

# RDFデザインパターン

情報システム研究機構 データサイエンス共同利用基盤施設 ライフサイエンス統合データベースセンター 川島秀一 2019/10/31

#### SPARQLを支える技術

仕様

RDF (Resource Description Framework)

RDFS (Resource Description Framework)

**OWL** (Web Ontology Language)

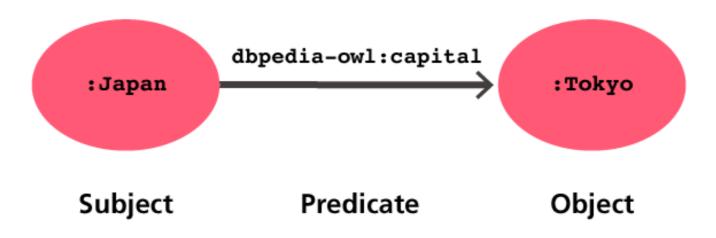
SPARQL (SPARQL Protocol and RDF Query Language)

ソフトウェア/ サービス RDFストア / トリプルストア SPARQLエンドポイント

## RDFの構造



## RDF



オープンデータ時代の標準WEB API SPARQL(インプレス)より

## URL, URI, IRI

#### IRI ed Resource lo

Internationalized Resource Identifier

Unicode

#### URI

Uniform Resource Identifier

Ascii code

#### URL

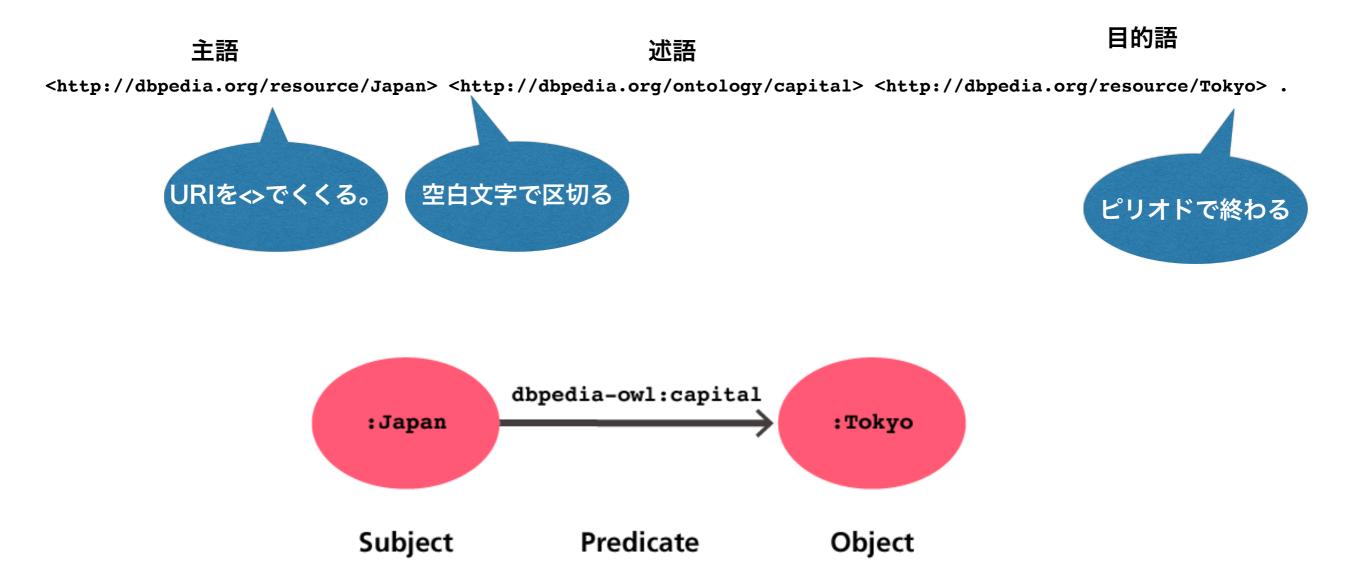
Uniform Resource Locator

#### URN

Uniform Resource Name

#### **RDF**

## RDFの例(N-tripleフォーマット)



#### RDF

#### turtle フォーマットでのRDFの例

@prefix : <http://dbpedia.org/resource/> .
@prefix dbpedia-owl: <http://dbpedia.org/ontology/> .

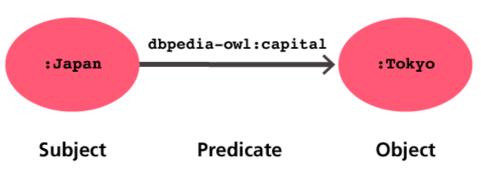
:Japan dbpedia-owl:capital :Tokyo .

prefix

接頭辞

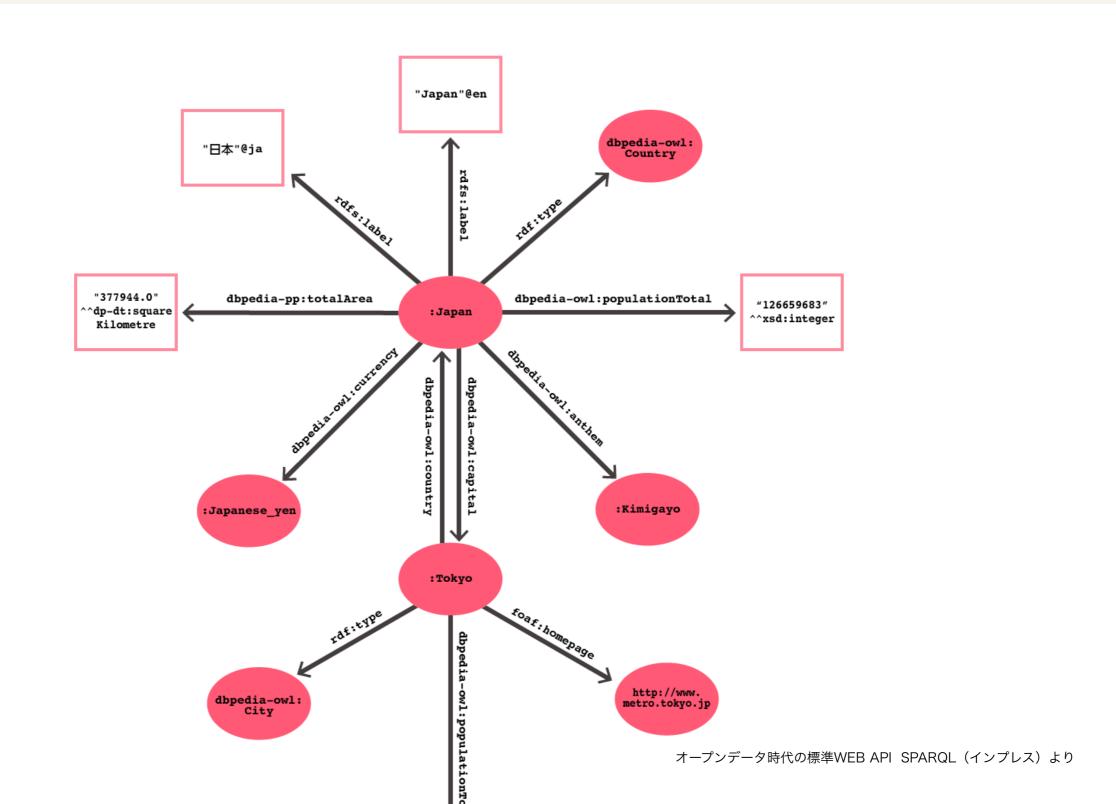
URIを、prefixとロー カルパートの組み合わ せで表現する | ローカルパート ローカル部分

URIのうち、prefix と して定義された以外の 残りの部分 ファイルの先頭の、 @prefix 行にprefix の定義がされている。



オープンデータ時代の標準WEB API SPARQL(インプレス)より

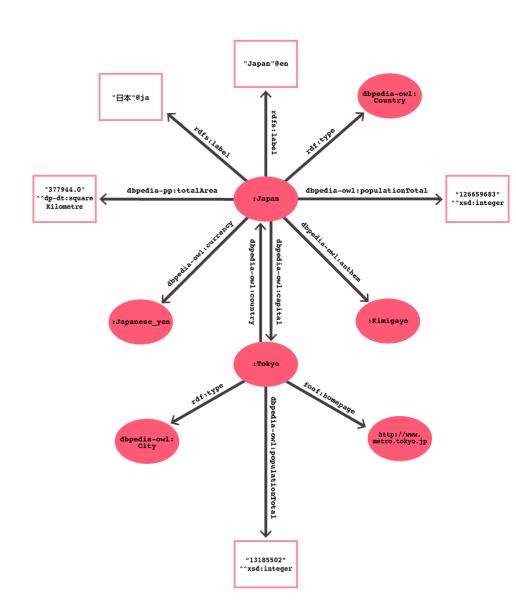
## RDFグラフ



#### RDF Turtle フォーマット

```
@prefix : <http://dbpedia.org/resource/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix dbpedia-owl: <http://dbpedia.org/ontology/> .
@prefix dbpedia-pp: <http://dbpedia.org/ontology/PopulatedPlace/> .
@prefix dbpedia-dt: <http://dbpedia.org/datatype/> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
:Japan dbpedia-owl:capital :Tokyo .
:Japan rdf:type dbpedia-owl:Country .
:Japan dbpedia-owl:anthem :Kimigayo .
:Japan dbpedia-owl:currency :Japanese yen .
:Japan rdfs:label "Japan"@en .
:Japan rdfs:label "日本"@ja .
:Japan dbpedia-owl:populationTotal 126659683 .
:Japan dbpedia-pp:totalArea "377944.0"^^dbpedia-dt:squareKilometre .
:Tokyo rdf:type dbpedia-owl:City .
:Tokyo dbpedia-owl:country :Japan .
:Tokyo dbpedia-owl:populationTotal 13185502 .
:Tokyo foaf:homepage <a href="http://www.metro.tokyo.jp">http://www.metro.tokyo.jp</a>.
```

右のグラフ図のRDF (Turtle フォーマット)



#### RDFリテラル 1

- ・リテラルは、URI(IRI)以外の値。
- ・文字列、日付、数値など。
- ・リテラルは、目的語の位置にだけくる。
- ・データ型のURIを用いて、リテラルの型を修飾することができる。
- ・文字列は、言語タグで修飾することができる。

:Japan rdfs:label "Japan" .

引用符 " でくくり、データ型で修飾され ていないと、文字列として扱われる

:Japan rdfs:label "Japan"^^xsd:string .

上の行と同じ意味。

:Japan rdfs:label "Japan"@en .
:Japan rdfs:label "日本"@ja .

言語タグで修飾することで、文字列の言語を指定することができる。

#### RDFリテラル 2

:Japan dbpedia-owl:populationTotal "126659683"^^xsd:decimal .

数値が、引用符 " でくくってあって、数値を表すデータ型で修飾されていると、数値として扱われる。126659683は数値として扱われる。

:Japan dbpedia-owl:populationTotal "126659683" .

データ型で修飾されていないと、126659683は、文字列として扱われる。

:Japan dbpedia-owl:populationTotal 126659683 .

数字が、引用符で括られていないと、数値として扱われる。

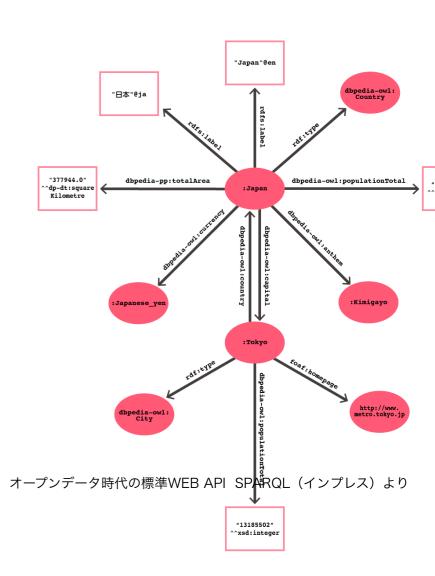
:Japan dbpedia-pp:totalArea "377944.0"^^dbpedia-dt:squareKilometre .

このように独自のURIで修飾することもできます。

### RDF Turtle フォーマット 主語の省略

```
@prefix : <http://dbpedia.org/resource/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix dbpedia-owl: <http://dbpedia.org/ontology/> .
@prefix dbpedia-pp: <http://dbpedia.org/ontology/PopulatedPlace/> .
@prefix dbpedia-dt: <http://dbpedia.org/datatype/> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
:Japan dbpedia-owl:capital :Tokyo ;
   rdf:type dbpedia-owl:Country ;
   dbpedia-owl:anthem :Kimigayo ;
   dbpedia-owl:currency :Japanese yen ;
   rdfs:label "Japan"@en ;
   rdfs:label "日本"@ja;
   dbpedia-owl:populationTotal 126659683;
   dbpedia-pp:totalArea "377944.0"^^dbpedia-dt:squareKilometre ;
   dbpedia-owl:country :Japan .
:Tokyo rdf:type dbpedia-owl:City;
   dbpedia-owl:country :Japan ;
   dbpedia-owl:populationTotal 13185502 ;
   foaf:homepage <http://www.metro.tokyo.jp> .
```

右のグラフ図のRDF(Turtle フォーマット) 同じ主語のトリプルをまとめて、2度めは省略する記法を使った例。 連続する行が同じ主語なので、2行目以降は省略してよい。その際、主語の書かれた1行目の末尾は、. (ピリオドでなく、セミコロンでおわる)また、同じ主語でまとめられた最後の行のみ、ピリオドで終わる。

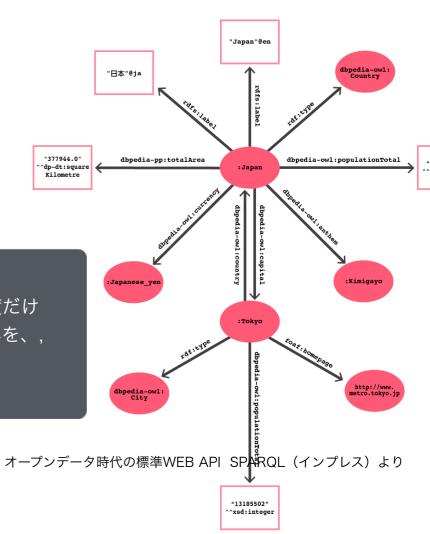


## RDF Turtle フォーマット 述語の省略

```
@prefix : <http://dbpedia.org/resource/> .
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix dbpedia-owl: <http://dbpedia.org/ontology/> .
@prefix dbpedia-pp: <http://dbpedia.org/ontology/PopulatedPlace/> .
@prefix dbpedia-dt: <http://dbpedia.org/datatype/> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
:Japan dbpedia-owl:capital :Tokyo ;
 a dbpedia-owl:Country ;
 dbpedia-owl:anthem :Kimigayo ;
 dbpedia-owl:currency :Japanese yen ;
 rdfs:label "Japan"@en, "日本"@ja;
 dbpedia-owl:populationTotal 126659683;
 dbpedia-pp:totalArea "377944.0"^^dbpedia-dt:squareKilometre ;
 dbpedia-owl:country: Japan .
:Tokyo a dbpedia-owl:City;
 dbpedia-owl:country :Japan ;
 dbpedia-owl:populationTotal 13185502 ;
 foaf:homepage <http://www.metro.tokyo.jp> .
```

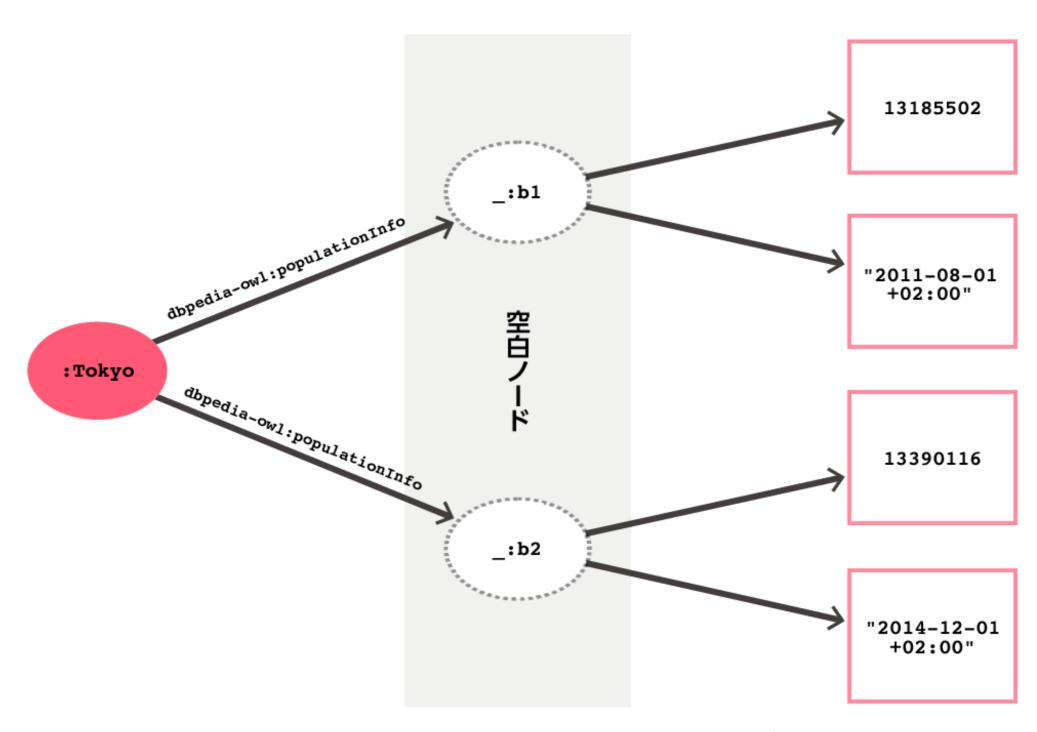
右のグラフ図のRDF(Turtle フォーマット) 省略記法を全て利用 Turtle フォーマットでは、rdf:type を、a の一文字で記述できる。 (この記法は、rdf:type が表すインスタンス関係が、この場合だと、Japan (is) a country . を意味していることからき

ている)



連続する行が同じ述語なので、一度だけ 書いて、次の行からは、目的語のみを、, (カンマ)区切って記述できる。

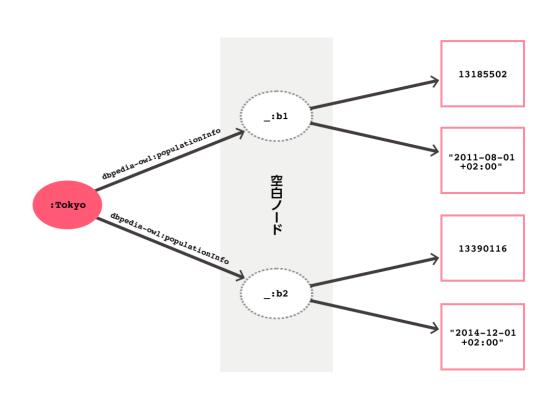
## RDF 空白ノードの利用



## RDF 空白ノードの利用

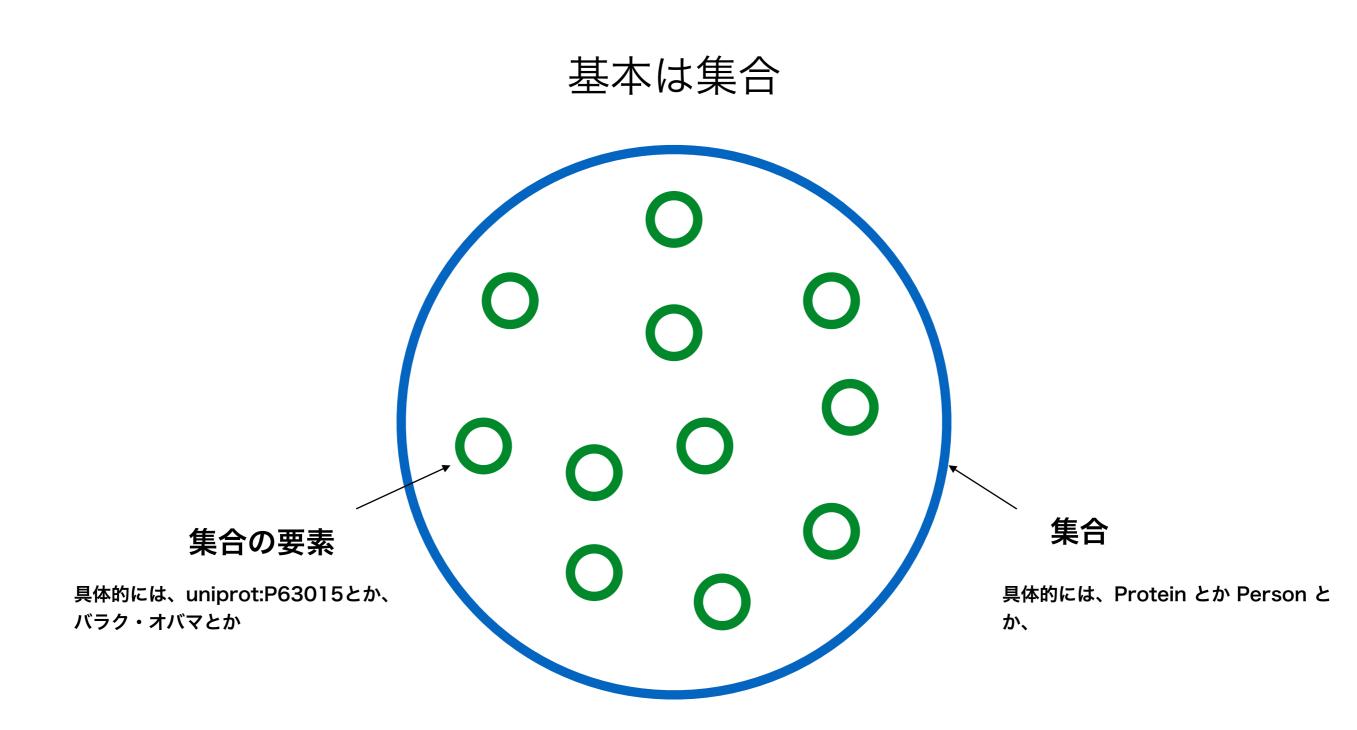
```
@prefix : <http://dbpedia.org/resource/> .
@prefix dbpedia-owl: <http://dbpedia.org/ontology/> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix sb: <http://sparqlbook.jp/book/> .
:Tokyo
   sb:populationInfo [
       dbpedia-owl:populationAsOf "2011-08-01"^^xsd:date ;
       dbpedia-owl:populationTotal 13185502
   ],[
       dbpedia-owl:populationAsOf "2014-12-01"^^xsd:date ;
       dbpedia-owl:populationTotal 13390116
   ] .
           右のグラフ図のRDF(Turtle フォーマット)
                空白ノードは、[]で表現する。
   a foaf:Person .
```

空白ノードで始まってもよい。 この例だと、ある匿名の人物がいました。を意味します。

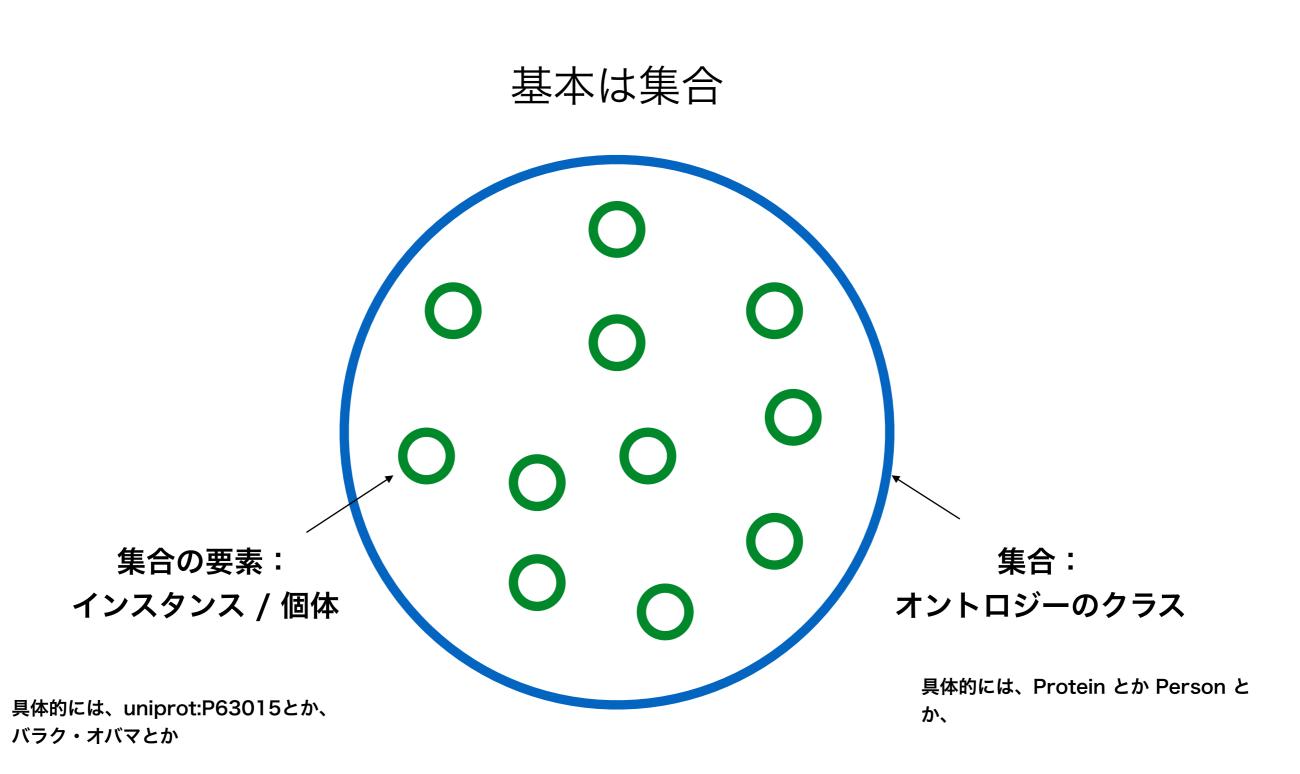


オープンデータ時代の標準WEB API SPARQL(インプレス)より

## RDFデータ+ OWLオントロジーの考え方

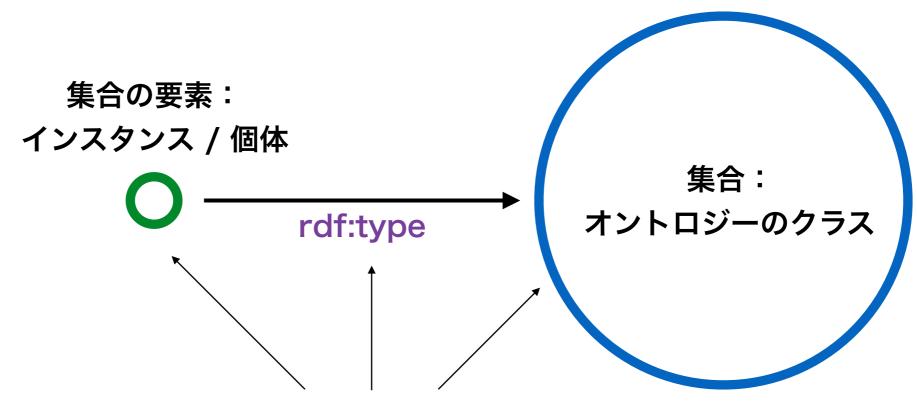


## RDFデータ+OWLオントロジーの考え方



## RDFデータ+OWLオントロジーの考え方

#### RDFで記述するには



3つともURIリソースとして記述する

dbpedia:David\_Bowie

rdf:type

foaf:Person

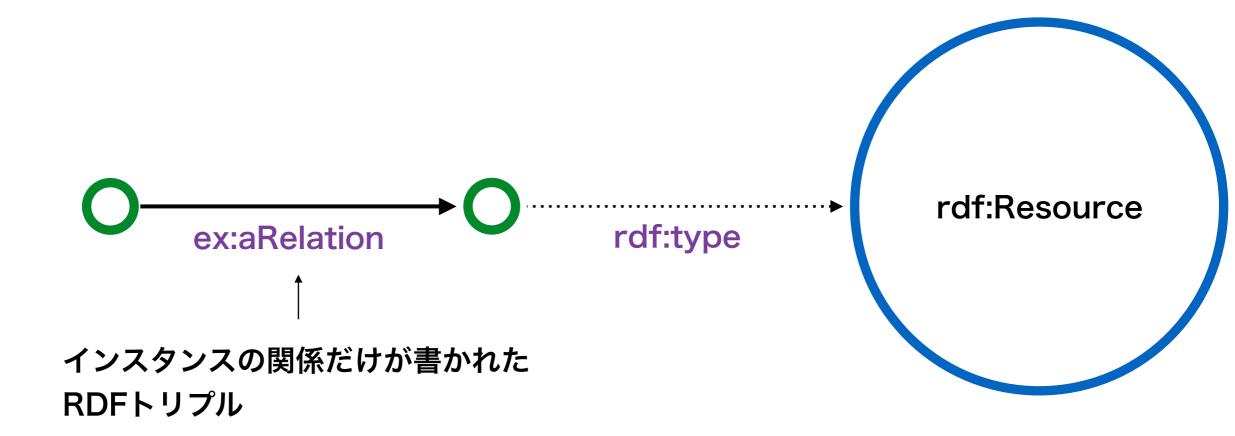
uniprot:P63015

rdf:type

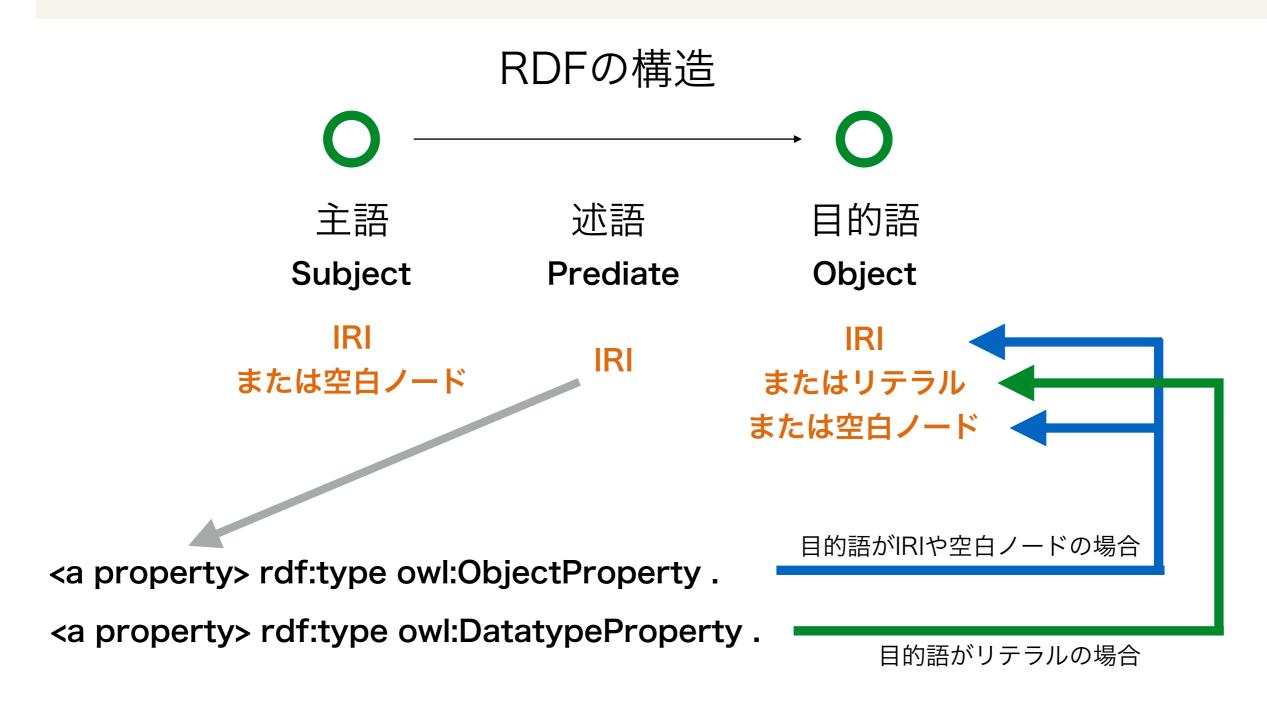
core:Protein

## RDFデータ+OWLオントロジーの考え方

### オントロジークラスがない場合



## RDF Property のタイプ



主語や目的語と同様に、述語もタイプが定義されています。

#### RDFとは

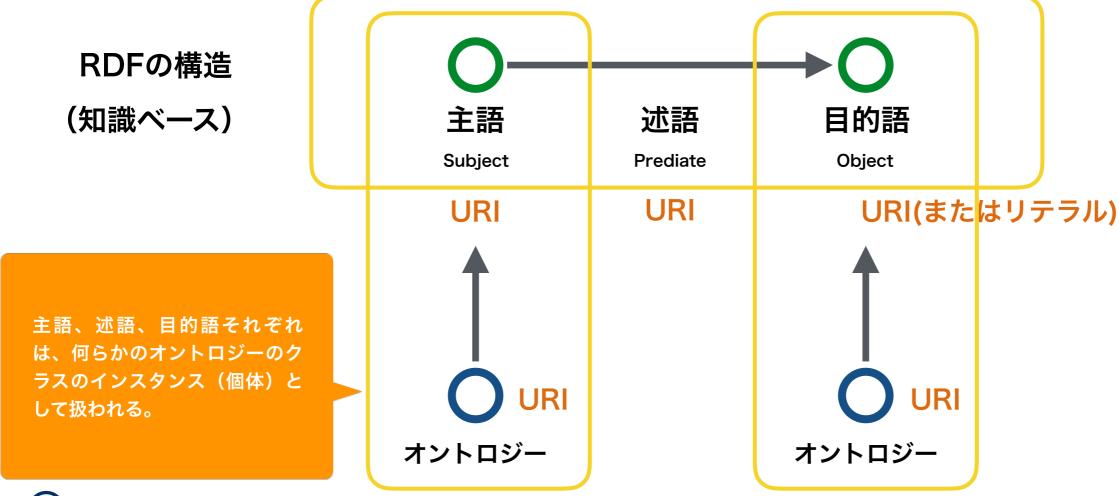


**RDF**: Resource Framework Description

Resource: ウェブ上にある(=URIを持つ)情報資源

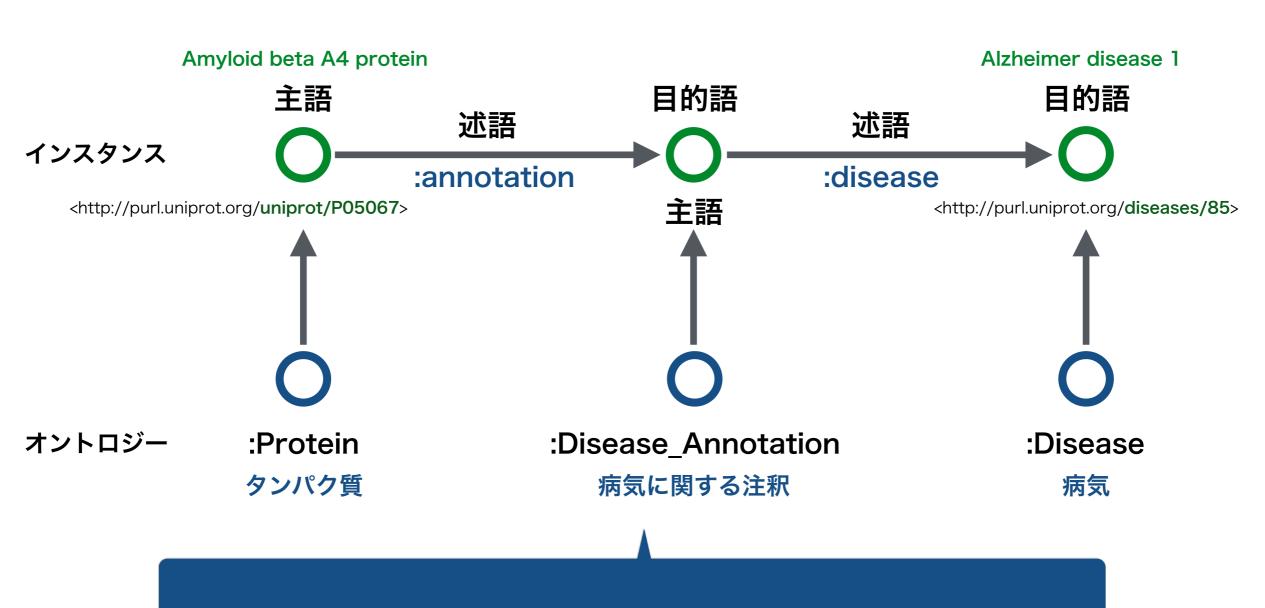
情報を3つのURIの組み合わせで記述するための枠組み

RDF 1.1 から正確にはIRI



#### RDFの例

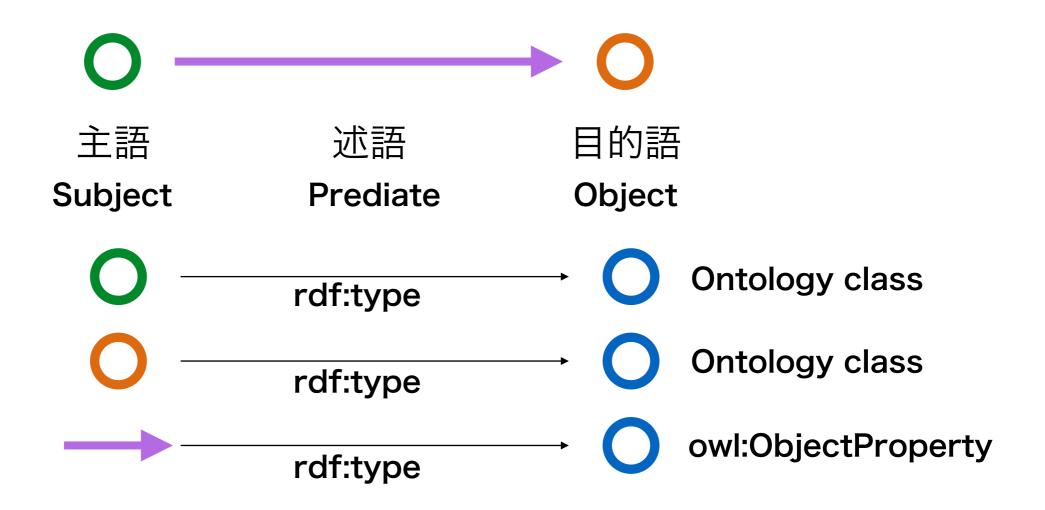




RDFとして、リソースの関係を記述していくことは、各リソースに対応するオントロジーのクラスをマップする必要がある

## RDF Property のタイプ

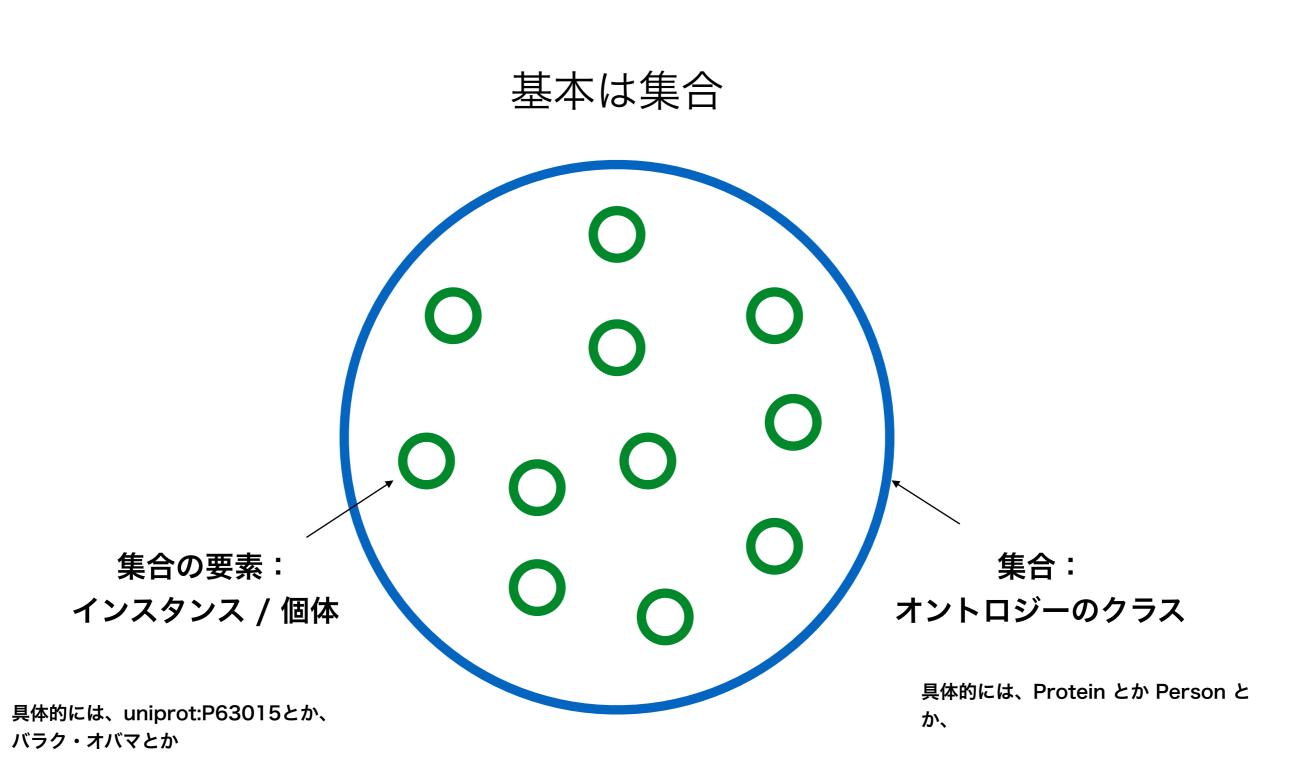
### RDFの構造



### 用語の整理

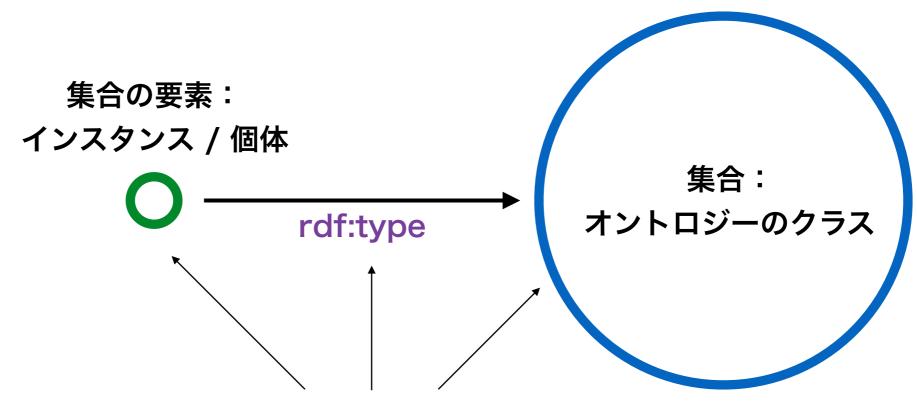
- ・セマンティック・ウェブ界隈では、同じ内容に対して、異なる用語が使われ ていることがあるので、初心者には分かりにくいことがある。
- ・URL ≅ URI ≅ IRI ≅ リソース
- ・インスタンス≅個体≅Individual
- ・述語 (predicate) ≅ プロパティ (property)
- ・知識表現の分野などで使うis-a関係は、rdfs:subClassOf
- ・instance-of関係は、rdf:type (turtleでは、rdf:typeを a と省略できる (is aから)。

## リソースはオントロジーのクラスのインスタンスにする



## リソースはオントロジーのクラスのインスタンスにする

#### RDFで記述するには



3つともURIリソースとして記述する

dbpedia:David\_Bowie

rdf:type

foaf:Person

uniprot:P63015

rdf:type

core:Protein

## アノテーション情報をできるだけ記載する

- リソースには、rdf:label プロパティを用いて、ラベルをつける
  - 検索結果の表示や、アプリケーションでの表示の際に、ラベルがある と便利
- 必要に応じてdcterms:description や skos:definition などを用いて、 リソースに関する説明を記述すると、データの意味がわかりやすくなる。
  - 特に、オントロジーのクラスにはつけるべき
- ローカルIDを、dcterms:identifierプロパティを用いて記述する
  - リソースURIから、ローカルパートを切り出す手間が省ける

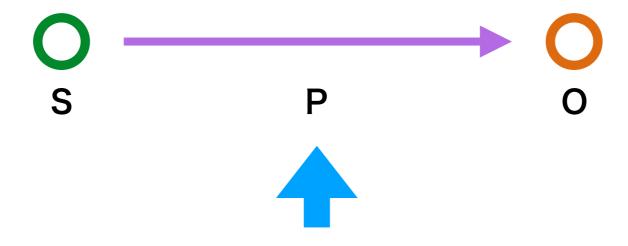
## URIのデザイン

- RDFを作成する際に、URIのデザインは重要なパートになります。
  - できるだけ将来にわたって変化しないURIがいいとされています
  - 例えば、CGI引数をとるようなURIは使わない
    - とはいえ、すでにあるデータベースをRDF化するような場合には難しい場合合もありますが
    - よい例 http://example.org/database/ID
  - データベースのバージョンや日付などが入っていない方がよい
  - httpとhttpsに気をつける

#### 参考

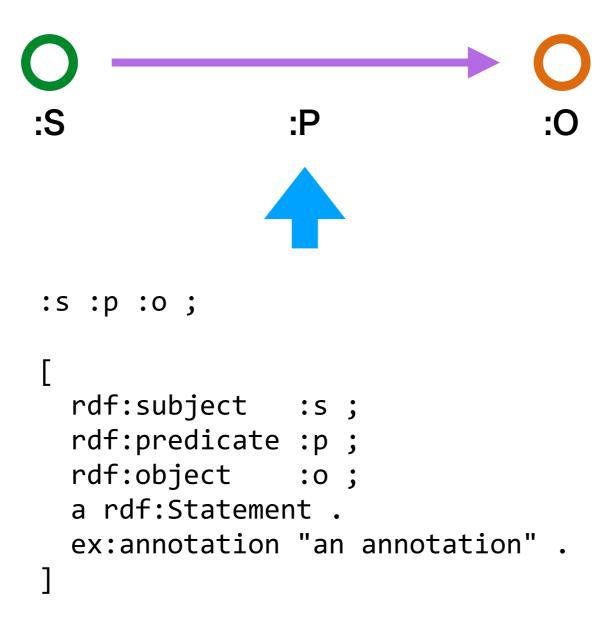
- 'Cool URIs for the Semantic Web'
  - https://www.w3.org/TR/cooluris/
- ・セマンティック Web のためのクールな URI
  - https://standards.mitsue.co.jp/resources/w3c/TR/2008/NOTEcooluris-20080331/
- ・クールなURIは変わらない
  - https://www.kanzaki.com/docs/Style/URI

## ステートメントに注釈を加えたい場合

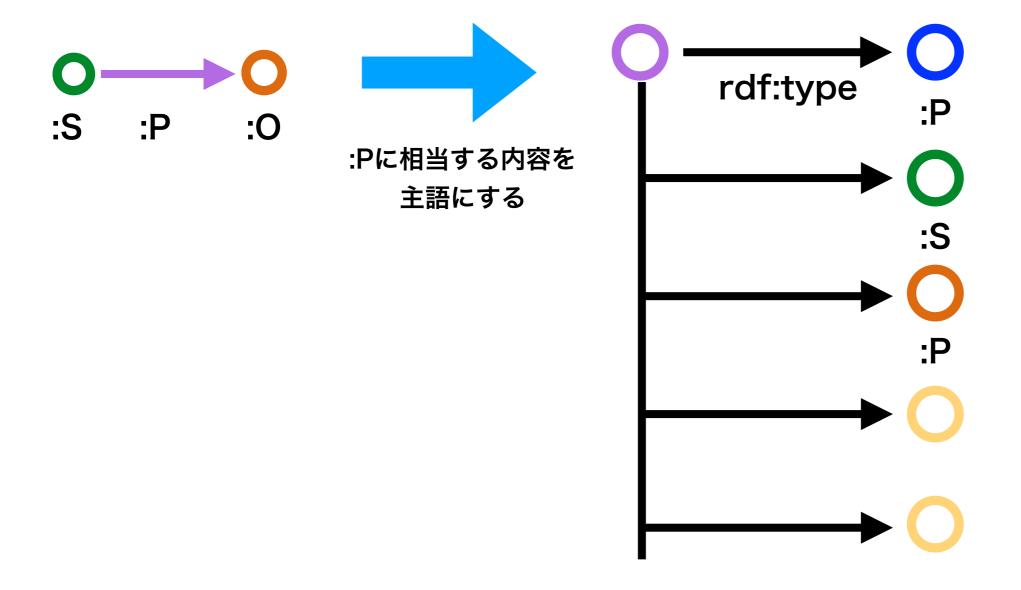


このSPOが述べている内容に情報を付加したい

## Reification (具体化)



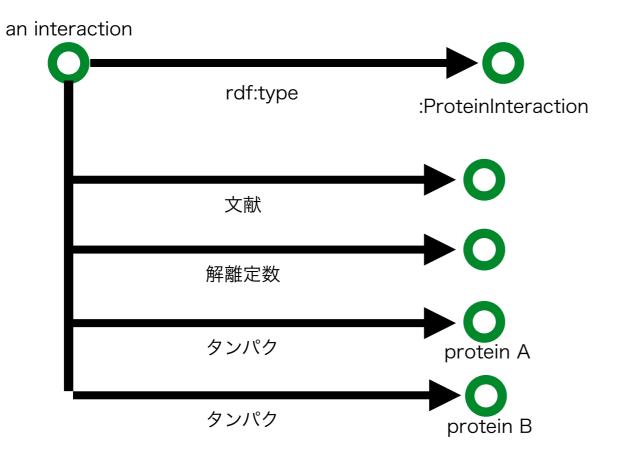
## 述語にあたるリソースをクラスとして明示的に定義する



#### タンパク間相互作用のRDF

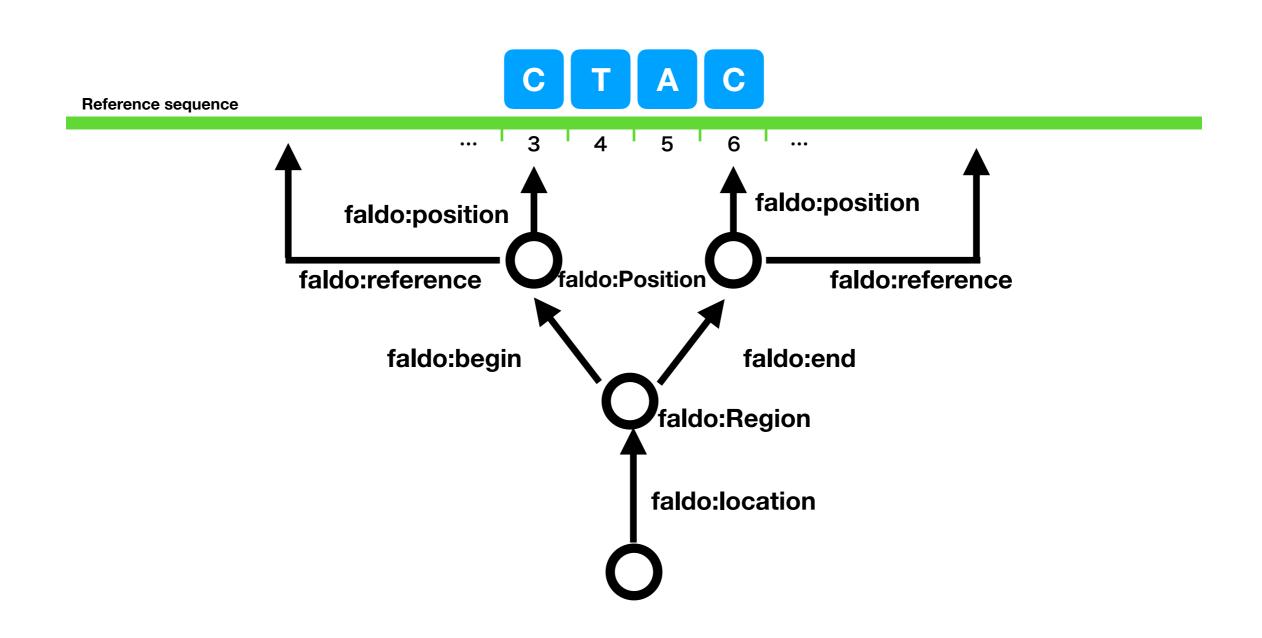


このトリプルだと、タンパク間相互作用に関する、付加情報(文献などの由来情報や、解離定数等)を追加したいときに、どのように記述すればよいかという問題がでてくる。

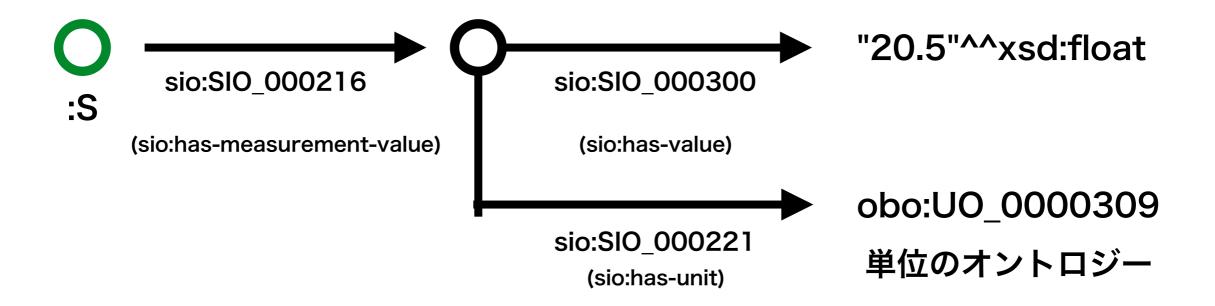


解決策: ProteinInteraction のようなクラスを定義して、そのインスタンスに、必要な情報を記述していく。

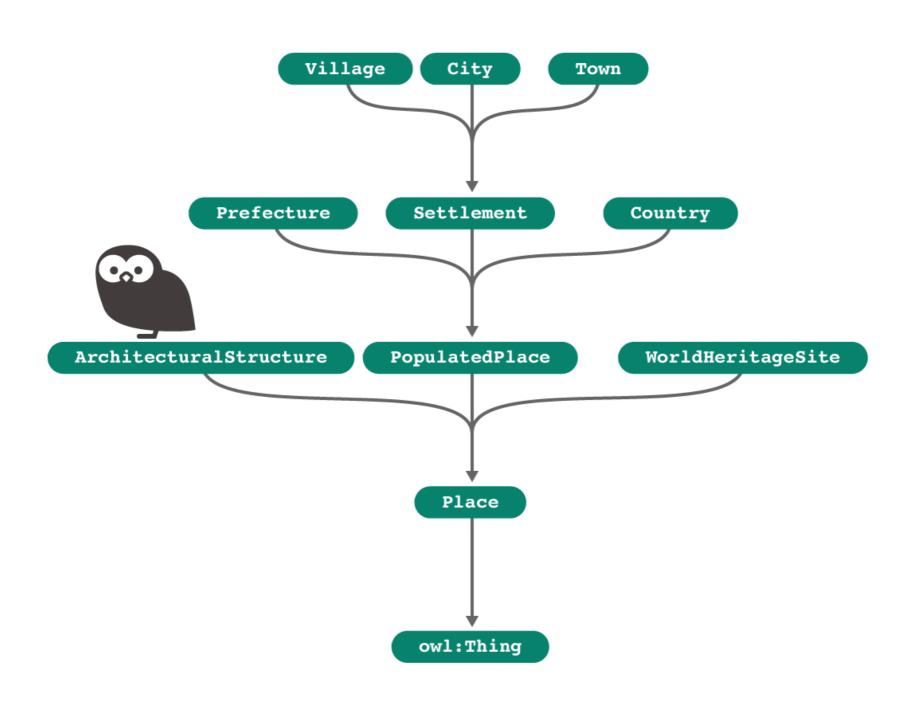
## ゲノム座標に注釈を記述する場合



## 単位のついた値を記述したい場合



## オントロジーを適切に使う

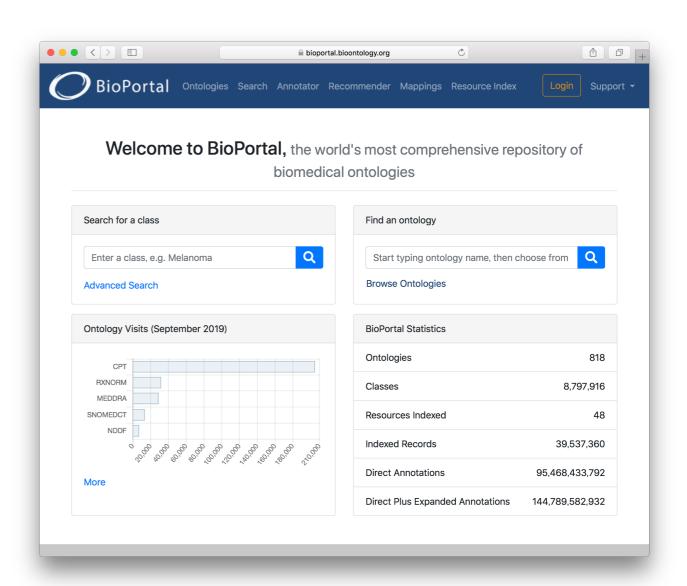


### オントロジー利用の注意点

- クラスとプロパティを間違えて使う方が時々おられますが、クラスをRDFの述語 部位に使うことはできません。
- ・オントロジーは、ターミノロジーではないので、クラス名やラベルだけで内容を 判断せず、そのクラスの定義を確認したり、階層関係(上位や下位の)を確認す ることが誤用を防ぐことにつながります。

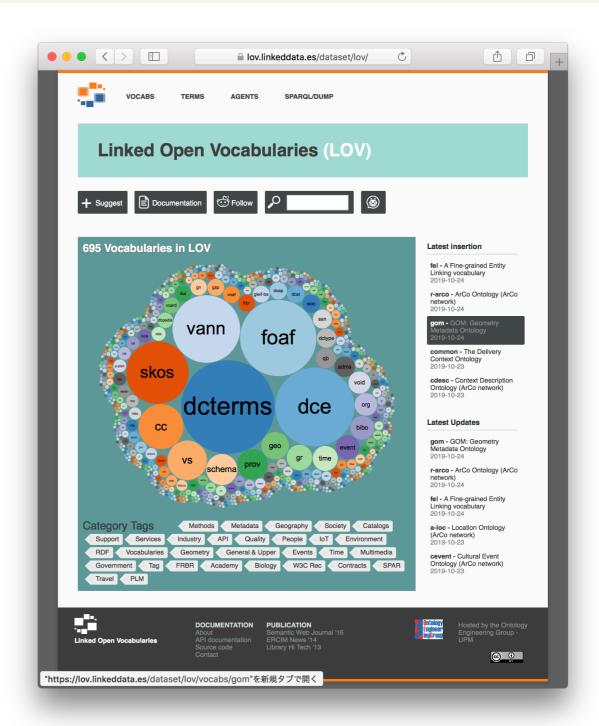
## オントロジーの探し方: BioPortal

- 生命科学系のオントロジーは、 BioPortalに多く収録されています。 (2019年10月時点で、818オントロジー)
- オントロジーおよび、クラスを検索 することができます。
- プロパティを検索する方法はありません(これは、特に初心者の方には難しいところです。)。



## オントロジーの探し方: LOV

Linked Open Vocabularies には、生命科学系オントロジー以外のオントロジーが多く掲載されています(2019年10月時点で、695オントロジー)



#### RDFガイドラインに紹介されている語彙・オントロジー



#### (生命科学ではない) 一般的な語彙・オントロジー

オントロジー名		主な内容	
RDF	Resource Description Framework	RDFの基本語彙	
RDFS	RDF Schema	RDFの構造を記述する語彙	
OWL	Web Ontology Language	オントロジーを記述	
DC	Dublin Core	基本メタデータ	
DC terms	DCMI Metadata Terms	DCの拡張語彙	
SKOS	Simple Knowledge Organization System	既存知識間のマッピング	
FOAF	Friend of a Friend	人間/組織の関係	
VoID	Vocabulary of Interlinked Datasets	データベース間の関係	
UO	Ontology of Units of Measurement	単位	
QUDT	Quantities, Units, Dimensions, and Types Ontology	単位、次元、量	
PROV-O	PROV Ontology	由来情報	
PAV	Provenance, authoring and versioning	由来情報、著者情報等	
XSD	W3C XML Schema Definition Language	データ型	

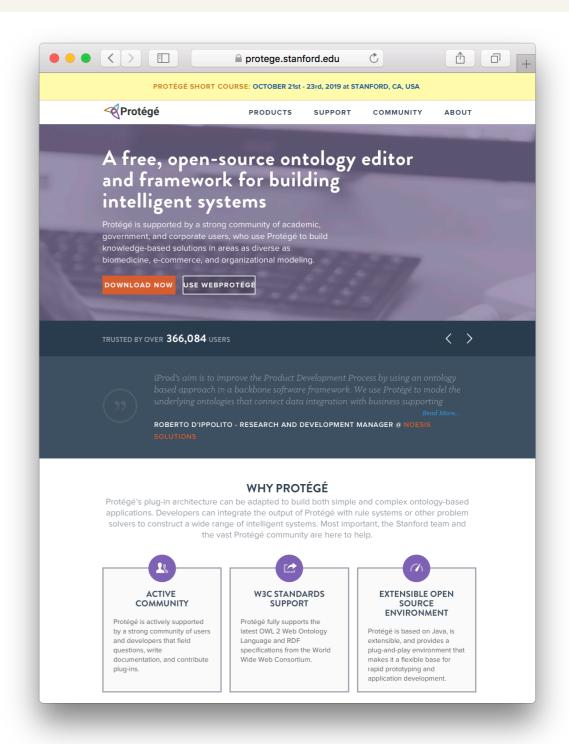


# オントロジー作成の方法

情報システム研究機構 データサイエンス共同利用基盤施設 ライフサイエンス統合データベースセンター 川島秀一 2019/10/31

# Protégé オントロジーエディタ

- スタンフォード大で開発されている、最 も広く使われているオントロジーエディ タ(登録ユーザ数36万人)
- https://protege.stanford.edu/からフリーでダウンロードできます。



# 例題:日本酒オントロジー

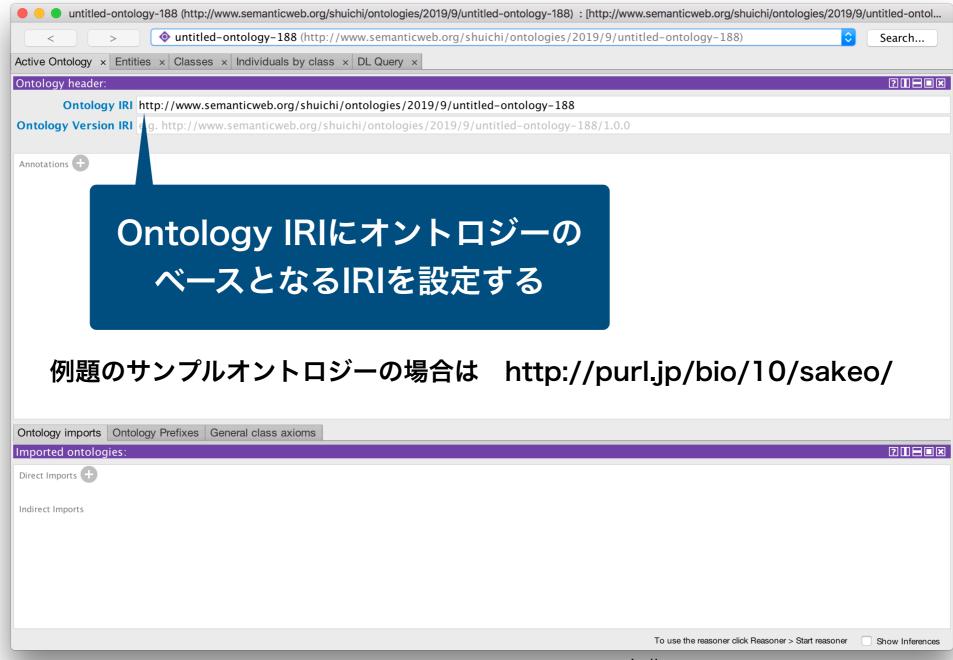
#### 特定名称の清酒の表示[10]

特定名称	使用原料	精米歩合	香味等の要件	こうじ米使用割合
本醸造酒	米、米こうじ、水、醸造アルコール	70%以下	香味、色沢が良好	
特別本醸造酒	木、木とりし、小、醸造アルコール	60%以下又は特別な製造方法(要説明表示)	香味、色沢が特に良好	
純米酒	米、米こうじ、水	_	香味、色沢が良好	
特別純米酒	木、木とりし、水	60%以下又は特別な製造方法(要説明表示)	香味、色沢が特に良好	15%以上
吟醸酒	米、米こうじ、水、醸造アルコール	60%以下	吟醸造り、固有の香味、色沢が良好	15%以上
純米吟醸酒	米、米こうじ、水	00%以下	内嵌色り、固有の音味、色水が及対	
大吟醸酒	米、米こうじ、水、醸造アルコール	50%以下	吟醸造り、固有の香味、色沢が特に良好	
純米大吟醸酒	米、米こうじ、水	50%以下	内瞰色り、四行の音味、巴沢が付に及好	

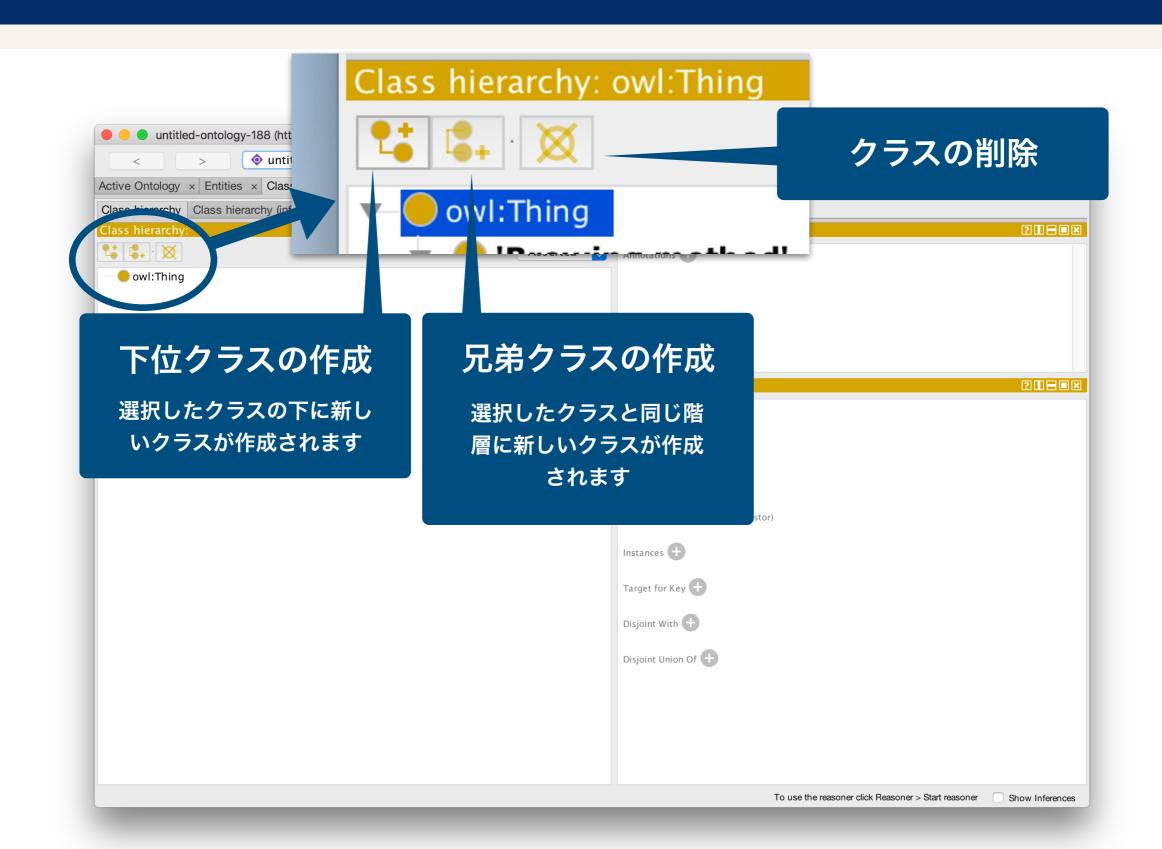
#### 特定名称の規則性を表化

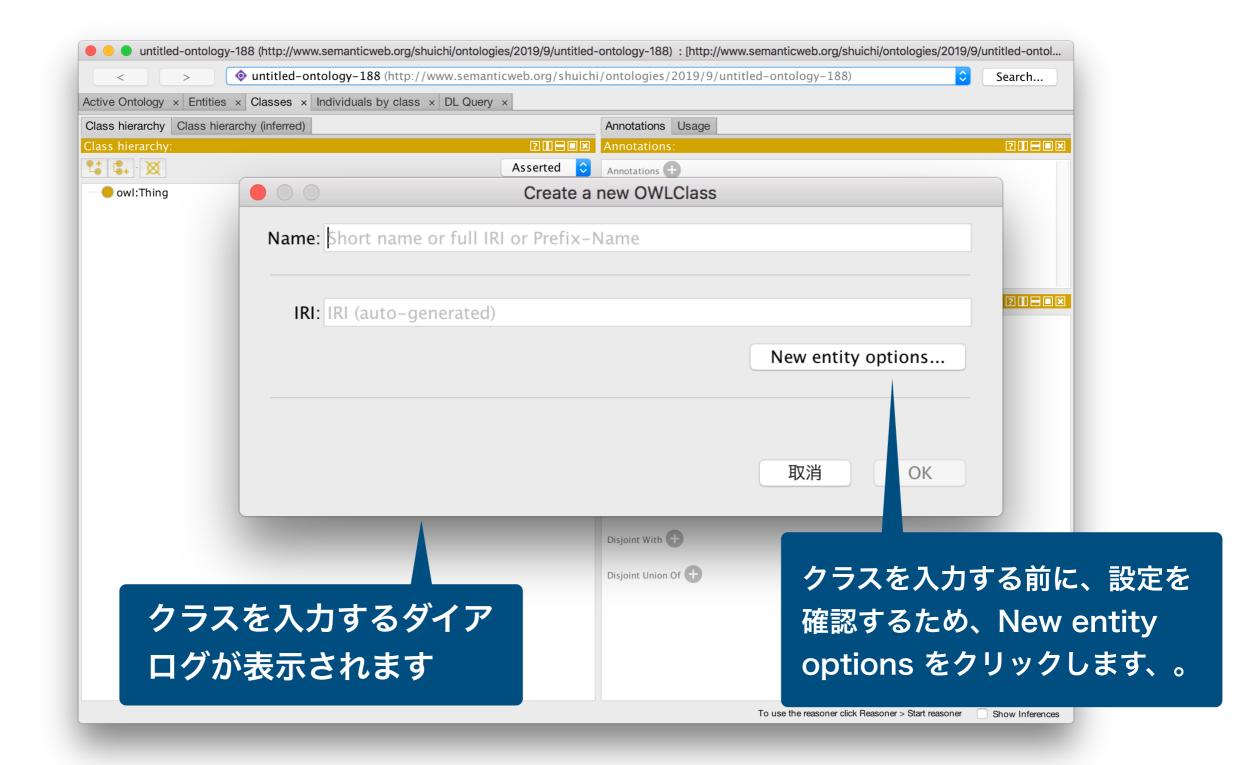
	吟醸造り	吟醸造り	特別な製造方法	
	かつ	かつ	又は	_
	精米歩合:50%以下	精米歩合:60%以下	精米歩合:60%以下	
醸造アルコール: <b>なし(</b> 純米)	純米大吟醸酒	純米吟醸酒	特別純米酒	純米酒
醸造アルコール: <b>あり</b> (アル添)	大吟醸酒	吟醸酒	特別本醸造酒	<b>本醸造酒</b> (精米歩合: <b>70%以下</b> )

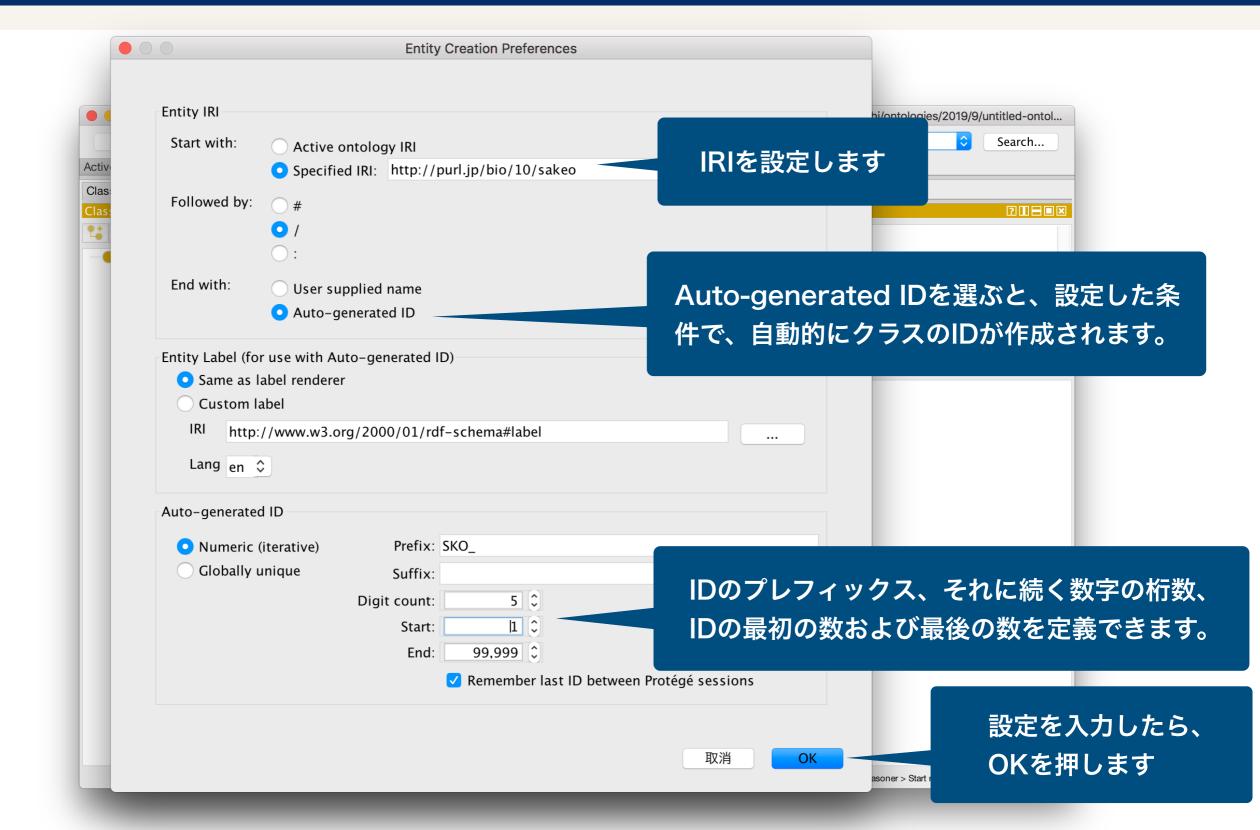
# Protégéを起動する

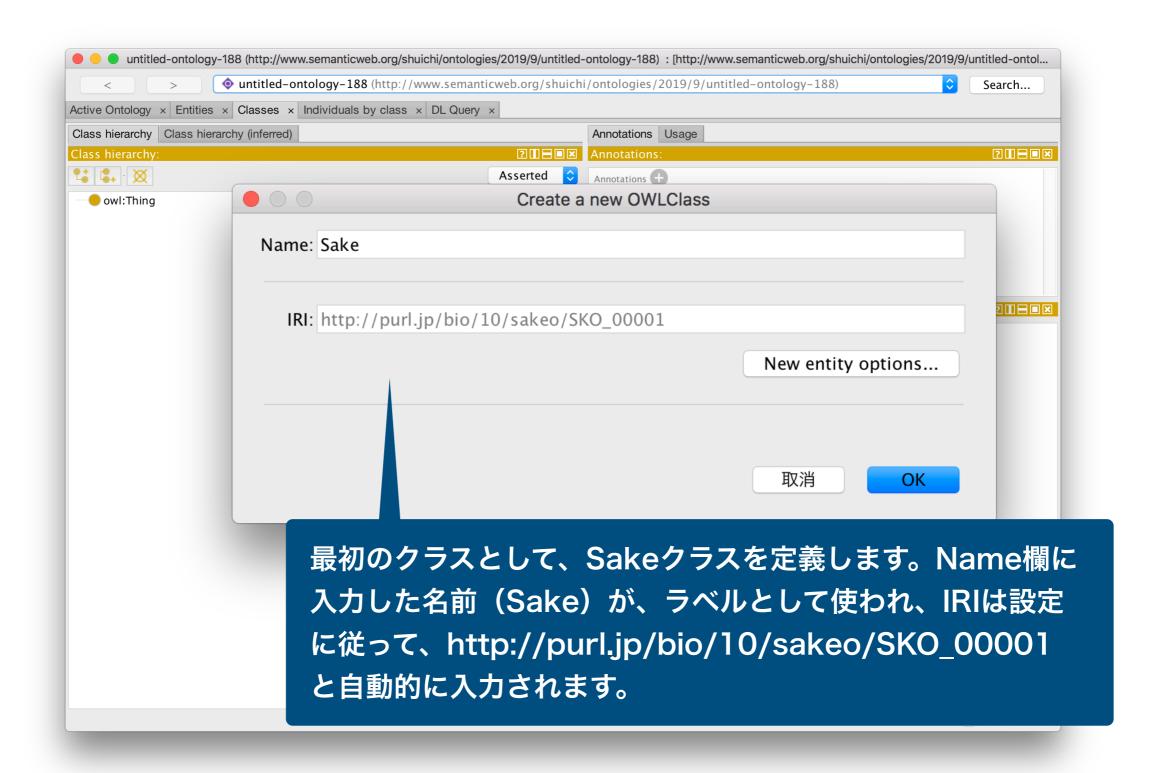


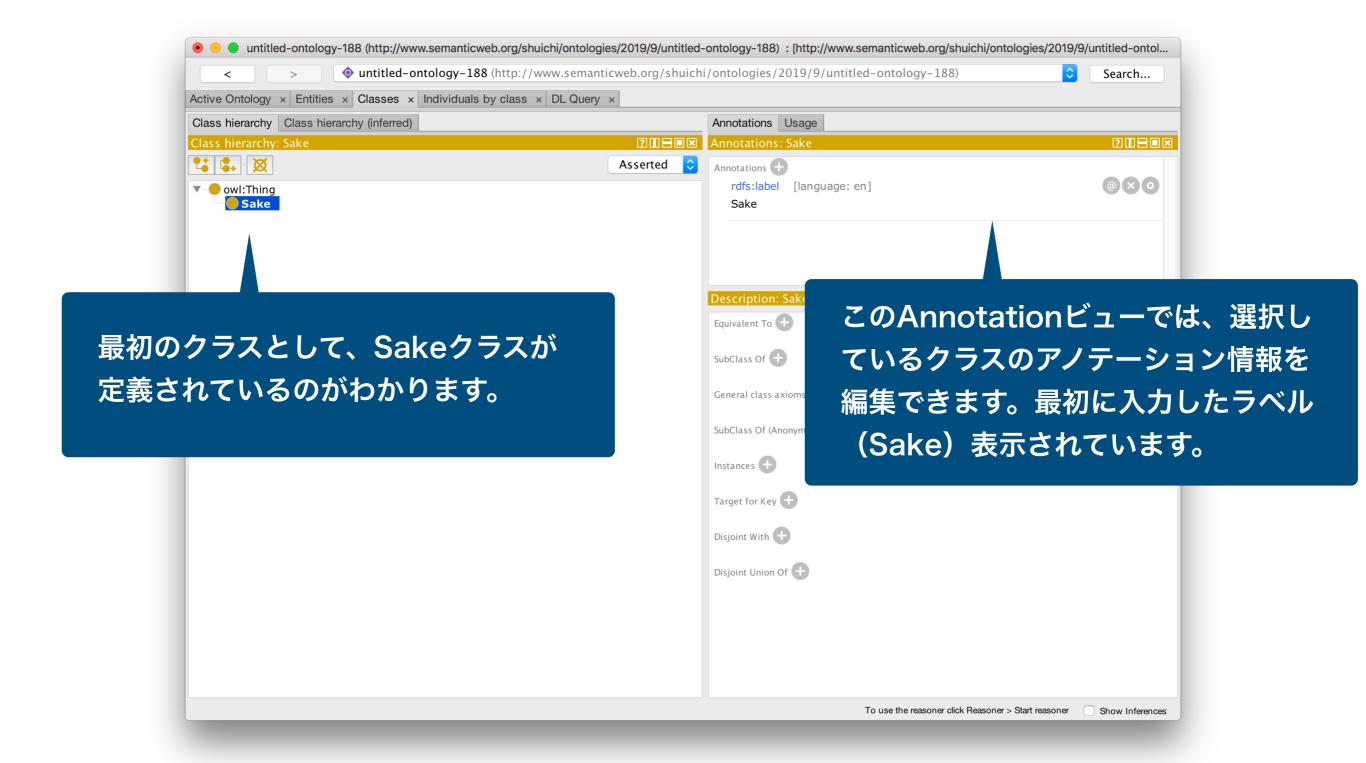


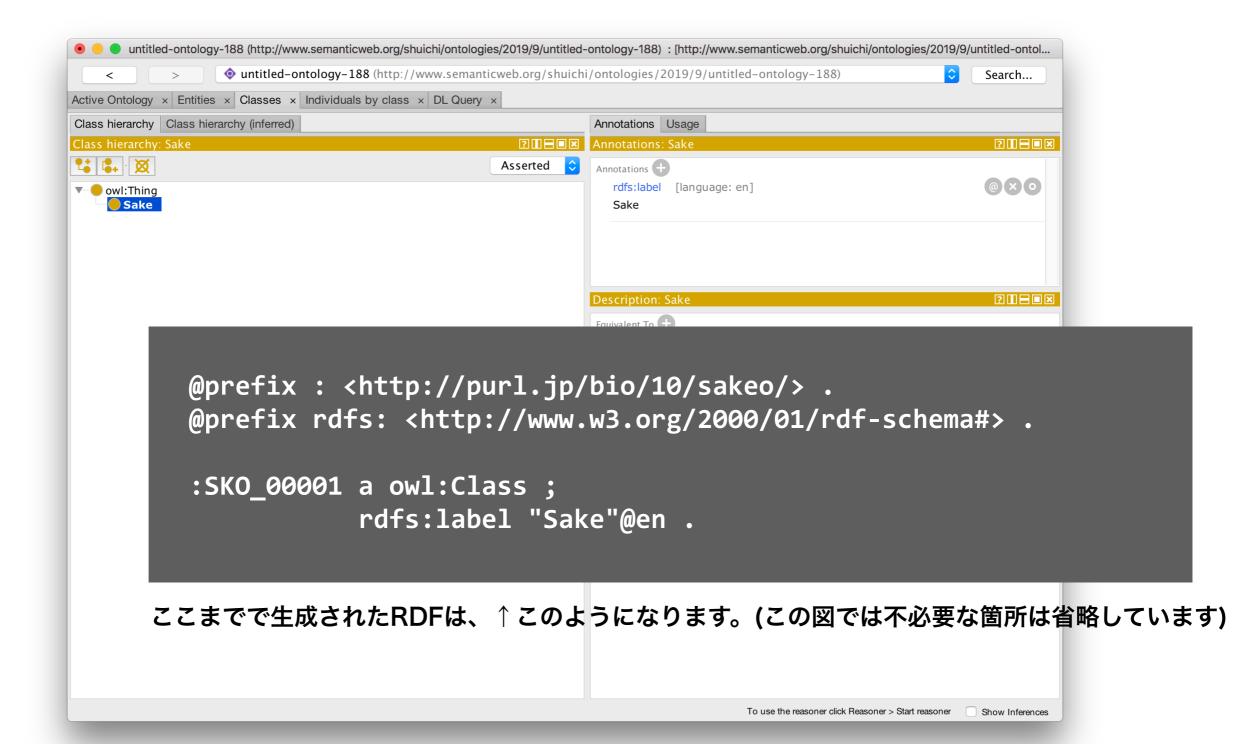


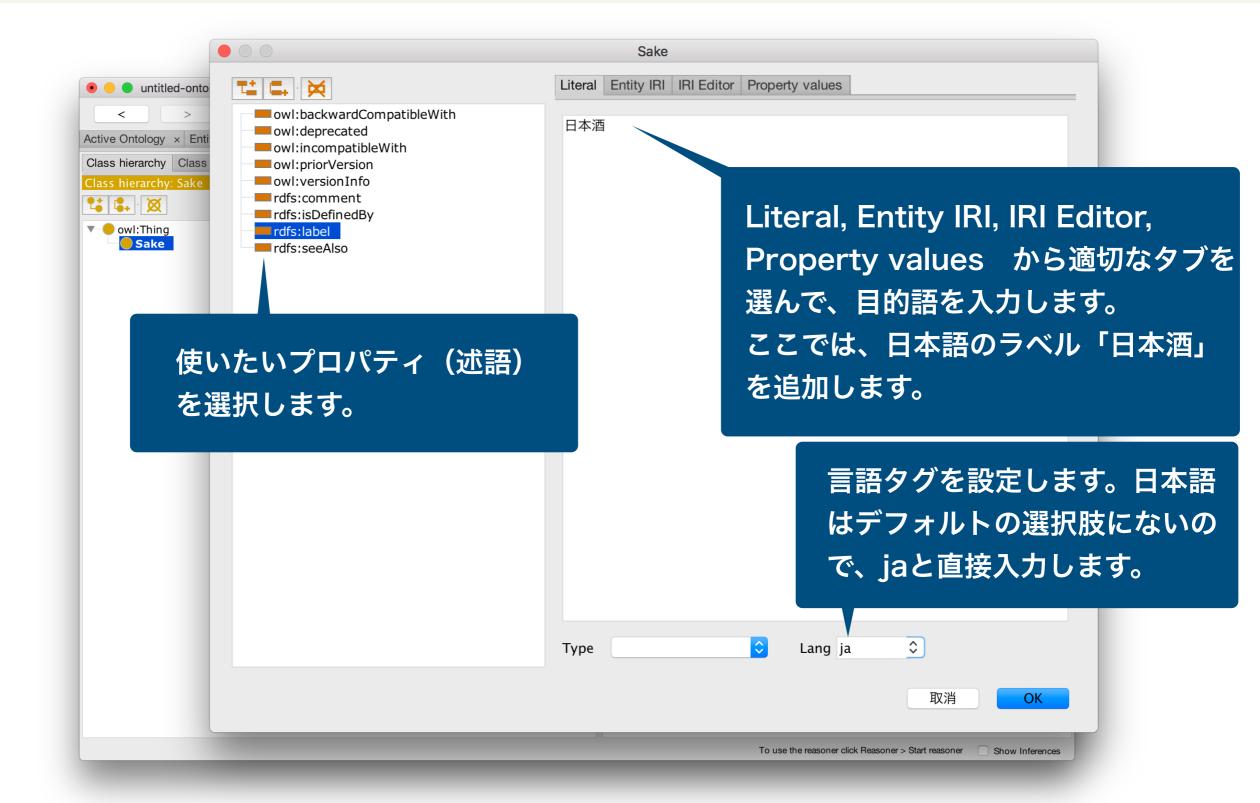


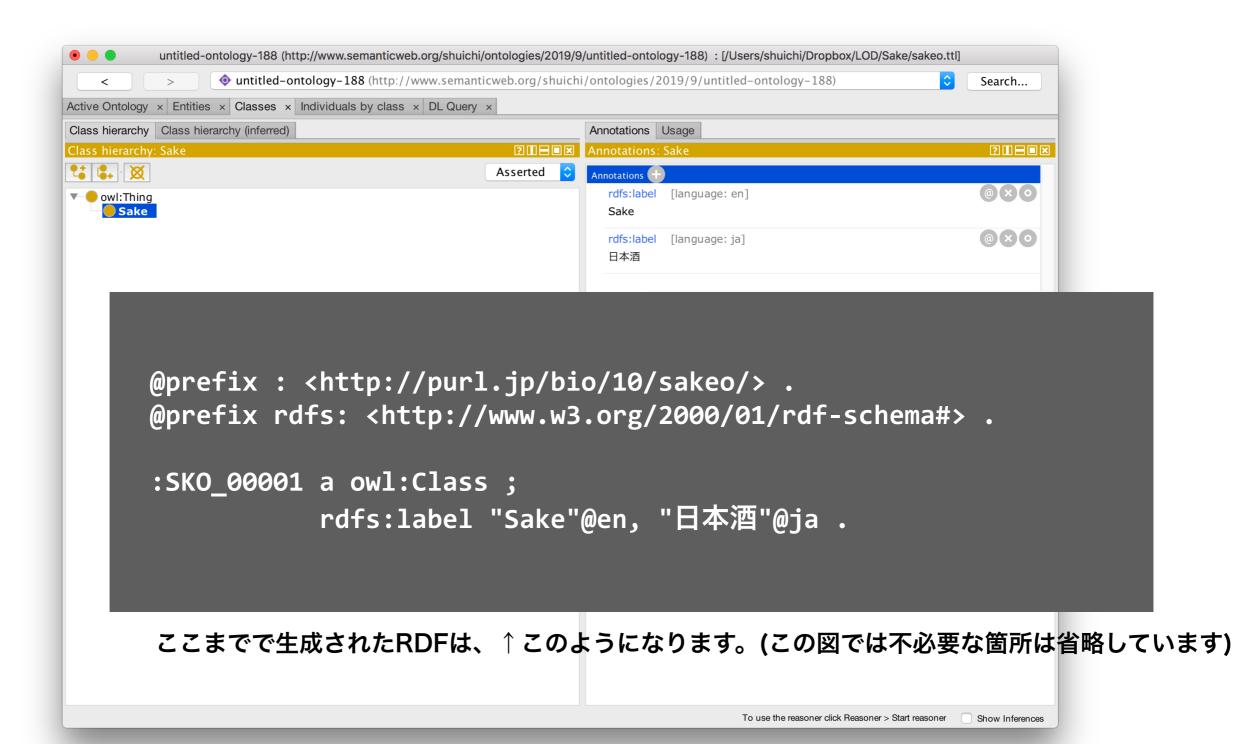


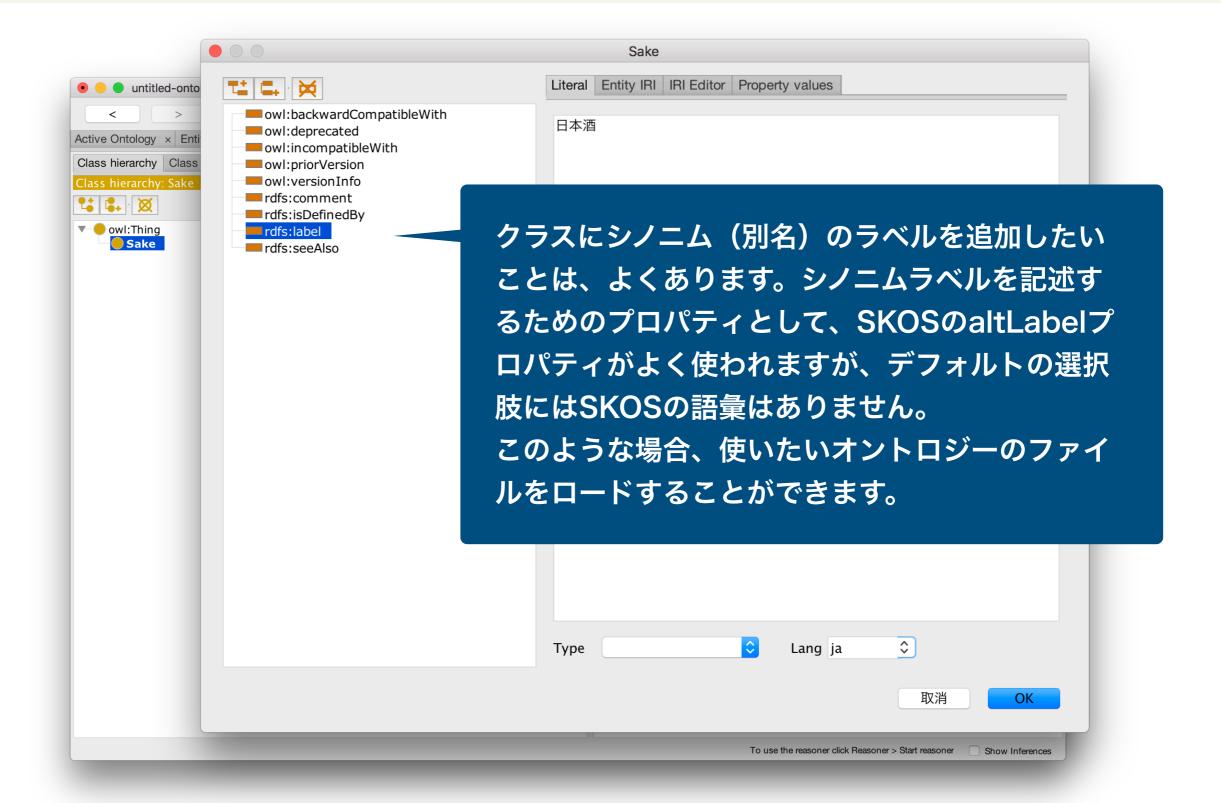


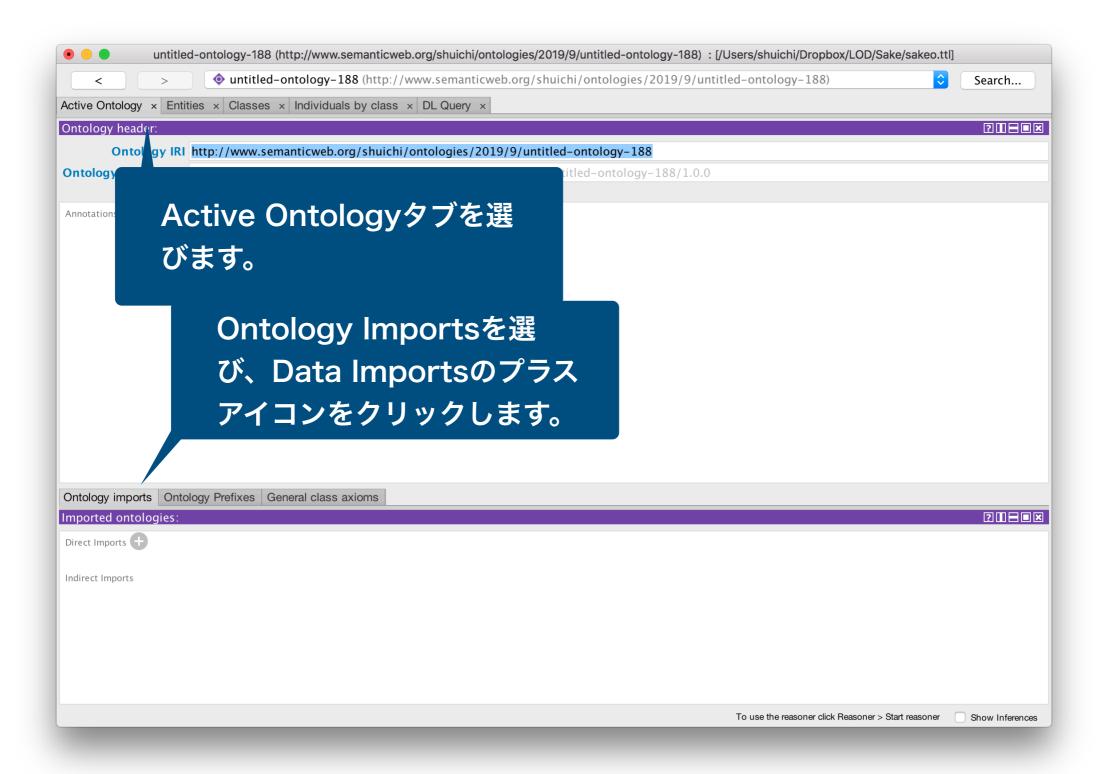


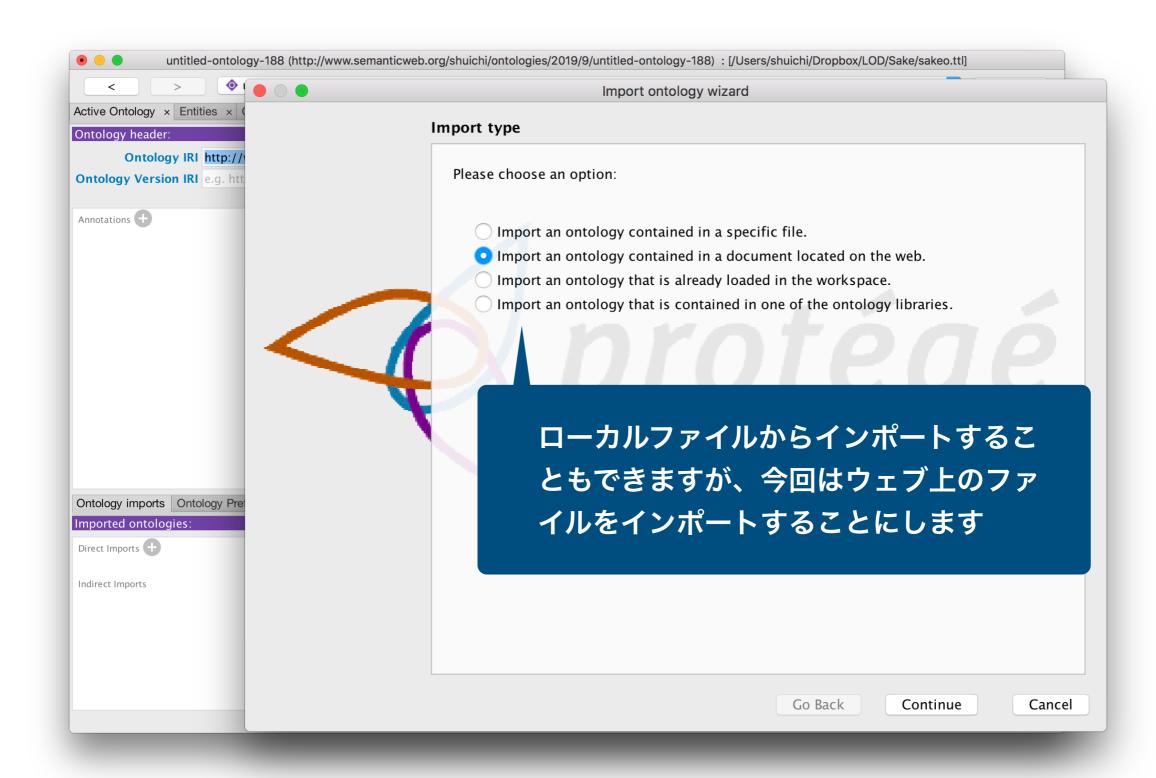


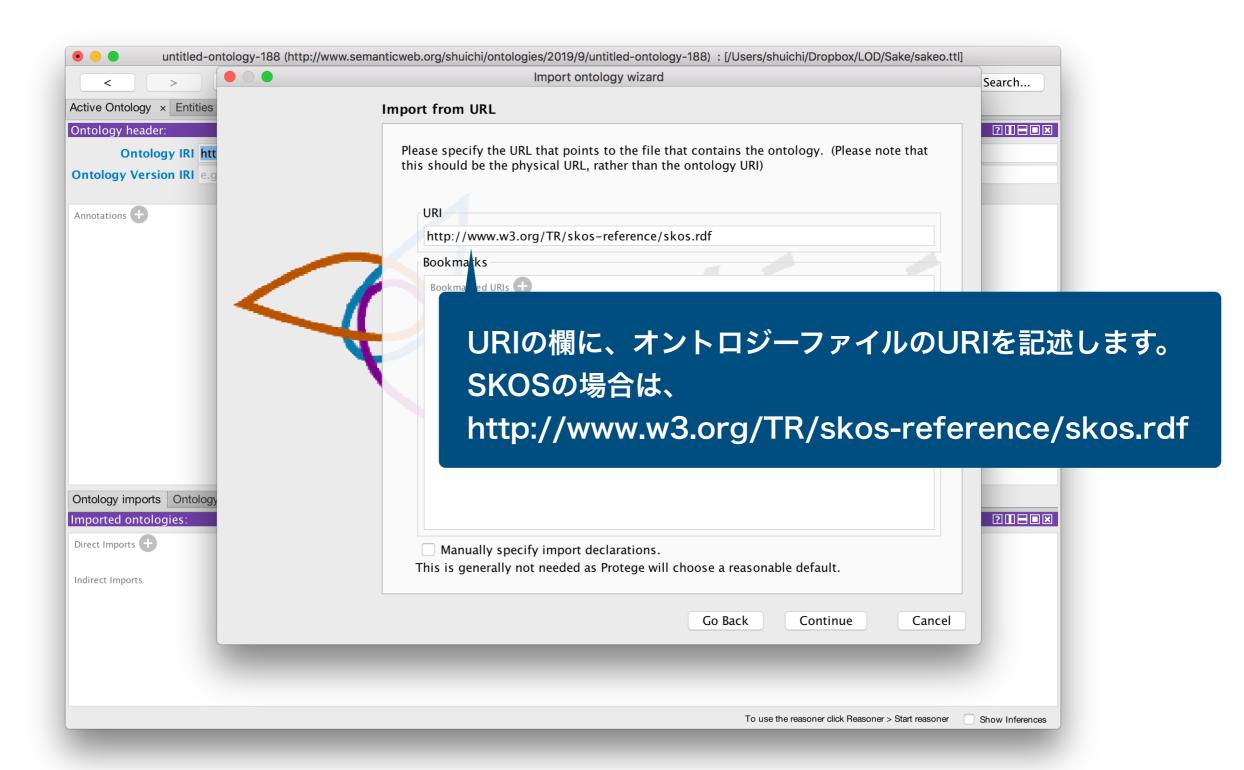


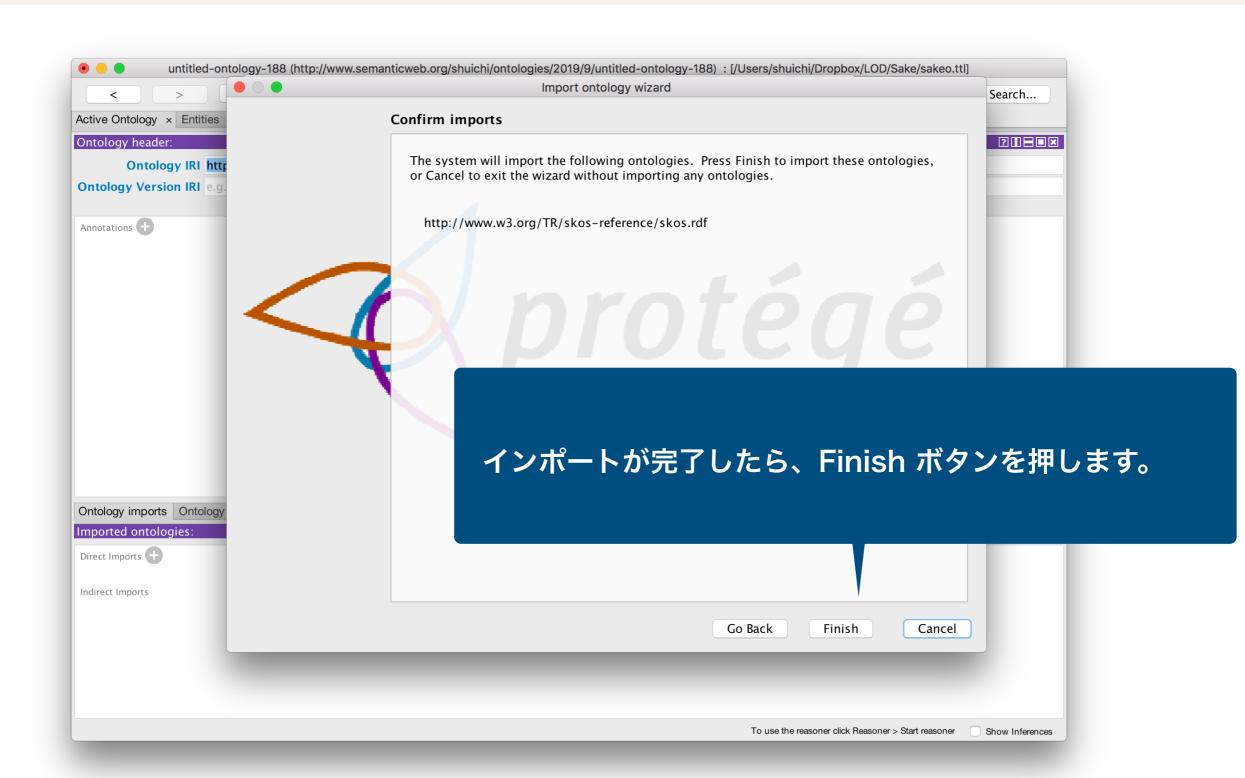


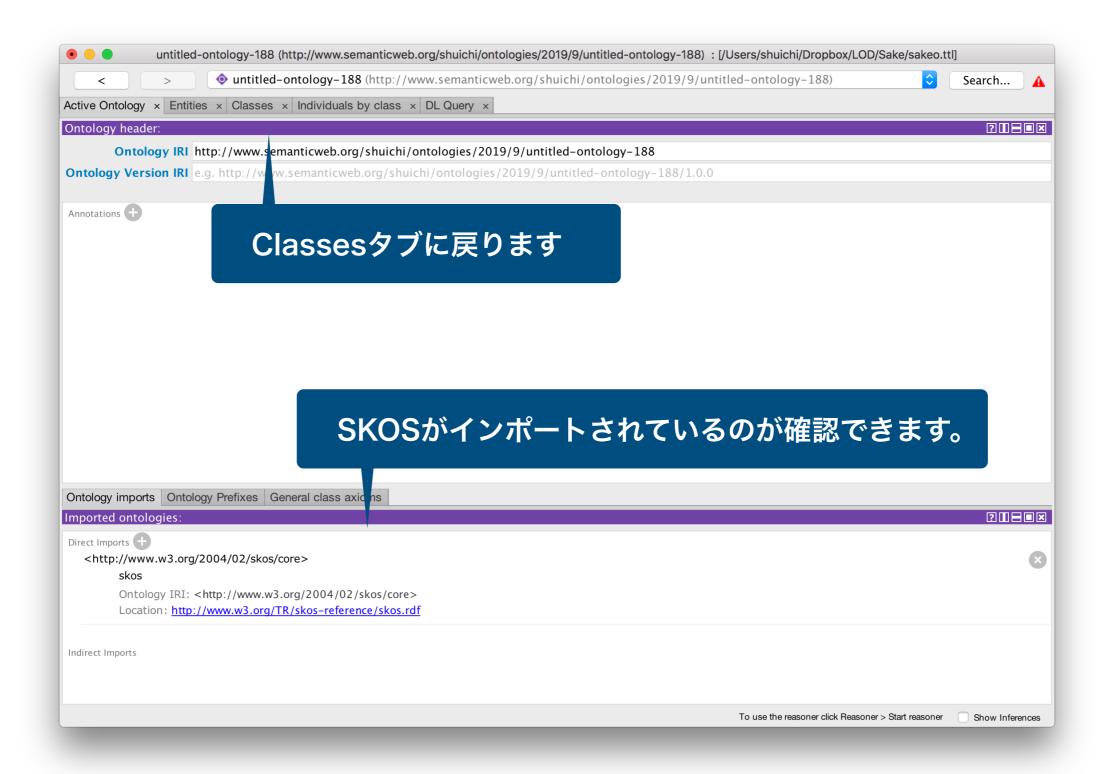


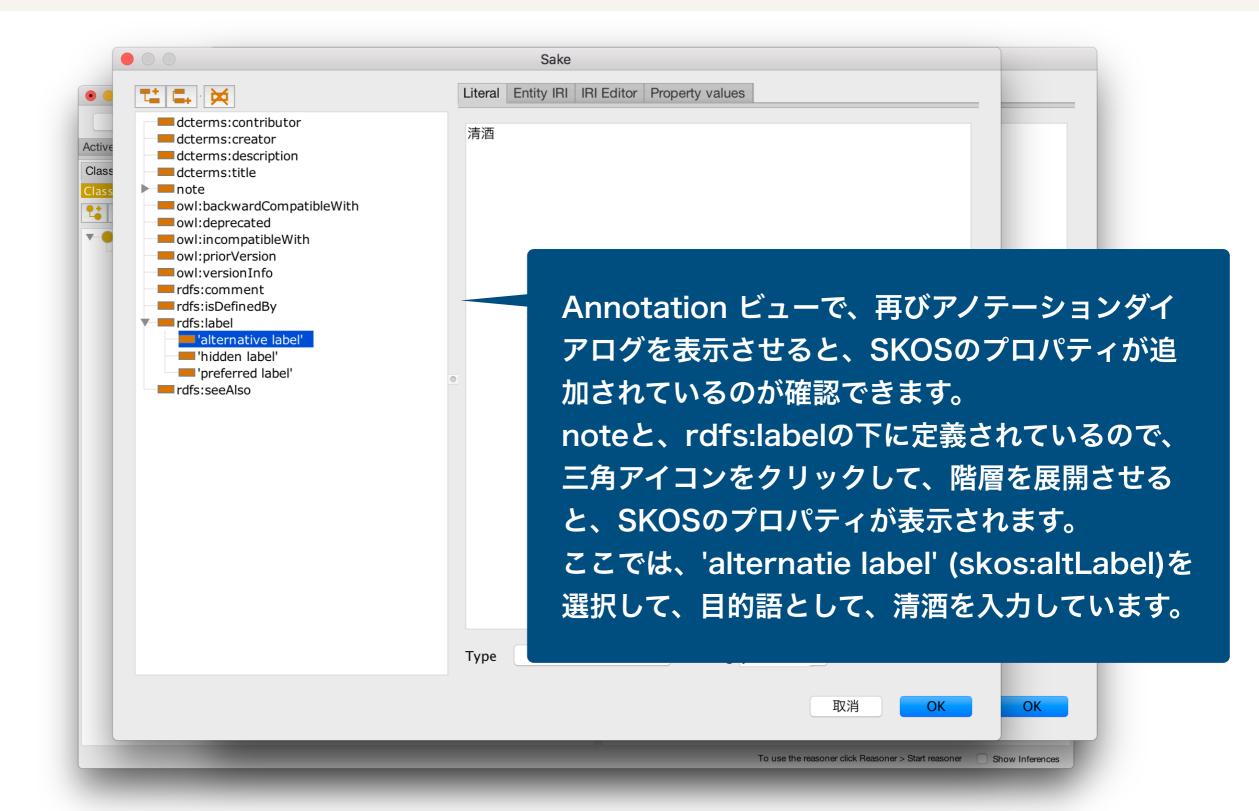


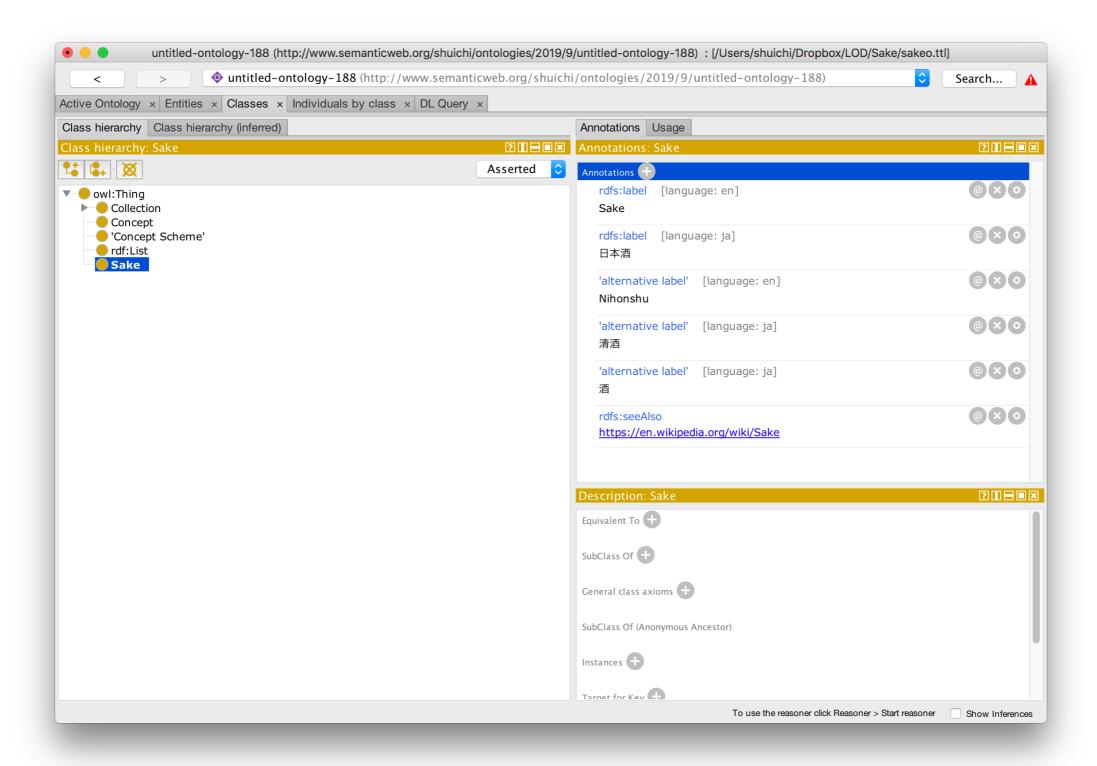


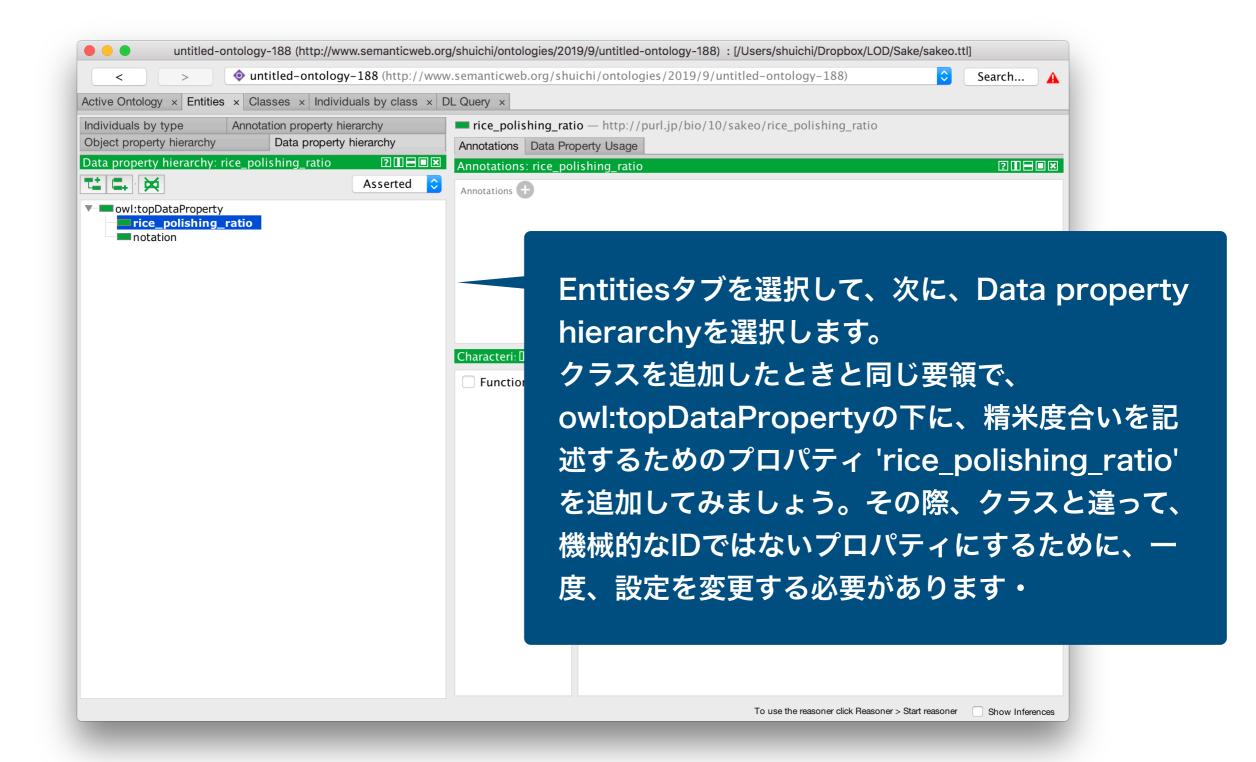


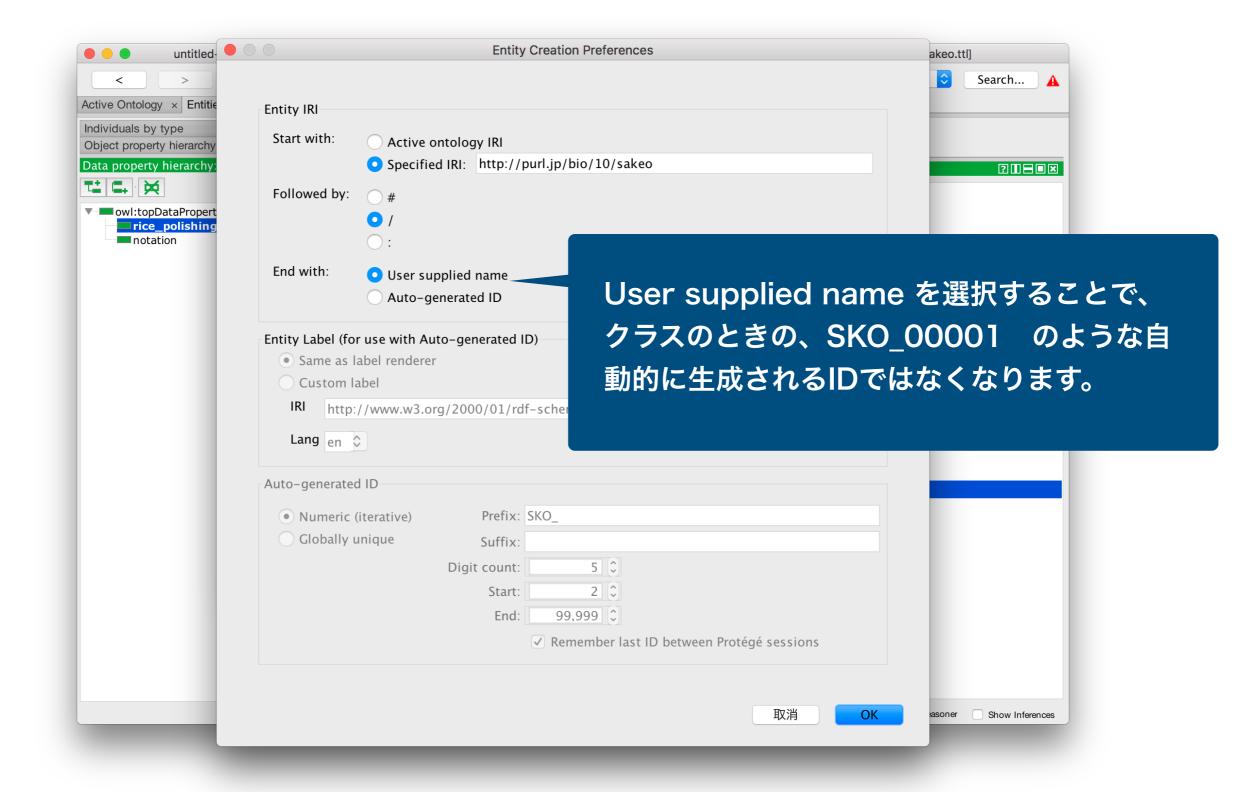


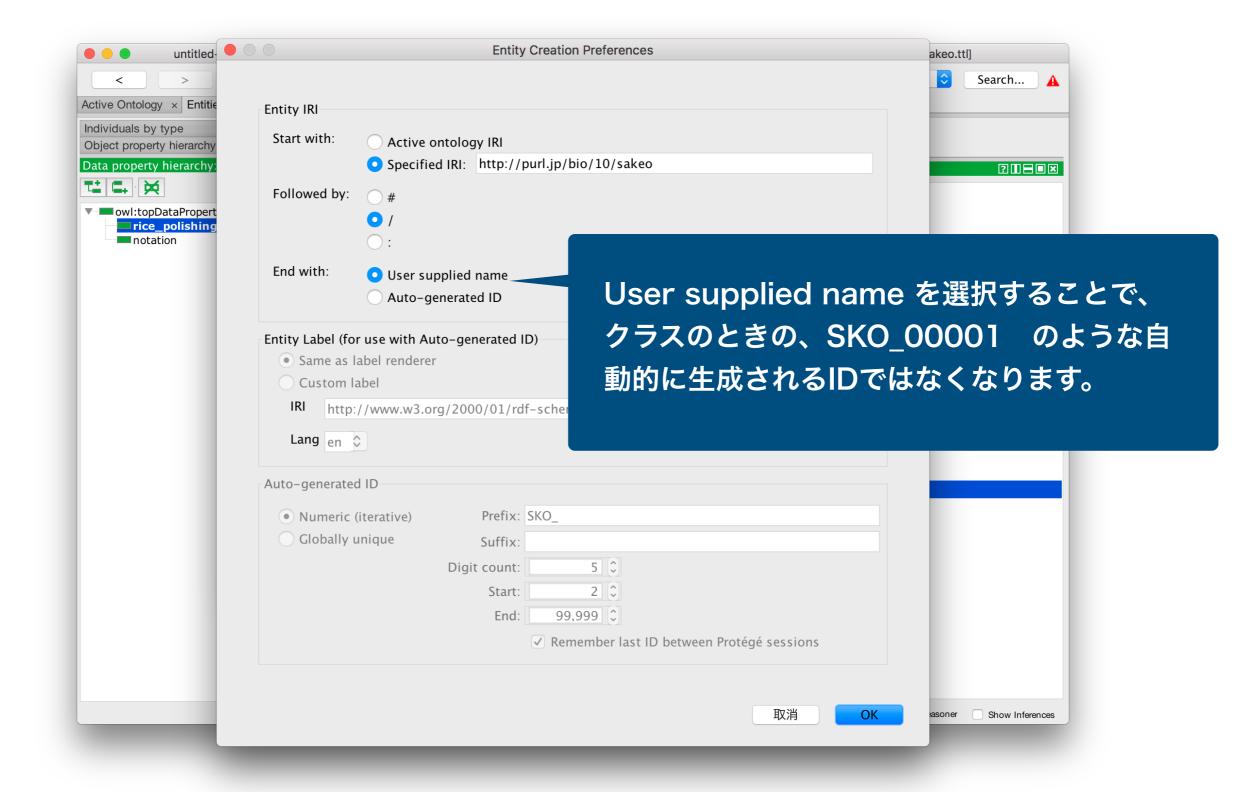


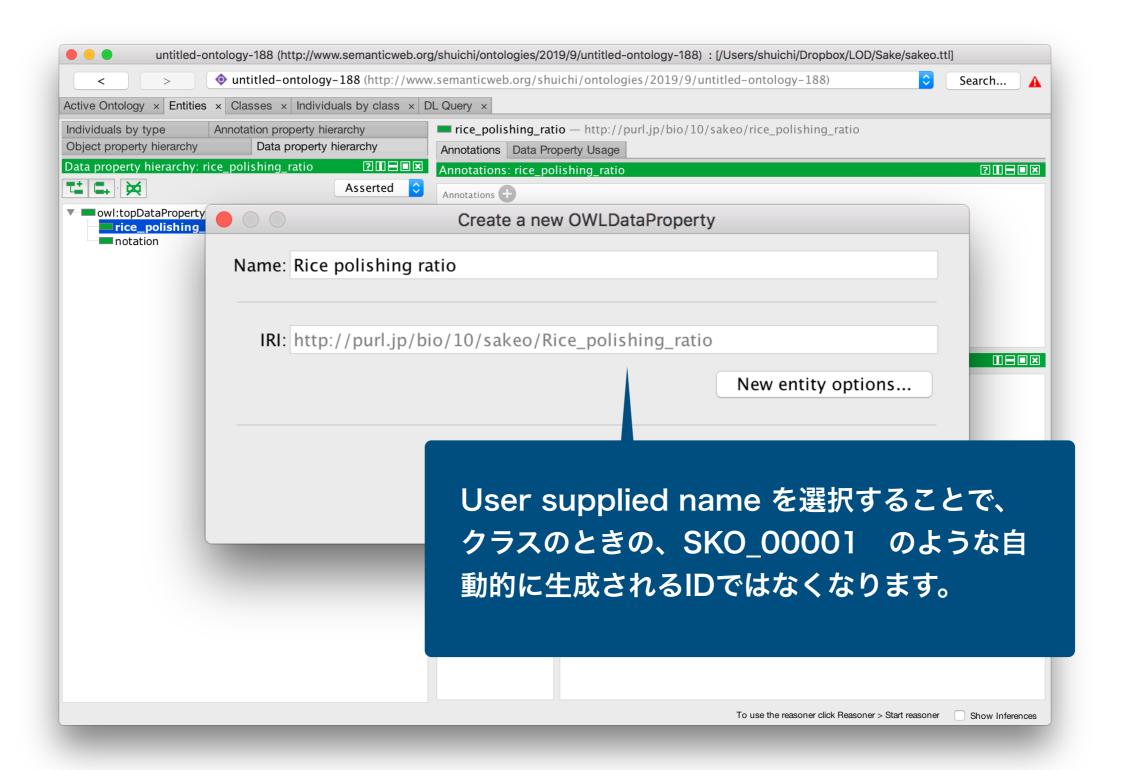




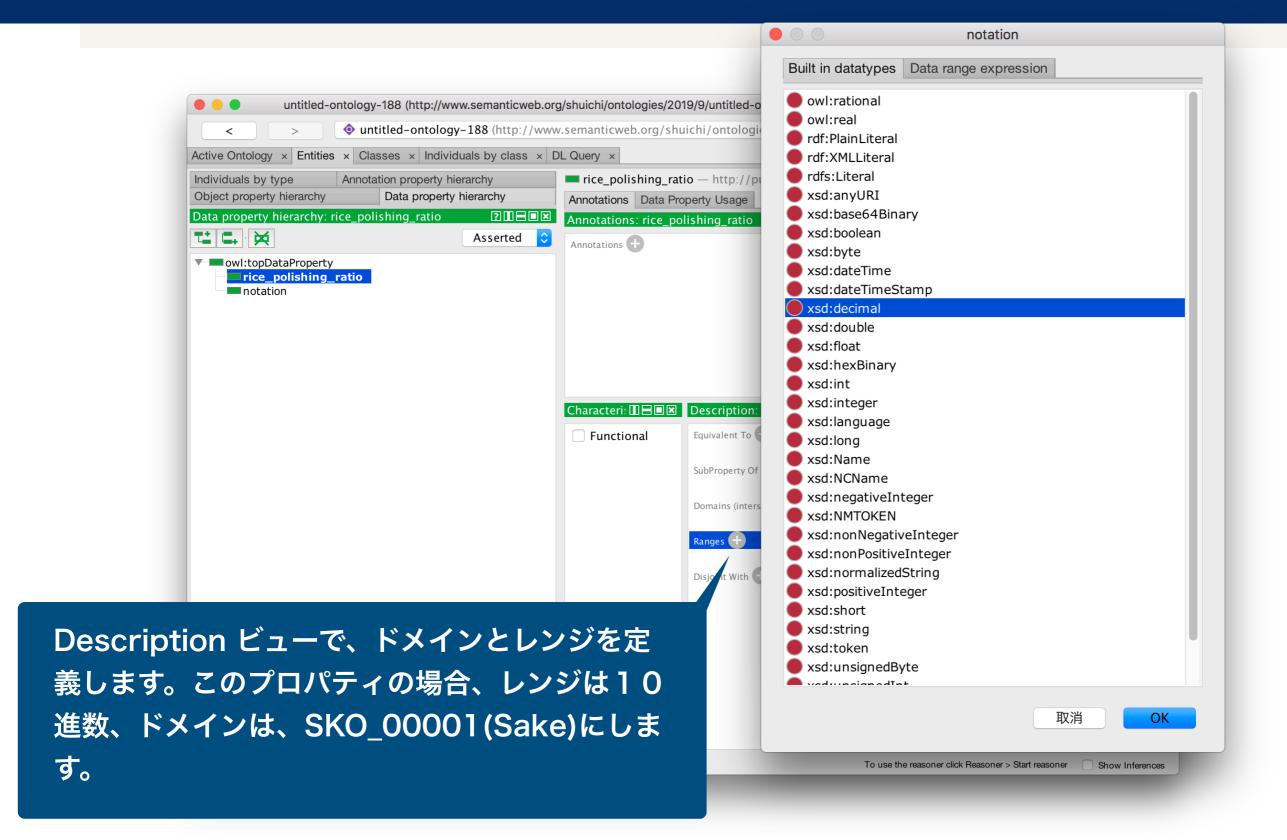






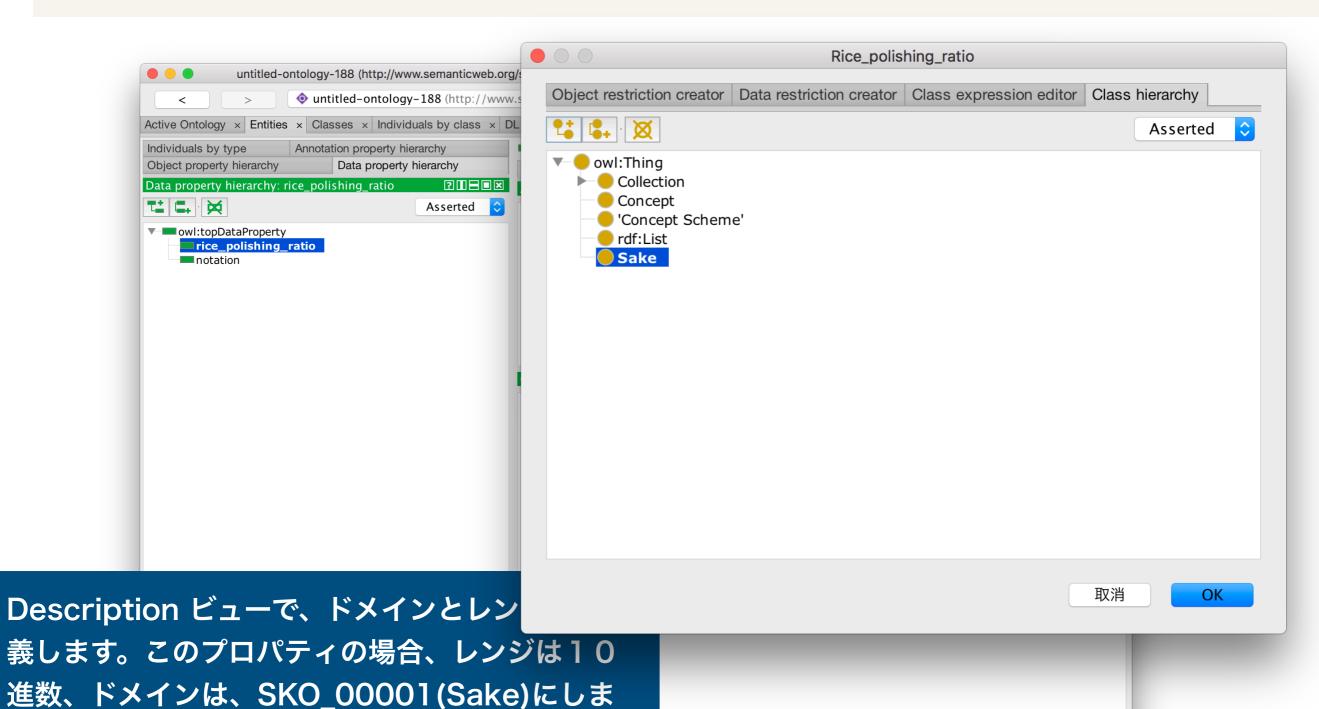


## ドメイン・レンジを定義する



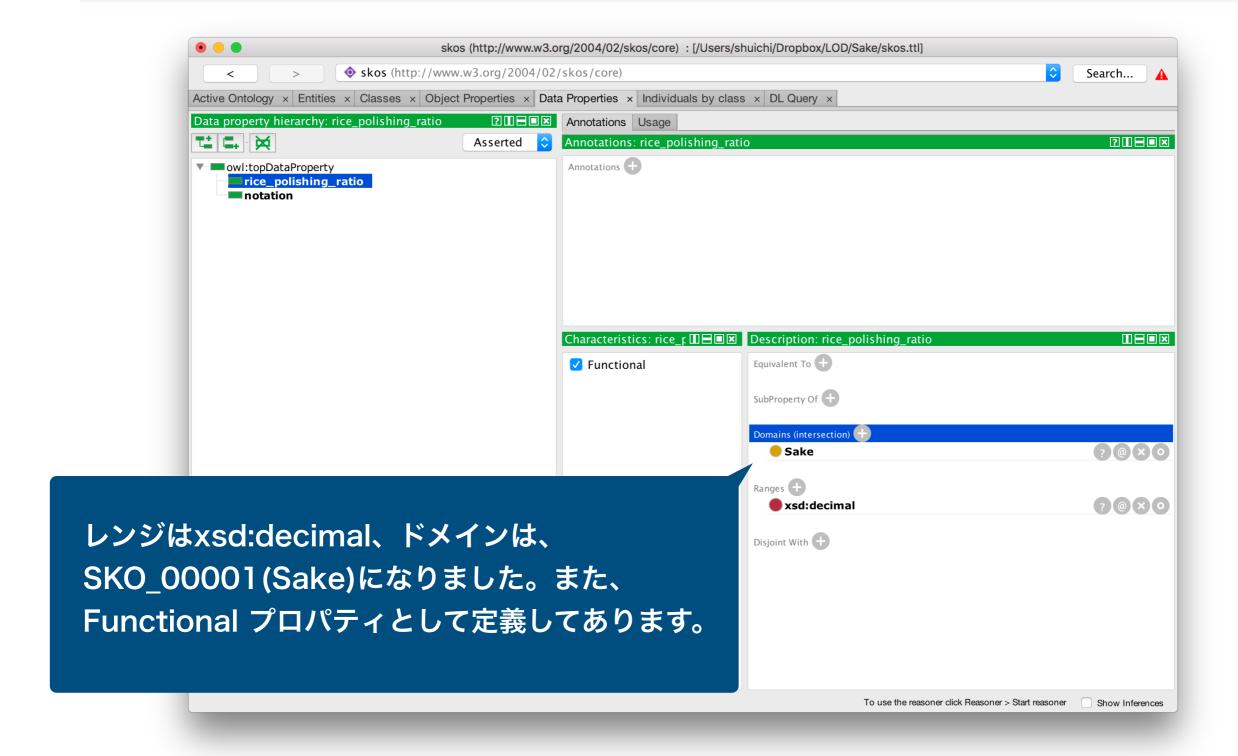
#### ドメイン・レンジを定義する

す。



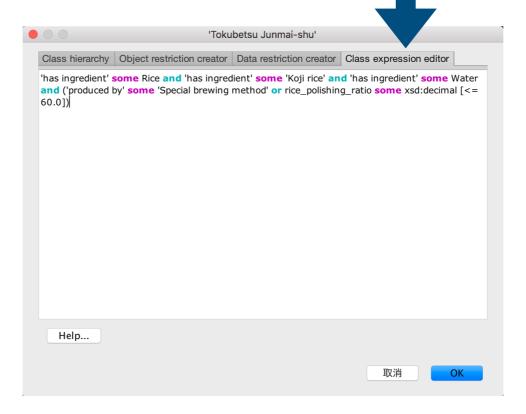
To use the reasoner click Reasoner > Start reasoner

# ドメイン・レンジを定義する



# Manchester OWL syntax

簡単なクラス表現はオントロジーエディタのGUIで記述できるが、複雑なプロパティ制約等になると、Mancherster OWL syntaxを用いて、マニュアルに記述する必要があります。



#### 参考

- http://ceur-ws.org/Vol-216/submission\_9.pdf
- https://protegewiki.stanford.edu/images/5/5f/Owled2008dc\_paper\_11.pdf

# Manchester OWL syntax

#### 例題の日本酒オントロジーで、特別純米酒のクラス表現

'has ingredient' some Rice and
'has ingredient' some 'Koji rice' and
'has ingredient' some Water and
('produced by' some 'Special brewing method' or rice\_polishing\_ratio some xsd:decimal [<= 60.0])

慣れないと難しい感じがしますが、これを直接 RDFで書くと右にのようになるので、これも大変

```
:SKO 00012
   a owl:Class;
   rdfs:label "Tokubetsu Junmai-shu"@en, "特別純米酒"@ja;
   rdfs:subClassOf :SKO 00002 ;
   owl:equivalentClass [
       a owl:Class ;
       owl:intersectionOf ([
               a owl:Class;
               owl:unionOf ([
                       a owl:Restriction;
                       owl:onProperty :produced by ;
                       owl:someValuesFrom :SKO_00020
                       a owl:Restriction;
                       owl:onProperty :rice polishing ratio ;
                       owl:someValuesFrom [
                           a rdfs:Datatype;
                           owl:onDatatype xsd:decimal ;
                            owl:withRestrictions ([
                                   xsd:maxInclusive 60.0
               a owl:Restriction;
               owl:onProperty :has ingredient ;
               owl:someValuesFrom :SKO 00006
               owl:onProperty :has_ingredient ;
               owl:someValuesFrom :SKO 00008
               a owl:Restriction;
               owl:onProperty :has ingredient ;
               owl:someValuesFrom :SKO 00009
   1.
```