

Guideliness for the Ontology Engineering work
Asun Gómez-Pérez
Oscar Corcho
Ontology Engineering Course 2022-23
Master on Artificial Intelligence

The goal of this work is (a) to develop an OWL ontology using the NeOn methodology, to implement it in an ontology development tool (i.e. Protégé), to evaluate it using Opps! and generate the documentation with Widoco. Anex I includes a set of topics in the domain of tourism.

To this end, the following tasks are recommended.

1. Introduction. High level overview of the ontology you plan to develop and brief description of the datasets and web pages that you plan to use.
2. Make a short overview of the NeOn methodology making explicit:
 - a. which scenarios you plan to follow
 - b. which activities from the glossary of activities you plan to execute in your ontology development project
3. For each of the activities involved in the selected scenarios include a section in your documentation. At least, you should include a section for the following activities
 - a. **Ontology specification.** You should include a complete ontology specification requirement document using the template explained during the lectures. The goal of the ontology should be clear enough. The document should include relevant competency questions and their answers.
 - b. **Ontology Schedule.** You should identify the ontology life cycle model you use in your ontology development. Activities should come from the glossary of terms. Include a Gantt chart with the planned activities.
 - c. **Search for existing top level and domain ontologies** that you plan to reuse when building your ontology. Explain if you need to reuse the ontology as a whole or if you need some modules or statements. Explain in which repositories you made the search, ontologies found, their relevance to your work and the criteria being used for selecting or withdrawing some of them.
 - d. **Search for non ontological resources and other terminologies** that could be transformed into ontologies. Keep track of the URLs where you found them. For the selected resources, do not forget to justify why you have selected them. Check if you have Access rights for using them within your hands-on assignment.
 - e. **Search for some ontology design patterns** in the ontology design pattern portal that could be reused in your development.
 - f. **Build a conceptual model** that integrates outcomes from the previous sections (c,d,e). This is the most important part of the work you are doing. Try to use:
 - i. Top level ontologies and other well-known ontologies. Classify in the pyramid of ontologies (figure use vs reuse) each of the ontologies that you reuse.
 - ii. Transform each non-ontology resource into an ontology by using the T-box, A-Box or Population.
 - iii. Select some Ontology Design Patterns (events, sequence, etc.)
 - iv. Build the conceptual model of your ontology by integrating the above sources. The conceptual model should have at least 40 concepts,

several subclass-of relations, disjoint, part-of (if needed), and ad-hoc relations.



- g. **Implement the ontology** in an ontology development tool, or other ontology editor, using OWL as ontology language
 - h. **Evaluate the ontology** with OOPS! and include in your report the pitfalls found. It is highly advisable to combine this evaluation with other ontology evaluation techniques.
 - i. **Improve your ontology** (conceptual model and implementation) taking into account the suggestions given by OOPS!. Iterate in these steps until the ontology pass most of the OOPS! recommendations.
 - j. **Document the ontology** with Widoco and use Ontoology if required
 - k. If possible, make some of your classes **multilingual**.
4. Conclusions

Anexo I

Prácticas MUIA Ingeniería Ontológica – 2022

Seleccionar una ontología, entre las siguientes:

Caso 1. Ontologías para datos abiertos de universidades

En el contexto de las iniciativas de datos abiertos que muchas administraciones públicas llevan adoptando desde hace una década, las universidades públicas han tardado más en comenzar a adoptar políticas de datos abiertos sobre sus datos administrativos, algo que contrasta con el empuje hacia la ciencia abierta y los datos abiertos de investigación que muchas están adoptando (por ejemplo, en Europa, en el contexto de la European Open Science Cloud).

Por ejemplo, la Universidad de Alicante tiene un portal de datos abiertos (<https://datos.ua.es/>) donde se publican diversos tipos de datos abiertos sobre la universidad.

La iniciativa Universidata (<https://www.universidata.es/>) ha surgido como una iniciativa público-privada para la definición de estos conjuntos de datos que se pueden hacer disponibles en los portales de datos abiertos de las universidades. Concretamente, es de interés el llamado “núcleo común” (<https://dimetrical.atlassian.net/wiki/spaces/UNC/pages/414154820/Inicio>), que define un conjunto de datasets que podrían ser publicados por las universidades.

El objetivo de este caso de estudio es el de generar al menos dos ontologías bien diseñadas relacionadas con algunos de los conjuntos de datos aquí identificados. Se adjunta link a un repositorio en GitHub donde se está realizando este tipo de trabajo de manera abierta, y se recomienda leer el TFG titulado “Creación de un Grafo de Conocimientos para Datos Abiertos Universitarios. Subconjunto de Datos 2.” realizado por Victor Morcuende y defendido en junio de 2022.

Caso 2. Ontologías para datos abiertos de ciudades

En el contexto de la iniciativa OpenCityData (<https://github.com/opencitydata/>) y del proyecto Ciudades Abiertas (<https://ciudadesabiertas.es/>) se han desarrollado un buen número de ontologías que permitan representar datos abiertos de ciudades, en muy diversas áreas (transporte, gestión económica, equipamientos, servicios, etc.). En la Web de vocabularios de Ciudades Abiertas (<https://vocab.ciudadesabiertas.es/>) se encuentran las ontologías que se consideran mejor diseñadas, de acuerdo con principios de reutilización de ontologías, patrones claros y bien utilizados, un gran número de ejemplos, e incluso vídeos explicativos.

El objetivo de este trabajo será el de seleccionar algún conjunto de datos que no haya sido aún tratado, de entre todos los conjuntos de datos identificados en la iniciativa OpenCityData y disponibles en <https://opencitydata.github.io/CatalogoFEMP/>, y crear una ontología y todos sus artefactos asociados siguiendo el mismo esquema que se utiliza en la página de vocabularios de Ciudades Abiertas. Pueden ser especialmente interesantes los conjuntos de datos sobre infraestructura verde de la ciudad.

Caso 3. Ontología para la representación de pólizas de seguros

Las pólizas de seguros se escriben normalmente como documentos con una gran cantidad de terminología legal y una estructura similar. La falta de unificación de un modelo concreto y específico para representar una póliza hace que sean muy complicadas de procesar automáticamente, para lo cual existen iniciativas que están tratando de aplicar técnicas de procesamiento de lenguaje natural para obtener el modelo de dichas pólizas.

El objetivo de este trabajo es el de partir de un conjunto reducido de pólizas de seguro (de hogar, automóvil o cualquier otro aspecto), seleccionado por los alumnos, y crear un modelo que permita representar las entidades y relaciones entre entidades principales de dichas pólizas, de una manera que sea suficientemente extensible como para poderla aplicar a otras pólizas.