



OntologySummit2014: Thursday 2014-03-13

Summit Theme: "Big Data and Semantic Web Meet Applied Ontology"

Summit Track Title: "Track B Making use of Ontologies: Tools, Services, and Techniques"

Session II

A Reuse-based Lightweight Method for Developing Linked Data Ontologies and Vocabularies

María Poveda-Villalón, Mari Carmen Suárez-Figueroa and Asunción Gómez-Pérez

Ontology Engineering Group. Departamento de Inteligencia Artificial.

Facultad de Informática, Universidad Politécnica de Madrid.

Campus de Montegancedo s/n.

28660 Boadilla del Monte. Madrid. Spain

{mpoveda, mcsuarez, asun}@fi.upm.es

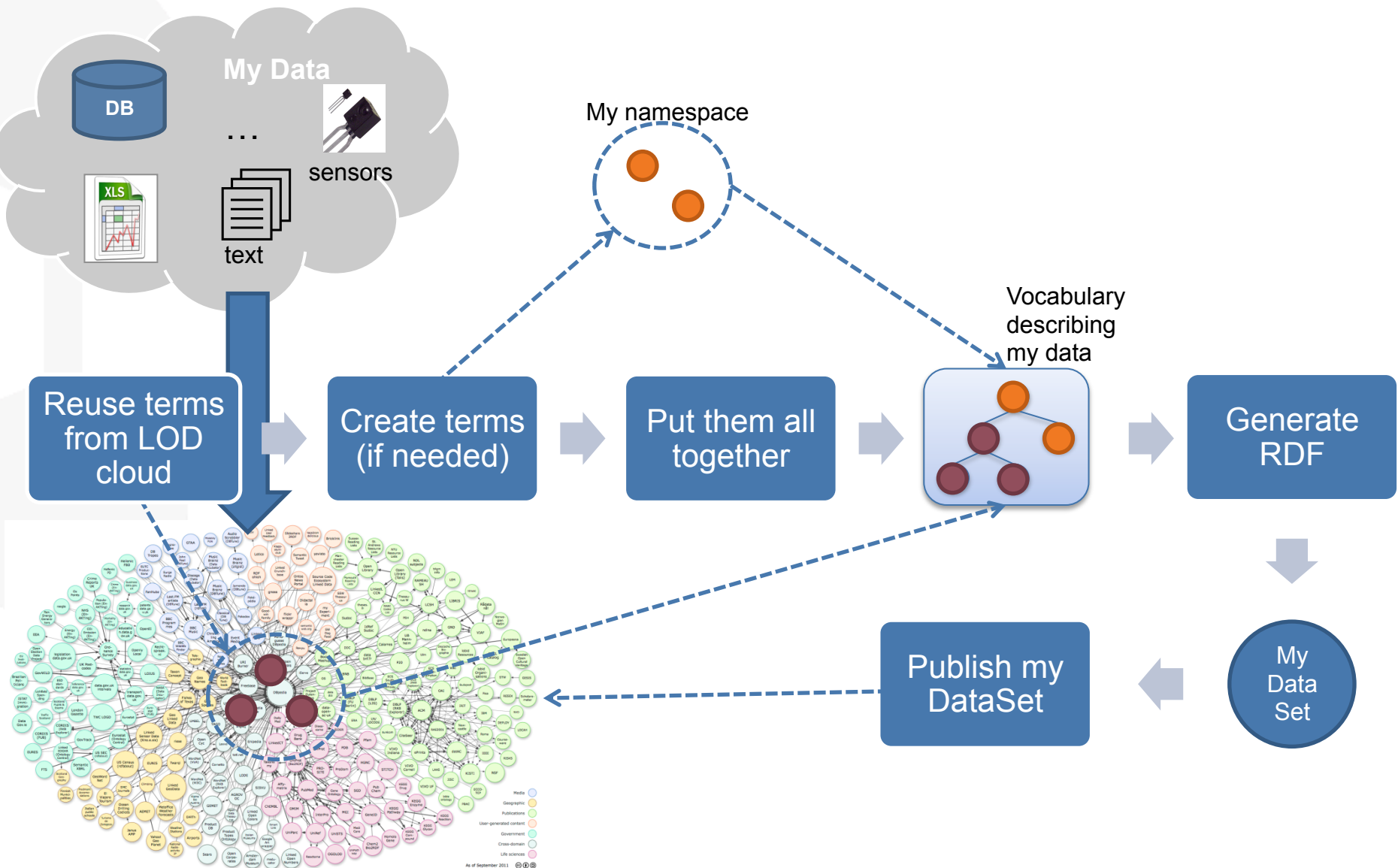
Disclaimer:

Along this talk the terms “ontology” and “vocabulary” will be used indistinctly.

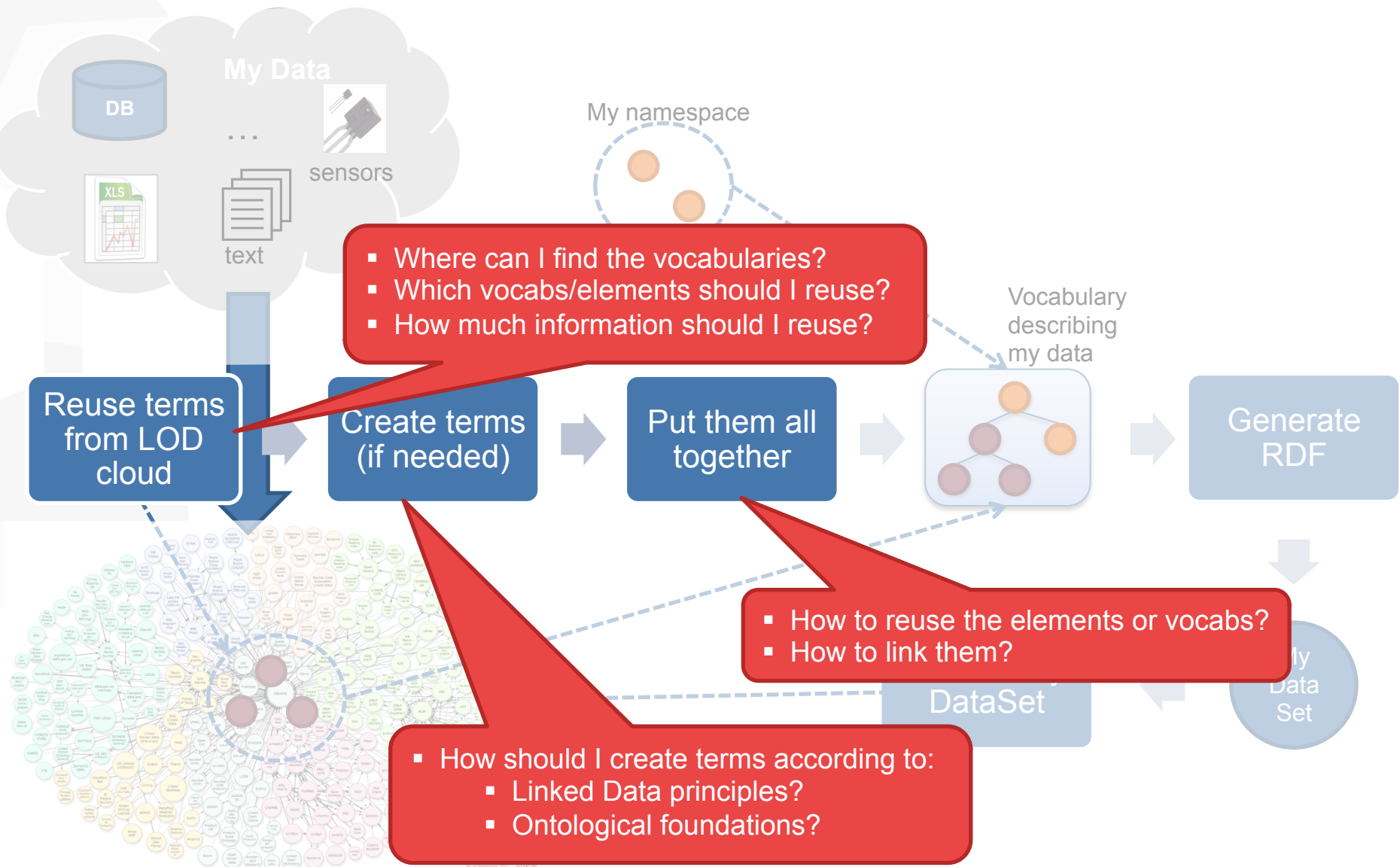
From <http://www.w3.org/standards/semanticweb/ontology>:

There is no clear division between what is referred to as “vocabularies” and “ontologies”. The trend is to use the word “ontology” for more complex, and possibly quite formal collection of terms, whereas “vocabulary” is used when such strict formalism is not necessarily used or only in a very loose sense.

- **Motivation**
- Proposed method overview
- Conclusions and future work



"Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch.
<http://lod-cloud.net/>"



"Linking Open Data cloud diagram, by Richard Cyganiak and Anja Jentzsch.
<http://lod-cloud.net/>"

Considerations about Linked Data developments:

- Agile, rapid developments
- Particular requirements: on the web, dereferenceability, w3c standards, linked
- Domain experts, not knowledge representation experts
- No fixed requirements

About the proposal:

- Lightweight: provide techniques, tools, workflows, and examples
 - Keeping the core activities
- Data driven: the starting point of the process is a list of terms extracted from the raw data
- Reuse based
 - Increase interoperability
 - Decrease cost
- Web oriented
 - Linked Data principles

Ontology development

LD development

Heavyweight methodologies

NeOn [1]

On-To-Knowledge [2]

DILIGENT [3]

Methodology [4]

- Time and resource consuming processes

- Based on competency question technique

- Non Linked Data principles grounded

Lightweight methodologies

eXtreme method [5]

- Set requirements (no maintenance)

XD methodology [6]

- ODPs reuse

RapidOWL [7]

- No reuse

LD guides

Linked Data Book [8]

- What to do but not how to do it

[1] Suárez-Figueroa, M.C. *PhD Thesis: NeOn Methodology for Building Ontology Networks: Specification, Scheduling and Reuse*. Spain. June 2010.

[2] S. Staab, H.P. Schnurr, R. Studer, Y. Sure. Knowledge Processes and Ontologies. *IEEE Intelligent Systems* 16(1):26–34. (2001).

[3] H. S. Pinto, C. Tempich, S. Staab. *DILIGENT: Towards a fine-grained methodology for Distributed, Loosely-controlled and evolInG Engineering of oNTologies*. In Proceedings ECAI 2004.

[4] A. Gómez-Pérez, M. Fernández-López, O. Corcho. *Ontological Engineering*. November 2003. Springer Verlag. ISBN 1-85233-551-3.



[5] Hristozova, M., Sterling, L. *An eXtreme Method for Developing Lightweight Ontologies*. CEUR Workshop Series, 2002.



[6] Presutti, V., Daga, E., Gangemi, A., Blomqvist E. *eXtreme Design with Content Ontology Design Patterns*. WOP 2009



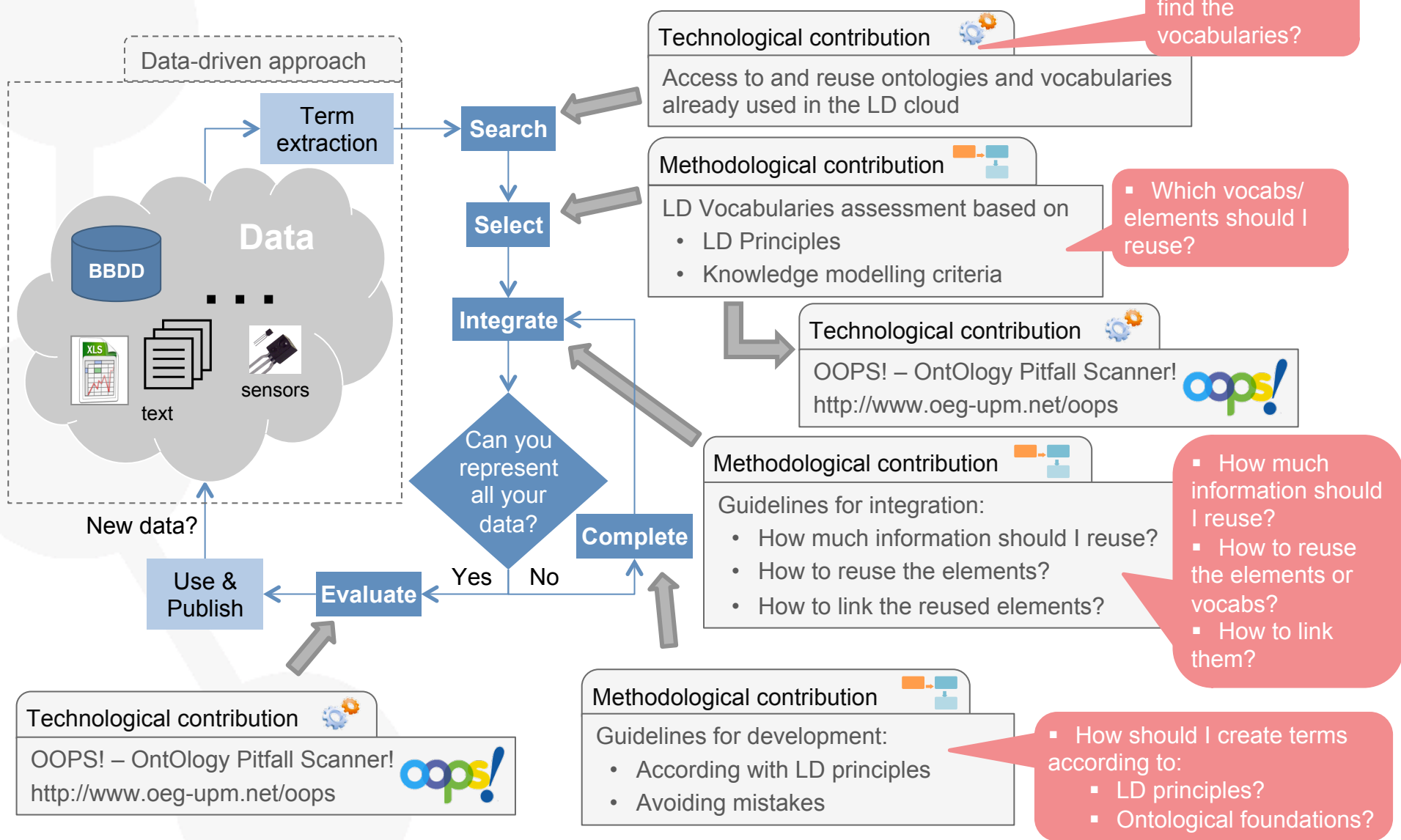
[7] Auer, S.: RapidOWL - an Agile Knowledge Engineering Methodology. In: STICA 2006, Manchester, UK (2006)

[8] Tom Heath and Christian Bizer (2011) *Linked Data: Evolving the Web into a Global Data Space* (1st edition). Synthesis Lectures on the Semantic Web: Theory and Technology, 1:1, 1-136. Morgan & Claypool.

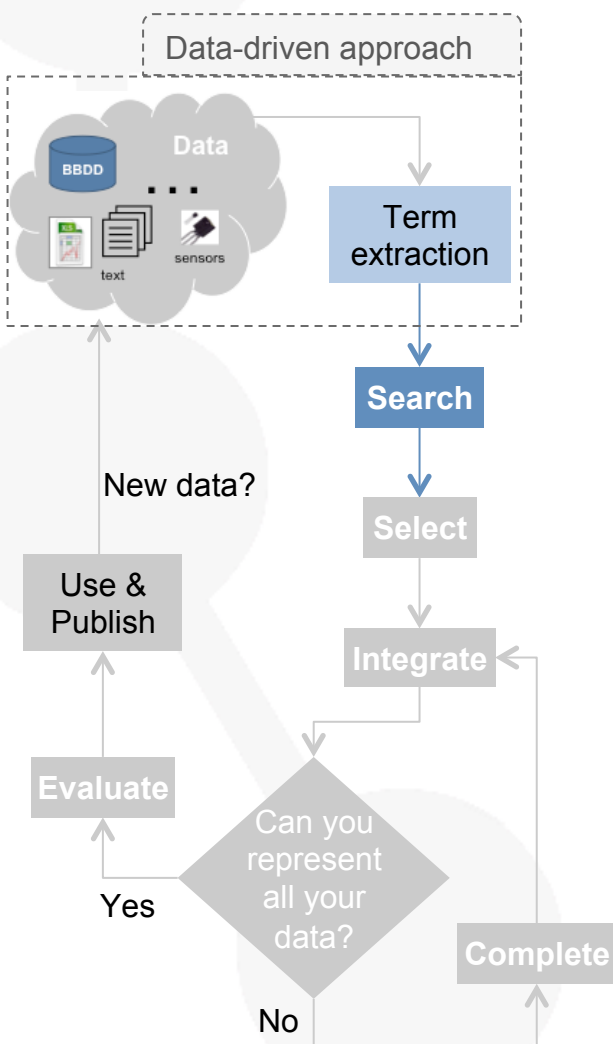
- Motivation
- **Proposed method overview**
- Conclusions and future work

Proposal: LOT – Linked Open Terms methodology

The proposal consists in a **lightweight method** for building ontologies and vocabularies



Ontology Search refers to the activity of finding candidate ontologies or ontology modules to be reused [1].



Focus:

- Terms used in LOD
- Terminology extracted directly from the data
 - Expert advise || Done by experts

Terminology extraction:

- Identify nouns, verbs, etc.
 - Tools: Freeling for free text
 - Tools for ddbb, csv, tables, xml, etc?

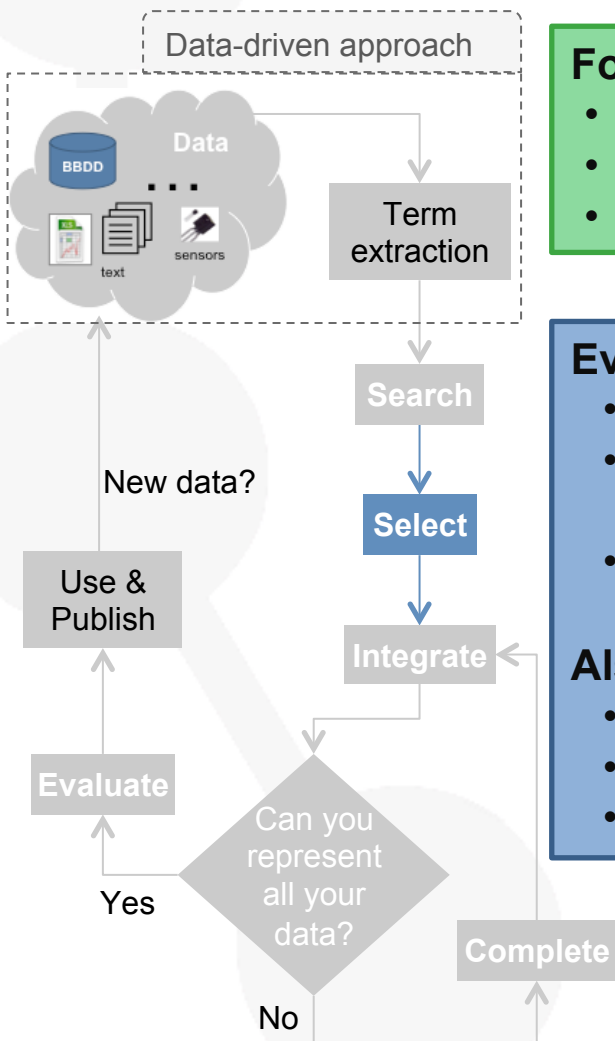
Search tools:

- General purpose:
 - LOV: <http://lov.okfn.org>
 - LOD2Stats: <http://stats.lod2.eu/vocabularies>
 - Others: ODP, Swoogle, Sindice, etc
- Domain base:
 - Bioportal: <http://bioportal.bioontology.org/>
 - Smartcity: <http://smartcity.linkeddata.es/>



[1] Suárez-Figueroa, M.C. *PhD Thesis: NeOn Methodology for Building Ontology Networks: Specification, Scheduling and Reuse*. Spain. June 2010.

Ontology Selection refers to the activity of choosing the most suitable ontologies or ontology modules among those available in an ontology repository or library, for a concrete domain of interest and associated tasks [1].



Focus:

- Assessment by Linked Data principles
- Modelling issues
- Domain coverage: data driven

Evaluation tools:

- OOPS! – OntOlogy pitfalls scanner [2] <http://www.oeg-upm.net/oops/>
- Triple checker <http://graphite.ecs.soton.ac.uk/checker/> (already included in OOPS!)
- Vapour <http://validator.linkeddata.org/vapour> (to be included in OOPS!)

Also it should be considered:

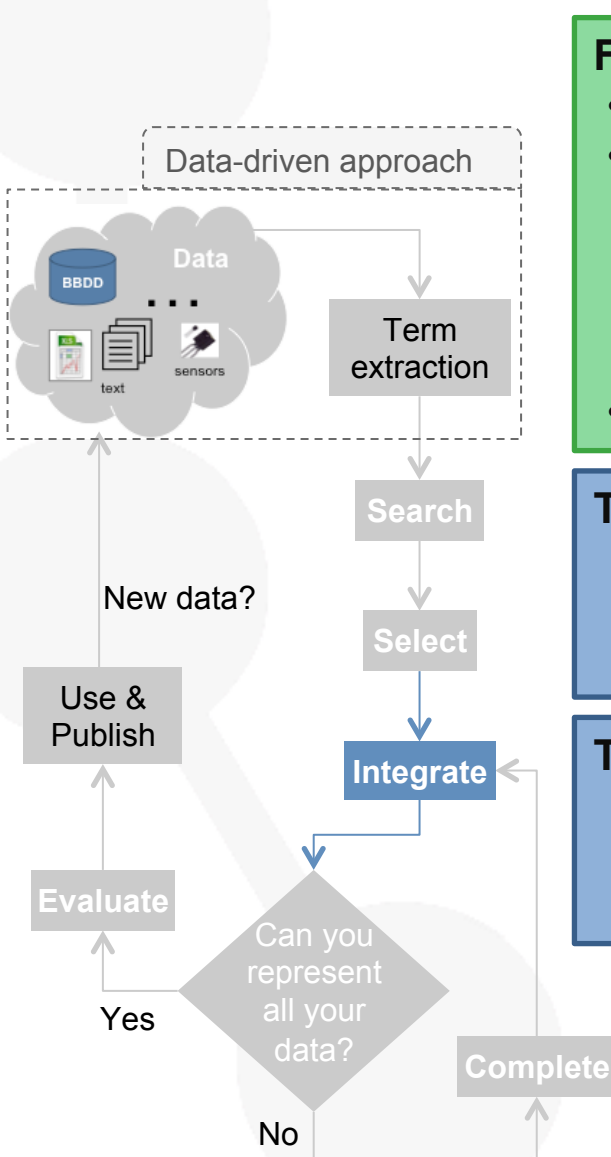
- Modelling issues (OOPS!, reasoners, manually review, etc.)
- Domain coverage (based on the data to be represented)
- Used in Linked Data (LOD2Stats, Sindice, etc)



[1] Suárez-Figueroa, M.C. *PhD Thesis: NeOn Methodology for Building Ontology Networks: Specification, Scheduling and Reuse*. Spain. June 2010.

[2] Poveda-Villalón, M., Suárez-Figueroa, M. C., & Gómez-Pérez, A. (2012). *Validating ontologies with oops!*. In *Knowledge Engineering and Knowledge Management* (pp. 267-281). Springer Berlin Heidelberg.

Ontology Integration. It refers to the activity of including one ontology in another ontology [1].



Focus:

- How much information should I reuse?
- How to reuse the elements or vocabs? Preliminary analysis [2]
 - Should I import another ontology?
 - Should I reference other ontology element URIs?
 - ... replicating manually the URI?
 - ... merging ontologies?
- How to link them?

Techniques:

- Import the ontology as a whole
- Reuse some parts of the ontology (or ontology module)
- Reuse statements

Tools:

- Ontology editors: Protégé, NeOn Toolkit, etc.
 - Plug-ins: Ontology Module Extraction and Partition
- Text editors for manual approach



[1] Suárez-Figueroa, M.C. *PhD Thesis: NeOn Methodology for Building Ontology Networks: Specification, Scheduling and Reuse*. Spain. June 2010.

[2] Poveda-Villalón, M., Suárez-Figueroa, M. C., & Gómez-Pérez, A. *The Landscape of Ontology Reuse in Linked Data*. 1st Ontology Engineering in a Data-driven World (OEDW 2012) Workshop at the 18th International Conference on Knowledge Engineering and Knowledge Management. Galway, Ireland, 9th October 2012. <http://www.slideshare.net/MariaPovedaVillalon/mpoveda-oedw2012v1>

Ontology Enrichment It refers to the activity of extending an ontology with new conceptual structures (e.g., concepts, roles and axioms) [1].

Focus:

- How should I create terms according to ontological foundations and Linked Data principles?

Ontology development:

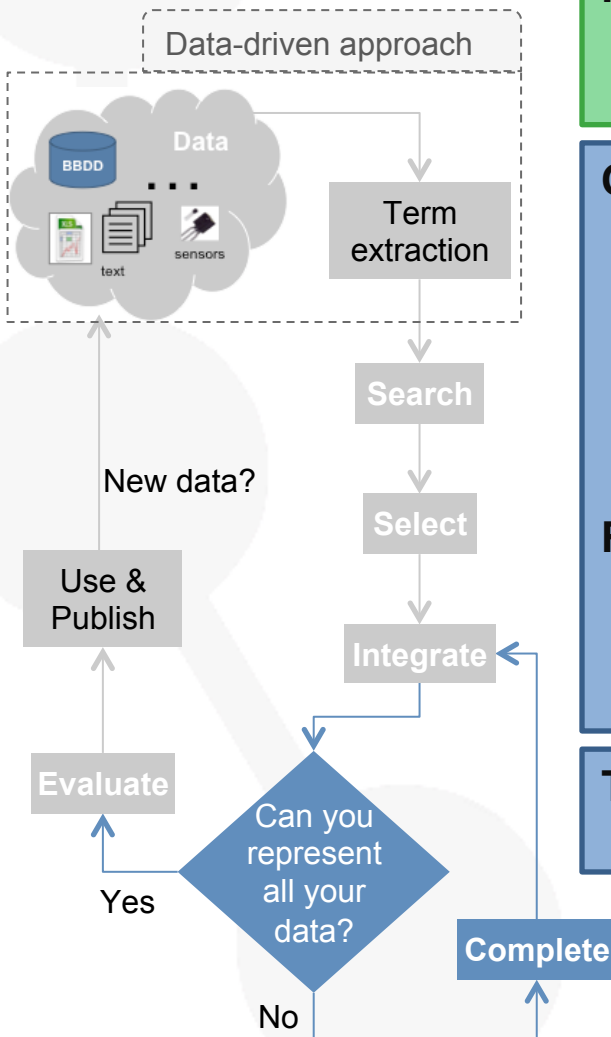
- Ontology Development 101: A Guide to Creating Your First Ontology [2]
- Ontology Engineering Patterns <http://www.w3.org/2001/sw/BestPractices/>
- Extracting Ontology conceptualization, formalization techniques from existing methodologies

Recommendation

- Link to existing entities
- Provide human readable documentation
- Keep the semantics of the reused elements

Tools:

- Ontology editors: Protégé, NeOn Toolkit, etc.



- [1] Suárez-Figueroa, M.C. *PhD Thesis: NeOn Methodology for Building Ontology Networks: Specification, Scheduling and Reuse*. Spain. June 2010.
- [2] Natalya F. Noy and Deborah L. McGuinness. *Ontology Development 101: A Guide to Creating Your First Ontology*. Stanford Knowledge Systems Laboratory Technical Report KSL-01-05 and Stanford Medical Informatics Technical Report SMI-2001-0880, March 2001.

Ontology Evaluation it refers to the activity of checking the technical quality of an ontology against a frame of reference. [1].

Focus:

- Assessment by Linked Data principles
- Modelling issues
- Domain coverage: data driven

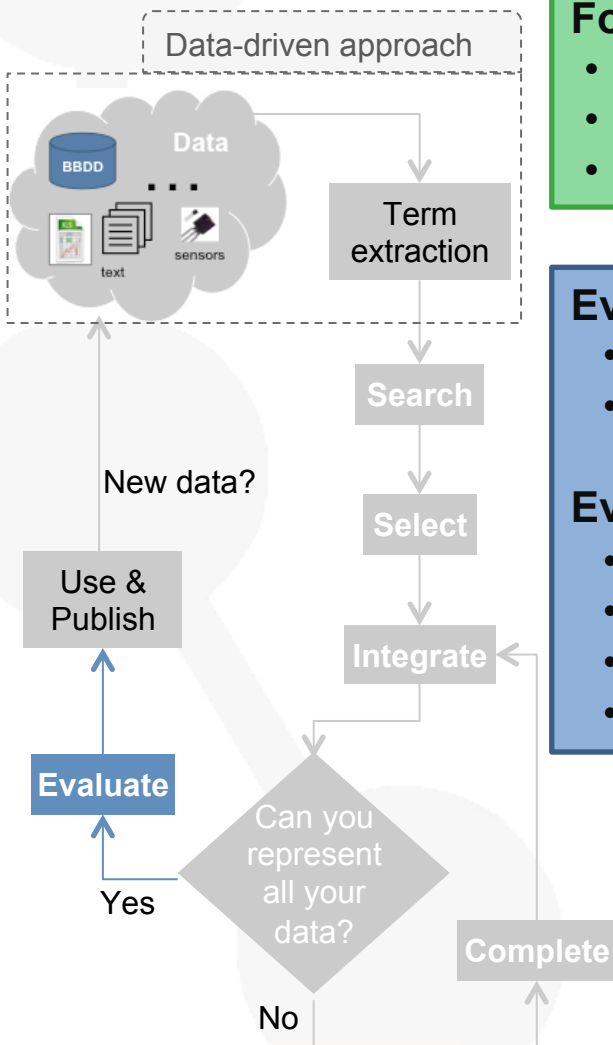


Evaluation tools related to Linked Data principles:

- OOPS! – OntOlogy pitfalls scanner [2] <http://www.oeg-upm.net/oops/>
- Triple checker <http://graphite.ecs.soton.ac.uk/checker/> (already included in OOPS!)

Evaluation tools/techniques other aspects:

- Modelling issues (OOPS!, reasoners, manually review, etc.)
- Domain coverage (based on the data to be represented)
- Application based (queries)
- Syntax issues: validators



[1] Suárez-Figueroa, M.C. *PhD Thesis: NeOn Methodology for Building Ontology Networks: Specification, Scheduling and Reuse*. Spain. June 2010.

[2] Poveda-Villalón, M., Suárez-Figueroa, M. C., & Gómez-Pérez, A. (2012). *Validating ontologies with oops!*. In *Knowledge Engineering and Knowledge Management* (pp. 267-281). Springer Berlin Heidelberg.

- Motivation
- Proposed method overview
- **Conclusions and future work**

Conclusions	<ul style="list-style-type: none">• There are a lot of tools and techniques that have to be connected• There are still difficulties to make more lightweight the ontology conceptualization and formalization activities
Future work	<ul style="list-style-type: none">• Provide detailed guidelines<ul style="list-style-type: none">• Include examples• Set thresholds to suggest different techniques for reuse• Provide detailed workflows for each activity• Test with users

Further information: M. Poveda-Villalón. *A Reuse-based Lightweight Method for Developing Linked Data Ontologies and Vocabularies*. PhD symposium at the 9th Extended Semantic Web Conference (ESWC2012). 27th – 31st May 2012. Heraklion, Greece. http://oa.upm.es/14479/1/ESWC2012-DS-Camera_Ready_v05.pdf

List of URLs alphabetical order:

- Bioportal
- Freeling
- LOD2Stats
- LOV
- NeOn Toolkit
- NeOn Toolkit Plug-in Ontology Module Extraction
- NeOn Toolkit Plug-in Ontology Module Partition
- ODP (Ontology Design Pattern) Portal
- Ontology Engineering Patterns
- OOPS!
- Protégé
- Sindice
- Smartcity ontology catalogue
- Swoogle
- Triplechecker
- Vapour

<http://bioportal.bioontology.org/>
<http://nlp.lsi.upc.edu/freeling/>
<http://stats.lod2.eu/vocabularies>
<http://lov.okfn.org>
<http://neon-toolkit.org/>
http://neon-toolkit.org/wiki/Ontology_Module_Extraction
http://neon-toolkit.org/wiki/Ontology_Module_Partition
<http://ontologydesignpatterns.org/>
<http://www.w3.org/2001/sw/BestPractices/>
<http://www.oeg-upm.net/oops/>
http://protegewiki.stanford.edu/wiki/Main_Page
<http://sindice.com/>
<http://smartcity.linkeddata.es/>
<http://swoogle.umbc.edu/>
<http://graphite.ecs.soton.ac.uk/checker/>
<http://validator.linkeddata.org/vapour>

Thanks!



OntologySummit2014: Thursday 2014-03-13

Summit Theme: "Big Data and Semantic Web Meet Applied Ontology"

Summit Track Title: "Track B Making use of Ontologies: Tools, Services, and Techniques"

Session II

A Reuse-based Lightweight Method for Developing Linked Data Ontologies and Vocabularies

María Poveda-Villalón, Mari Carmen Suárez-Figueroa and Asunción Gómez-Pérez

Ontology Engineering Group. Departamento de Inteligencia Artificial.

Facultad de Informática, Universidad Politécnica de Madrid.

Campus de Montegancedo s/n.

28660 Boadilla del Monte. Madrid. Spain

{[mpoveda](mailto:mpoveda@fi.upm.es), [mcsuarez](mailto:mcsuarez@fi.upm.es), [asun](mailto:asun@fi.upm.es)}@fi.upm.es