

SKOS Parsers for the OAEI 2007 alignment tasks

Miklos Nagy (nagy@pt.lu)

Introduction

The purpose of the SKOSParser and SKOSChunkParser is to create OWLontology from the SKOS files that have been provided for the different OAEI alignment tasks. SKOSParser loads the whole SKOS file and converts it into one OWLontology. It is recommended to use when the size of the source file is less than 5 MB e.g. environment SKOS. SKOSChunkParser can be used to iterate through a large SKOS file and create OWLontology fragments with predefined concept/class size. In order to avoid memory problems both parsers use the StAX parsing model.

Conversion assumptions

We have recognised that there is no simple one to one mapping between SKOS and OWL. In order to create OWL from SKOS we have made the following conversion assumptions:

- *skos:Concept* is converted into *owl:Class*
- *skos:prefLabel* is converted into *rdfs:label*
- *skos:altLabel* is converted into *rdfs:label*
- *skos:broader* is converted into *rdfs:subClassOf*
- *skos:definition* is converted into *rdfs:comment*
- *skos:scopeNote* is converted into *rdfs:comment*

Comparison with library, food, environment OWL

The provided conversions are nearly identical to our solution. There is a difference with the *skos:related* and *skos:hiddenLabel* statements which we do not convert into OWL.

- *skos:Concept* is converted into instances of *owl:Class*
- *skos:prefLabel*, *skos:altLabel* and *skos:hiddenLabel* statements are converted to *rdfs:label* statements, which removes the subtle distinctions that exist between these different properties (in GTT for instance, many altLabels are not synonyms at all)
- various kind of *skos:notes* are converted to *rdfs:comments*
- *skos:broader* statements are converted into *rdfs:subClassOf* statements;
- *skos:related* statements are converted into *rdfs:seeAlso* statements.

Limitations of the parsers

Both the SKOSParser and SKOSChunkParser do not load all language label elements from the SKOS file. The default language which will be loaded is only the English. For both parsers a second constructor is provided where the language element is defined. Additionally the parsers have been developed for the OAEI SKOS therefore by default it can parse and convert correctly the provided SKOS files.

Usage

The zip file contains all dependent libraries including the StAX parser. However it is recommended to use together with the alignment api. As a starting point one can refer to the `uk.ac.open.UsageTest` class, which provides examples of the usage. Additionally a javadoc contains further detailed explanations of the particular methods and classes. In case of any further questions do not hesitate to contact me on the following e-mail: nagy@pt.lu.