

UNIVERSIDADE DO MINHO

MESTRADO INTEGRADO EM ENGENHARIA INFORMÁTICA

4º ANO, 2º SEMESTRE, 2019/2020

Scripting no Processamento de Linguagem Natural

Ontologia para HTML



Ana Pereira
A81712

June 26, 2020

Índice

1	Introdução	2
2	Ontologia Simplificada	3
3	Termos Conhecidos	4
4	Conversão para HTML	4
5	Conclusão	5

1 Introdução

Este trabalho prático tinha como objetivo criar uma ferramenta que transformasse ontologias inseridas no domínio de histórias de família em páginas HTML simples.

Para além dessa transformação também foi efetuada a conversão da ontologia simplificada numa ontologia expandida, na sintaxe Turtle, referindo, quando aplicável, classes e propriedades presentes em ontologias conhecidas, como FOAF e Schema.

Este relatório explica sucintamente as estratégias e decisões implementadas na criação da ferramenta.

2 Ontologia Simplificada

O propósito da ferramenta é criar uma visualização simples dos dados presentes numa ontologia. No entanto, os formatos destas (RDF, Turtle) contém detalhes que, mesmo sendo necessários, afetam a sua perceção. Então, o nosso programa recebe uma ontologia simplificada, sem ser necessário adicionar prefixos/namespaces.

De seguida são apresentados excertos da ontologia exemplo, demonstrando a sintaxe simplificada das classes, propriedades e indivíduos.

```
Ontology    FamilyTest.

Person  a    Class.
Image   a    Class;
        subClass    MediaObject.

gender  a    DatatypeProperty;
        domain    Person;
        range     String.

biography  a    DatatypeProperty;
        range   String.

image      a    ObjectProperty;
        range   Image.

worksFor   a    ObjectProperty;
        domain  Person;
        range   Institution.

ChronologyID1 a Individual;
              a Chronology;
              hasEvent Marriage1;
              hasEvent BirthOfAna;
              hasEvent BirthOfJose.
```

Exemplos da sintaxe simplificada da ontologia *input* podem ser consultados através da função de ajuda, usando a *flag* -h ao correr o comando: `python3 ontology2html.py -h`.

Também é referido que, se a ontologia apresentar imagens, é necessário definir a propriedade `contentUrl`, correspondente ao *path* da imagem a mostrar.

3 Termos Conhecidos

Tendo efetuado o *parsing* da ontologia simplificada, esta é guardada num dicionário. De modo a convertê-la numa ontologia expandida com a sintaxe do Turtle normal, foi necessário compilar um ficheiro, **knownDict**, com equivalências de predicados, classes e propriedades com elementos já conhecidos de outras ontologias.

Excerto de equivalências presentes no ficheiro:

Class	owl:Class
domain	rdfs:domain
range	rdfs:range
String	xsd:string

Para além disso também tem correspondências de algumas classes e propriedades (em português e inglês) do domínio familiar com classes e propriedades presentes nas ontologias **FOAF** e **Schema**. Para as propriedades é indicado o *domain* e o *range* das mesmas.

Exemplo de 2 **datatype properties** e 2 **object properties**:

age	foaf:age	Person	Int
gender	foaf:gender	Person	String
hasChild	schema:children	Person	Person
hasEvent	schema:event	Thing	Event

Assim, carregando a informação deste ficheiro para o programa é possível gerar a ontologia no formato **.ttl**, adicionando os prefixos e namespaces necessários, bem como classes e propriedades do FOAF e do Schema.

4 Conversão para HTML

É possível ver o que está presente na ontologia em páginas HTML. Para tal foi usada a *stylesheet* da **w3schools**, presente na pasta **html**, adicionando *css* às páginas para uma melhor visualização dos dados.

A página principal da ontologia, **mainPage.html**, contém uma listagem dos elementos presentes na mesma, sendo que cada um dos elementos tem a sua página individual com os respetivos triplos.

Os indivíduos que sejam referentes a imagens, isto é, pertençam a uma das seguintes classes: **Image**, **Imagem**, **Picture**, **Photo** ou **Foto**; se tiverem a propriedade **contentUrl** são apresentados *inline*.

5 Conclusão

Apesar de inicialmente ter tido algumas dificuldades a entender o que era suposto implementar com esta ferramenta, sinto que realizei um bom trabalho, atingindo os objetivos do mesmo.

Alguns aspectos que poderiam ser melhorados seriam: a apresentação das páginas referentes a cada indivíduo, mostrando-se mais personalizadas; e a utilização de *libraries* como o **owlready2** ou o **rdflib** para efetuar a expansão da ontologia para o formato Turtle.