

第2回Linked Open Data活用ワークショップ in 神戸

2017年9月22日 (金)@理化学研究所融合連携イノベーション推進棟



オープンデータのLOD化

大阪大学産業科学研究所

古崎 晃司

Kozaki@ei.sanken.osaka-u.ac.jp

■ 目標

- CSV形式で公開されている「オープンデータ」をLOD化する基本的な方法を身につける

■ 実施内容

1. CSVをRDFに変換

- できるだけ、他のLODと「共通の語彙」を用いる
- 他LODとのリンクをつける

2. RDFファイルを公開

- SPARQLエンドポイント(API)の利用
- DBpediaと同じような形(参照解決可能なURI)での公開

既存データをRDF化する方法



■ RDFデータを作成するツールを使用

- **Open Refine** (<http://openrefine.org/>) + **RDF Refine** (<http://refine.deri.ie/>)

- データを整備・公開するための多様な機能をサポート

- **LinkData** (<http://linkdata.org/>)

LODチャレンジ2011
アプリケーション部門最優秀賞

- テーブルデータをRDFに変換して公開することができるサイト

- **SparqlEPCU** (<http://lodcu.cs.chubu.ac.jp/SparqlEPCU/>)

- LODの作成・活用のための支援サイト
- CSVファイルをRDFとしての公開する機能もあり

LODチャレンジ2012
アプリケーション部門最優秀賞

- **StatLD** (<http://satolab.tiu.ac.jp/statld/>)

- 統計Linked Dataの活用を中心としたツールを公開
- 一般のLinked Dataの作成にも使用可能

■ 独自のプログラムでRDFデータを作成

本日，利用するツール



■ CSV2LOD～RDF変換支援ツール～

- <http://lodosaka.jp/tool/CSV2LOD/>
- **CSV形式のデータをRDFに変換する**
- プロトタイプなので，ご意見大歓迎！

■ なんでもリンク

- <http://link.lodosaka.jp/>
- **オープンデータに外部LOD (DBpedia) へのリンクを追加する**

■ Simple LODI

- <https://github.com/uedayou/simplelodi>
- **DBpediaのようにLinked Open Data(LOD)を公開する**
＝参照解決可能＋コンテンツネゴシエーションに対応した
LODを公開する.

使用するツール①

CSV2LOD～RDF変換支援ツール～



CSV2LOD～RDF変換支援ツール～

CSV選択

CSVファイル(元データ)を選択してください。

☐ ファイルから選択 選択されていません

☐ URLを指定(※準備中)

☐ テキストボックスに入力

※サンプル(リンク先を保存してお使い下さい):
[大阪市の「施設情報ポイントデータ\(官公庁\)」](#)(CC-BY:大阪市), [神戸市の「市役所・区役所」](#)(CC-BY:神戸市)

テンプレート作成

メタデータの入力

ベースIRI(※)を入力してください。

※変換したRDFファイルを公開するURLに相当します。

☐ 一つのファイルで公開(#)

☐ データベース(トリプルストア)を用いて公開(/)

ライセンス情報を入力してください。

クレジットで表示すべき名前

※オープンデータとして公開するには「CC-BY」または「パブリックドメイン」が望まれます。

☐ クリエイティブ・コモンズ(CC)を利用

☐ その他のライセンスを利用

<http://lodosaka.jp/tool/CSV2LOD/>

- **Linked Dataの基本原則に沿ったLODを, 誰でも, (ある程度)簡単に作成できるツールを提供したい**
- **元データ(CSV)の編集を, 極力, 減らしたい**
 - データが更新されたら, すぐに再変換したい
 - →元データと変換情報のテンプレートは別ファイルに
- **足りない機能があれば, 気軽に拡張したい**
 - 小規模なプログラムで開発+オープンソース化により改変を自由に
- **初心者から上級者まで使いやすいものを**
 - 汎用性の高いテンプレート(Turtleライクな形式)により, 複雑なRDFモデルにも対応
 - 初心者向けには, 語彙選択支援など, 簡単にテンプレートが作れる
- **動作環境**
 - インストール不要, かつ, ネットがなくても使える→JavaScriptのみで開発
 - IE8にも, できれば対応したかったが...(現時点は×)

使用するツール② なんでもリンク



The screenshot shows a web browser window with the address bar displaying 'link.lodosaka.jp'. The page title is 'なんでもリンク'. The main content area contains the following text:

文字列をDBpediaのURIに割り当てます

与えられた文字列に対応するURIを対象エンドポイントに問い合わせして割り当てます。結果はCSV形式でダウンロードできます。一度実行するとリロードするまで再実行は出来ません。本アプリケーションはGitHubで公開しています。

SPARQLエンドポイント:

以下のいずれかの方法で割り当てます。

1. 割り当てたい文字列(一行一文字列)が書かれたファイルを以下にドロップしてください。自動で処理が始まります。



2. 割り当てたい文字列(一行一文字列)を入力してください。入力後、「割当実行」をクリックしてください。

Below the second instruction is a large empty text input field.

<http://link.lodosaka.jp/>

使用するツール③

Simple LODI



<https://github.com/uedayou/simplelodi>

- RDFデータベースなしで、「参照解決可能なLOD」の公開が可能
- 一般的なWebサイト用のレンタルサーバのスペック(PHP 5.4以上, mod_rewrite)で動作

Simple LODI : Simple Linked Open Data Interface

コンテンツネゴシエーションや拡張子変更により、RDFファイルをさまざまなフォーマットで出力ができるLODフロントエンドプログラムです。対応する入力フォーマットは、RDF/XML(拡張子:xml)、N-Triples(拡張子:nt)、Turtle(拡張子:ttl)です。

対応出力フォーマット(拡張子)

- HTML (html)

■ 「CSV形式」のファイルを利用

- 最初の1行目に「データの項目」
2行目以降に各データが記入されているもの
- 不要な行など、が入っているものは使えません
- [CSV2LODの既知のバグ]
"1列目に空白セルが含まれる"と、正しく変換できません.

■ ハンズオンに利用するオープンデータ

- 神戸市オープンデータ: 観光施設情報

<https://data.city.kobe.lg.jp/data/dataset/dataset-00000005/resource/369be024-e7c2-46be-98ad-a6b2519f12e3>

利用可能なデータ例



tourism_od2810.csv - Excel														サインイン	
ファイル ホーム 挿入 ページレイアウト 数式 データ 校閲 表示 ACROBAT															
A1 : fx 名称															
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	
1	名称	名称のルビ	郵便番号	カテゴリ	住所	緯度	経度	電話番号	説明文	営業時間	定休日	URL	バリアフリー対応		
2	香りの家オランダ館(旧ヴォルビ		'6500002	'2401	神戸市中	34.7022	135.1907	078-261-8	北野のオ	9:00~18:0	なし	http://ww	なし		
3	デンマーク館(旧ヨハン・フラウベ		'6500002	'2401	神戸市中	34.70235	135.1905	078-261-8	バイキング	9:00~18:0	なし	http://ww	なし		
4	ウィーン・オーストリアの家(旧W		'6500002	'2401	神戸市中	34.70241	135.1908	078-261-8	平成4年、	9:00~18:0	なし	http://ww	なし		
5	竹中大工道具たけなかにだいくど		'6510056	'2401	神戸市中	34.70686	135.1975	078-242-0	竹中大工	火~日 9月(祝日の		http://dougukan.jp			
6	風見鶏の館 かざみどりのやか		'6500002	'2401	神戸市中	34.7014	135.1895	078-242-8	風見鶏の	9:00~18:0	6月・2月の	http://www.kobe-kazamidori.co			
7	ラインの館 ※工事のため休館中		'6500002	'2401	神戸市中	34.70113	135.1912	078-222-8	大正4年(1	9:00~18:0	メンテナンスのため休館中です。H.284.1~				
8	うろこの家・うろこ美術館		'6500002	'2401	神戸市中	34.70296	135.1907	0120-888-	神戸で最	4月~9月	年中無休	http://kot	なし		
9	異人館 パラスティン邸		'6500002	'2401	神戸市中	34.70093	135.1906	078-241-1	ロシア人	11:00~18	火(祝日の	http://ww	なし		
10	萌黄の館 もえぎのやかた		'6520002	'2401	神戸市中	34.70098	135.1893	078-222-8	萌黄の館	(9:00~18:0	2月の第3	http://www.city.kobe.lg.jp/cultur			
11	山手八番館		'6500002	'2401	神戸市中	34.70279	135.1906	0120-888-	山手八番	4月~9月	年中無休	http://kot	なし		
12	北野外国人倶楽部		'6500002	'2401	神戸市中	34.70313	135.1915	0120-888-	北野外国	4月~9月	年中無休	http://kobe-i-jinkan.net			
13	坂の上の異人館(旧中国領事館		'6500002	'2401	神戸市中	34.70317	135.1916	0120-888-	1940年、	4月~9月	年中無休	http://kot	なし		
14	ベンの家		'6500002	'2401	神戸市中	34.70093	135.1915	0120-888-	英国の貴	4月~9月	年中無休	http://kot	なし		
15	洋館長屋(仏蘭西館)		'6500002	'2401	神戸市中	34.70086	135.1914	0120-888-	元・外国	4月~9月	年中無休	http://kot	なし		
16	英国館		'6500002	'2401	神戸市中	34.70078	135.1912	0120-888-	イギリス	4月~9月	年中無休	http://kobe-i-jinkan.net			
17	神戸トリックアート不思議な領事		'6500002	'2401	神戸市中	34.70092	135.1909	0120-888-	アミューズ	4月~9月	年中無休	http://kot	あり		
18	イタリア館(プラトン装飾美術館)		'6500002	'2401	神戸市中	34.70243	135.1921	078-271-8	イタリア	9:30~17:0	年中無休		なし		
19	神戸北野美術館		'6500002	'2401	神戸市中	34.70125	135.1915	078-251-8	神戸北野	9:30~17:0	不定休(要・問合せ)		なし		
20	北野工房のまち		'6500004	'2401	神戸市中	34.69556	135.1872	078-221-8	トアロード	9:30~18:0	不定休、日	http://kite	車イス、オストメイト、		
21	SPACE11 DARVISH MUSEUM		'6500003	'2401	神戸市中	34.69889	135.1912	078-291-8	大阪羽曳	10:00~18:0	年末年始	http://ww	有り		
22	花と緑のまち推進センター		'6500006	'2401	神戸市中	34.69461	135.1794	078-351-8	神戸市の	9:00~17:0	月曜日(市	http://ww	ユニバーサルトイレ有		
23	神戸ドルミュージアム		'6500021	'2402	神戸市中	34.68938	135.1906	078-327-8	神戸ドル	10:00~19	水(祝日の	http://www.timeroman.com/			
24	神戸市役所1号館24階展望ロビ		'6508570	'2402	神戸市中	34.68949	135.1957	078-331-8	フラワー	口平日 8:15	12月29日~1月3日は休み、設備点検日年				
25	神戸市危機管理センター防災展		'6500033	'2402	神戸市中	34.68976	135.1945					http://www.city.kobe.lg.jp/safet			
26	神戸市立相楽こうべしつそうら		'6500004	'2403	神戸市中	34.69239	135.1827	078-351-8	相楽園は、	金~水 9月	木(祝祭日	http://www.sorakuen.com/			
27	KOBEとんぼこうべとんぼだま		'6500034	'2404	神戸市中	34.68858	135.1937	078-393-8	KOBEとん	10:00~19	なし	http://www.lampwork-museum.c			
28	神戸市立博物館こうべしつそうら		'6500034	'2404	神戸市中	34.68727	135.1929	078-391-8	昭和10年	(9:30~17:0	日(祝祭日	http://www	対応、ア、イ、オ、		

■ オープンデータのLOD化①

- 1. 「CSV2LOD」を使いRDFに変換する
- 2. LODの公開
 - ①RDFファイル(Turtle形式)をWebサイトに公開して閲覧
 - ②データをRDFデータベース(Fuseki)に格納する

■ オープンデータのLOD化②【外部リンク付き】

- 3. 「なんでもリンク」を使いWikidataへのリンク情報をCSVファイルに追加する
- 4. リンク情報付きのRDFをRDFデータベース(Fuseki)に格納して, Federateクエリを試す

■ Simple LODIを用いたLODの公開

- 5. 「Simple LODI」を使い公開

0. CSVファイルの準備

1. CSVファイルを準備する

- 練習では、神戸市オープンデータ:観光施設情報にWikidataへのリンクを追加したもの
https://github.com/KnowledgeGraphJapan/LODws2nd/blob/master/tourism_od2810%2Bwd.csv を利用

2. Google Sheets (またはExcel等) でCSVファイルを開く

- Google Sheetsでは「File」→「Open」→「Upload」
- エクセル等を利用してもよいが「UTF-8」のファイルを開く際には、注意が必要

3. 形式に問題がないかを確認する

1. 1行目が項目, 2行目以降がデータ(値)になっているか?
2. 1行目の項目と2行目以降にずれがないか?
3. 1列目に空白セルが含まれていないか?
4. 不要な行がないか?

1. CSV2LODの作業手順

■ 作業の手順

1. 元データ(**CSV形式**)を選択して読み込む
2. 変換に必要な情報を入力
 - URIに関する基本情報
 - 利用するプロパティ(関係の名称)
3. **RDF変換用のテンプレート**を生成
4. **CSVをRDFに変換**(今回は**Turtle形式のみ**)
5. **RDFファイルをダウンロード**

CSVファイルの読み込み

1. 「ファイルを選択」ボタンで、あらかじめ用意しておいたCSVファイルを選択する
2. 「CSVファイル読み込み・プレビュー」ボタンを押し、正しく読み込めることを確認する

CSV2LOD～RDF変換支援ツール～

CSV選択

CSVファイル(元データ)を選択してください。

☒ ファイルから選択 **ファイルを選択** 選択されていません

☐ URLを指定(※準備中)

☐ テキストボックスに入力

(1) CSVファイル読み・プレビュー

※サンプル(リンク先を保存してお使い下さい):

[大阪市の「施設情報ポイントデータ\(官公庁\)」](#)(CC-BY:大阪市), [神戸市の「市役所・区役所」](#)(CC-BY:神戸市)

CSVのプレビュー結果（例）

CSV2LOD～RDF変換支援ツール～

CSV選択

CSVファイル(元データ)を選択してください。

- ☒ ファイルから選択
- ☐ URLを指定(※準備中)
- ☐ テキストボックスに入力

※サンプル(リンク先を保存してお使い下さい):
[大阪市の「施設情報ポイントデータ\(官公庁\)」](#)(CC-BY:大阪市), [神戸市の「市役所・区役所」](#)(CC-BY:神戸市)

施設ID	施設名	郵便番号	住所	方書	電話番号	FAX番号	緯度	経度	行政区コード	カテゴリリスト
1	神戸市役所	650-8570	中央区加納町6-5-1		078-331-8181(代)		34.689454	135.195739	中央区	市役所・区役所
2	東灘区役所	658-8570	東灘区住吉東町5-2-1		078-841-4131(代)	078-811-4901	34.720263	135.265452	東灘区	市役所・区役所
3	灘区役所	657-8570	灘区桜口町4-2-1		078-843-7001(代)	078-843-7034	34.712348	135.239685	灘区	市役所・区役所
4	中央区役所	651-8570	中央区雲井通5-1-1		078-232-4411(代)	078-242-3599	34.695102	135.197783	中央区	市役所・区役所

- 文字コードは自動判定されます
 - サンプルデータについては動作確認済み
- ファイルサイズが大きい(200KB以上?)と読み込めない場合があります
 - テキストエディタ等で開いて、「テキストボックスに入力」の欄にコピー&ペースとして読み込むと、ある程度のサイズのデータでも読み込めます。
- 「ファイルを選択」がうまく動かないとき
 - 「テキストボックスに入力」を利用して下さい

テンプレート作成① メタデータの入力

■ 変換するデータに関する基本情報を入力する

The screenshot shows a web browser window with the URL `lodosaka.jp/tool/CSV2LOD/`. The page title is "テンプレート作成" (Template Creation). The main heading is "メタデータの入力" (Input of Metadata). Below this, it says "ベースIRI (※) を入力してください。" (Please input the Base IRI (※)).

A red callout box points to the "ベースIRI" section, containing the text: "ベースIRI : 作成するRDFデータでIDに用いるIRI (URI) の設定" (Base IRI: Setting of IRI (URI) used for ID in the created RDF data).

A red box highlights the "公開" (Publication) options, which include a text input field for the URL and three radio buttons: "一つのファイルで公開 (#)" (Publish with one file (#)), "データベース (トリプルストア) を用いて公開 (/)" (Publish using database (triple store) (/)), and "SimpleLODIを用いて公開 (/)" (Publish using SimpleLODI (/)).

A blue callout box points to the "ライセンス情報" (License Information) section, which includes a text input field for the name, a note about open data, and radio buttons for "クリエイティブ・コモンズ (CC) を利用" (Use Creative Commons (CC)) and "その他のライセンスを利用" (Use other licenses).

The "クリエイティブ・コモンズ (CC) を利用" option is selected, and a dropdown menu shows "PD(パブリック・ドメイン)" (PD (Public Domain)).

ライセンス情報

ベースIRI

: 作成するRDFデータでIDに用いるIRI (URI) の設定

メタデータの入力

ベースIRI (※) を入力してください。

※変換したRDFファイルを公開するURLに相当します。

- ☒ 一つのファイルで公開 (#)
- ☐ データベース (トリプルストア) を用いて公開 (/)
- ☐ SimpleLODIを用いて公開 (/)

ライセンス情報を入力してください。

クレジットで表示すべき名前

※オープンデータとして公開するには「CC-BY」または「パブリックドメイン」が望まれます。

- ☒ クリエイティブ・コモンズ (CC) を利用
- ☐ その他のライセンスを利用

■ ベースIRI

- <http://data.lodosaka.jp/lodws/> **ファイル名**とする(**拡張子はつけない**)
- ファイル名は, 練習では「**作業者の名前(例:kozaki)**」とする.
(実際の公開時には, 公開するURLを元に検討する)
 - **日本語ファイル名は避けたほうが無難(?)**
→ **環境によっては, 上手く認識しない場合あり.**
- オプションは「**データベース(トリプルストア)を用いて公開(/)**」を選択

■ ライセンス

- 神戸市のオープンデータは, すべてCC-BYで公開されているものなの, 下記のように選択・入力する
 - クレジットで表示すべき名前: **神戸市, 作業者**
 - ライセンスは, クリエイティブ・コモンズ(CC)の「**CC-BY**」

ベースIRI（URI）とは、

■ ベースIRIとは、

- RDFにおいてIDとするIRIを決める際に基準とするIRI
- 例) Wikidataの場合は
<http://www.wikidata.org/entity/Q41765>
↑この部分に相当
- RDFファイルでは、Prefixがついていないリソース（データ）のIRIは、行頭にベースIRIが付加されるものと扱われる
- 例) ベースIRIが<http://test.data/>とすると
そのRDFファイルで<1>というリソースのIRIは
<<http://test.data/>1>となる
- ※Turtle形式の@baseには、#のIRIは使えない。

ベースIRI（URI）の決め方



- 各データへはベースIRIによって決まるIRIでアクセスされるので,
 - データの公開者が管理している(できる)IRI
 - データを公開する際に用いるIRI(URL)を用いることが望ましい
- # と / の使い分けは, 一般に,
 - 1つRDFファイルでデータを公開するときは
http://test/data#001 のように#を
 - データサイズが大きく, RDFデータベースを用いて公開するときは
http://test/data/001 のように/を使うことが多い

テンプレート作成② プロパティ（語彙）の設定

プロパティ(語彙)の設定

使用する共通語彙を選択してください。

※「CSVの項目名」に対応させる「語彙」を選択します。

<input checked="" type="checkbox"/>	Schema.org	schema:	http://schema.org/
<input checked="" type="checkbox"/>	IPA共通語彙基盤	ic:	http://imi.ipa.go.jp/ns/core/rdf#
<input checked="" type="checkbox"/>	ダブリンコア(DC)	dc:	http://purl.org/dc/elements/1.1/
<input checked="" type="checkbox"/>	Geo-vocabulary	geo:	http://www.w3.org/2003/01/geo/wgs84_pos#

(2) プロパティ選択表示

「プロパティ選択表示」ボタンを押すと、
RDF変換で用いる語彙の候補が推薦される

推薦に用いる
語彙の選択

CSVファイルの各「項目名」に対応させるプロパティ(語彙)の設定

ID 列	項目名	プロパティ名	データ型
● 施設ID		<input checked="" type="checkbox"/> bp:施設ID	文字列(言語指定なし) ▼
		<input type="checkbox"/> その他	文字列(言語指定なし) ▼
		<input checked="" type="checkbox"/> bp:施設名	文字列(言語指定なし) ▼
		<input type="checkbox"/> rdfs:label(ラベル(日本語))	文字列(日本語) ▼
● 施設名		<input type="checkbox"/> rdfs:label(ラベル(英語))	文字列(英語) ▼
		<input type="checkbox"/> rdfs:label(ラベル)	文字列(言語指定無し) ▼
		<input type="checkbox"/> schema:name(名前(日本語))	文字列(日本語) ▼
		<input type="checkbox"/> schema:name(名前(英語))	文字列(英語) ▼
			文字列(言語指定無し) ▼
			文字列(日本語) ▼
			文字列(言語指定無し) ▼
			文字列(言語指定なし) ▼

推薦された語彙

CSVの1行目
の項目名

■ ID列

- CSVファイル内で「一意のID」となる項目があれば、「ID列」として選択する
- なければ「IDを自動付与」を選択する

■ プロパティ

- 適切な語彙が推薦されていたら選択
- 今回、最低限、下記のプロパティを用いる
(他のプロパティが使われる場合もある)
 - 名称 `rdfs:label`
 - 説明 `rdfs:comment`
 - Webサイト `foaf:homepage`
 - 緯度 `geo:lat`
 - 経度 `geo:long`
 - ~~外部LODとのマッピング情報~~ `rdfs:seeAlso`

バグがあって
正しく動かないため、
選択しないで下さい。

■ プロパティ(続き)

- それ以外の語彙を利用する場合「その他」欄に入力
 - 「推薦に用いる語彙」の一覧にあるprefixは利用可能
 - それ以外の場合は, 語彙のIRIをすべて記載
→よく使われる語彙の一覧は次スライド参照.
- CSVの項目名を利用するときは「bp:項目名」を選択
 - bp: は「ベースIRI+/property#」を表すprefix
 - データ型は, できれば適切なものを選択する

※プロパティを複数選択することも可

→複数の語彙に対応したいときに利用
(例: 共通語彙基盤 + Schema.org)

Schema.org <http://schema.org/docs/schemas.html>

Schema.org(日本語訳サイト) <http://schema-ja.appspot.com/>

共通語彙基盤(IMI) <http://imi.ipa.go.jp/>

Linked Open Vocabularies (LOV)

<http://lov.okfn.org/dataset/lov/>

DBPedia

http://mappings.dbpedia.org/index.php/Main_Page

Dublin Core

<http://dublincore.org/documents/dcmi-terms/>

→解説 <http://www.kanzaki.com/docs/sw/dublin-core.html>

Friend of a Friend (FOAF)

<http://xmlns.com/foaf/spec/>

→解説 <http://www.kanzaki.com/docs/sw/foaf.html>

- 現状では、推薦できる語彙は「ごく一部」のものに限られています
 - 共通語彙基盤には、ごく一部のプロパティのみ、試験的に対応.
- 今後、対応する語彙は順次増やす予定です

RDFファイルへの変換

RDFファイルへの変換

- (3) テンプレート生成/更新 上記で設定した情報でRDF変換テンプレートを生成します
- テンプレートのダウンロード 上記のRDF変換テンプレートをダウンロードします
- (4) 「CSV→RDF」の変換実行 下記のRDF変換テンプレートを用いてCSVファイルをRDFファイルに変換します

RDF変換テンプレート(※下記を直接編集することもできます)

```
##baseIRI=http://data.lododaka.jp/iodc-kobe2015/kobe/ institution01_20141128#
##ID=[[施設ID]]
##cc:attributionName=神戸市, 古崎
##cc:license=http://creativecommons.org/licenses/by/4.0/deed.jp
##mapping-rules-start
ID bp:施設ID [[施設ID]] ;
bp:施設名 [[施設名]] ;
rdfs:label [[施設名]] ;
bp:郵便番号 [[郵便番号]] ;
bp:方書 [[方書]] ;
bp:電話番号 [[電話番号]] ;
bp:FAX番号 [[FAX番号]] ;
geo:lat [[緯度]]^^xsd:double ;
geo:lng [[経度]]^^xsd:double ;
bp:行政区コード [[行政区コード]] ;
bp:カテゴリリスト [[カテゴリリスト]]
##mapping-rules-end
```

- 「**テンプレート生成/更新**」: 設定した情報に基づいてCSVをRDFに変換するテンプレートを作成する
- 「**テンプレートのダウンロード**」で, 作成したテンプレートをダウンロードして再利用できる
- 「**CSV→RDFの変換実行**」で, CSVファイルをRDFに変換する

既存のRDF変換テンプレートの読み込み

- ☒ ファイルから選択 選択されていません
- ☐ URLを指定
-

RDFファイルへの変換結果

■ 変換結果は下記のように表示される

- 「ファイル名」を入力し「RDFファイルのダウンロード」ボタンでRDFをダウンロードする。→ファイル名は「**作業者名(練習時)**」か「**元ファイル名**」+**.ttl**とする
- うまくダウンロードできない場合は、コピー&ペーストを利用する

RDFファイルへの変換結果

ファイル名:

```
@base <http://data.lodosaka.jp/ldc-kobe2015/kobe/institution01_20141128#> .
@prefix bp: <http://data.lodosaka.jp/ldc-kobe2015/kobe/institution01_20141128/property#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix cc: <http://creativecommons.org/ns#> .
@prefix geo: <http://www.w3.org/2003/01/geo/wgs84_pos#> .

<> cc:attributionName "神戸市, 古崎" ;
    cc:license <http://creativecommons.org/licenses/by/4.0/deed.ja> .
<http://data.lodosaka.jp/ldc-kobe2015/kobe/institution01_20141128#1> bp:施設ID "1" ;
    bp:施設名 "神戸市役所" ;
    rdfs:label "神戸市役所" ;
    bp:郵便番号 "650-8570" ;
    bp:電話番号 "078-331-8181 (代)" ;
    geo:lat "34.689454"^^xsd:double ;
    geo:lng "135.195739"^^xsd:double ;
    bp:行政区コード "中央区" ;
    bp:カテゴリリスト "市役所・区役所" .
<http://data.lodosaka.jp/ldc-kobe2015/kobe/institution01_20141128#2> bp:施設ID "2" ;
    bp:施設名 "東灘区役所" ;
    rdfs:label "東灘区役所" ;
```

(5) RDFファイルのダウンロード ※うまくダウンロードできない場合は[こちら](#)をご覧ください

- ダウンロードした「RDF変換テンプレート」を読み込んで利用することも可能
 - 同じ形式のCSVファイルが複数ある場合に有効
 - 設定画面へは反映されないため「テンプレート生成/更新」ボタンを押すと、読み込んだテンプレートが破棄されるので注意
- テンプレートは、直接、画面上で編集することも可能
 - 読み込んだテンプレートの修正や、設定からは生成できない複雑なRDF変換用のテンプレートを生成するのに利用できる

既存のRDF変換テンプレートの読み込み

☒ ファイルから選択 template_20...23 (1).txt

☐ URLを指定

2. LOD公開

- 作成したRDFファイルをWebサイトで公開する
 - 今回は, イベント用のWebサイトに公開
 - 公開するURLは,
`http://data.lodosaka.jp/lodws/ファイル名`
 - アップする前に, RDFのファイル名がベースIRIで指定したもの+.ttlになっているか確認
例) `http://data.lodosaka.jp/lodws/kozaki`
ならば `kozaki.ttl`

1. FTPクライアントへログイン

- <https://webftp.heteml.jp/> を利用
 - ユーザー名：**koujikozaiki_lodws**
 - パスワード：**kobe2017**
- FTPクライアントを利用する場合は
 - ホスト名(アドレス) **ftp187.heteml.jp**
 - ※FTPS 接続 の場合は **ssl187.heteml.jp** (推奨)

2. FTPでアップロード

- 「ファイルを選択」ボタン
→作成したRDF(****.ttl)を選択してアップロード
※上書き時には、「上書き」にチェック

■ ①Webブラウザでの閲覧

- WebブラウザのURL欄に
- <http://data.lodosaka.jp/lodws/>ファイル名
- のように, IRIを入れるとWebブラウザで閲覧可能

■ ②LODブラウザでの閲覧

- 下記のLODブラウザのIRIの欄に入力する
- **Quick and Dirty RDF browser**
※日本語IRIは文字化けする
<http://graphite.ecs.soton.ac.uk/browser/>
- **(Yet Another) Linked Data Browser**
※フォーマットでTurtleを選択する必要あり
<http://www.kanzaki.com/works/2014/pub/ld-browser>

■ Fusekiのインストール

- 「トリプルストアの導入」のハンズオンの方法を参考に, Fusekiをインストール

<https://goo.gl/fUtv5m>

■ Fusekiで新しいDBを作成

- Manage datasets > add dataset

■ 作成したDBにRDFファイルをアップロード

- upload data

■ SARQLクエリを試す

■ 外部のLODと連携してのクエリの実用例

PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>

PREFIX wdt: <http://www.wikidata.org/prop/direct/>

SELECT ?s ?wd ?l

WHERE {

?s bp:参照 ?wd

こちらは、手元のDBから検索

SERVICE <https://query.wikidata.org/sparql>

{

?wd rdfs:label ?l.

FILTER(lang(?l)="en")

}

}

LIMIT 25

WikidataのSPARQLエンドポイント(機械用)から、英語のラベルを取ってくるクエリを実行して、結果をまとめて返す

3. 外部LODとのリンク



今回は省略
(前でデモ)

1. なんでもリンクにアクセス
2. オープンデータ(CSV)の「名称」に相当する列をコピー
3. なんでもリンクの入力欄にペースト
4. 「SPARQLエンドポイント」欄に
<https://query.wikidata.org/sparql> を入力
5. 「割当実行」DBpediaのデータと一致するデータの一覧が表示される(最初の10件)
6. 「ダウンロード」ボタンで結果をダウンロード
7. エクセル等で開き, 元のオープンデータの最終列に
マッピング情報をコピー&ペースト
 - 項目名は「参照」としておく

マッピング情報に関する注意



今回は省略
(前でデモ)



- マッピング情報はUTF8形式となっているので、
以下のように扱う
 - **Google Sheets**では
「File」→「Open」→「Upload」
 - **Excel**の場合は、
「データ」→「挿入」→「外部データの取り込み」→「テキストファイル」
- 以下の作業は、作業1, 2と同様
 - マッピング情報に用いる「プロパティ」の選択のみ気を付ける

5. Simple LODIの設定



今回は省略
(前でデモ?)

- 1. **Simple LODI**をダウンロード(**ZIP**形式が簡単)して, 解凍する.
 - <https://github.com/uedayou/simplelodi> から
- 2. 「**simplelodi-master**」のフォルダ名を, **LOD**を公開するフォルダ名に変更
 - 例) <http://data.lodosaka.jp/lodws/kozaki>
ならば「**kozaki**」
- 3. 「2のサブフォルダ**data**」に「5. で分割した. ttlファイルを格納」する
- 4. 「2のフォルダー式」を
<http://data.lodosaka.jp/lodws/>
にアップロードする
※**FTP**ソフトを利用する(例:**FFFTP**, など)

- 公開した**LOD**は,
<http://data.lodosaka.jp/lodws/フォルダ名/ファイル名.ttl>
のアドレスでアクセスできる.
- 拡張子を変更することで, 様々な形式で取得可
 - .ttl → Turtle形式
 - .nt → N-Triple形式
 - .xml → RDF/XML形式
 - .html → HTML形式
 - .json → JSON形式
 - .jsonld → JSON-LD形式