http://kgc.knowledge-graph.jp/data/CrookedMan</a>
9
                    WHERE {
                                 ?s kgc:hasPredicate <http://kgc.knowledge-graph.jp/data/predicate/meet> .
```

#### 性質・状態(hasProperty)の例

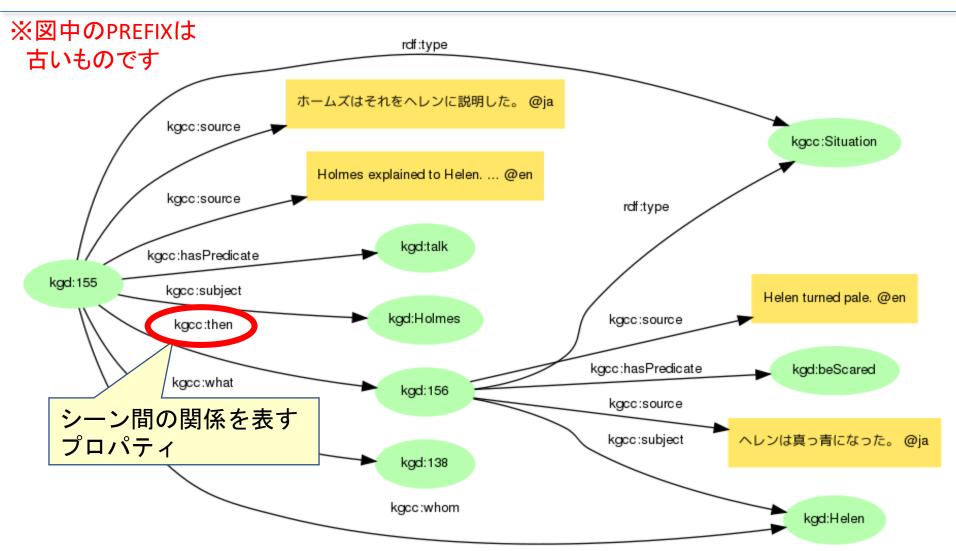
#### ※図中のPREFIXは 古いものです



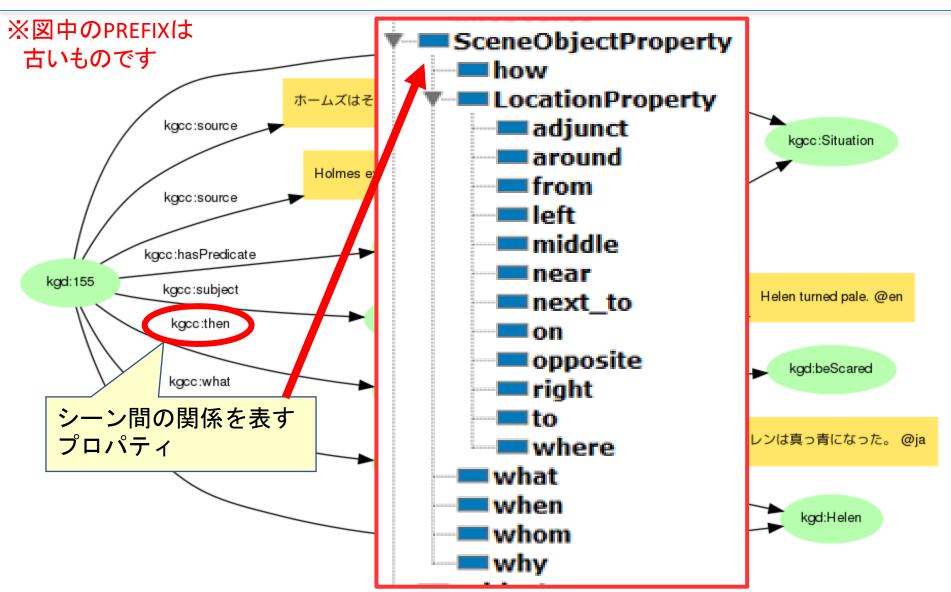
→ 「まらだのひも」で用いられているhasProperty(その場面の内容を表す状態)の一覧を取得する

```
1  PREFIX kgc: <http://kgc.knowledge-graph.jp/ontology/kgc.owl#>
2  
3  SELECT DISTINCT ?o
4  FROM <http://kgc.knowledge-graph.jp/data/SpeckledBand>
5  WHERE {
6     ?s kgc:hasPredicate ?o .
7  }
```

## シーン間の関係



# シーン間の関係の種類





# 場面のつながりの検索

## →「まらだのひも」で場面間のつながりの一覧を取得する

```
1 ▼ PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
                                                      PREFIX rdf: <a href="mailto:rdf">rdf: <a hre
                                                      PREFIX kgc: <a href="http://kgc.knowledge-graph.jp/ontology/kgc.owl#">http://kgc.knowledge-graph.jp/ontology/kgc.owl#>
           4
                                                      SELECT DISTINCT *
                                                      FROM <a href="FROM">FROM</a> <
                                                    WHERE !
                                                                               ?s rdf:type/rdfs:subClassOf kgc:Scene .
                                                                              ?o rdf:type/rdfs:subClassOf kgc:Scene ...
                                                                            ?s ?p ?o .
10
```

# 各語彙のIRIの決め方

- → 対象とする小説が増えたことで、異なる小説を横断した 処理ができるようにIRIの決め方を統一
- ◆ 各小説ごとに「ベースIRI」および「グラフIRI」
  - > < http://kgc.knowledge-graph.jp/data/小説名/>
  - ➤ 例:「踊る人形」のシーン100は <a href="http://kgc.knowledge-graph.jp/data/DancingMen/100">http://kgc.knowledge-graph.jp/data/DancingMen/100</a>>
- → 小説をまたいで共通化して参照する語彙のIRI

  - > 固有名詞(例:ホームズ, 地名)
    - →<http://kgc.knowledge-graph.jp/data/YYY>

# SPARQLの参考資料(1)



- 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 <a href="http://sparqlbook.jp/">http://sparqlbook.jp/</a>
- **SPARQLに関する解説記事** 
  - 山本 泰智,知識グラフ,セマンティックウェブを構成する RDFと問い合わせ言語SPARQL,情報の科学と技術, 2020,70 巻,8号,p. 392-398, <a href="https://doi.org/10.18919/jkg.70.8\_392">https://doi.org/10.18919/jkg.70.8\_392</a>
  - 古崎 晃司, ウェブの情報資源活用のための技術: ナレッジグラフとしてのLOD活用, 情報の科学と技術, 2020, 70 巻, 6 号, p. 303-308 https://doi.org/10.18919/jkg.70.6\_303



http://sparqlbook.jp/より

- Wikidataを例としたSPARQL演習スライド+解説動画(by古崎)
  - https://github.com/KnowledgeGraphJapan/LOD-ws-2020/blob/master/SPARQL/SPARQL-Practice-Wikidata.md

# SPARQLの参考資料(2)



#### ■ 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/2oDPl0Q
- Wikidataを使った日本の政治家の出身大学ランキング <a href="http://bit.ly/2PBt8fn">http://bit.ly/2PBt8fn</a>

#### ■ 技術書典

- 技術書典7,8にてSPARQL,LOD関係の本を頒布
  <a href="https://techbookfest.org/event/tbf08/circle/5687828398735360?fbclid=IwAR30949R49r1giptnzA66f\_BGB52T46JrU\_qGhM\_c9KIizmUA2YYdPzyyyA">https://techbookfest.org/event/tbf08/circle/5687828398735360?fbclid=IwAR30949R49r1giptnzA66f\_BGB52T46JrU\_qGhM\_c9KIizmUA2YYdPzyyyA</a>
- 技術書典9(9/12-22)にても配本予定 https://techbookfest.org/event/tbf09