# Trendy Practices and Tools in Ontological Engineering

**María Poveda Villalón**
**Ontology Engineering Group**
**Universidad Politécnica de Madrid, Spain**

mpoveda@fi.upm.es
@MariaPovedaV

7th September 2018
Zaragoza

CC BY NC SA

# About OEG

Directors: Asunción Gómez-Pérez, Oscar Corcho

Position: 8º ranking UPM (200 groups)

Research group (30 people)

- 3 Full Professors
- 5 Associate Professors
- 3 Assistant Professors
- 7 Senior Postdocs
- 6 PhD Students
- 2 MSc and BSc Students
- 2 software engineers
- 1 system administrator
- 2 project managers

170+ Collaborations

50+ Visitors

http://www.oeg-upm.net/

https://github.com/oeg-upm

@oeg-upm

# Ontology Engineering Group at a glance

- Created in1995
- World-wide known in the research areas
  - Ontologies
  - Semantic Web and Linked Data
  - Multilingual linked Data
  - Open Data
  - eScience
- **Projects (> 12M€)**
  - 27 EU projects (7 as coordinator)
  - 54 National Projects
  - 27 contracts with companies
- **Publications**
  - > 100 journal papers
  - > 400 International conferences and book chapters
  - 7 Books
- **Impact of publications**
  - Asunción Gómez-Pérez (h:58, 20,000 citations)
  - Oscar Corcho García (h: 44, 11,000 citations)
- **Services to the community**
  - **Host esDbpedia**
  - **Host linkeddata.es**
  - **Ontology development services**

- Awards and Prizes
  - Ada Byron
  - Aritmel
  - Juan López de Peñalver
  - Fujitsu, Open data, ISWC, ESWC
  - SUR Awards Watson for Tech. Watch
- Supervision of students
  - 28 Ph.D thesis (9 awarded best thesis prize)
  - >150 MS.C thesis and BS.C
- Events organization
  - 11 editions of the International Summer School on Ontological Engineering and the Semantic Web
  - > 50 WS and tutorials
- Standardization activities
  - >25 @ W3C, ISO, OASIS, etc.
- Mobility
  - PhD students: 3-6 months abroad
  - Postdocs: 1 month every 2 years
- Visibility
  - Program chairs of ESWC, ISWC, KCAP, EKAW, TKE, TIA
  - Editorial board of Journals
  - Invited talks at conferences and events
  - Programme Committee presence

# Participation in standardization activities

## W3C- World Wide Web Consortium

- Web of Things Working Group
- Web of Things Interest Group
- LBD Community Group
- BPMLOD: Best Practices for Multilingual Linked Open Data Community Group
- CSV on the Web Working Group
- eGovernment Interest Group
- Library Linked Data Incubator Group
- LD4LT: Linked Data for Language Technologies Community Group
- LDP: Linked Data Platform Working Group
- Media Annotations Working Group
- ODRL Community Group
- Ontolex: Ontology Lexica Community Group
- prov: Provenance Working Group
- prov-xg: Provenance Incubator Group
- Research Object for Scholarly Communication Community Group
- RDB2RDF Working Group
- Semantic Sensor Networks Community Group
- SPARQL Working Group
- SDW: Spatial Data on the Web Working Group
- RDF Stream Processing Community Group
- Open Linked Education Community Group
- Bioschemas for Lifesciences Community Group

## OASIS

- OSLC: Open Services for Lifecycle Collaboration

## ETSI

- SmartM2M
- ISG CIM

## AENOR

- Comité Técnico de Normalización CTN178 Ciudades Inteligentes. Norma UNE 178301

## Open Knowledge Foundation

- Working Group on Open Data in Linguistics

## ISO -International Organization for Standardization

- ISO/TC 19150 - Geographic Information – Ontology
- Ad-hoc Group Linked Data – ISO/TC211ISO/TC 37/SC 1 Principles and methods
- ISO/TC 37/SC 3 Systems to manage terminology, knowledge and content
- ISO/TC 37/SC 4 Language resource management

## Dublin Core Metadata Initiative (CMI)

- DCMI Metadata Provenance Task Group
- DCMI Vocabulary Management Community
- DCMI Bibliographic Metadata Task Group

## Europeana Network

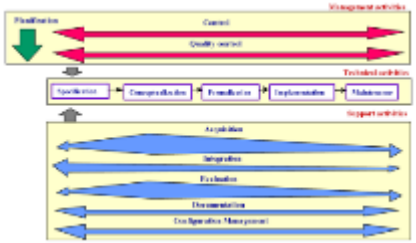
- Technology Developer and Knowledge Partner / Expert

## International Federation of Library Associations (IFLA)

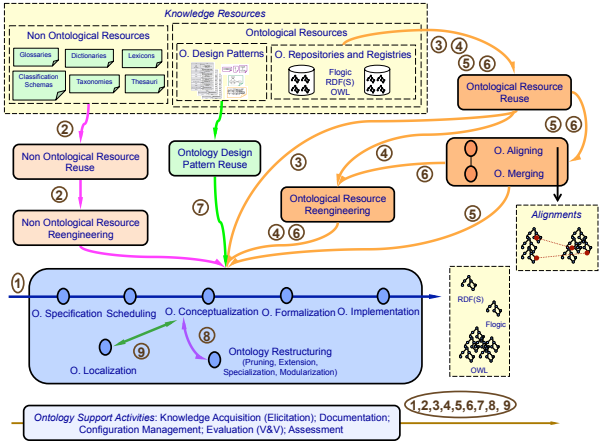
- Semantic Web Special Interest Group

- METHONTOLOGY and NeOn methodologies for building Networks of Ontologies
  - World-wide used

METHONTOLOGY (1997)



NeOn (2010)




## Ontologies

- Metadata Vocabularies (@W3C)
- Provenance (@W3C)
- Internet of Things (@W3C @ETSI)
- Geography (@IGN, CENIG)
- (Multi)-Media (@W3C)
- Knowledge Management
- Software Engineering (OSLC@OASIS)
- Tourism
- User experience
- AENOR vocab. for smart cities
- Libraries (@BNE)
- Health (@WHO)
- e-Science
- …

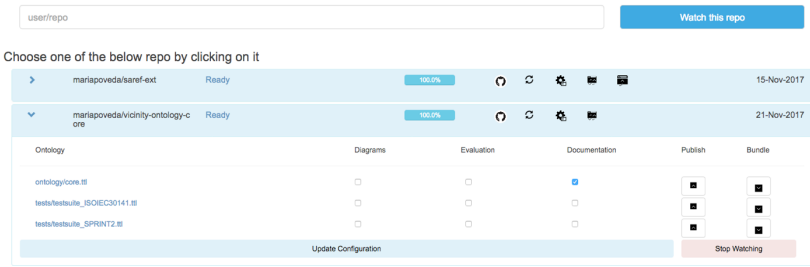
# Ontology engineering tools

- Ontology **evaluation**
  - OOPS! OntOlogy Pitfall Scanner http://oops.linkeddata.es/

- Vocabulary **documentation**
  - HTML: Widoco https://github.com/dgarijo/Widoco/
  - Diagrams: AR2DTool http://ar2dtool.linkeddata.es/

- Vocabulary **registry**
  - OEG vocabularies http://vocab.linkeddata.es/
  - Smart Cities http://smartcity.linkeddata.es/

- Vocabulary **distributed development**
  - OnToology http://ontoology.linkeddata.es/



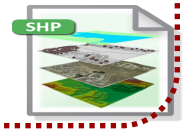
# Linked data Technologies and Projects

Morph

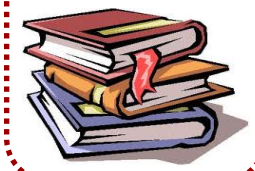
NOR$_2$O

Geometry2RDF

shp2RDF




Marimba

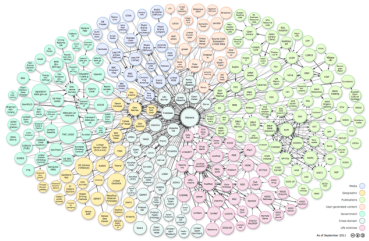


SPARQL -Stream

Sem4Tags

geo REST service annotation

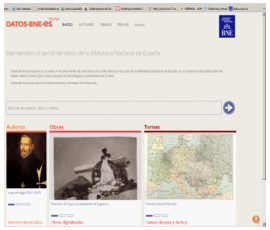
## RDF Generation and Linking



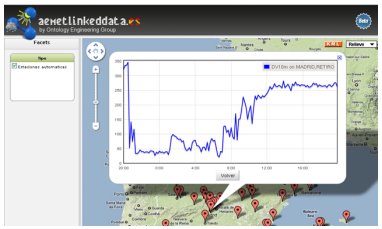
## Visualization

Map4RDF

Linked Library Data Visualisation



Sensor Data Visualisation




## Access

- Agora
- LDP4j
- Conditional Access
- Diagnose and repair

# Me at the OEG

# About me

**María Poveda-Villaón, PhD**
Ontological Engineering

mpoveda@fi.upm.es
@MariaPovedaV
mariapoveda
thepetiteontologist
MariaPovedaVillalon


Development | Evaluation | Conceptualization
Semantic Web | Linked Open Data | Publication

## Education

**2009**
Computer Science (Eng)

**2010**
MsC Artificial Intelligence

**2016**
PhD Artificial Intelligence

## Publications

- 5 Journal papers (4 indexed)
- 7 Conference papers
- 3 Book chapters
- 25 Workshop & demo papers
- 1 Journal editor
- 2 WS proc. editor

816 citations | H-index 15

## Ontologies

- IoT: SAREF, VICINITY
- Web Of Things
- Video games
- Scientific reviews
- Dataset profiling: agri
- Meteorology
- Patient safety

## Software

oops!
WOMBAT
Ontology

## Organization

- 4 Workshops
- 3 Tutorials, VoCamp
- 2 Open Data Day
- 19 PC member Conf & WS

## Skills

- W3C participation
- OWL, RDF, JENA, JAVA
- OpenRefine
- Public speaking

## Training

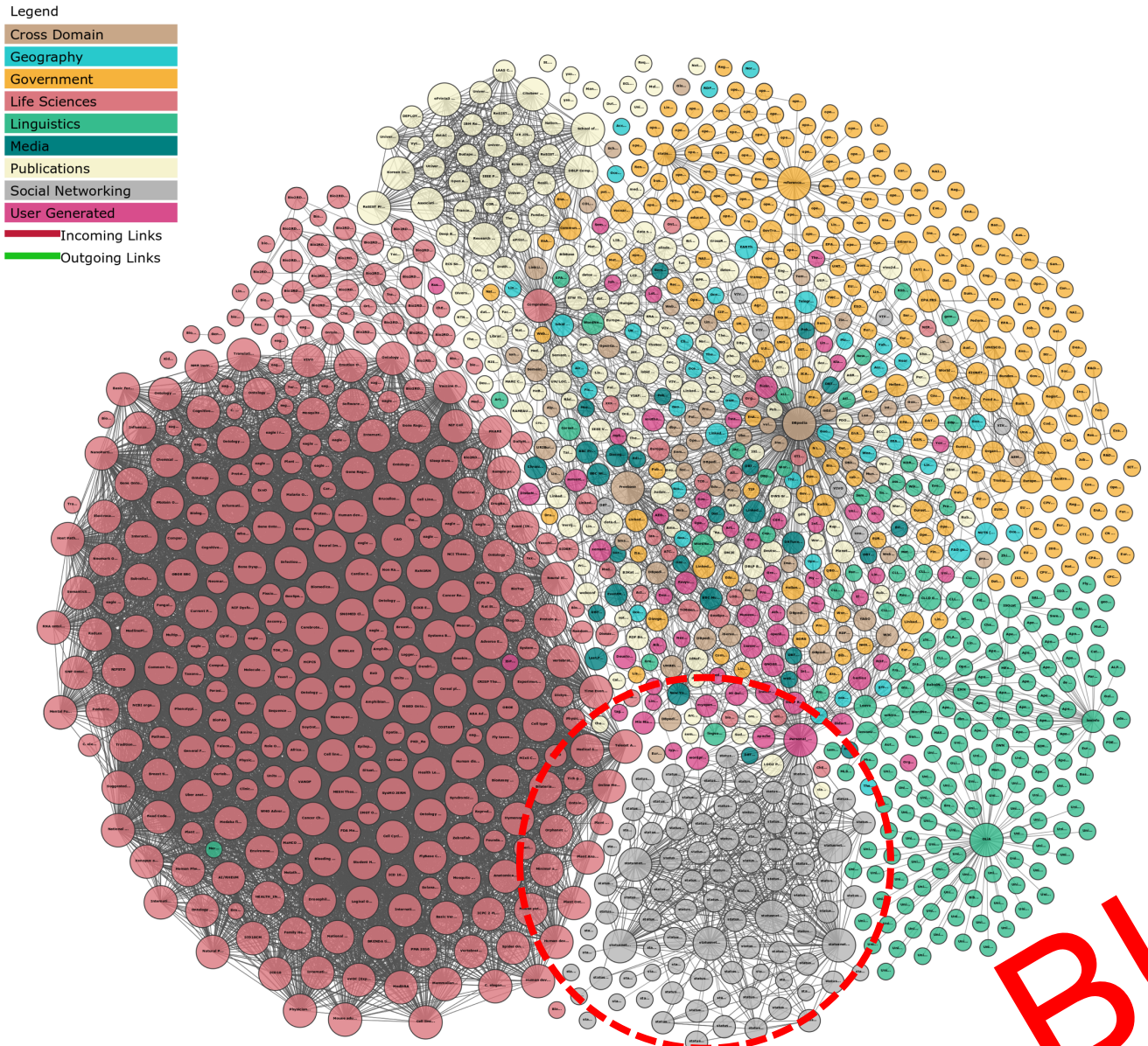
- MOOCs, SPOCs,
- Ad-hoc courses

## Projects

- 5 European
- 5 Spanish

# New trends in ontological engineering, practices and tools

# Linked Open Data cloud




BUT!

Image taken from https://lod-cloud.net/

# Breaking news!

Google Dataset Search Beta

Buscar conjuntos de datos

Probar boston education c

Follow researchinp

## Research in Progress

R.I.P.

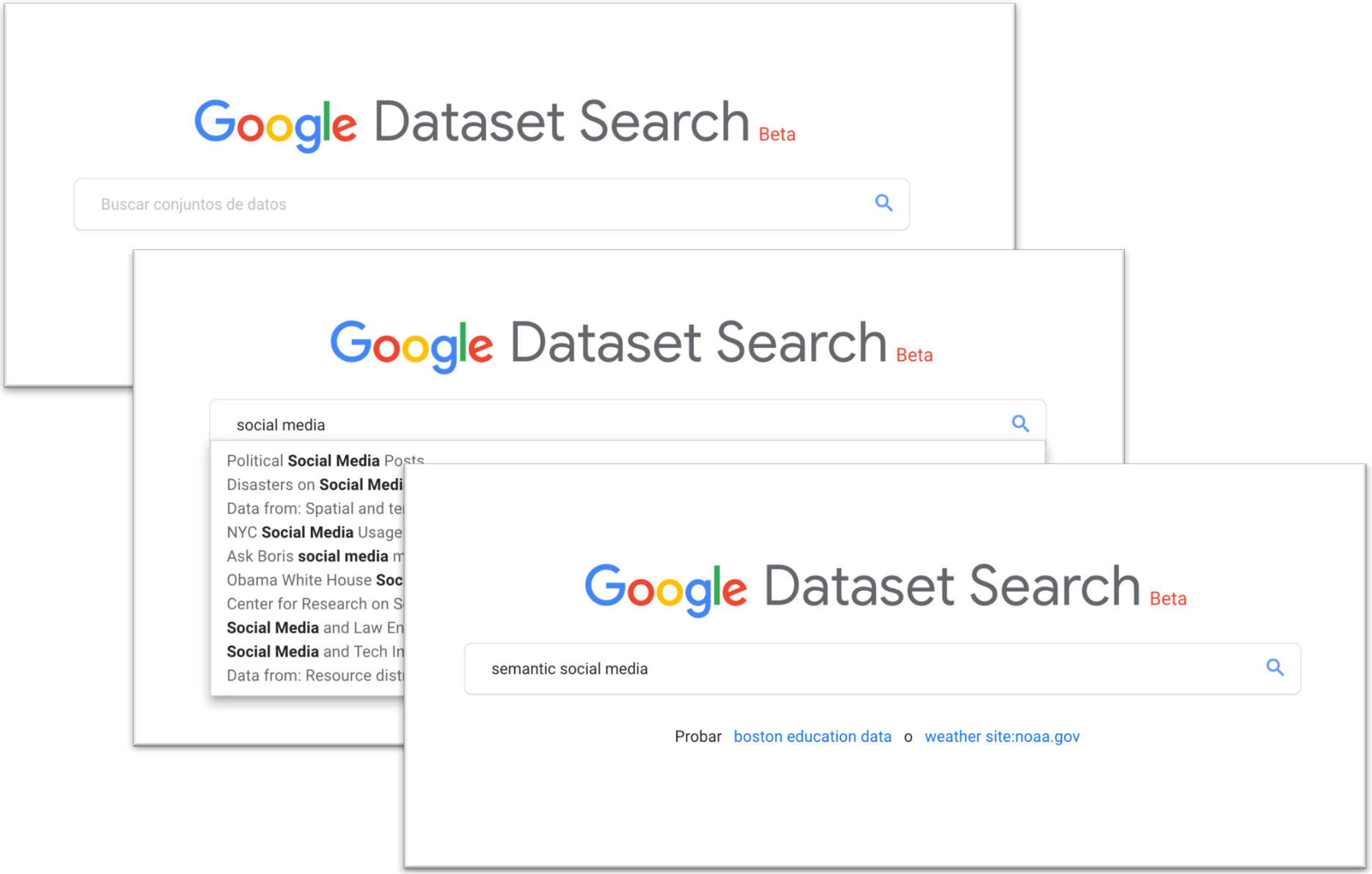
Things about the research in progress.

### Google: We are happy to announce our new web site on [your thesis topic]

by *Didac & Emin*

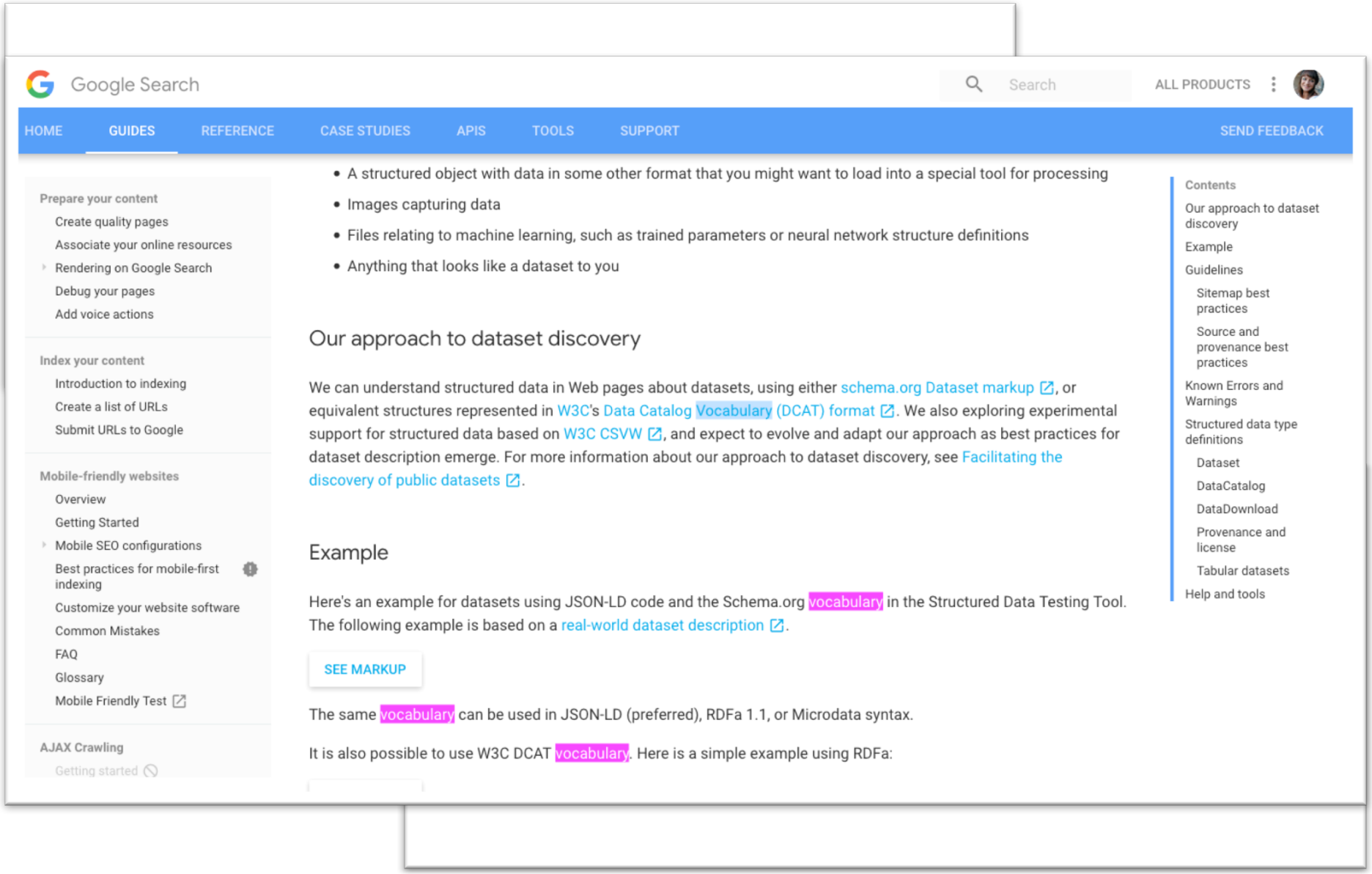
http://researchinprogress.tumblr.com/post/37821860151/google-we-are-happy-to-announce-our-new-web-site

# Breaking news!




https://toolbox.google.com/datasetsearch

# Breaking news!




https://toolbox.google.com/datasetsearch

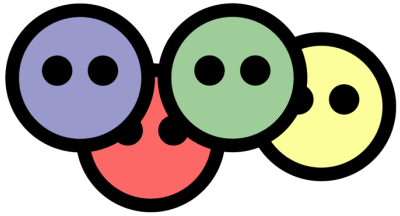
**Vocabularies** define the **concepts** and **relationships** used to **describe** and represent an area of concern.

Definition taken from: http://www.w3.org/standards/semanticweb/ontology

# We (might) all agree on

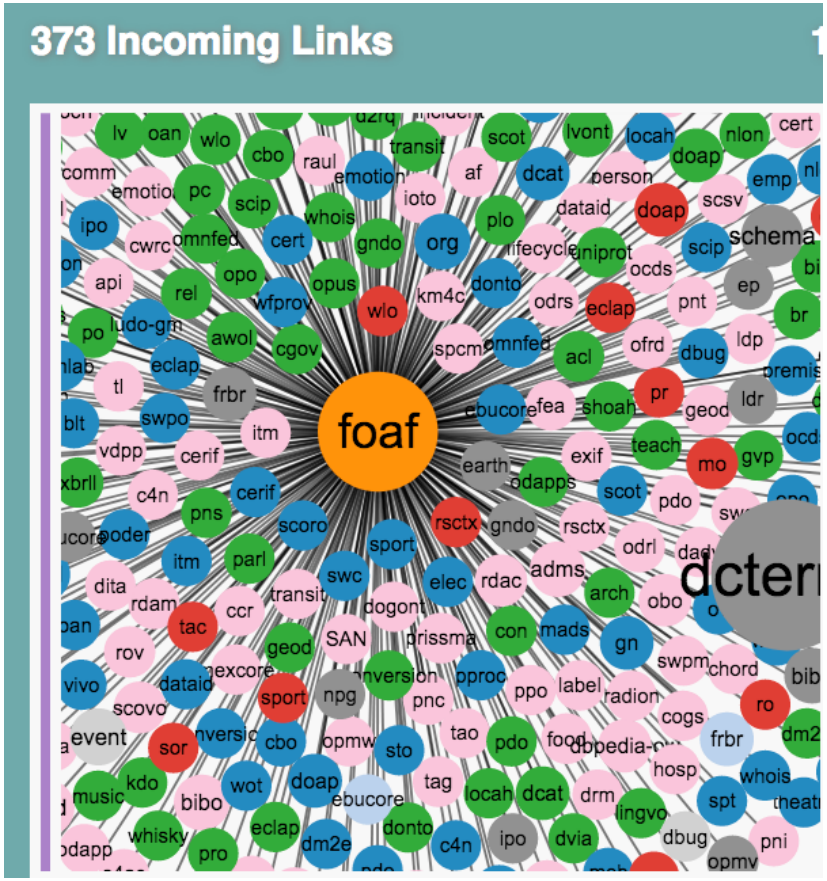
- Ontologies offer many benefits
  - Semantics
  - Interoperability
    - Languages
    - Granularity
    - Formats
    - Schemas
  - Disambiguation
  - Reasoning
  - **REUSE**

# And there are some famous ontologies in social media

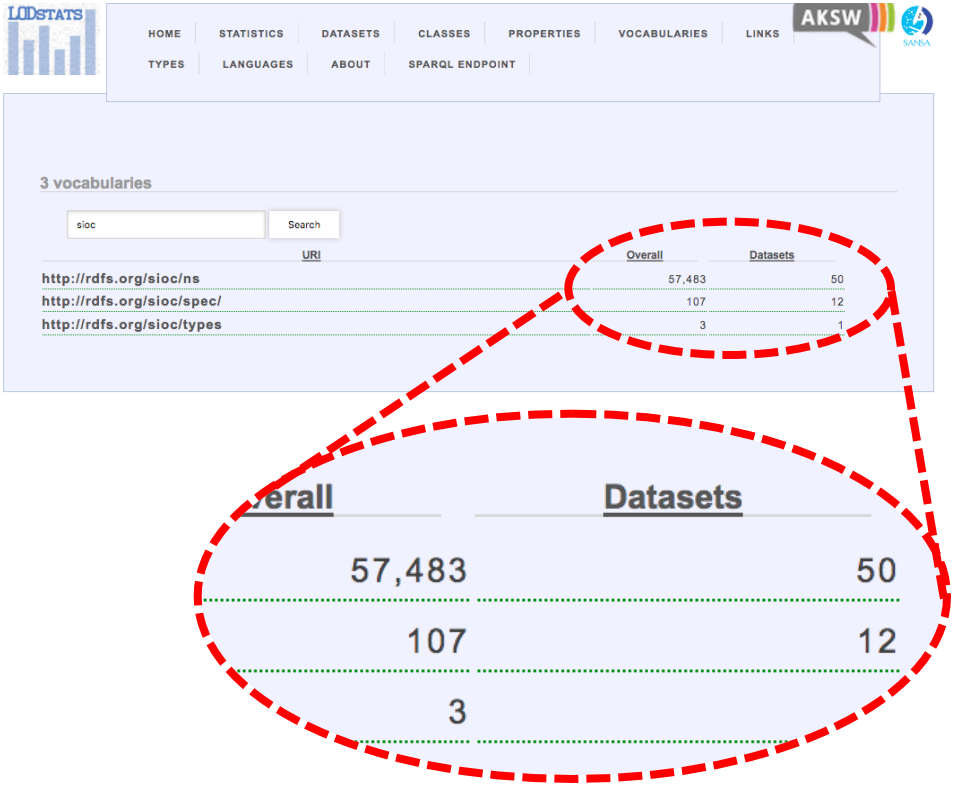


http://xmlns.com/foaf/0.1/

http://rdfs.org/sioc/ns#

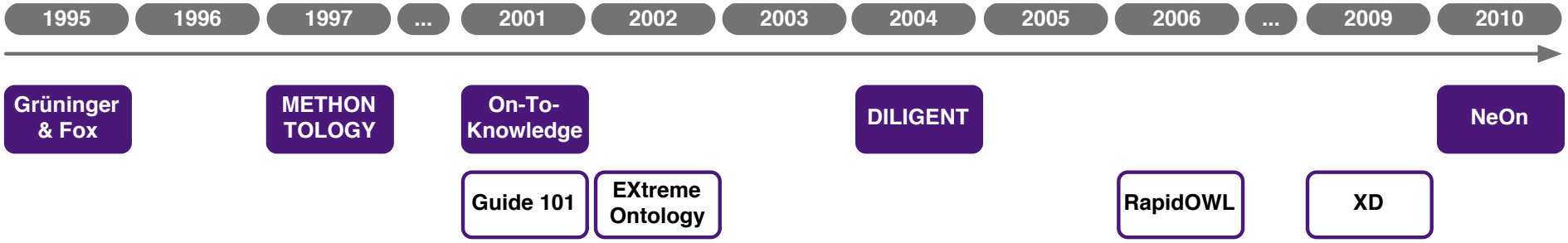



https://lov.linkeddata.es/dataset/lov/vocabs/foaf




http://lodstats.aksw.org/vocabularies?search=sioc

# Ontology Development Methodologies

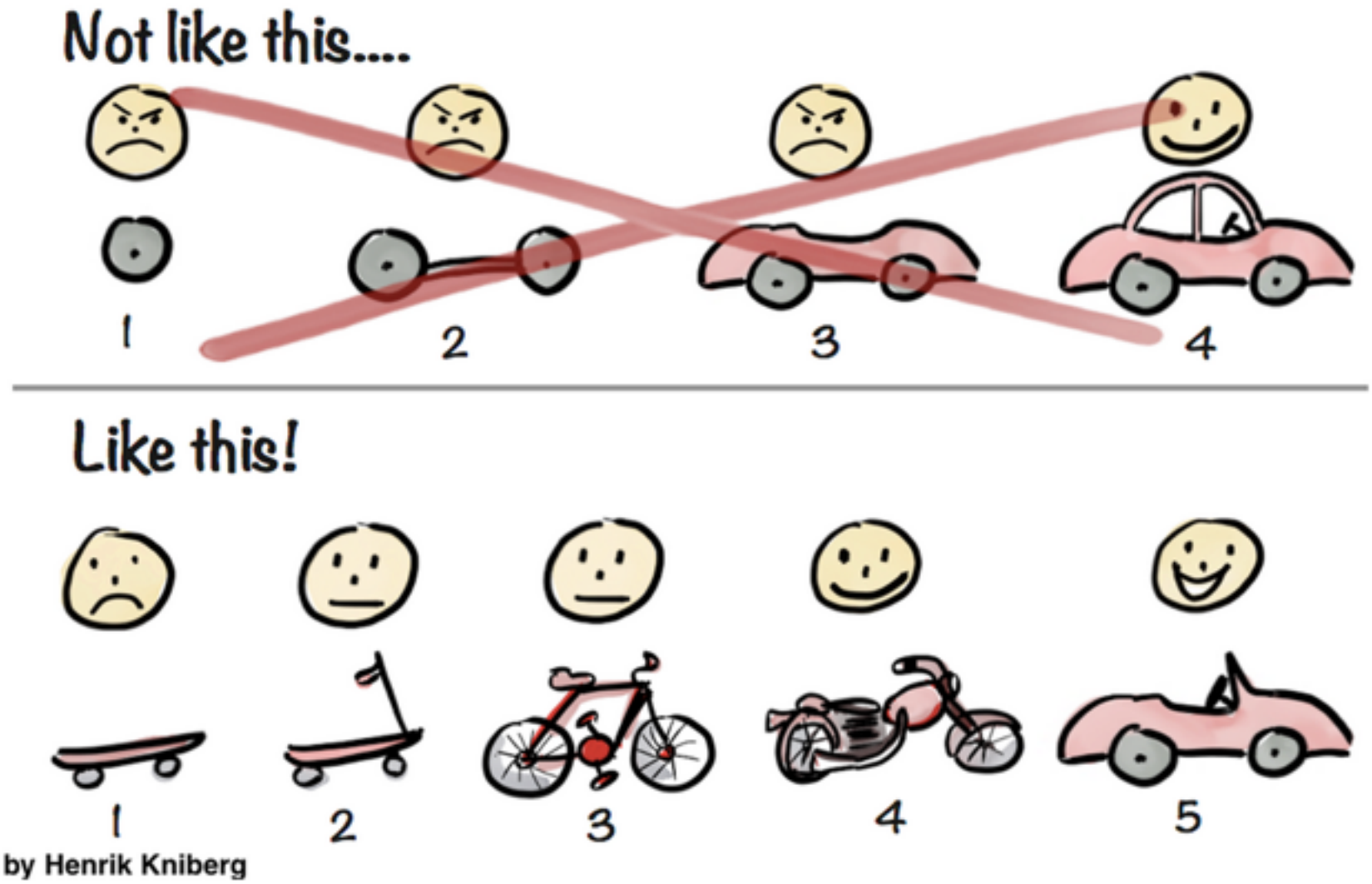



- Towards lightweight and agile processes
- Inspiration from software development practices
- Coupling Software and ontology development

# What is agile?

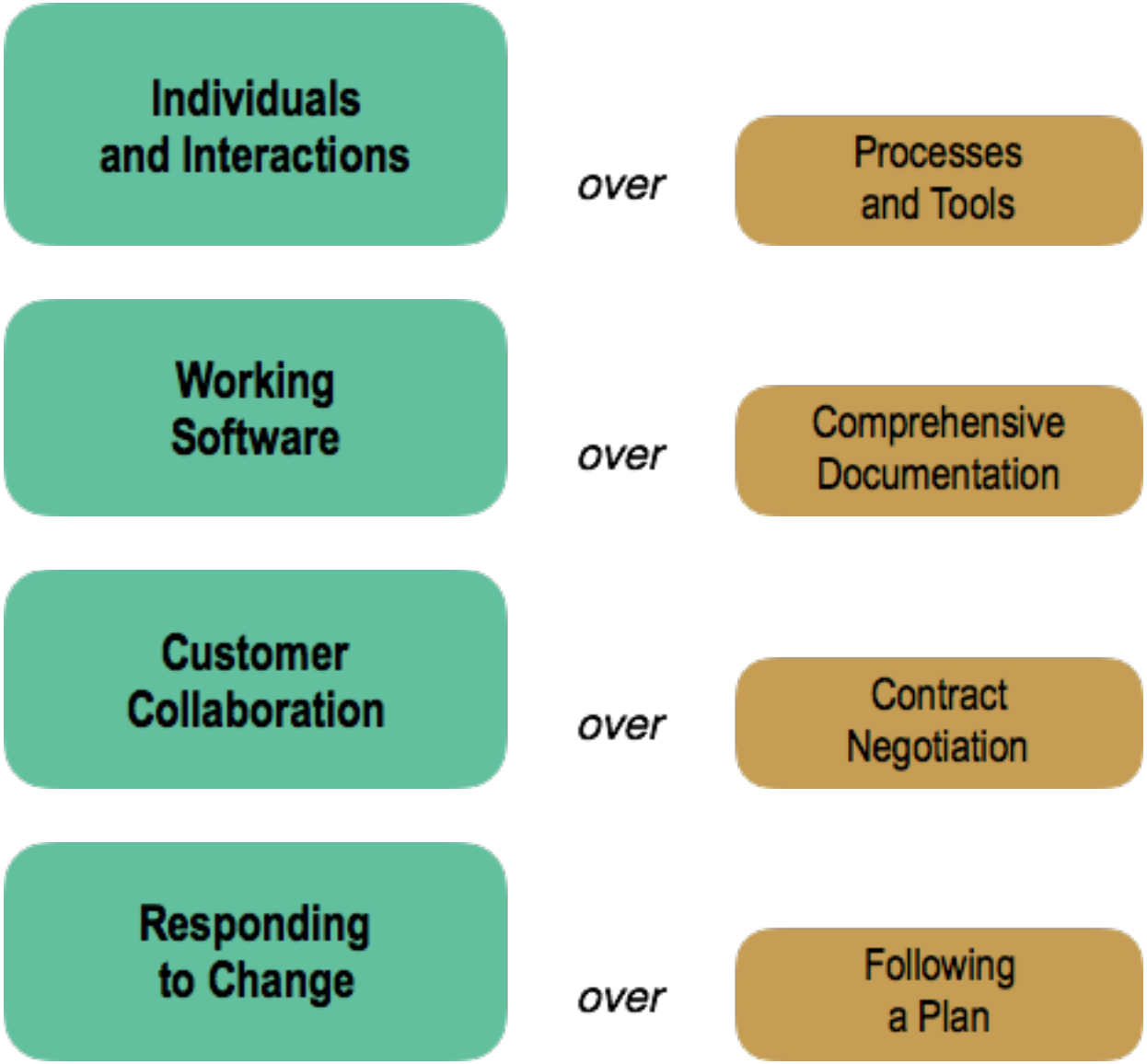
Slide from Carlos Badenes

# What is agile?




Slide from Carlos Badenes

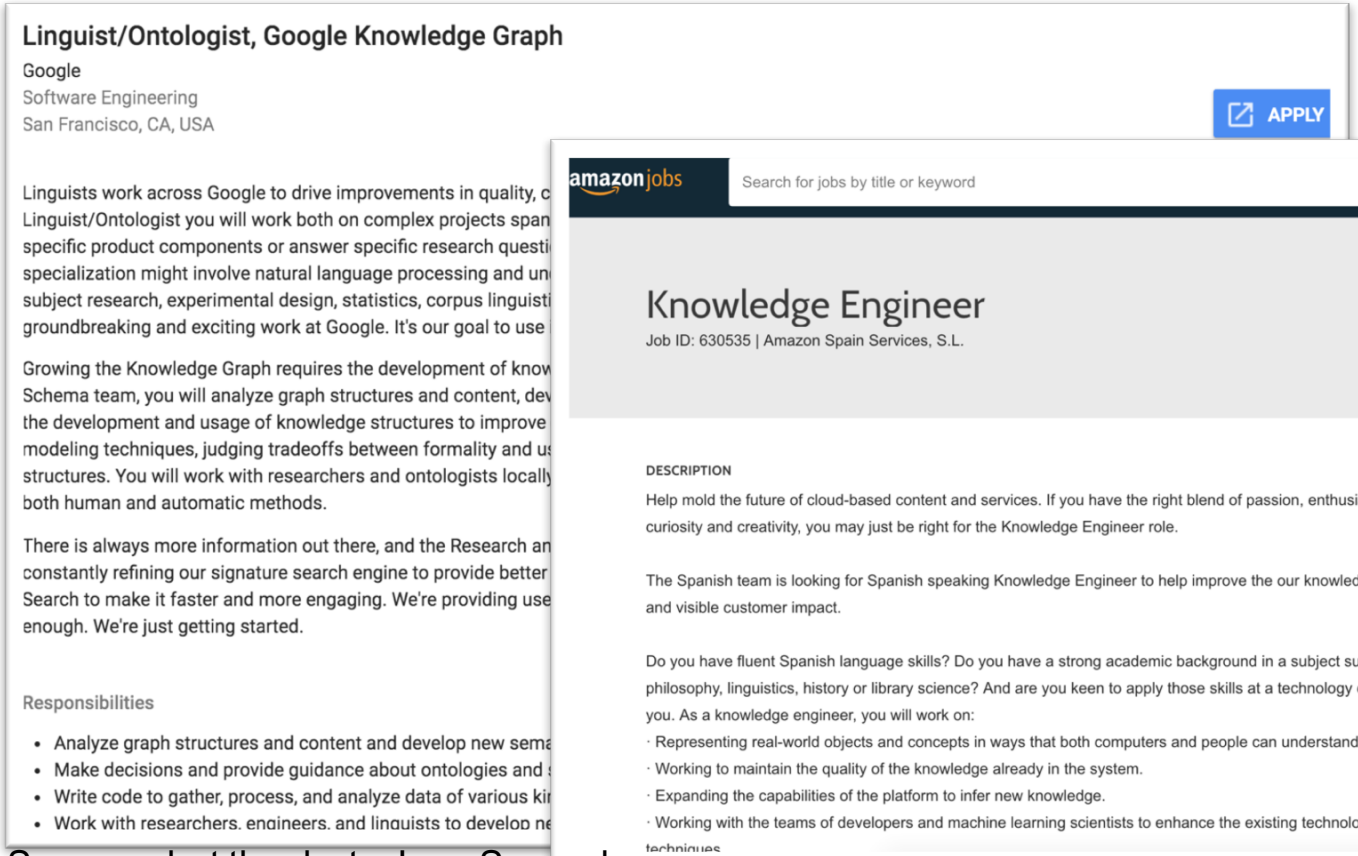
# Agile Manifesto




Slide from Carlos Badenes

# And how does it look like for ontological engineering?

- Ontologies are not often the final product. Do not get the customer attention. Like databases?
- Software development are including ontologies as part of the project



**Linguist/Ontologist, Google Knowledge Graph**

Google
Software Engineering
San Francisco, CA, USA

APPLY

Linguists work across Google to drive improvements in quality, c
Linguist/Ontologist you will work both on complex projects span
specific product components or answer specific research questi
specialization might involve natural language processing and un
subject research, experimental design, statistics, corpus linguisti
groundbreaking and exciting work at Google. It's our goal to use

Growing the Knowledge Graph requires the development of know
Schema team, you will analyze graph structures and content, dev
the development and usage of knowledge structures to improve
modeling techniques, judging tradeoffs between formality and us
structures. You will work with researchers and ontologists locally
both human and automatic methods.

There is always more information out there, and the Research an
constantly refining our signature search engine to provide better
Search to make it faster and more engaging. We're providing use
enough. We're just getting started.

Responsibilities

- Analyze graph structures and content and develop new sema
- Make decisions and provide guidance about ontologies and
- Write code to gather, process, and analyze data of various ki
- Work with researchers, engineers, and linguists to develop ne



amazon jobs | Search for jobs by title or keyword | Location

**Knowledge Engineer**

Job ID: 630535 | Amazon Spain Services, S.L.

Apply now

DESCRIPTION

Help mold the future of cloud-based content and services. If you have the right blend of passion, enthusiasm, customer obsession, curiosity and creativity, you may just be right for the Knowledge Engineer role.

The Spanish team is looking for Spanish speaking Knowledge Engineer to help improve the our knowledge base & reach with direct and visible customer impact.

Do you have fluent Spanish language skills? Do you have a strong academic background in a subject such as mathematics, philosophy, linguistics, history or library science? And are you keen to apply those skills at a technology company? This role is for you. As a knowledge engineer, you will work on:

· Representing real-world objects and concepts in ways that both computers and people can understand.
· Working to maintain the quality of the knowledge already in the system.
· Expanding the capabilities of the platform to infer new knowledge.
· Working with the teams of developers and machine learning scientists to enhance the existing technology and invent new techniques

Job details

Madrid, Spain

Editorial, Writing, & Content Management

Share this job

Related jobs

Digital Music Content Manager
ES, Madrid

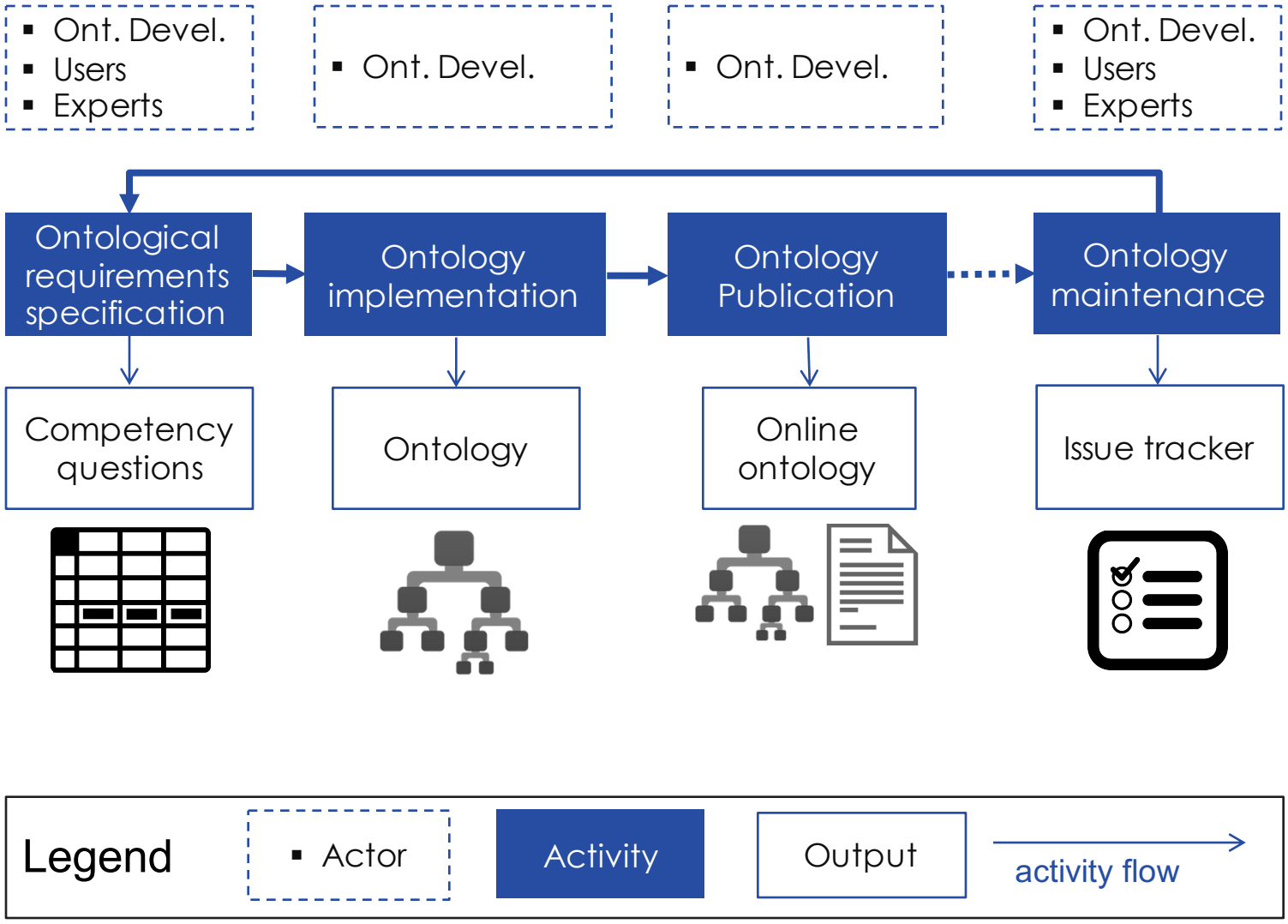
Screeenshot thanks to Juan Sequeda

# And how does it look like for ontological engineering?

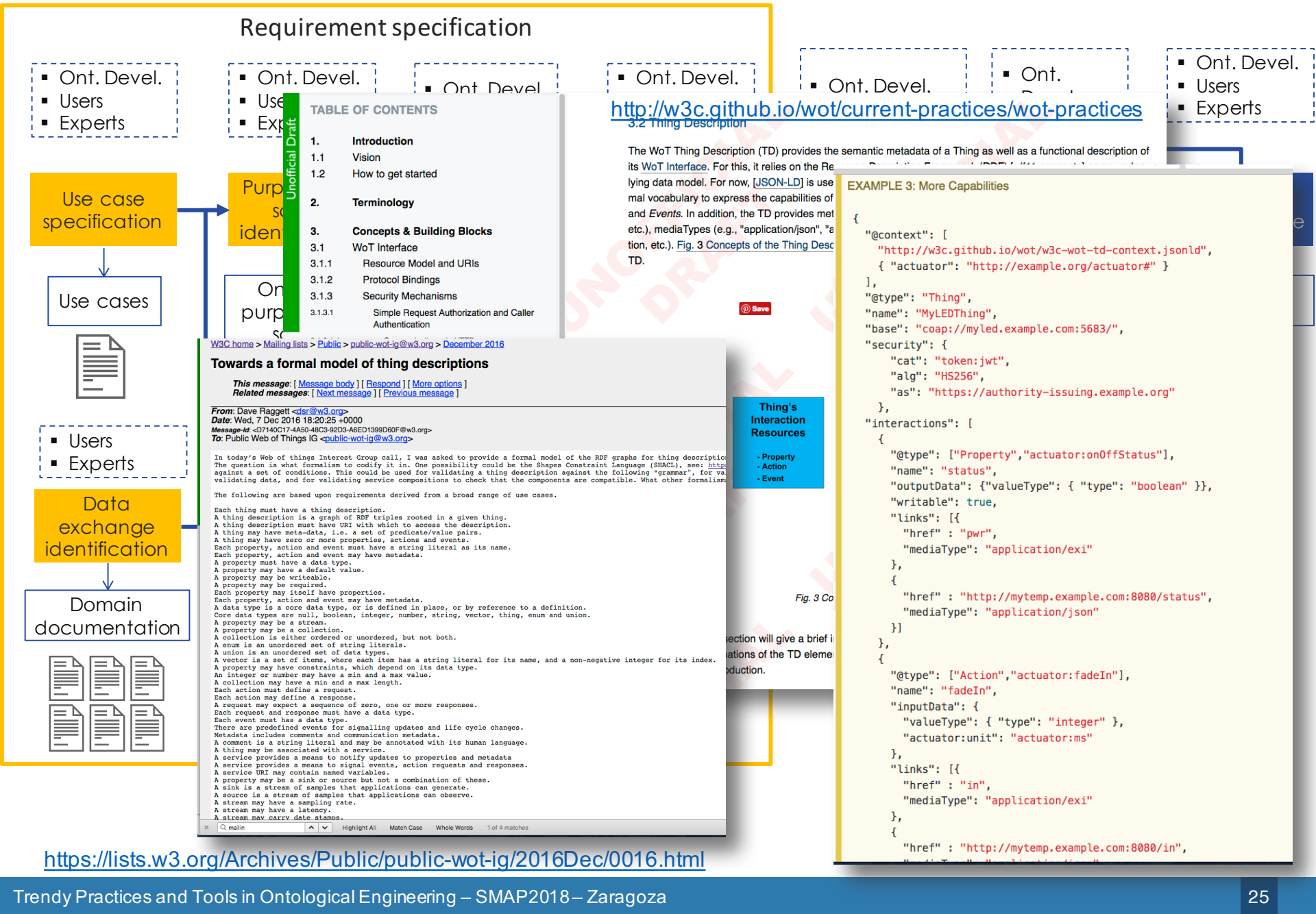
- Ontologies are not often the final product. Do not get the customer attention. Like databases?
- Software development are including ontologies as part of the project

- **Let's try to sneak our processes in SW practices!**

# Ontology development process overview

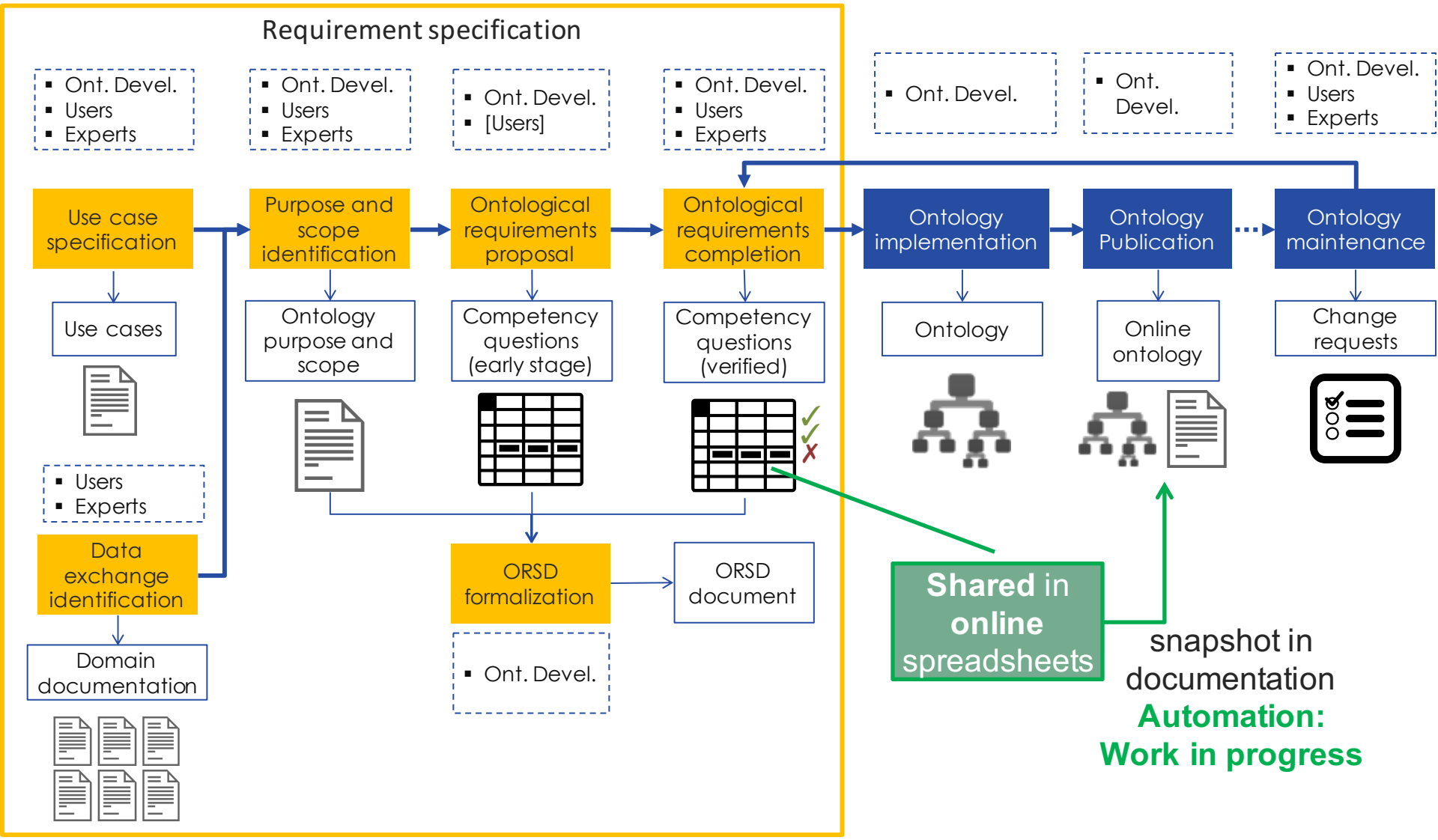


# Requirement specification

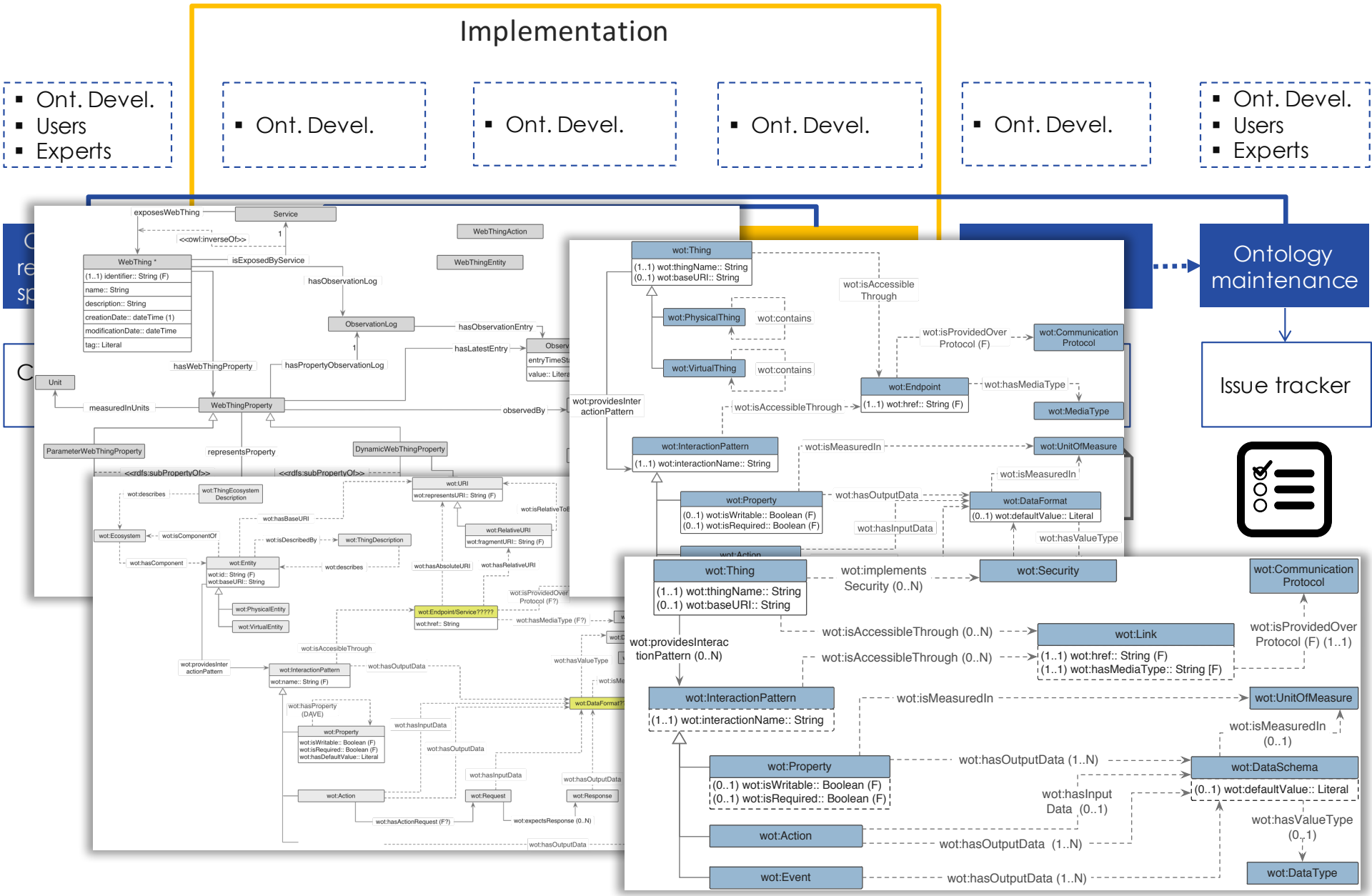



https://lists.w3.org/Archives/Public/public-wot-ig/2016Dec/0016.html

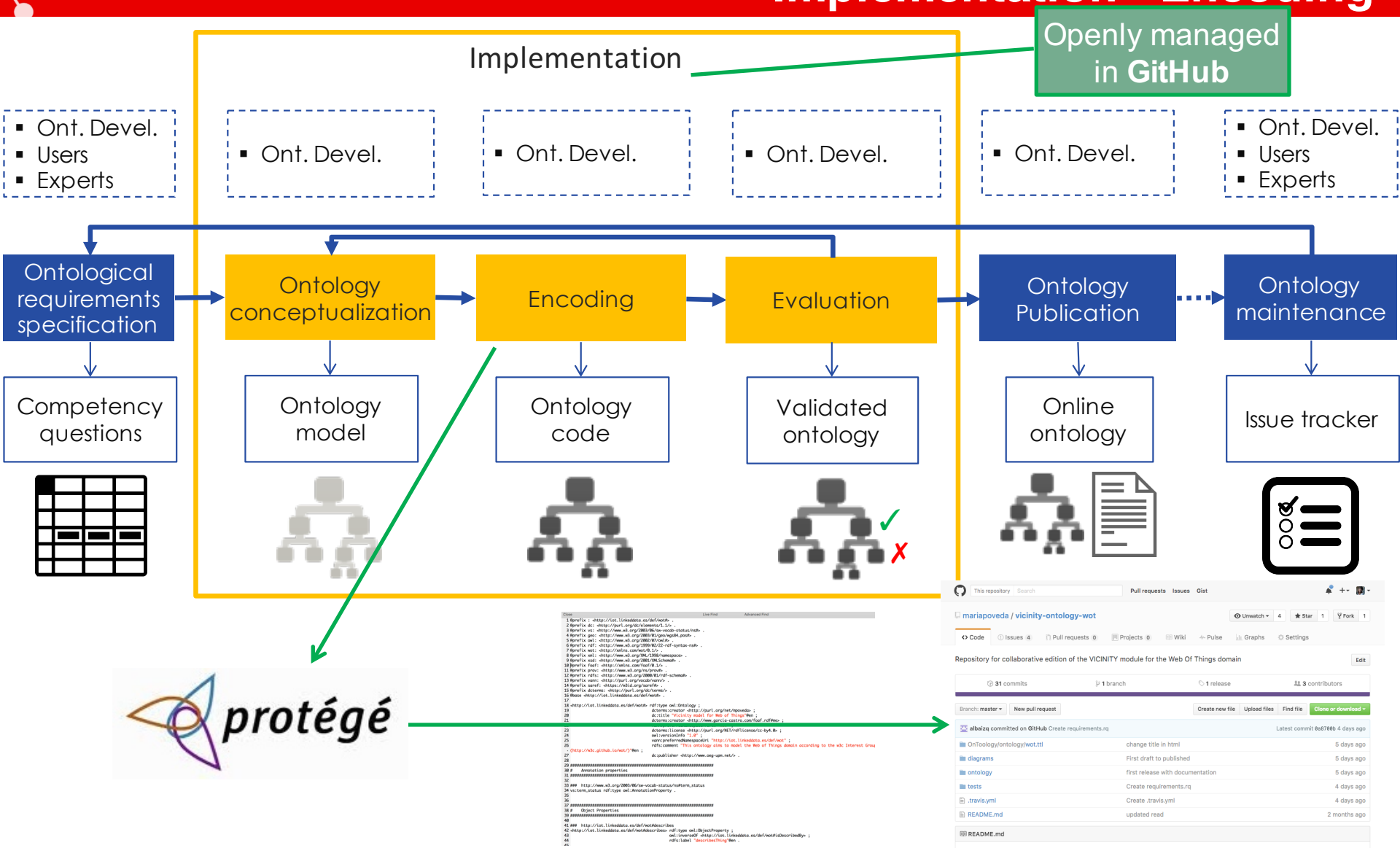
# Requirement specification



# Implementation - Conceptualization



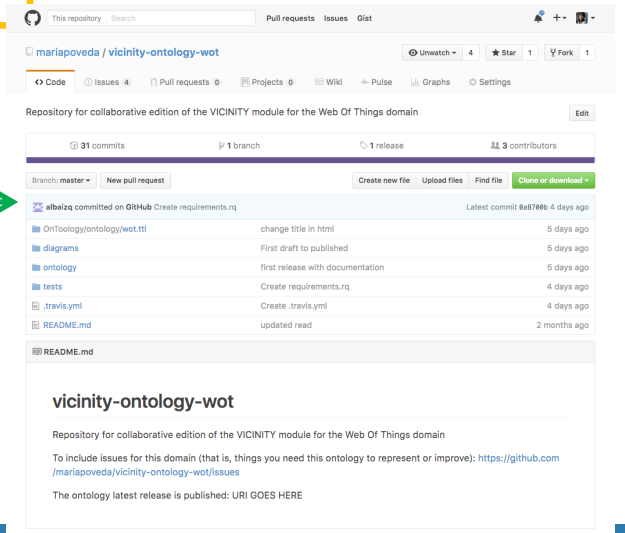
# Implementation - Encoding







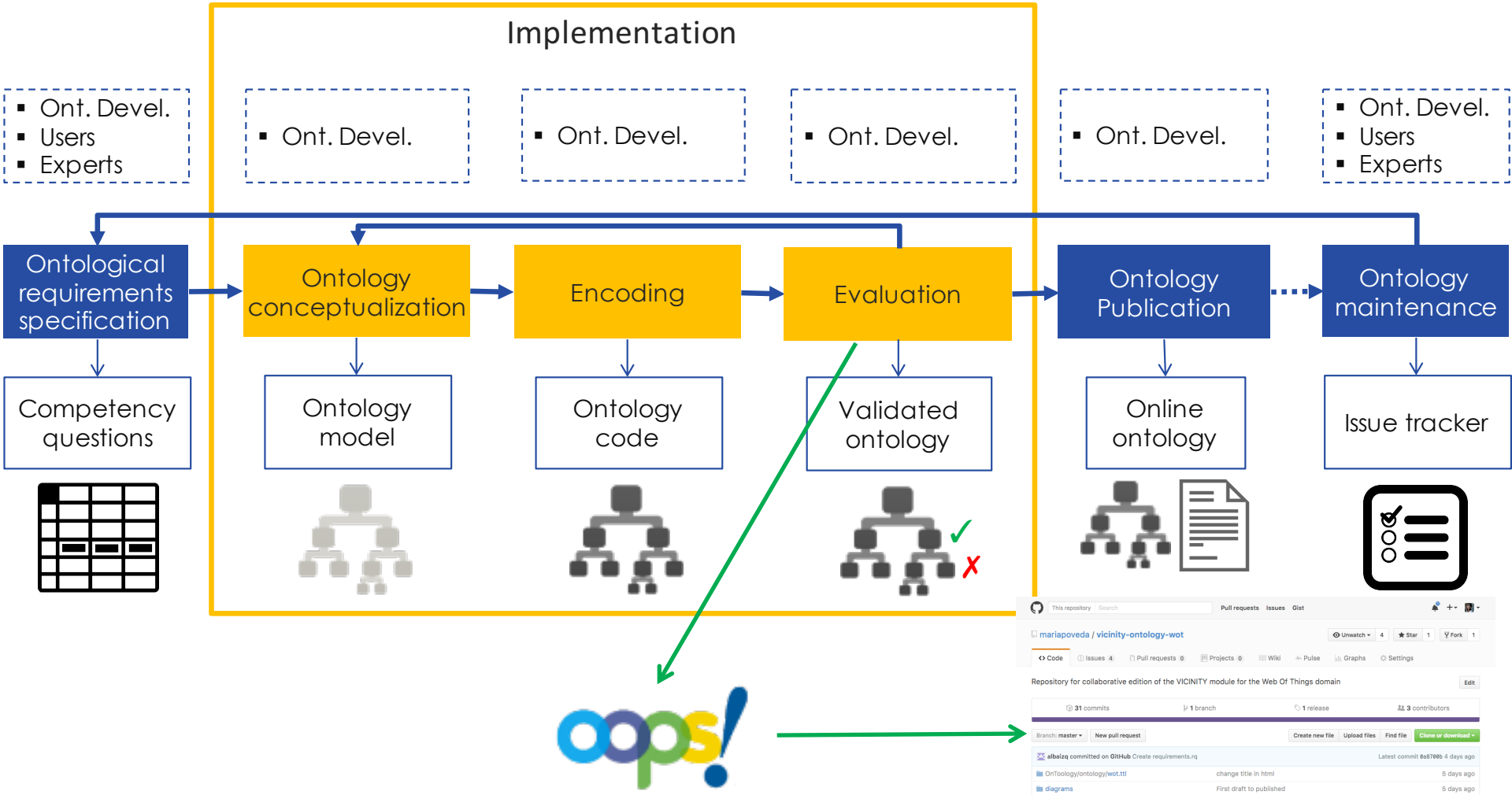


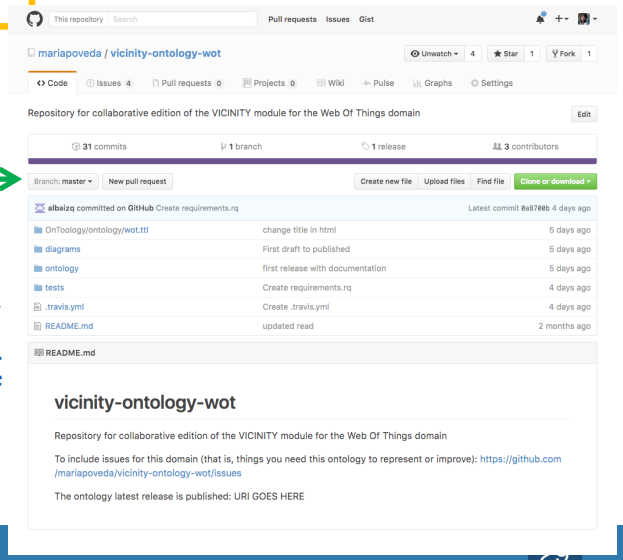



**GitHub repository**

https://github.com/mariapoveda/vicinity-ontology-wot

# Implementation - Evaluation



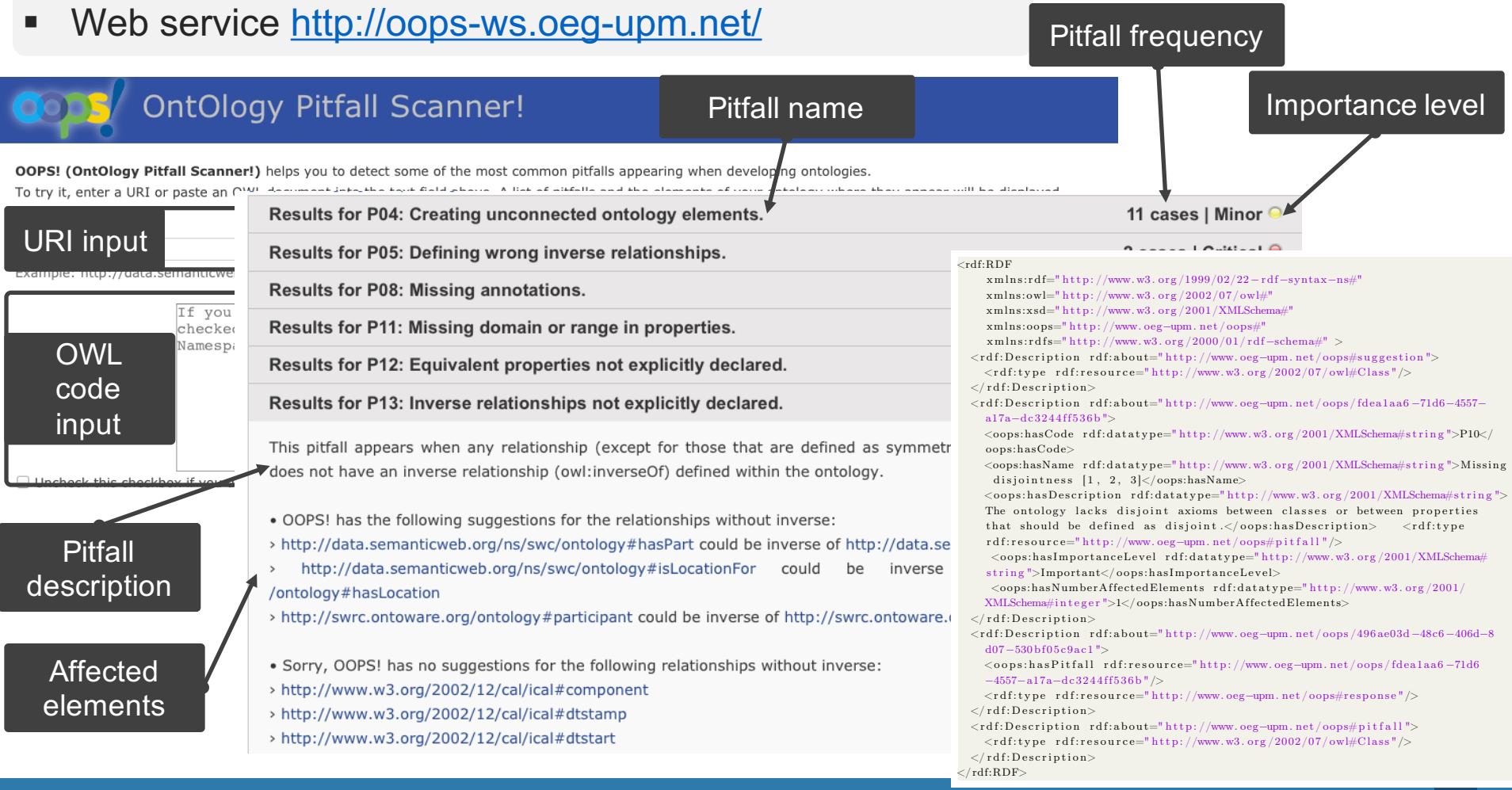




**Online** and **notifications** in **GitHub** repository
https://github.com/mariapoveda/vicinity-ontology-wot

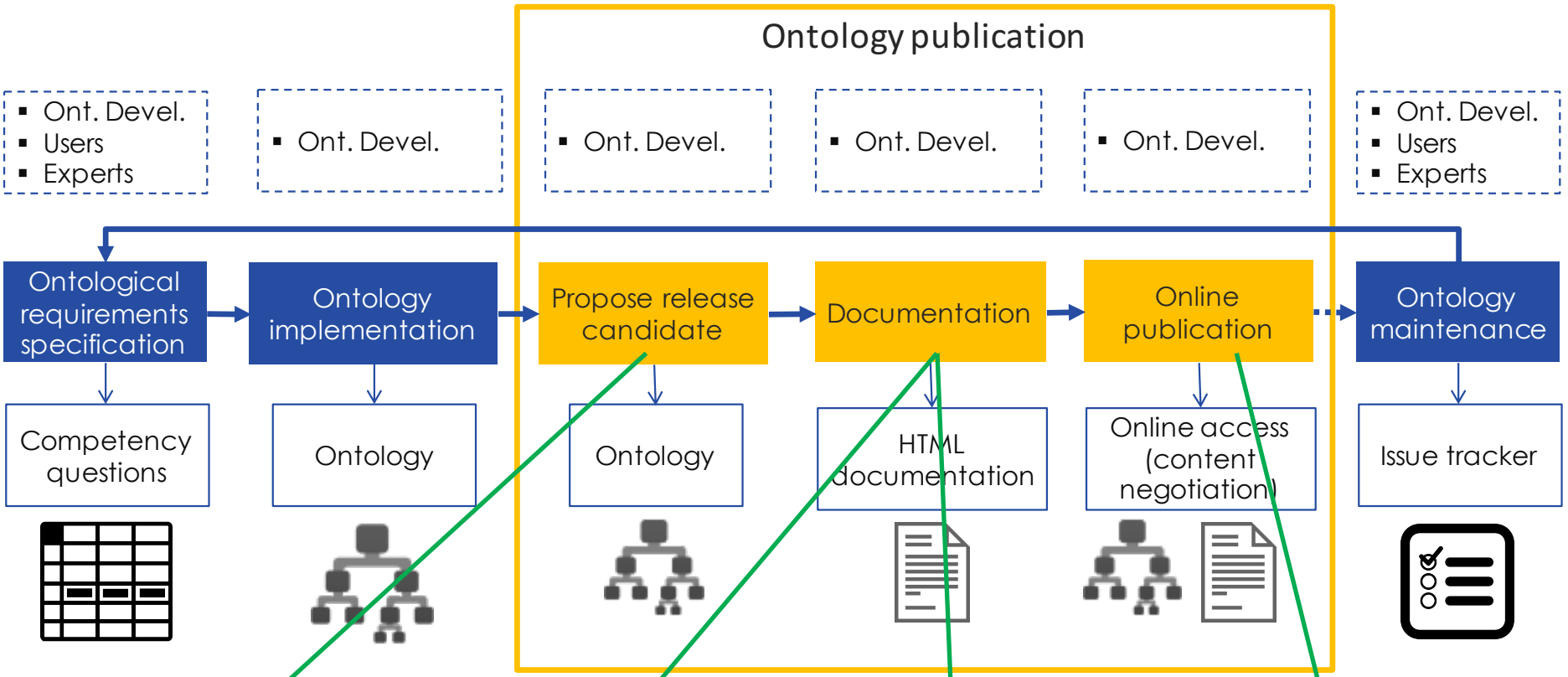
Ongoing work: tests from requirements

# Evaluation - OOPS! – OntOlogy Pitfall Scanner!

- Implements the **48** detection methods for **33 pitfalls**
  - Pitfalls selection
  - Selection by dimensions and aspects
- Web user interface http://oops.linkeddata.es/
- Web service http://oops-ws.oeg-upm.net/



# Publication



WIDOCO

- HTML generation from OWL code
- Multilingual
- Separated sections

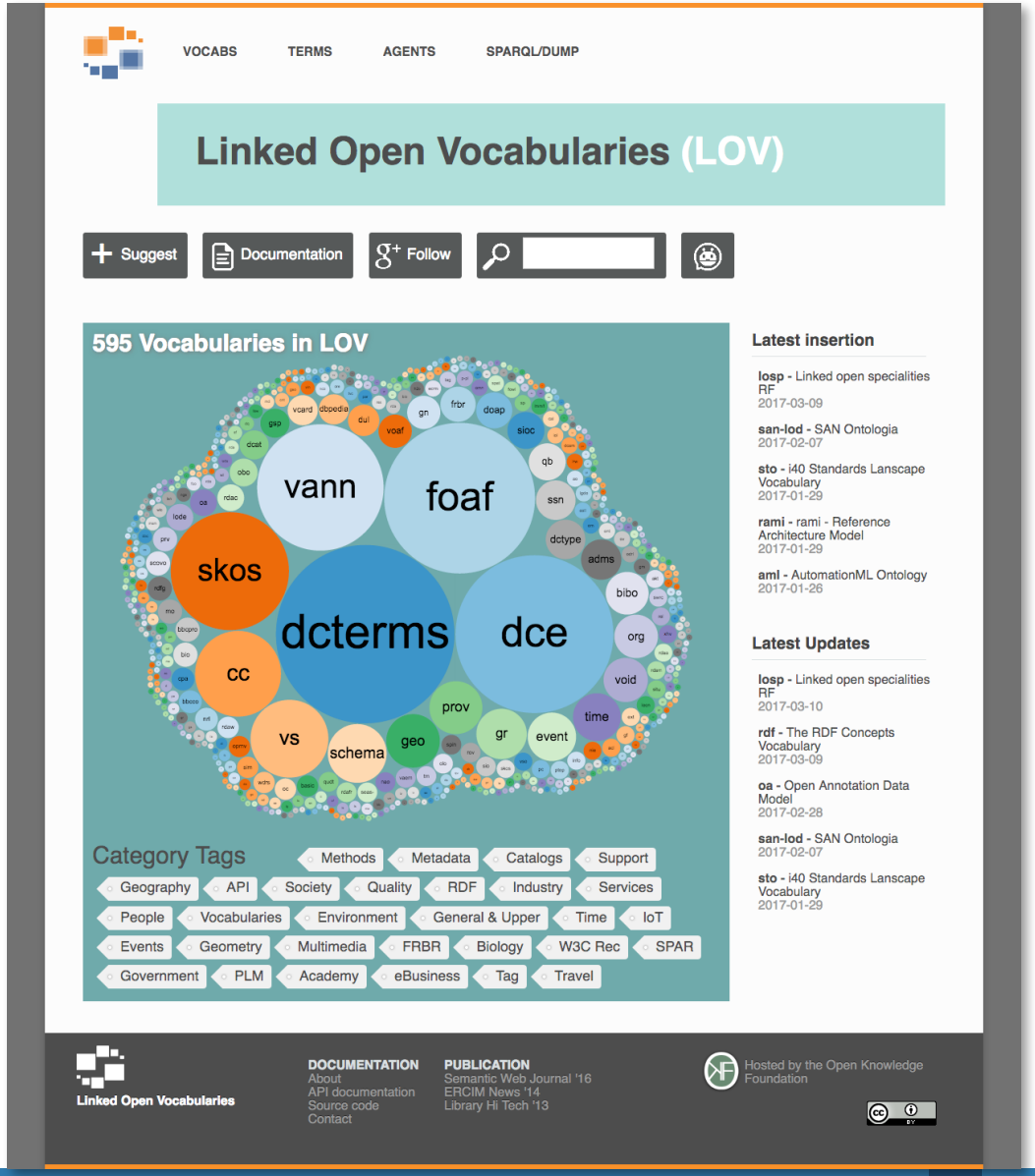
+ Diagrams (Some ideas: https://bit.ly/2GWhlaI)
+ Descriptions
+ Examples

- Own URI
- purl, w3id, etc.
- Content negotiation
- Registry

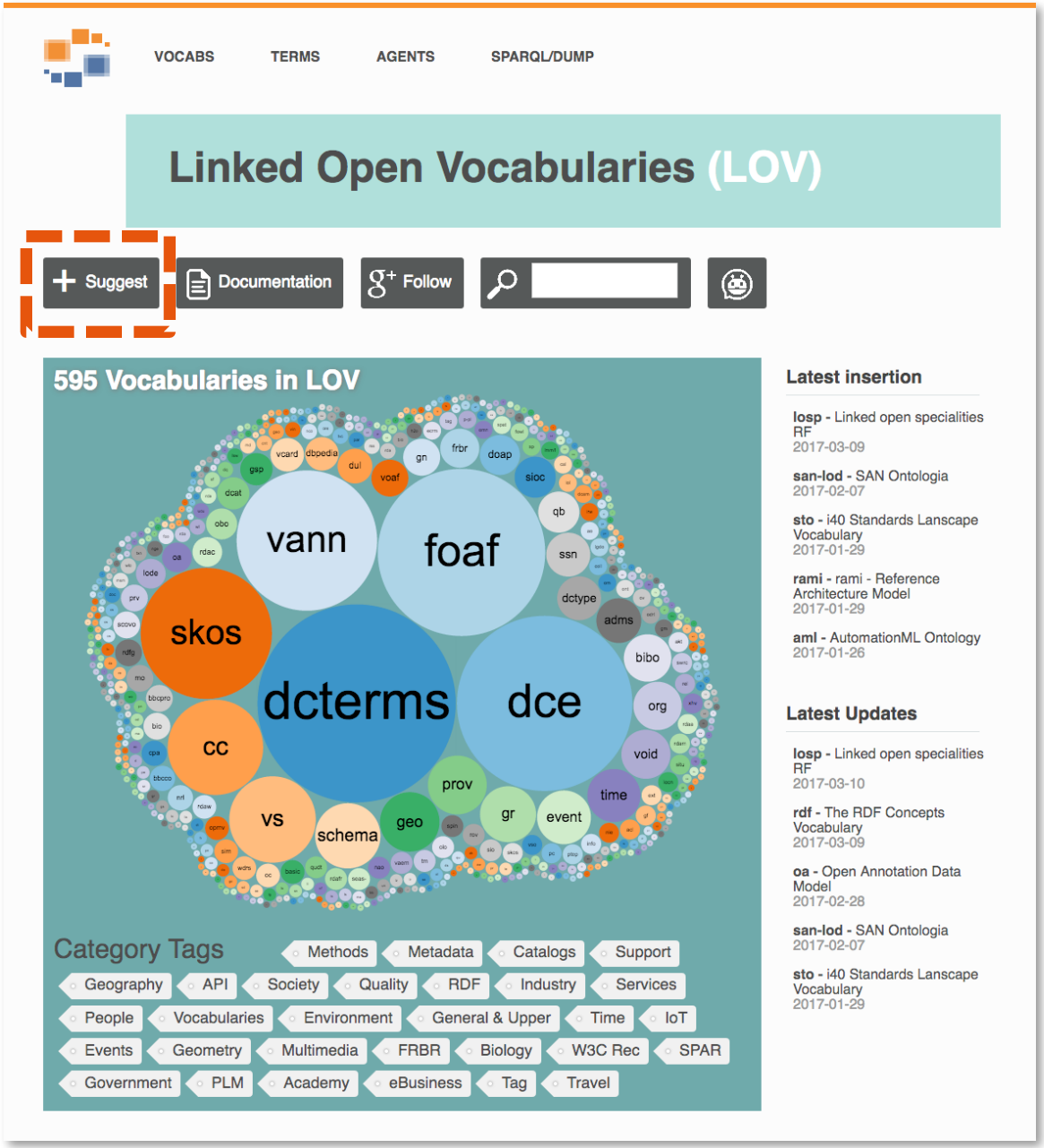
https://lov.linkeddata.es

- **Mission: promote and facilitate the reuse of well documented vocabularies in the Linked Data ecosystem**

- Vocabularies registry and index
- Datalift
  - http://datalift.org/
- Started at 2011
- Hosted by OEG



# Share your ontology



# Share your ontology

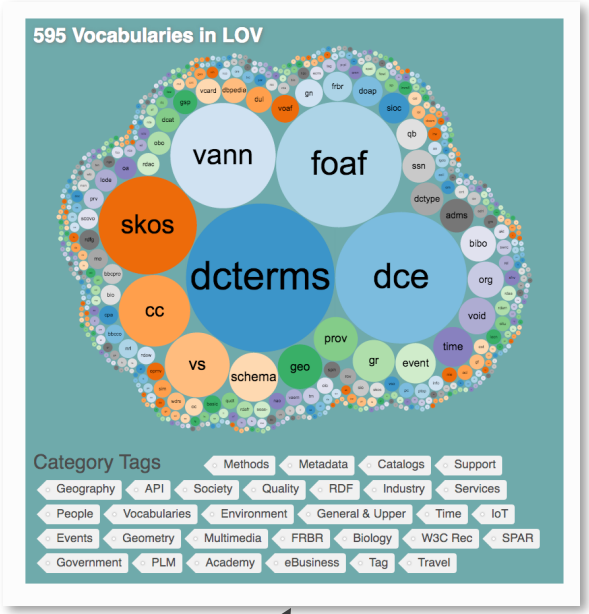
SUGGEST | Enter the vocabulary URI to recommend

Recommendations for vocabulary metadata description

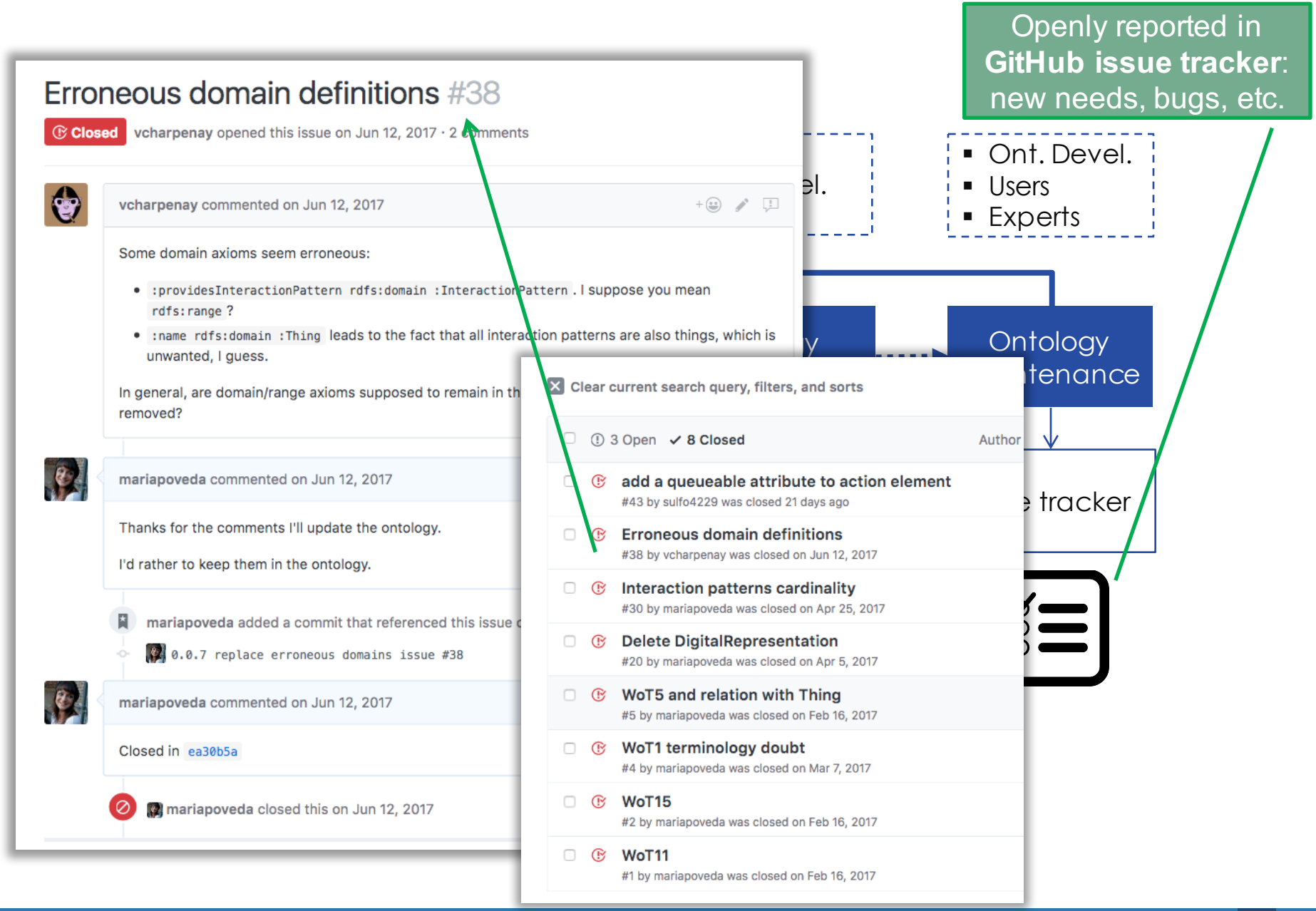
Authors

## LOV curation team

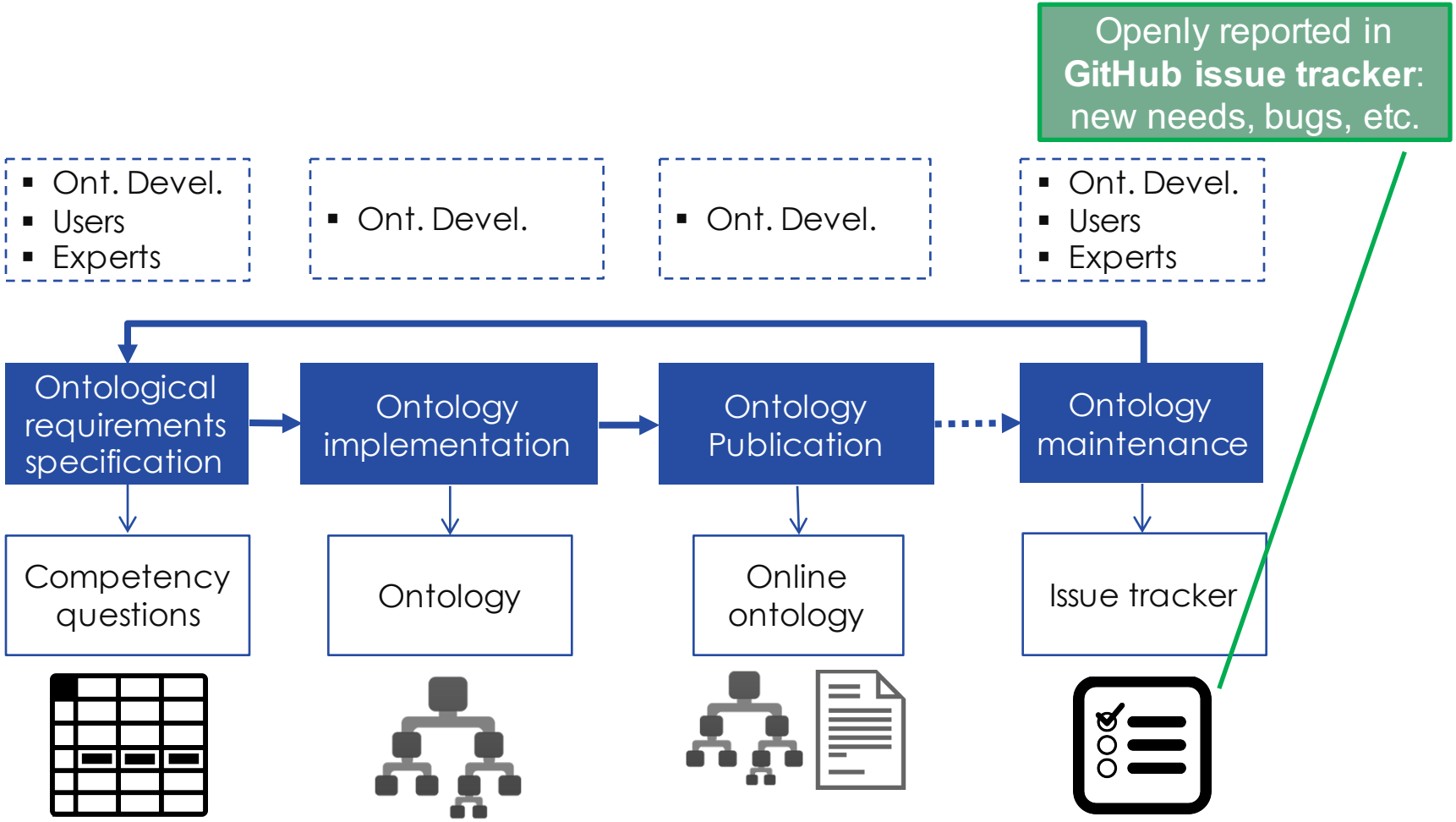
- **Metadata**: title, description, authors, dates, vann, version, etc.
- **Human readable**: labels and comments
- **Publication**
- **Licence**



# Ontology development process overview

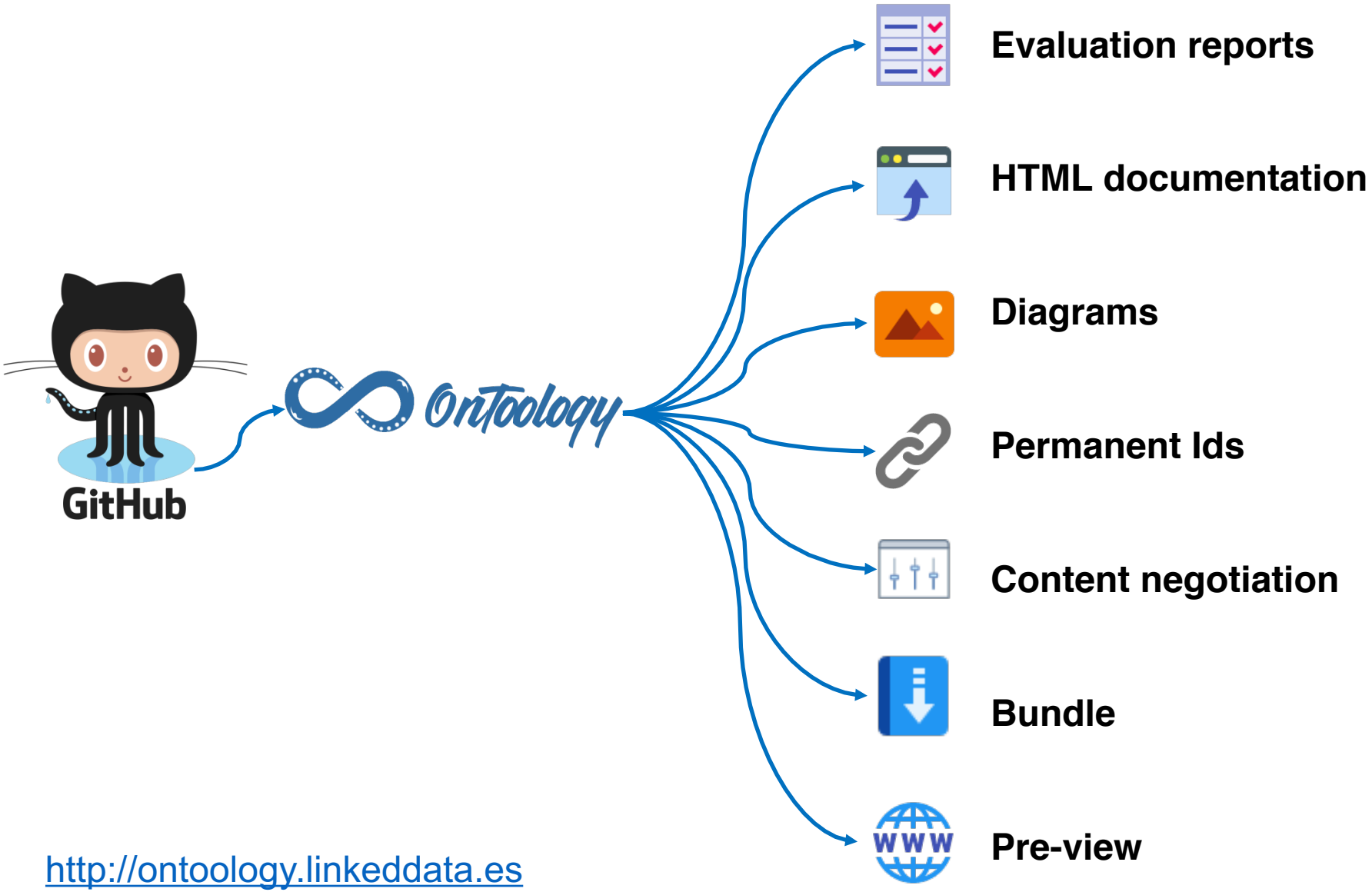


# Ontology development process overview



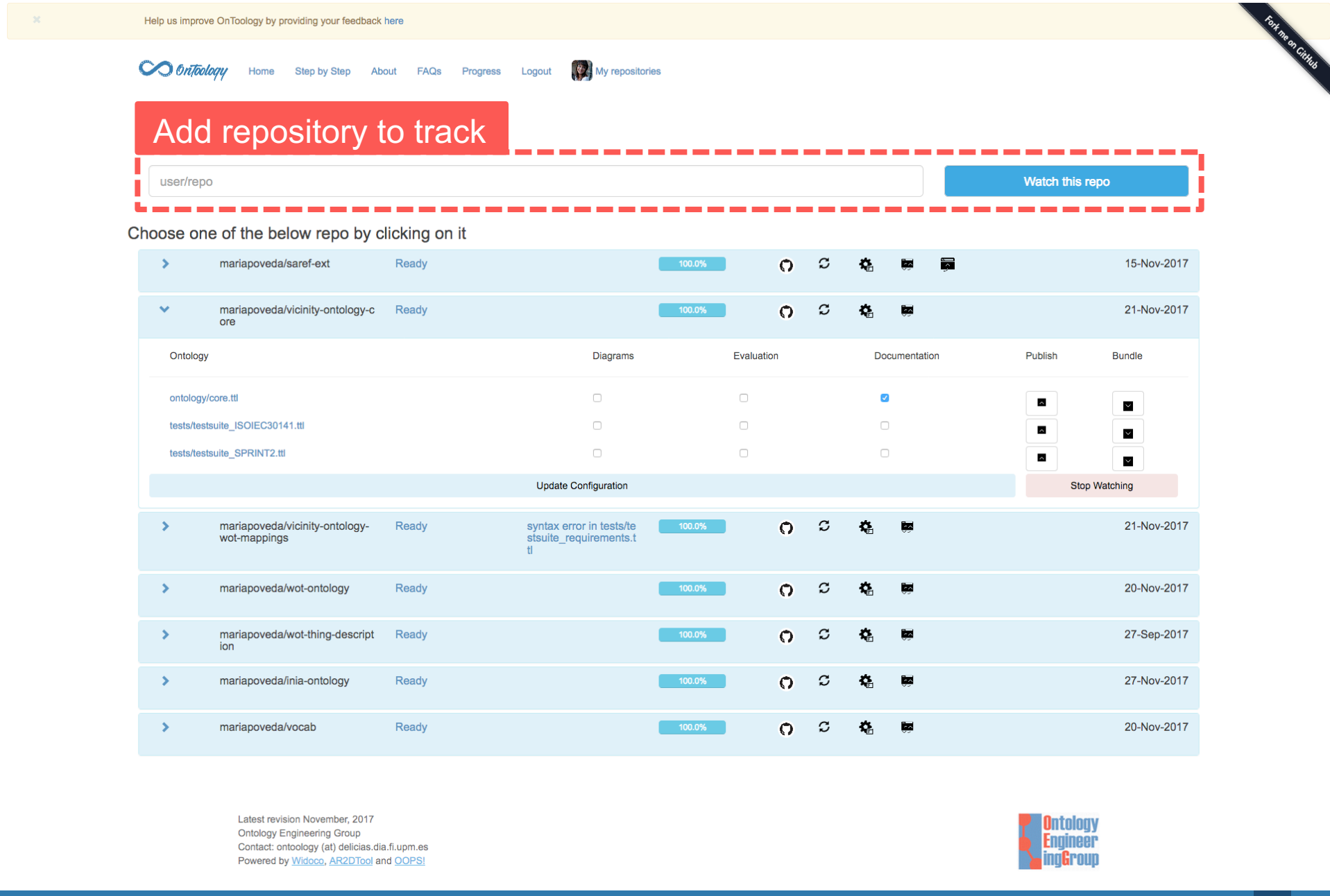

Aiming at bringing all this together…

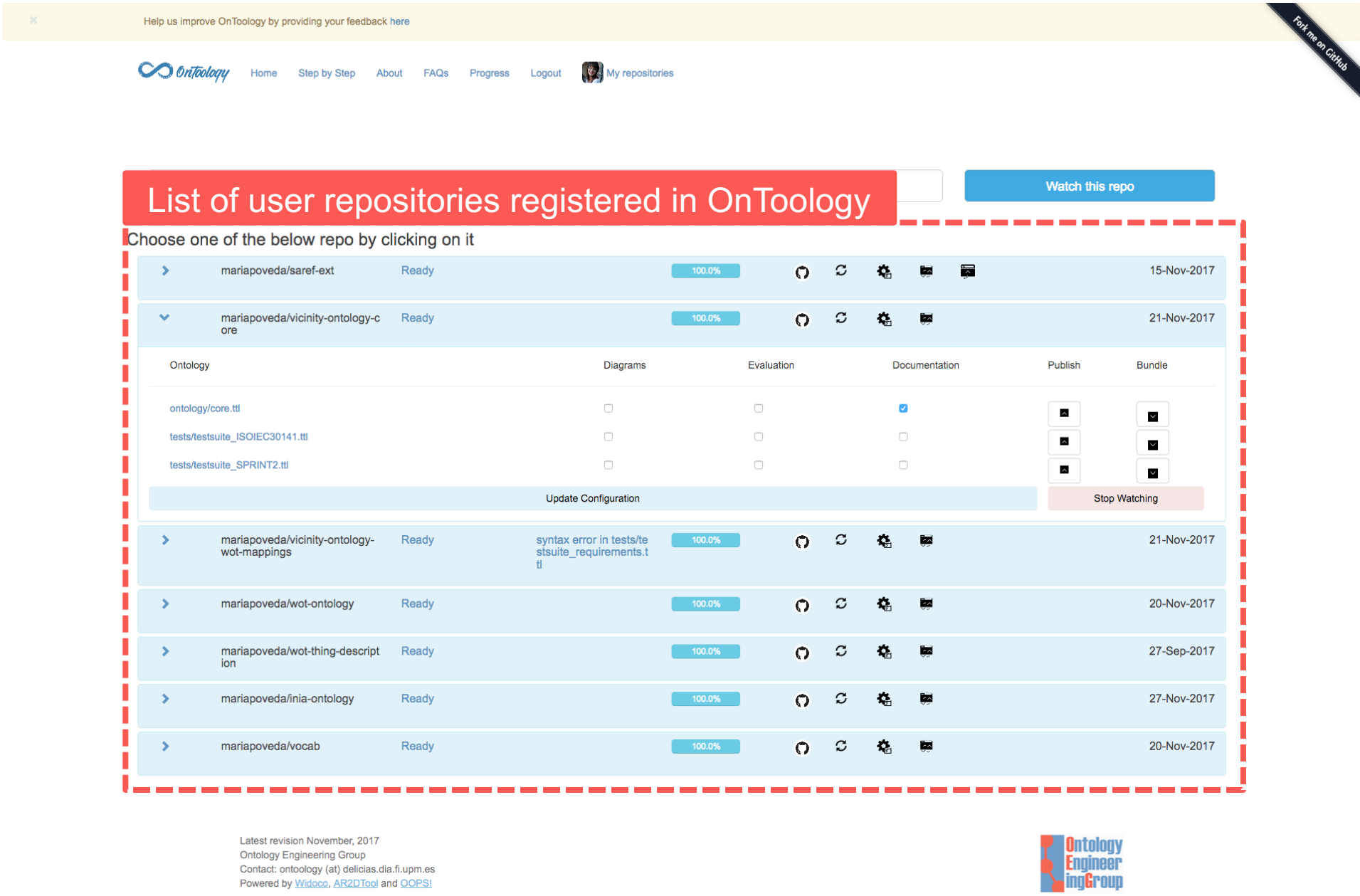
# Handle versions and distributed environments



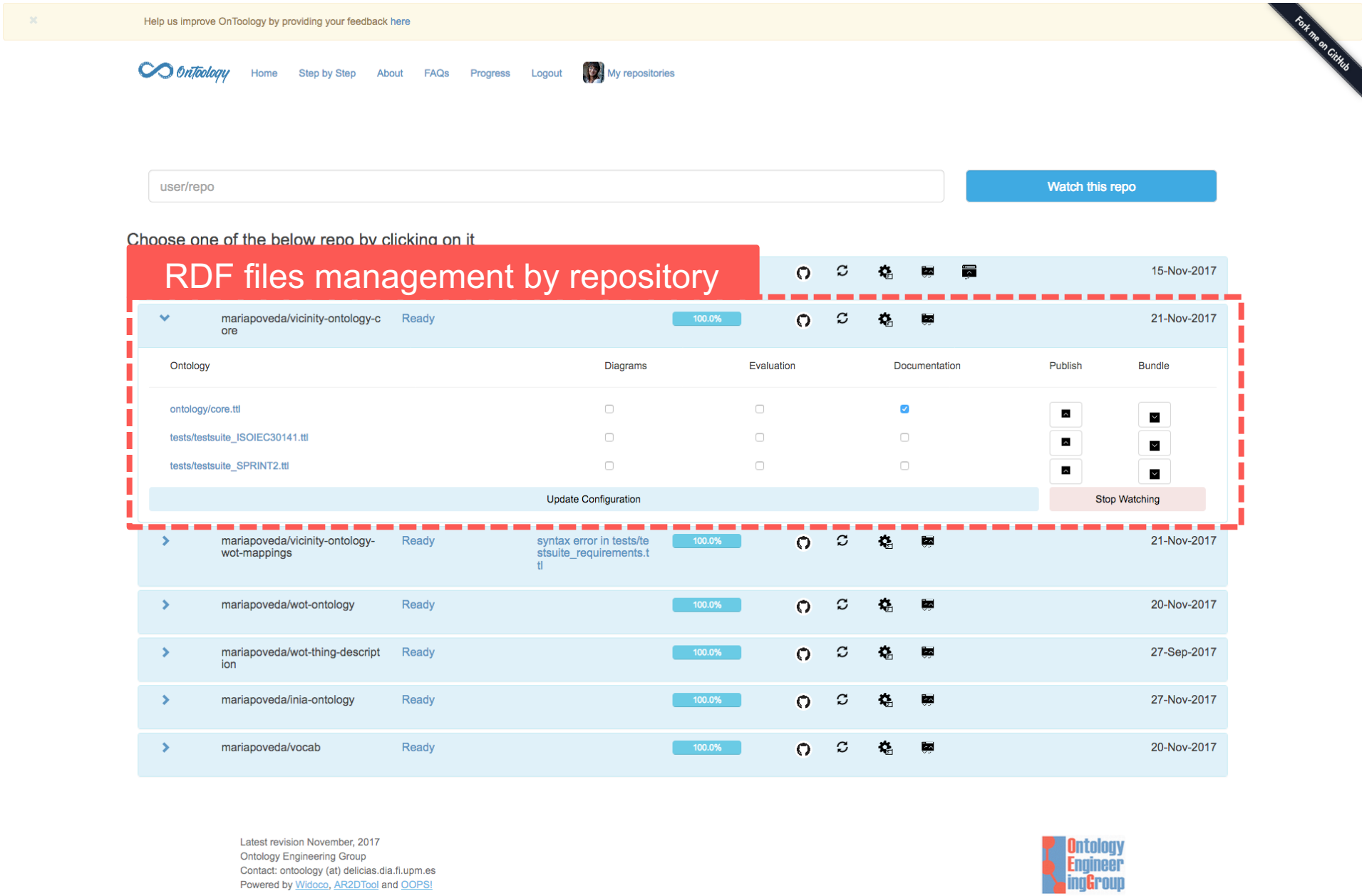

http://ontoology.linkeddata.es

# Using OnToology

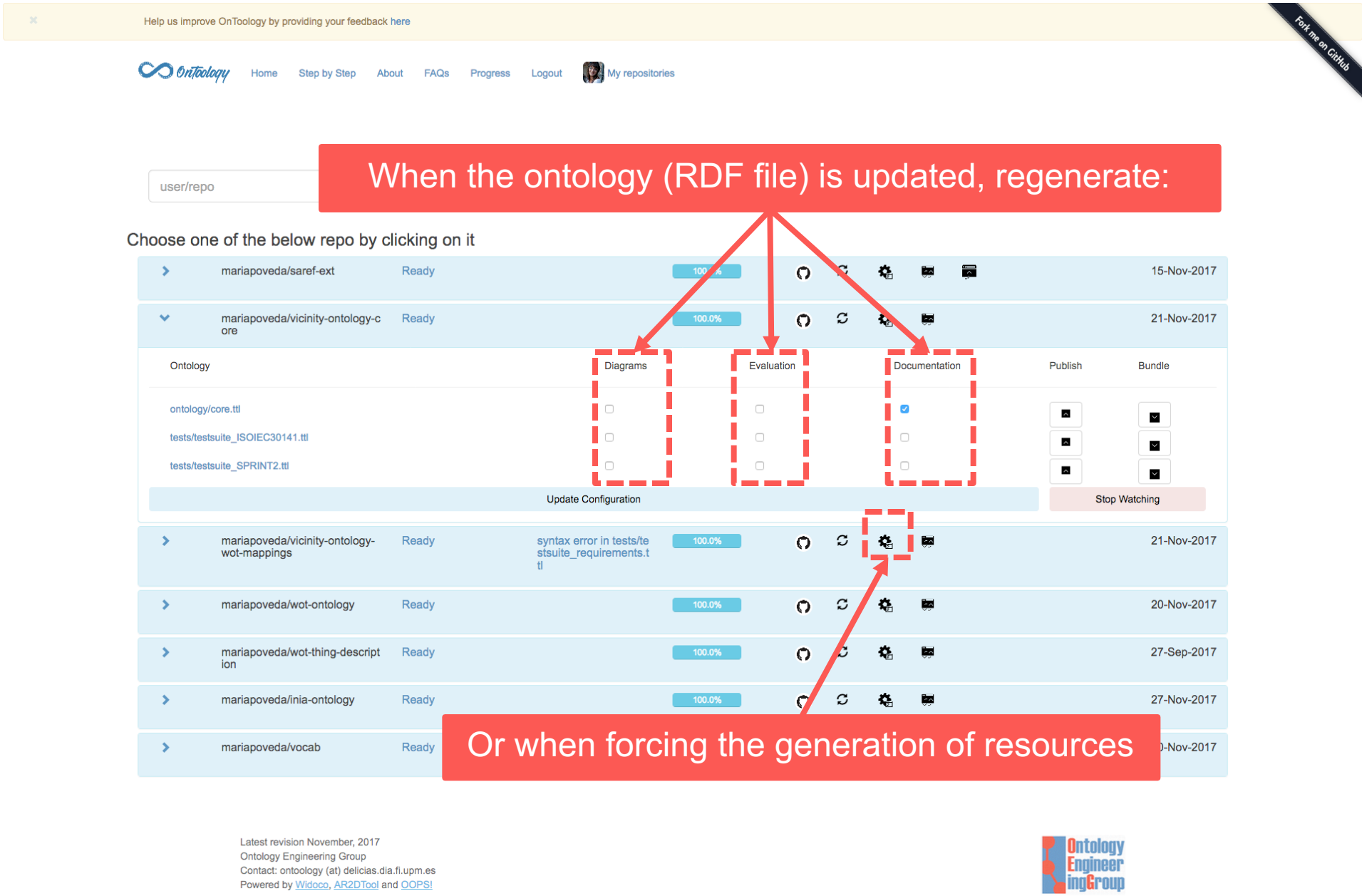



Help us improve OnToology by providing your feedback here
Fork me on GitHub
OnToology
Home
Step by Step
About
FAQs
Progress
Logout
My repositories
List of user repositories registered in OnToology
Watch this repo
Choose one of the below repo by clicking on it
mariapoveda/saref-ext
Ready
100.0%
15-Nov-2017
mariapoveda/vicinity-ontology-core
Ready
100.0%
21-Nov-2017
Ontology
Diagrams
Evaluation
Documentation
Publish
Bundle
ontology/core.ttl
tests/testsuite_ISOIEC30141.ttl
tests/testsuite_SPRINT2.ttl
Update Configuration
Stop Watching
mariapoveda/vicinity-ontology-wot-mappings
Ready
syntax error in tests/testsuite_requirements.ttl
100.0%
21-Nov-2017
mariapoveda/wot-ontology
Ready
100.0%
20-Nov-2017
mariapoveda/wot-thing-description
Ready
100.0%
27-Sep-2017
mariapoveda/inia-ontology
Ready
100.0%
27-Nov-2017
mariapoveda/vocab
Ready
100.0%
20-Nov-2017
Latest revision November, 2017
Ontology Engineering Group
Contact: ontology (at) delicias.dia.fi.upm.es
Powered by Widoco, AR2DTool and OOPS!
Ontology Engineering Group

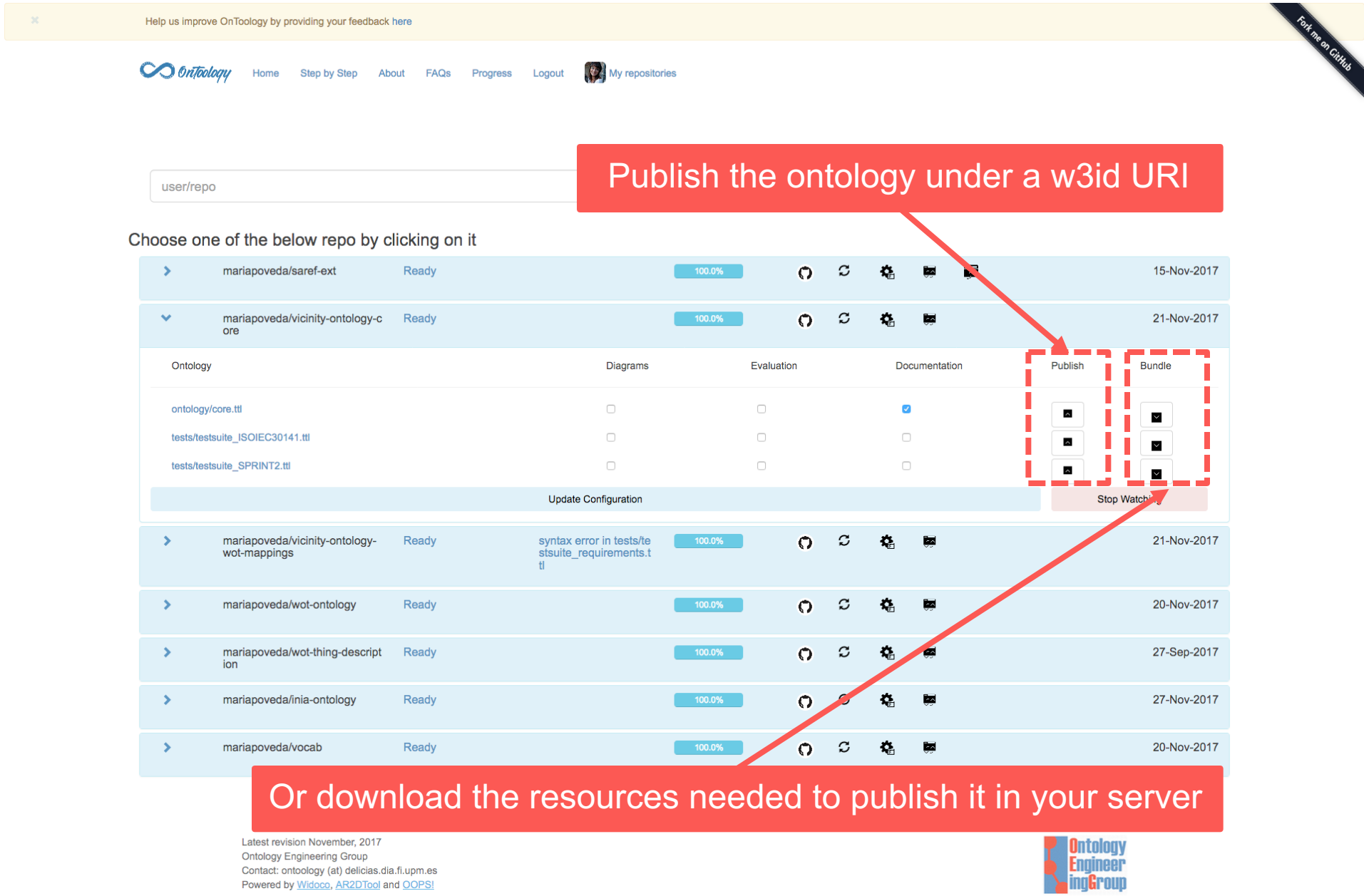
# Using OnToology



# Using OnToology



# Using OnToology



# OnToology users

- Since 2015
- **531** ontologies from **113** repositories
- Some examples:
  - opencitydata/medio-ambiente-contaminacion-acustica
  - opencitydata/medio-ambiente-calidad-aire
  - mariapoveda/wot-ontology
  - vcharpenay/wot-ontology
  - jpcik/medred
  - marianofl1971/dul-es
  - GeorgFerdinandSchneider/bot
  - …

http://vicinity.iot.linkeddata.es/

This slide has been taken from Raúl García Castro presentation at EMSE

# We still need to think about certain things

- What is delivered in OE at the end of each iteration?
  - Different products in different phases
- How do we validate if an iteration is complete in OE?
  - Different type of tests in different phases
  - Difficult to automate validation, mostly for human interaction need

| **Title** | P15. Using "some not" in place of "not some" | **Importance level** | Critical |
|---|---|---|---|
| **Aspects** | Wrong inference | **Affects to** | Classes |

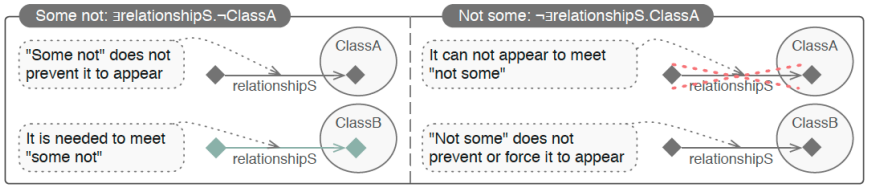
**Description**

The pitfall consists in using a "some not" structure when a "not some" is required. This is due to the misplacement of the existential quantifier (`owl:someValuesFrom`) and the negative operator (`owl:complementOf`).

(a) When to use a "some not" structure ($\exists relationshipS.\neg ClassA$): to state that there is at least one individual acting as object of the relationship "relationshipS" and such individual do not belong to class "ClassA". This implies that there must be at least one instantiation of the relationshipS whose target does not belong to "ClassA". This does not prevent instances from ClassA acting as objects of the relationship.

(b) When to use a "not some" structure ($\neg\exists relationshipS.ClassA$): to state that no individuals in class "ClassA" act as objects of the relationship "relationshipS". This does not imply the existence of individuals that do not belong to ClassA acting as objects of the relationship.

This pitfall is explained in more detail in [1]. See figure below for more details about situations using "some not" or "not some".



# But…

- Where is the common point between OE and?
  - Surely there are successful and not so successful adaptations
  - Dependant on the abstraction or formalism level?

https://es.pngtree.com/freepng/round-landscape-icon_2119094.html https://es.pngtree.com/freepng/cartoon-computer_2732591.html

# Thanks for your attention!


mpoveda@fi.upm.es

mpovedavillalon

thepetiteontologist

@MariaPovedaV

mariapoveda

MariaPovedaVillalon

# Trendy Practices and Tools in Ontological Engineering

**María Poveda Villalón**
**Ontology Engineering Group**
**Universidad Politécnica de Madrid, Spain**

mpoveda@fi.upm.es
@MariaPovedaV

5th April 2018
MINES Saint-Étienne

CC BY NC SA