The OntoEnrich platform: using workflows for quality assurance and axiomatic enrichment of ontologies

Manuel Quesada Martínez,
Jesualdo Tomás Fernández-Breis,
Robert Stevens,
Nathalie Aussenac-Gilles,
Daniel Karlsson

UK Ontology Network Meeting

Newcastle, April 14th 2016

UNIVERSIDAD DE MURCIA

IMIB
Instituto Murciano de Investigación Biosanitaria

MANCHESTER 1824
The University of Manchester

IRIT
CNRS - INPT - UPS - UT1 - UTM

LiU LINKÖPINGS UNIVERSITET

OntoEnrich

What is the OntoEnrich framework?

- Systematic analysis and exploitation of lexical regularities in ontologies labels.

- Principles of the analysis

  - Systematic naming.

  - Re-use of content.

- Web platform

E.g. *Taxonomic Relation suggested by the Systematic Naming*
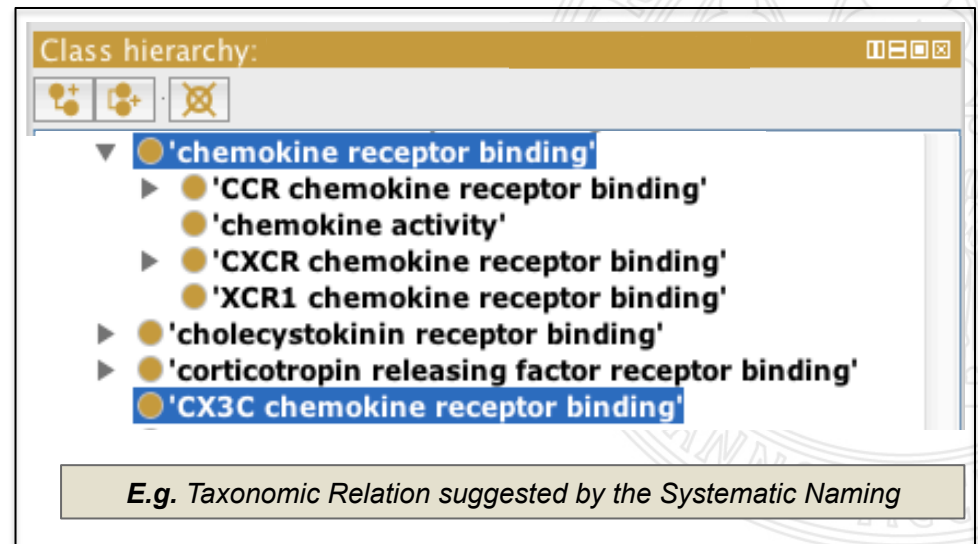
Foundations of lexical regularities

What is a lexical regularity?

A **lexical regularity (LR)** is a group of consecutive ordered words that appear in more than one class of an ontology.

properties

- **Length** and **frequency**.

- **Alignments** with other classes or fragments of classes: decompositions.

- ...

Lexical regularity	Length	Num Labels	Is a clas	Explore regularity
transmembrane transporter	2	395	false	👁
transmembrane transporter activity	3	393	true	👁
dehydrogenase activity	2	361	false	👁
receptor binding	2	339	true	👁
receptor activity	2	339	true	👁
synthase	1	324	false	👁

Insert the text to filter first column

⏮ ◀ 11-20 of 1.208 ▶ ⏭

Explore lexical regularity one by one　　　Click here to open

| Lexical regularity	congenital stenosis	Is a class
Labels exhibiting the regularity:	44 labels	
Least common subsumers:	<SNOMED CT Concept (SNOMED RT+CTV3)>	
Pos-tagging information:	{(JJ,44)} {(NN,44)}	

Labels view　　Axiom view　　**Hierarchical tree view**　　Metrics panel　　ODPs panel

▼ (
- ⊜ Congenital stenosis of pulmonary artery (disorder)
- ● Congenital esophageal ring (disorder)
▼ ⊜ Congenital stenosis (disorder)
- ● Congenital stenosis of nasolacrimal duct (disorder)
▶ ⊜ Congenital stenosis of cardiac valve (disorder)
...

* Image generated using the OntoEnrich web: SNOMED CT July 2015 International released ontology (right) and GO MF ontology (left)
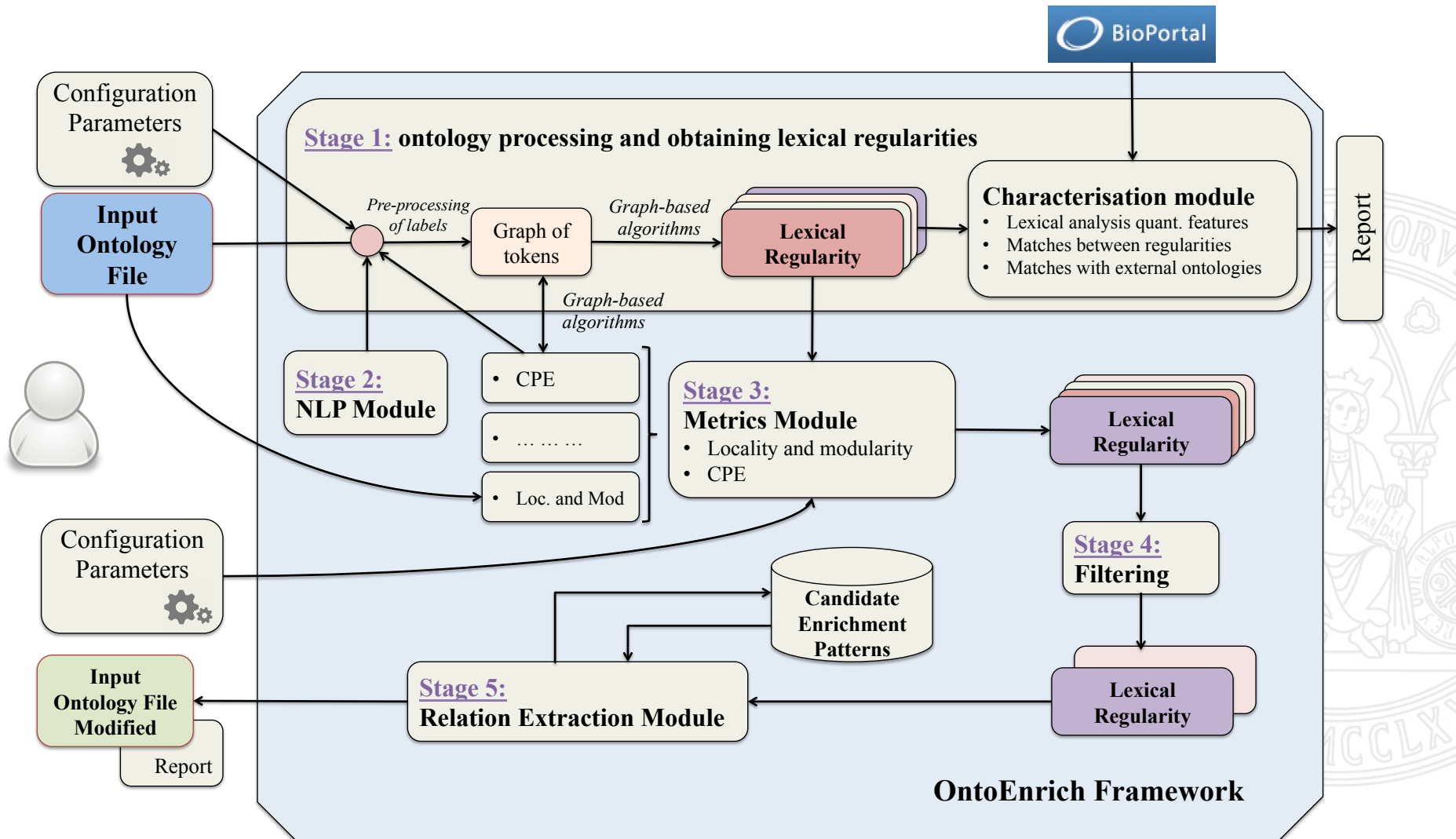
Workflows in OntoEnrich
Based on previous experiments and recommendations

An **OntoEnrich workflow** combines different types of filters, metrics and steps to support the user in the inspection of LRs, and in deciding how interesting they are

- E.g. **WF1** - *Lexical Regularity as common sub summer*:

    – Calculate the lexical analysis.

    – Filter 1: remove sub-regularities.

    – Filter 2: remove LRs that are not classes.

    – Metric 1: apply the LCS metric.

    – Exploration of LRs guided by the previous metric.

http://sele.inf.um.es/ontoenrich-beta/documentation/videosWorkflow/Workflow1_Alta.mp4

Workflows in OntoEnrich

Based on previous experiments and recommendations

- E.g. **WF2** – *POS tagging and lexical suggest logically define metric*

  - Calculate the lexical analysis.

  - Filter 1: select LRs that are or contains adjectives.

  - Manual inspection of LRs:

    - Metric 1: modularity.

    - Asserted vs inferred model.

    - Metric 2: lexical suggest logically define metric.

  - Calculate Metric 2 for the whole set of LRs and sort the set by it.

  - Exploration of LRs guided by the previous metric.
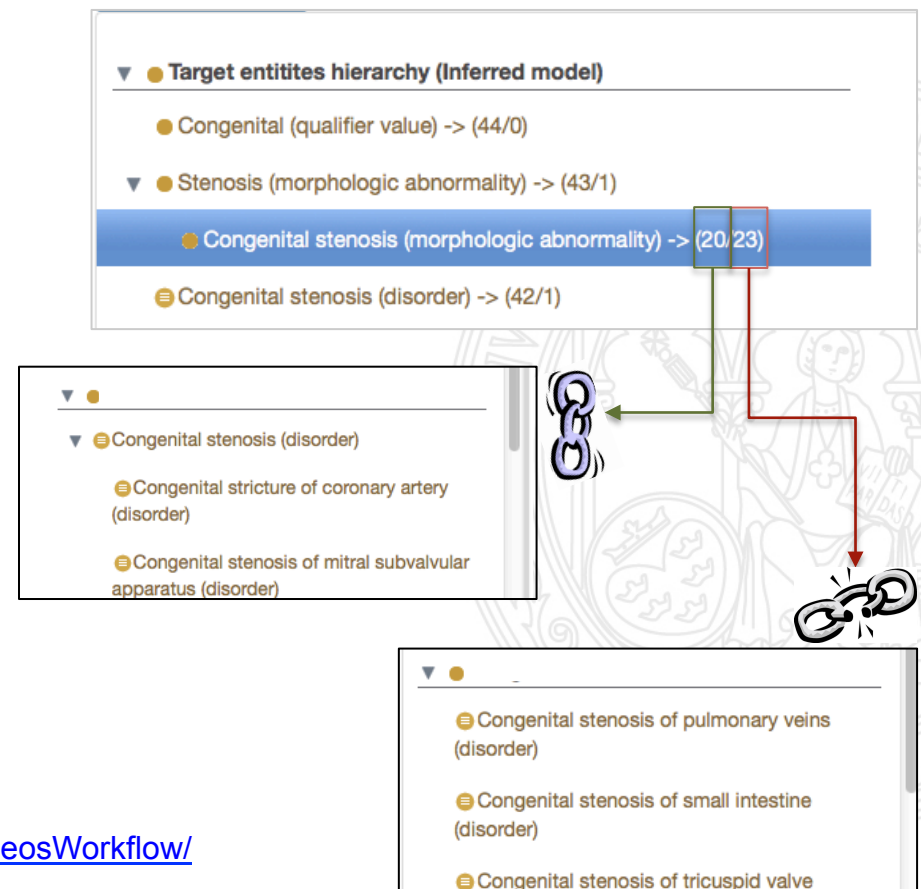
  http://sele.inf.um.es/ontoenrich-beta/documentation/videosWorkflow/Workflow3_Alta.mp4

See you in the demo session

Try online with your own ontology at:

http://sele.inf.um.es/ontoenrich

Manuel Quesada Martínez ✉ **manuel.quesada@um.es**

Thank you for your attention