

TP – RDF Resource Description Framework

Exercice 2: Le langage Turtle

Parmi les nombreux langages de sérialisation de RDF (N-Triple, RDF/XML, JSON-LD,), il est important de connaître Turtle, ceci pour plusieurs raisons:

- C'est un format concis et lisible, relativement facile à manipuler (éditer) comparativement aux autres formats de sérialisation du RDF,
- La syntaxe du langage SPARQL utilisé pour interroger des jeux de données RDF est basée sur Turtle pour représenter des motifs de graphes (*graph patterns*). Connaître Turtle vous facilitera donc l'apprentissage de SPARQL.

2.1 Lire du Turtle

Soit le fichier Turtle suivant :

```
@prefix rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#> .
@prefix rdfs: <http://www.w3.org/2000/01/rdf-schema#> .
@prefix xsd: <http://www.w3.org/2001/XMLSchema#> .
@prefix dbr: <http://dbpedia.org/resource/> .
@prefix geodata: <http://sws.geonames.org/> .
@prefix foaf: <http://xmlns.com/foaf/0.1/> .
@prefix schema: <http://schema.org/> .
@prefix ex: <http://example.net/people#> .

ex:Amelie_Ponet a foaf:Person ;
    foaf:givenname "Amélie";
    foaf:familyName "Ponet";
    foaf:knows ex:Ivan_Noel.

ex:Ivan_Noel a foaf:Person, schema:Person ;
    foaf:givenname "Ivan", "Iván"@hu ;
    foaf:familyName "Noel" ;
    foaf:birthday "1985-02-24"^^xsd:date ;
    schema:address [ a schema:Address ;
        schema:addressLocality "Grenoble" ;
        schema:url dbr:Grenoble;
        rdfs:seeAlso geodata:3014728
    ] ;
    foaf:knows ex:Amelie_Ponet,
        <https://www.w3.org/People/Berners-Lee/card#i> .
```

Ce fichier décrit des personnes en utilisant en particulier les vocabulaires [FOAF \(Friend Of A Friend\)](#) et [schema.org](#).

Question 1 : à quelles URIs correspondent `foaf:Person` et `schema:Person` ?

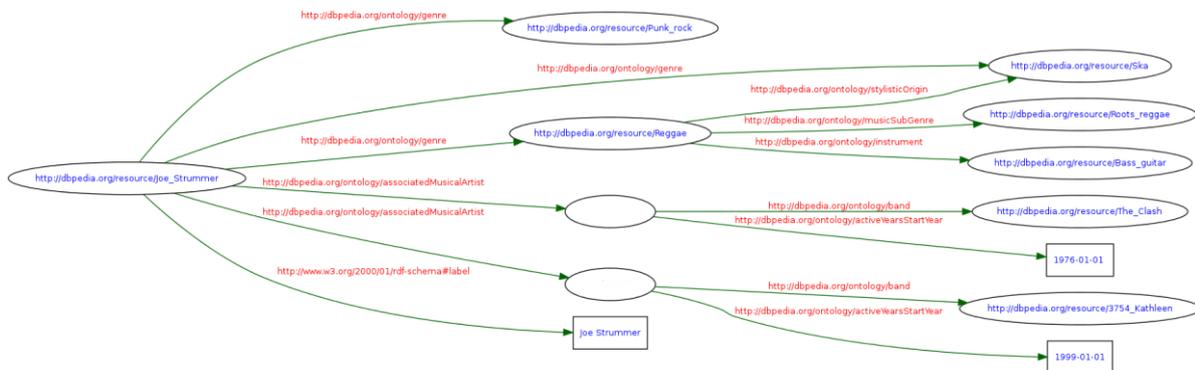
Tapez ces URIs dans votre navigateur. Que se passe-t-il ?

Pouvez-vous obtenir une représentation RDF de ces ressources ?

Question 2 : donnez une représentation graphique du graphe RDF défini par le fichier turtle ci-dessus.

2.2 Ecrire du Turtle

Le graphe ci-dessous représente des données extraites du graphe de connaissances DBpedia concernant le chanteur/compositeur britannique Joe Strummer.



Graphe de données sur Joe Strummer extrait de DBpedia.

Question : écrivez une représentation de ce graphe en Turtle et vérifiez que celle-ci est syntaxiquement correcte.

Pour effectuer cette validation, vous pouvez utiliser soit des outils en ligne soit des outils en ligne de commande. Parmi ceux-ci en voici quelques-uns :

- Outils en ligne.
 - [IDLab Turtle Validator](#) basé sur le package NodeJS [TurtleValidator](#).
 - [RDF Translator](#), outil de conversion pour les différents format de sérialisation de RDF (Turtle, RDF/XML, RDFa....) basé sur la bibliothèque Python [RDFLib](#).
 - [EASYRdf Converter](#), outil de conversion similaire à RDF Translator basé sur la librairie PHP [EASYRDF](#)
- Outils en ligne de commande.
 - [TurtleValidator](#) : validateur écrit en JavaScript (NodeJS).
 - Le framework Java [Jena](#) de la fondation [Apache](#), en particulier les [outils en ligne de commande](#) permettant lire et écrire des fichiers RDF.

Vous pouvez vous reporter à [ce document](#) pour plus d'informations sur l'installation de Jena sur votre poste de travail.

ATTENTION : quand vous utilisez le lien de ce document, recharger bien la page pour être sûr d'avoir la version mise à jour par Philippe le 30/09/2024.

Exercice 3: Modéliser en RDF

La table suivante montre un échantillon des données que la librairie ARTEMIS Bookstore conserve sur les livres qu'elle possède en stock. Ces données sont pour le moment stockées dans un tableur, il s'agit de les représenter en utilisant le modèle de données RDF.

	A	B	C	D	E	F
1	LastName	FirstName	BookTitle	PagesNb	Isbn	PublisherName
2	Abbey	Edward	Black Sun (Edward Abbey novel)	159	978-1-55566-286-8	Simon & Schuster
3	Abbey	Edward	Fire on the Mountain (Abbey novel)	211	0-8263-0457-5	Dial Press
4	Abbey	Edward	Hayduke Lives	352	ISBN 0-316-00411-1 (first edition,	Little, Brown and Company
5	Abbey	Edward	Jonathan Troy	374	1-131-40684-2	Dodd, Mead and Company
6	Abbey	Edward	The Monkey Wrench Gang	352	ISBN 0-397-01084-2 (hardback ed	Lippincott Williams & Wilkins
7	Abbey	Edward	The Brave Cowboy	277	0-8263-0448-6	Dodd, Mead and Company
8	Abbey	Edward	The Fool's Progress	485	0-8050-0921-3	Henry Holt and Company
9	Acito	Marc	How I Paid for College: A Novel of	288	ISBN 0-7679-1841-X (first edition,	Broadway Books
10	Acker	Kathy	Blood and Guts in High School	165	ISBN 0-8021-3193-X (Paperback e	Grove Press
11	Adler	Warren	The War of the Roses (novel)	263	0-446-51220-6	Hachette Book Group
12	Adler	Warren	The War of the Roses (novel)	263	0-446-51220-6	Hachette Book Group

artemisBookstoreData.xlsx

Question 1: dessinez un graphe RDF qui reflète le contenu de la table (vous ne vous intéresserez qu'aux données contenues dans les deux premières lignes de la table).

Pour les URIs vous pouvez utiliser les préfixes suivants:

- <http://www.artemisbookstore.com/ontology#> pour les ressources qui correspondent à une propriété ou une classe,
- <http://www.artemisbookstore.com/resource/> pour les autres ressources.

Question 2: Créez un fichier texte `artemis.ttl` dans le lequel vous écrirez une sérialisation de votre graphe en Turtle. Comme pour l'exercice 2.2 vérifiez que la syntaxe est correcte.

Exercice 4: Extraire des données RDFa

La page <http://www.ebusiness-unibw.org/> de l'E-Business and Web Science Research Group de l'[Université Militaire de Munich](http://www.ebusiness-unibw.org/), contient des données structurées (décrites à l'aide de RDFa).

Question 1: Que contiennent ces données ? Représentez ces données en Turtle et dessinez le graphe RDF correspondant.

Question 2: Google search recommande d' [utiliser JSON-LD](#) plutôt que du RDFa pour annoter des documents web. Produisez une sérialisation en JSON-LD que vous pourriez substituer au code RDFa dans la page <http://www.ebusiness-unibw.org/> (pensez à utiliser un convertisseur pour produire la sérialisation en JSON-LD).

Exercice 5 : Générer des données RDF.

Question 1: Ecrivez un programme Java qui lit le fichier CSV [artemisBookstoreData-v1.csv](#) des données de la librairie ArtemisBookstore vues à la [question 1 de l'exercice 3](#) et produit un fichier RDF au format turtle (`artemisBookstoreData-v1.ttl`).

Vérifiez que votre fichier rdf en utilisant l'un des outils de validation vus dans les exercices précédents.

optionel: vous pouvez éventuellement modifier votre programme pour, lorsque vous parsez votre fichier, créer un modèle RDF avec l'API core de Jena et proposer une sortie dans différents formats de sérialisation (RDF/XML, Turtle, JSON-LD...).

Pour vous aidez vous pouvez vous référer à l'introduction au framework Jena accessible depuis la page [TPs](#) du site web du cours.

Exercice 6 : Manipuler du RDF avec l'API Jena

Cet exercice est optionnel

Question 1: En utilisant l'API Jena Core écrire un programme Java qui lit le fichier RDF ((artemisBookstoreData-v1.ttl)) au format turtle que vous avez produit à l'exercice précédent et qui étant donné le nom d'un auteur trouve et affiche le titre de tous les livres dont il est l'auteur.