NCI Thesaurus OWL to OWL2: Conversion and Differences

This page explains the differences between NCI Thesaurus in OWL1 and OWL2.

- Removal of double-typing and ranges
- Creation of user-defined datatypes
- Transformation of complex properties to owl:Axioms
- Meaningless identifiers used in the IRI fragments, derived from concept codes

Removal of double-typing and ranges

In OWL1, properties were declared as an AnnotationProperty with type DatatypeProperty (or vice-versa). When asserted on a class, their ranges (string or Literal) would persist. Associations were declared as an AnnotationProperty with type ObjectProperty (or vice-versa). Example 1

In OWL2, properties are declared as just an AnnotationProperty. Their assertions have no range (making them untyped literals). Associations are declared as just an AnnotationProperty with a range of anyURI. Roles are still just an ObjectProperty. Example 2

Creation of user-defined datatypes

In OWL1, properties such as Concept_Status and Semantic_Type had a range of values asserted on their declaration. Example 3

In OWL2, these properties have a range of a user-defined datatype, an rdfs:Datatype declared elsewhere in the file. Example 4

Transformation of complex properties to owl:Axioms

In OWL1, properties such as FULL_SYN and DEFINITION (with range Literal) had values of embedded xml to represent all sub-properties ("qualifiers") and their respective values. Example 5

In OWL2, these sub-properties and their values exist in owl:Axioms. In the owl:Axiom, the annotatedSource links to the class ID, annotatedProperty links to the class ID of the property, and annotatedTarget links to the filler value of the property. The filler value of the property is the term-name for FULL_SYN, def-definition for DEFINITION and ALT_DEFINITION, go-term for GO_Annotation, and Target_Term for Maps_To. Example 6

Meaningless identifiers used in the IRI fragments, derived from concept codes

Our primary distribution up to Spring 2017 has been an OWL1 format using semantically meaningless identifiers for the IRIs; the fragment identifiers used for the IRIs were the values of the "code" property from each concept. We also continued publishing a deprecated version with semantically meaningful identifiers; rather than a code value, text resembling the preferred name of the concept was used, e.g. "Neoplasm_By_Site". This deprecated distribution is not going to be carried forward in OWL2. In the examples below, all IRI fragment identifiers utilize the concept's (or property's) code.

Example 1

```xml
<!-- http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#Preferred_Name -->
<owl:DatatypeProperty rdf:about="#P108">
  <rdf:type rdf:resource="http://www.w3.org/2002/07/owl#AnnotationProperty"/>
  <rdfs:label rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Preferred_Name</rdfs:label>
  <P97 rdf:parseType="Literal"><ncicp:ComplexDefinition><ncicp:def-definition>The word or phrase that NCI uses by preference to refer to the concept. NCI</ncicp:def-definition><ncicp:def-source>NCI</ncicp:def-source></ncicp:ComplexDefinition></P97>
  <P90 rdf:parseType="Literal"><ncicp:ComplexTerm><ncicp:term-name>Preferred Name</ncicp:term-name><ncicp:term-group>SY</ncicp:term-group><ncicp:term-source>NCI</ncicp:term-source></ncicp:ComplexTerm></P90>
  <P90 rdf:parseType="Literal"><ncicp:ComplexTerm><ncicp:term-name>Preferred Term</ncicp:term-name><ncicp:term-group>SY</ncicp:term-group><ncicp:term-source>NCI</ncicp:term-source></ncicp:ComplexTerm></P90>
  <P90 rdf:parseType="Literal"><ncicp:ComplexTerm><ncicp:term-name>Preferred_Name</ncicp:term-name><ncicp:term-group>PT</ncicp:term-group><ncicp:term-source>NCI</ncicp:term-source></ncicp:ComplexTerm></P90>
  <P106 rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Conceptual Entity</P106>
  <code rdf:datatype="http://www.w3.org/2001/XMLSchema#string">P108</code>
  <P107 rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Preferred Name</P107>
  <P108 rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Preferred_Name</P108>
</owl:DatatypeProperty>
```
Example 2

<!-- http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#Preferred_Name -->

<owl:AnnotationProperty rdf:about="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#P108">
  <P97>The word or phrase that NCI uses by preference to refer to the concept.</P97>
  <P107>Preferred Name</P107>
  <P90>Preferred Name</P90>
  <P90>Preferred Term</P90>
  <P106>Conceptual Entity</P106>
  <NHCO>P108</NHCO>
  <rdfs:label>Preferred_Name</rdfs:label>
</owl:AnnotationProperty>

<owl:Relation rdf:about="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#P108">
  <owl:announcedType rdf:resource="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#P108"/>
  <owl:announcedProperty rdf:resource="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#P97"/>
  <owl:announcedTarget>The word or phrase that NCI uses by preference to refer to the concept.</owl:announcedTarget>
  <P380>060127</P380>
  <P379>DEFAULT_Review</P379>
  <P378>NCI</P378>
</owl:Relation>

<owl:Relation rdf:about="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#P108">
  <owl:announcedType rdf:resource="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#P108"/>
  <owl:announcedProperty rdf:resource="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#P90"/>
  <owl:announcedTarget>Preferred Name</owl:announcedTarget>
  <P383>SY</P383>
  <P384>NCI</P384>
</owl:Relation>

<owl:Relation rdf:about="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#P108">
  <owl:announcedType rdf:resource="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#P108"/>
  <owl:announcedProperty rdf:resource="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#P90"/>
  <owl:announcedTarget>Preferred Term</owl:announcedTarget>
  <P383>SY</P383>
  <P384>NCI</P384>
</owl:Relation>

<owl:Relation rdf:about="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#P108">
  <owl:announcedType rdf:resource="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#P108"/>
  <owl:announcedProperty rdf:resource="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#P90"/>
  <owl:announcedTarget>Preferred Name</owl:announcedTarget>
  <P383>PT</P383>
  <P384>NCI</P384>
</owl:Relation>

<owl:Relation rdf:about="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#P108">
  <owl:announcedType rdf:resource="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#P108"/>
  <owl:announcedProperty rdf:resource="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#P90"/>
  <owl:announcedTarget>Preferred Name</owl:announcedTarget>
  <P383>PT</P383>
  <P384>NCI</P384>
</owl:Relation>

<owl:Relation rdf:about="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#P108">
  <owl:announcedType rdf:resource="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#P108"/>
  <owl:announcedProperty rdf:resource="http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#P90"/>
  <owl:announcedTarget>Preferred Name</owl:announcedTarget>
  <P383>PT</P383>
  <P384>NCI</P384>
</owl:Relation>
Example 3
An NCI Thesaurus property used to indicate the standing of a concept in relation to currently accepted classifications and concepts. In NCI Thesaurus concept status subtype indicates concepts with unusual and problematic characteristics that should be evaluated by people and/or programs before those concept are used.

Example 4
An NCI Thesaurus property used to indicate the standing of a concept in relation to currently accepted classifications and concepts. In NCI Thesaurus concept status subtype indicates concepts with unusual and problematic characteristics that should be evaluated by people and/or programs before those concept are used.

**Concept Status**

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP</td>
<td>Concept_Pending_Approval</td>
</tr>
<tr>
<td>DC</td>
<td>Deprecated_Concept</td>
</tr>
<tr>
<td>HC</td>
<td>Header_Concept</td>
</tr>
<tr>
<td>OC</td>
<td>Obsolete_Concept</td>
</tr>
<tr>
<td>PC</td>
<td>Provisional_Concept</td>
</tr>
<tr>
<td>RC</td>
<td>Retired_Concept</td>
</tr>
</tbody>
</table>

The `Concept_Status` is an enumerative data type containing the following values:

- Concept_Pending_Approval
- Deprecated_Concept
- Header_Concept
- Obsolete_Concept
- Provisional_Concept
- Retired_Concept

This data type is defined in the NCI Thesaurus ontology and can be accessed at [http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#Concept_Status-enum](http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#Concept_Status-enum).
Example 5

<!-- http://ncicb.nci.nih.gov/xml/owl/EVS/Thesaurus.owl#Algorithms -->

<owl:Class rdf:about="#C16275">
  <rdfs:label rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Algorithm</rdfs:label>
  <rdfs:subClassOf rdf:resource="#C20181"/>
  <P108 rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Algorithm</P108>
  <P366 rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Algorithms</P366>
  <P207 rdf:datatype="http://www.w3.org/2001/XMLSchema#string">C0002045</P207>
  <code rdf:datatype="http://www.w3.org/2001/XMLSchema#string">C16275</code>
  <P106 rdf:datatype="http://www.w3.org/2001/XMLSchema#string">Intellectual Product</P106>
</owl:Class>

Example 6
A defined procedure for solving a problem. Applied to a problem-solving procedure implemented in software to be executed by a computer.