Modelling, Learning and Populating Ontologies for the Semantic Web

Marco Rospocher

UNIVERSITÀ
di VERONA

Dipartimento
di **LINGUE**
E LETTERATURE STRANIERE

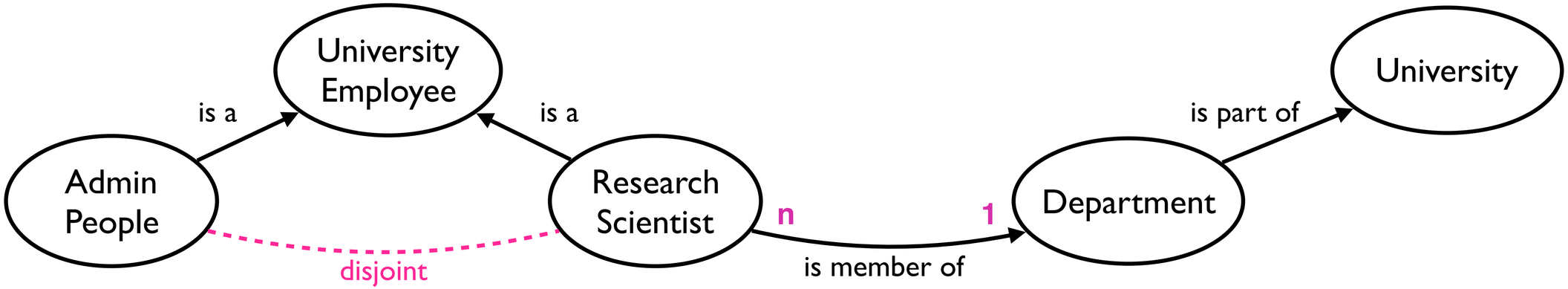
Ontologies

- An ontology is a **formal**, **explicit** specification of a **shared conceptualisation**

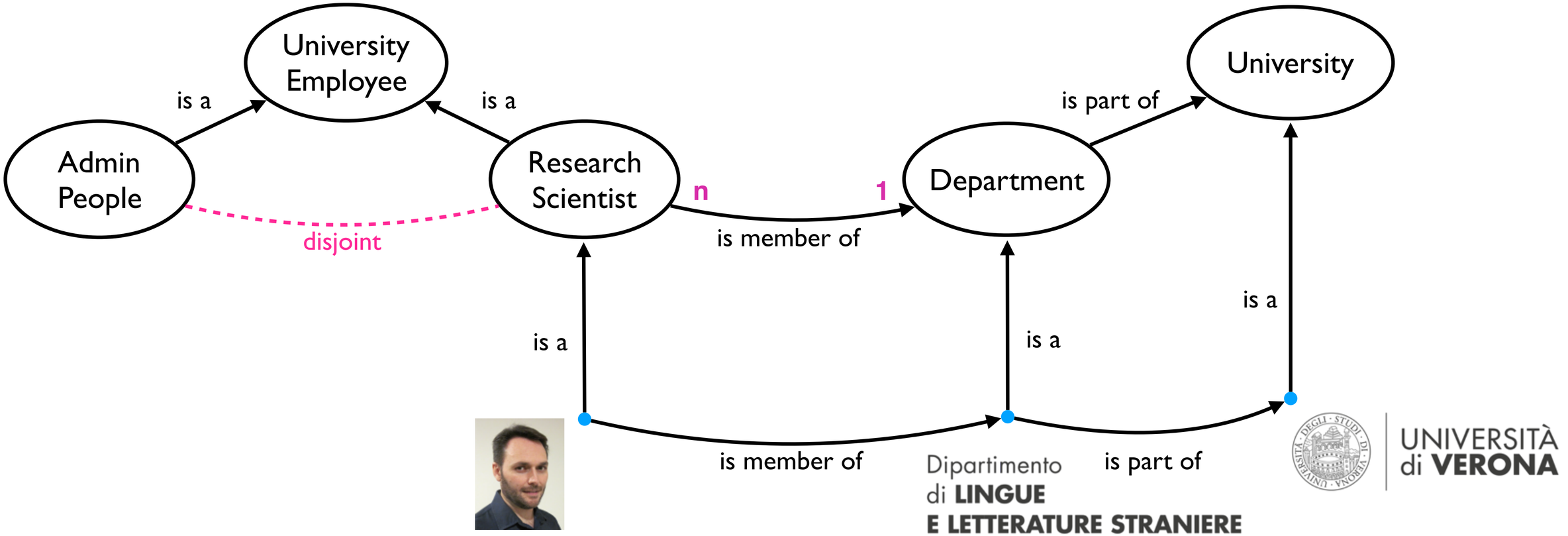
T.R. Gruber, A Translation Approach to Portable Ontology Specifications, *Knowledge Acquisition* 5, 2, 1993, 199-221.

  - conceptualisation:
    - abstract model: objects, concepts, and other entities that are assumed to exist in some area of interest and the relationships that hold among them
  - shared:
    - all the stakeholders should understand the primitive terms in the appropriate way
  - explicit:
    - all elements used in the conceptualisation must be defined
  - formal:
    - must be machine understandable

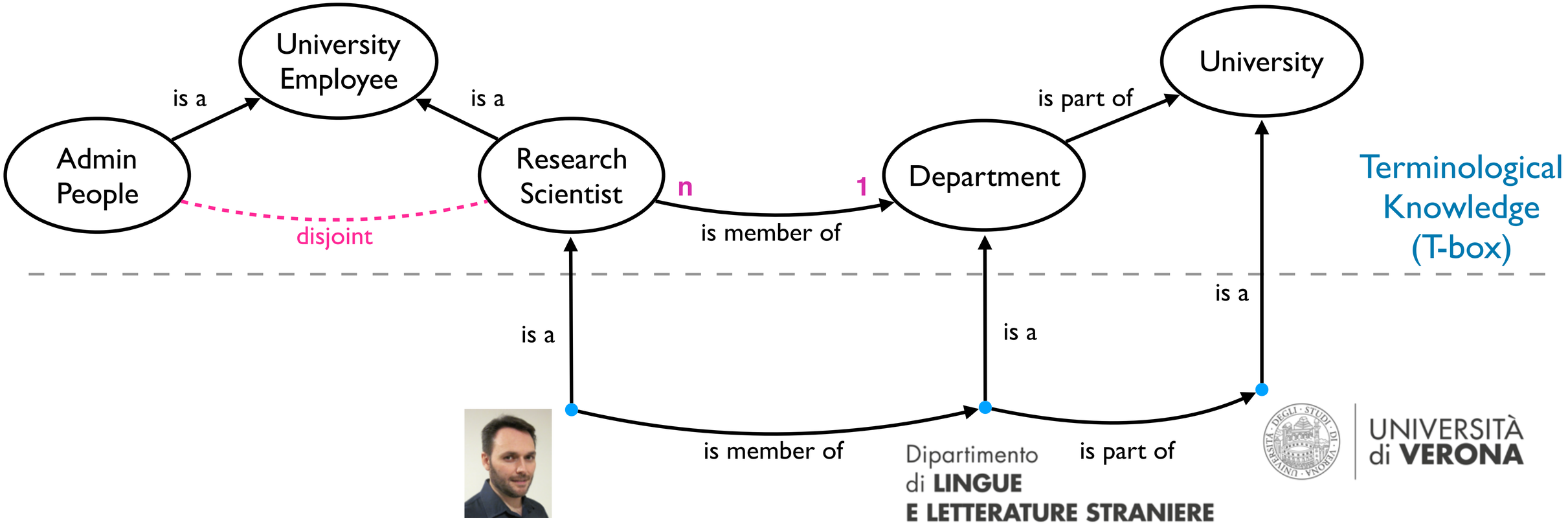
Ontologies



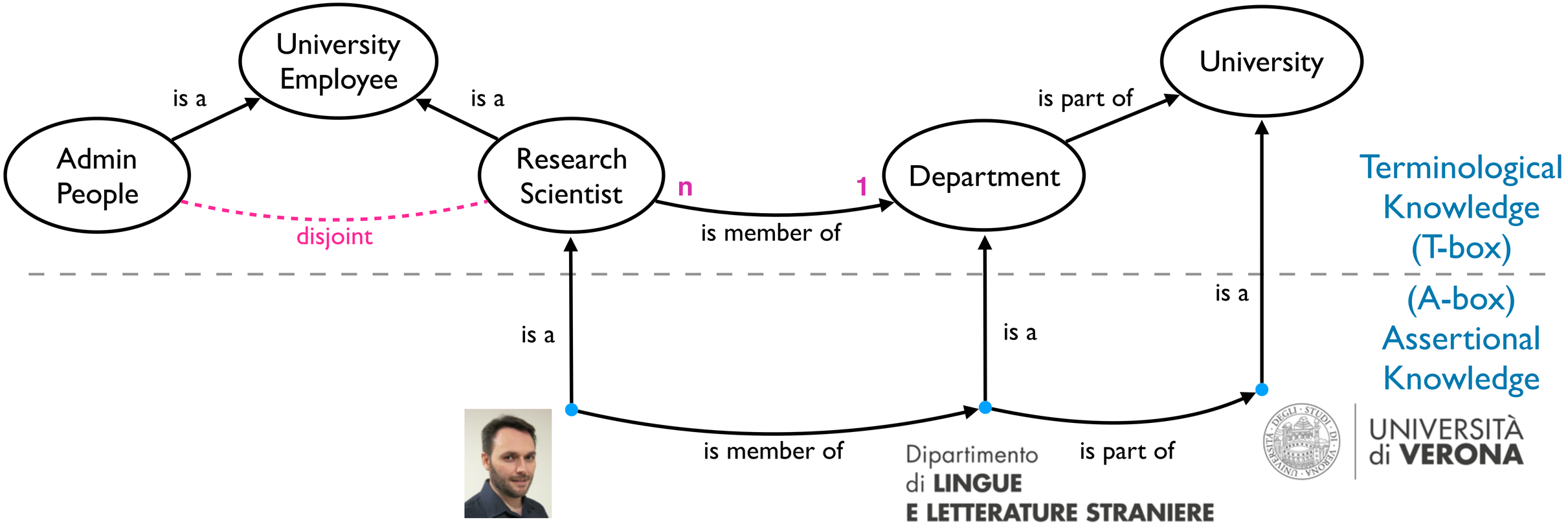
Ontologies


University Employee
is a
is a
Admin People
Research Scientist
n
1
Department
is part of
University
disjoint
is member of
is a
is a
is a
is member of
is part of
Dipartimento di LINGUE E LETTERATURE STRANIERE
UNIVERSITÀ di VERONA

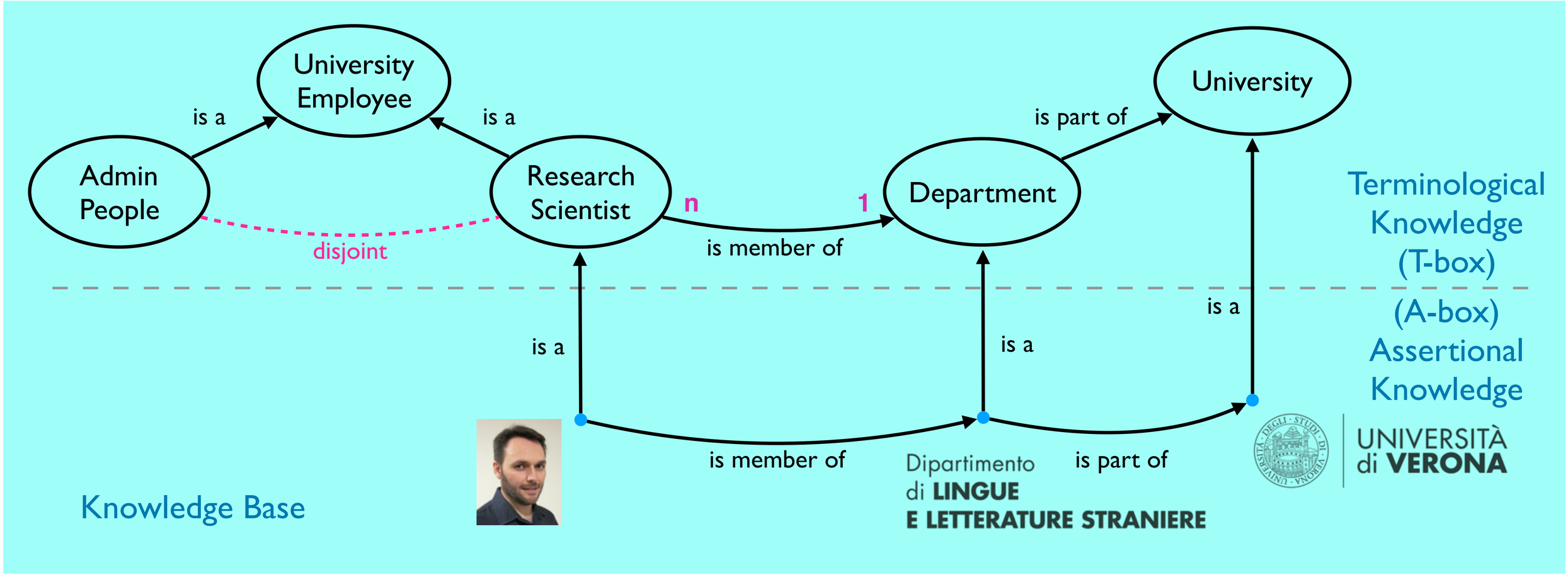
Ontologies


University Employee
is a
is a
Admin People
Research Scientist
n
1
Department
is part of
University
disjoint
is member of
Terminological Knowledge (T-box)
is a
is a
is a
is member of
is part of
Dipartimento di LINGUE E LETTERATURE STRANIERE
UNIVERSITÀ di VERONA

Ontologies


University Employee
is a
is a
Admin People
Research Scientist
disjoint
n
1
is member of
Department
is part of
University
is a
is a
is a
is member of
is part of
Terminological Knowledge (T-box)
(A-box) Assertional Knowledge
Dipartimento di LINGUE E LETTERATURE STRANIERE
UNIVERSITÀ di VERONA

Ontologies


University Employee
is a
is a
Admin People
Research Scientist
disjoint
n
1
Department
is member of
is part of
University
Terminological Knowledge (T-box)
(A-box) Assertional Knowledge
is a
is a
is a
is member of
is part of
Knowledge Base
Dipartimento di LINGUE E LETTERATURE STRANIERE
UNIVERSITÀ di VERONA

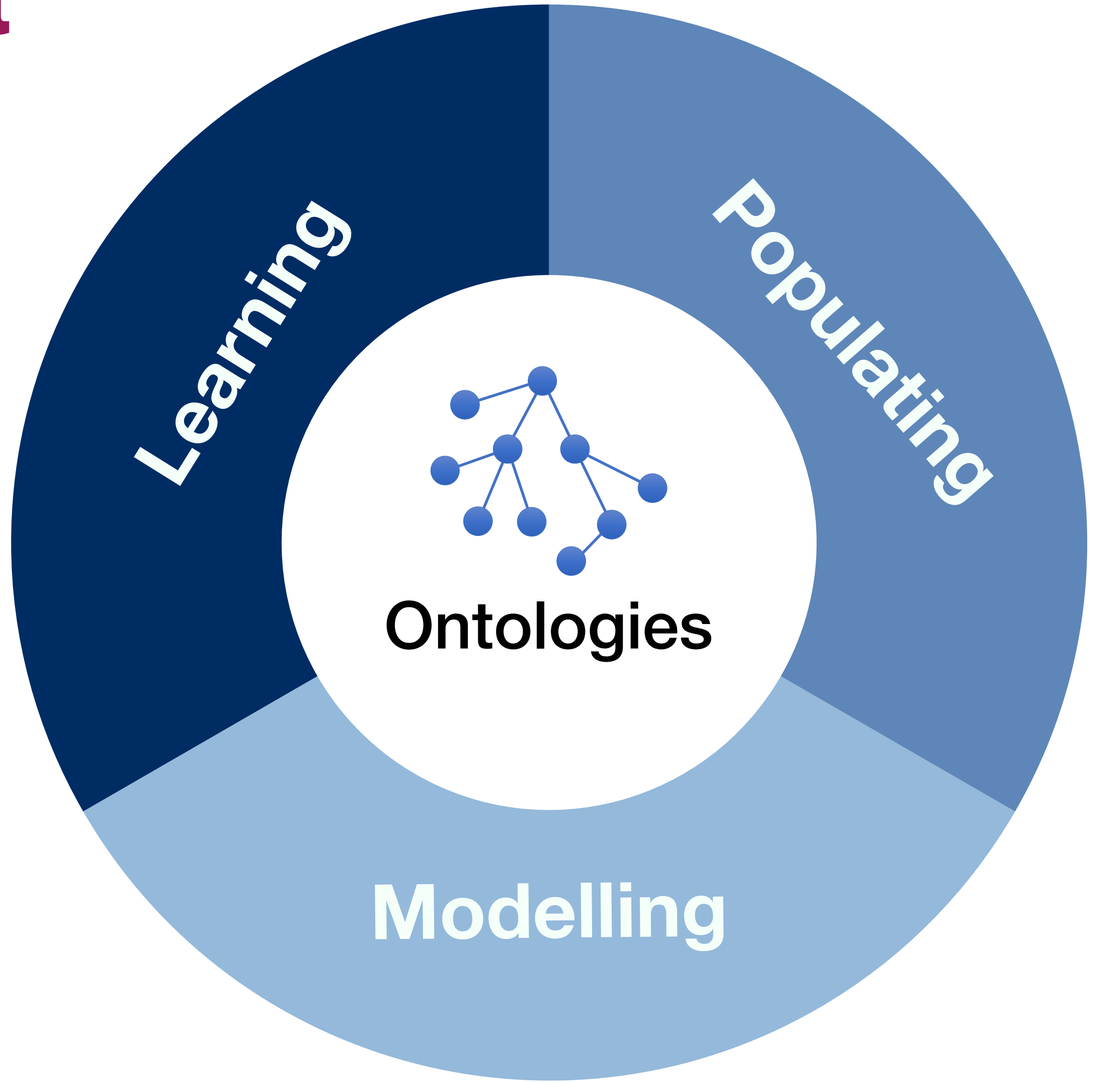
Ontologies

- One of the building blocks of the Semantic Web

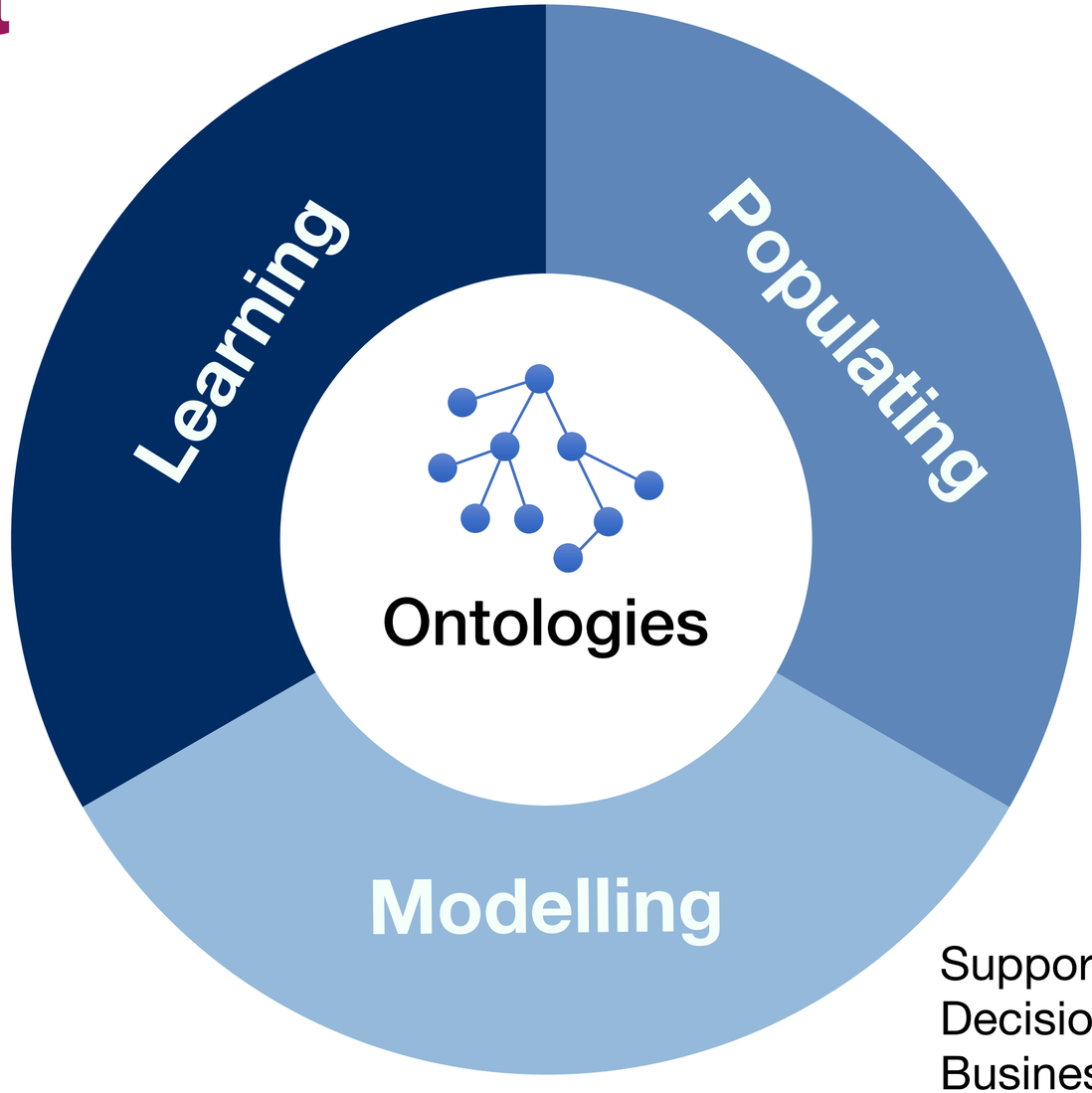
“I have a dream for the Web [in which computers] become capable of analyzing all the data on the Web – the content, links, and transactions between people and computers. A "Semantic Web", which makes this possible, has yet to emerge, but when it does, the day-to-day mechanisms of trade, bureaucracy and our daily lives will be handled by machines talking to machines. The "intelligent agents" people have touted for ages will finally materialize.”
[TBL, 1999]

- Knowledge representation languages for authoring ontologies: RDF, OWL
  - Formal semantics based on Description Logics (hence, we can infer new knowledge!)
  - Open world assumption and no Unique name assumption

Agenda



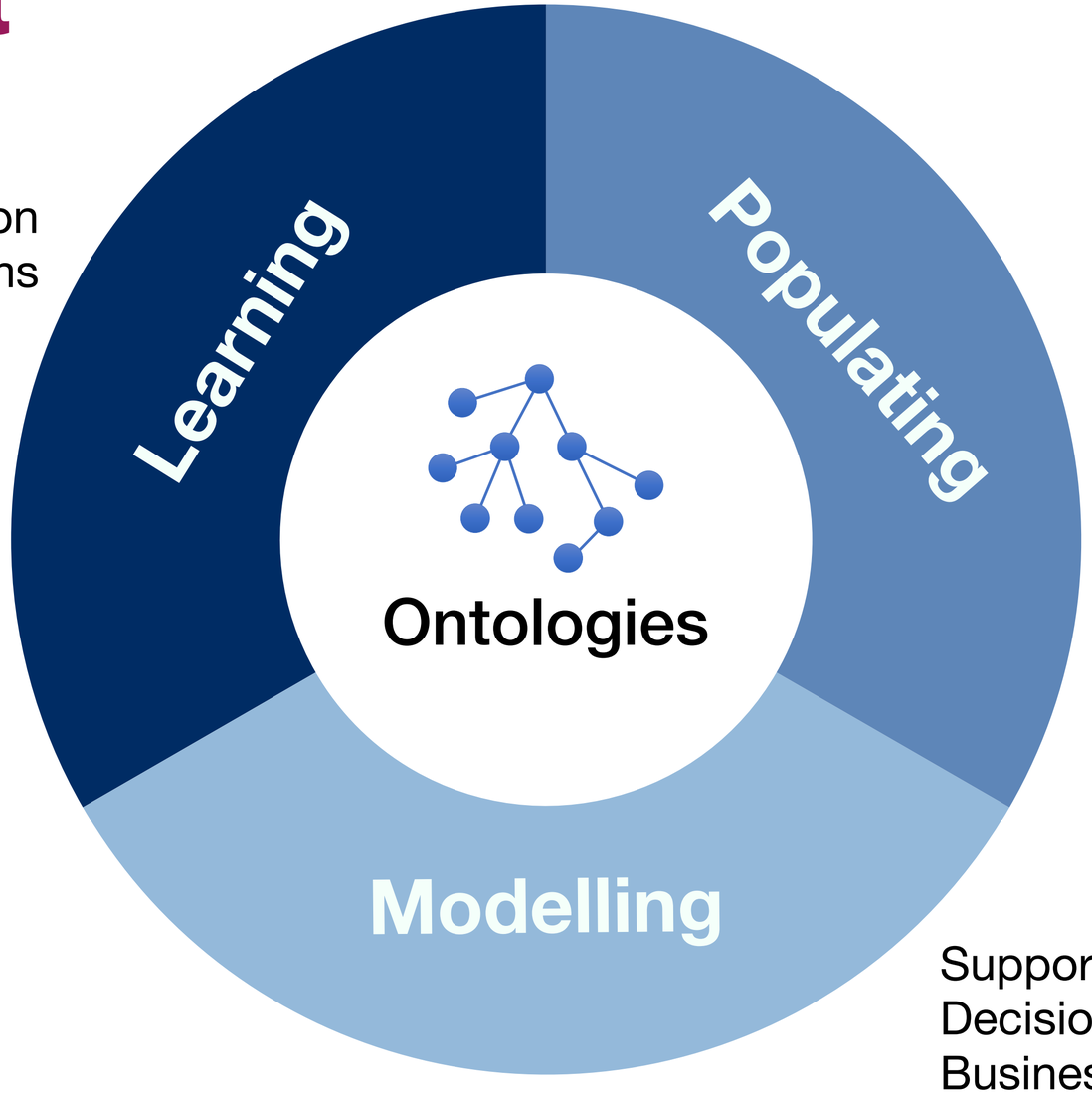
Agenda



Supporting collaborative modelling
Decision Support Systems
Business Processes

Agenda

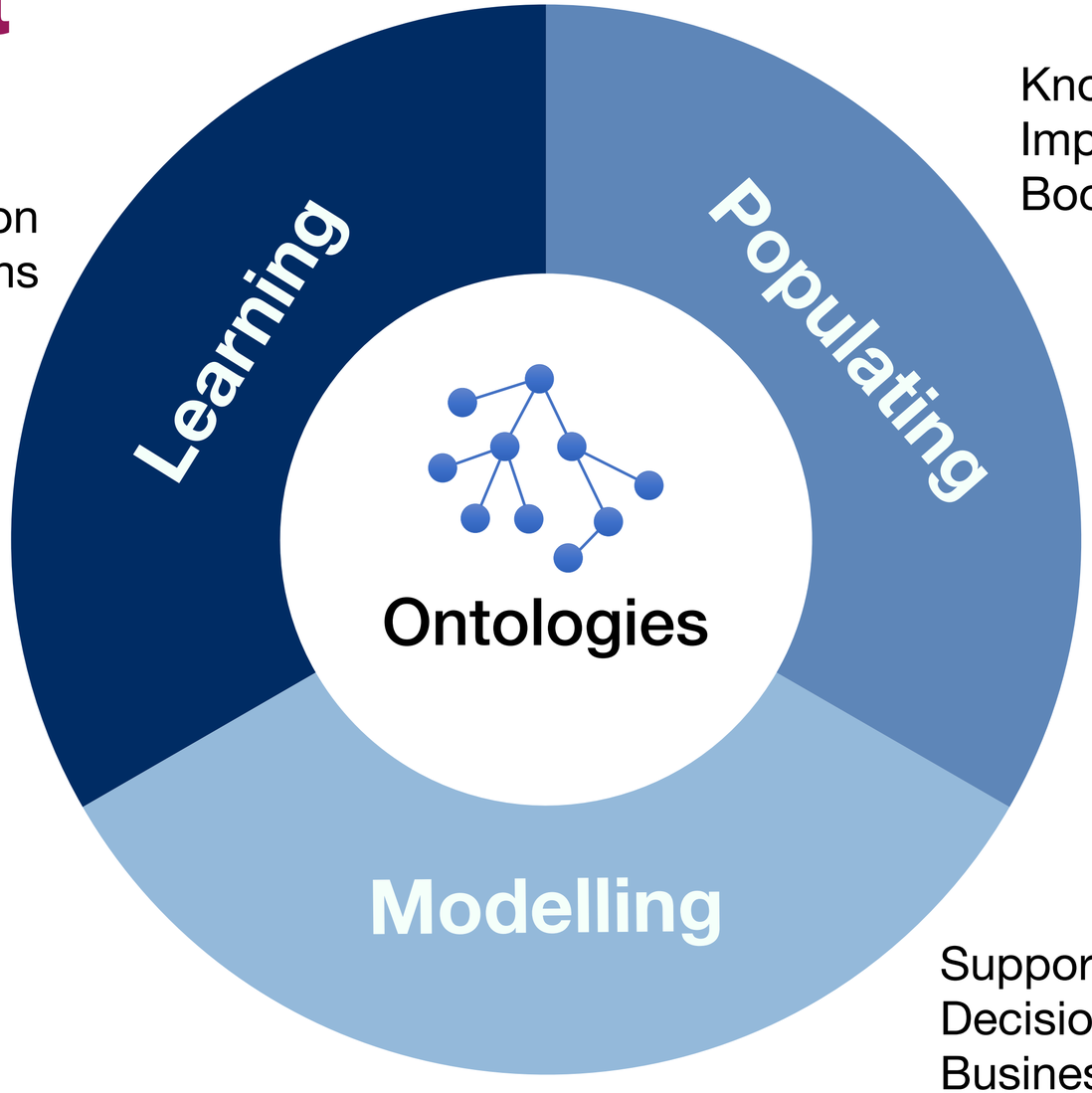
Terminological Extraction
Learning expressive axioms



Supporting collaborative modelling
Decision Support Systems
Business Processes

Agenda

Terminological Extraction
Learning expressive axioms

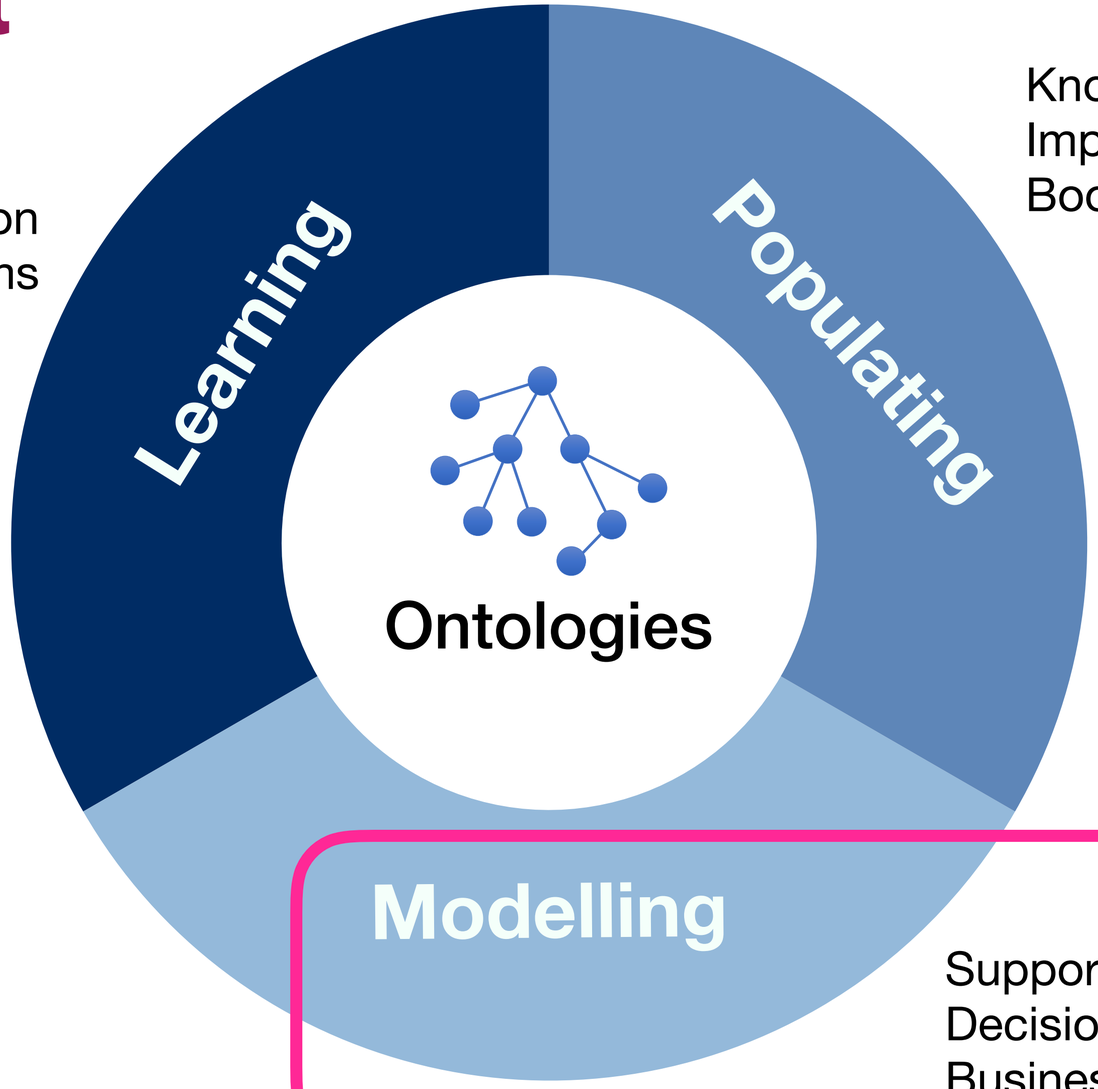


Knowledge Extraction
Improving NLP tasks
Boosting Information Retrieval

Supporting collaborative modelling
Decision Support Systems
Business Processes

Agenda

Terminological Extraction
Learning expressive axioms

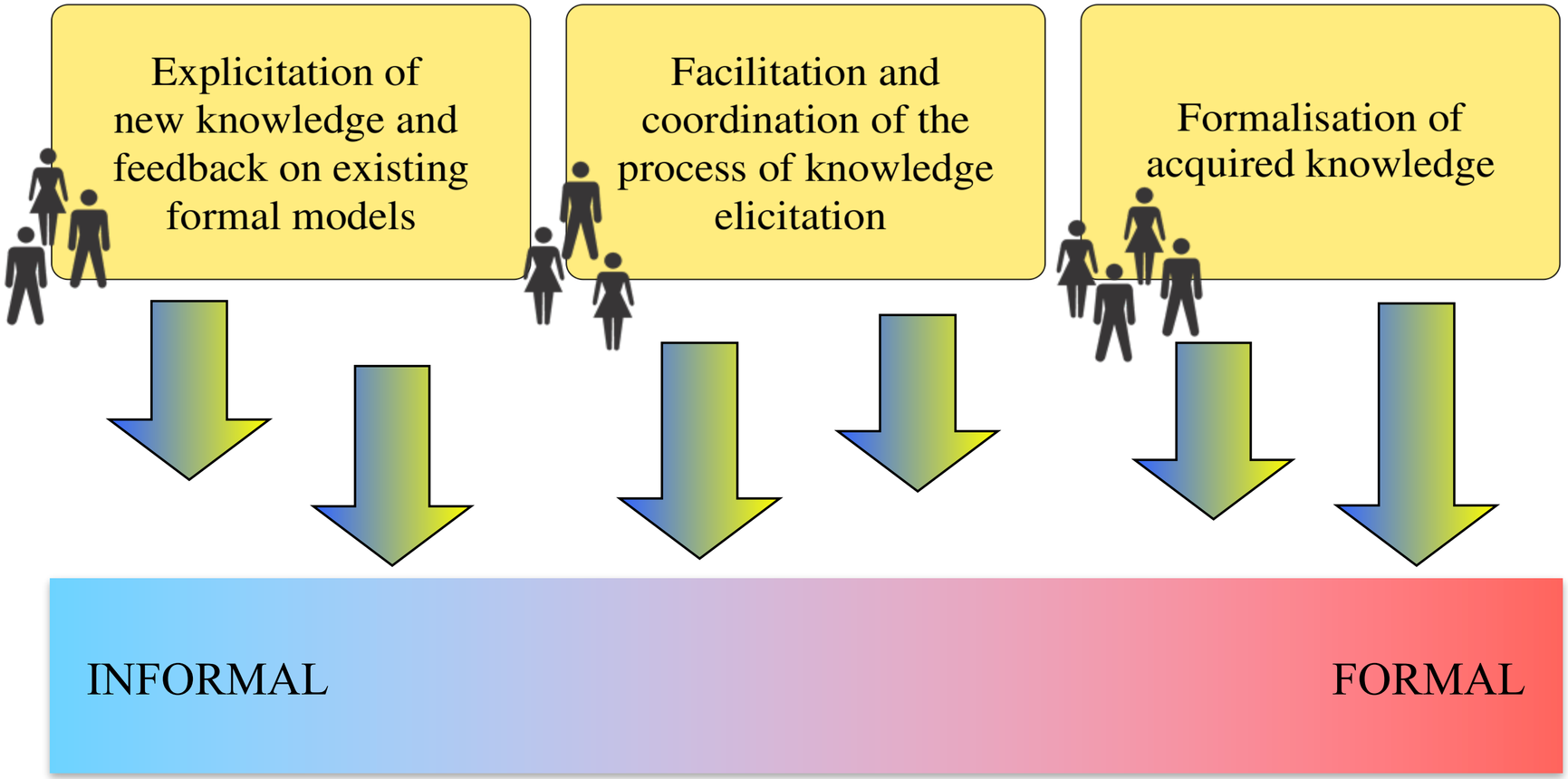


Knowledge Extraction
Improving NLP tasks
Boosting Information Retrieval

Supporting collaborative modelling
Decision Support Systems
Business Processes

Collaborative modelling

aposdle
learn @ work
EU-FP6



MoKi
the Modelling WiKi ---

aposdle
learn @ work
EU-FP6

Wikis offer many collaborative features
revision history, discussions, watchlist, notifications

1 page, 1 element

Concept “Mountain” ⟺

Mountain

A **mountain** is a large landform that stretches above the surrounding land in a limited area usually in the form of a peak. A mountain is generally steeper than a **hill**.

The highest mountain on earth is the Mount Everest

Chiara Ghidini, **Marco Rospocher**, Luciano Serafini:
Modeling in a Wiki with MoKi: Reference Architecture, Implementation, and Usages Int. J. On Advances in Life Sciences, vol. 4(3&4):111–124 (2012)

Chiara Di Francescomarino, Chiara Ghidini, **Marco Rospocher**:
Evaluating Wiki Collaborative Features in Ontology Authoring. IEEE Trans. Knowl. Data Eng. 26(12): 2997-3011 (2014)

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE
Dipartimento di Eccellenza 2018/2022

MoKi
the Modelling WiKi ---

aposdle
learn @ work
EU-FP6

Wikis offer many collaborative features
revision history, discussions, watchlist, notifications

1 page, 1 element

Concept “Mountain” ⟺

Mountain

A **mountain** is a large landform that stretches above the surrounding land in a limited area usually in the form of a peak. A mountain is generally steeper than a **hill**.

The highest mountain on earth is the Mount Everest

(unstructured content)

$\sqsubseteq Landform$

$\sqsubseteq \neg Hill \sqcap \neg Plain$

$\sqsubseteq \forall madeOf(Earth \sqcup Rock)$

$\sqsubseteq \exists height. \geq_{2500}$

$Mountain(Mt.Everest)$

$Mountain(Mt.Kilimanjaro)$

(structured content)

Unstructured + Structured content

Chiara Ghidini, **Marco Rospocher**, Luciano Serafini:
Modeling in a Wiki with MoKi: Reference Architecture, Implementation, and Usages Int. J. On Advances in Life Sciences, vol. 4(3&4):111–124 (2012)

Chiara Di Francescomarino, Chiara Ghidini, **Marco Rospocher**:
Evaluating Wiki Collaborative Features in Ontology Authoring. IEEE Trans. Knowl. Data Eng. 26(12): 2997-3011 (2014)

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE
Dipartimento di Eccellenza 2018/2022

MoKi

the Modelling WiKi ---

aposdle learn @ work

EU-FP6

Different access modes

Wikis offer many collaborative features

revision history, discussions, watchlist, notifications

1 page, 1 element

Concept “Mountain”

Mountain (unstructured view)

A **mountain** is a large landform that stretches above the surrounding land in a limited area usually in the form of a peak. A mountain is generally steeper than a **hill**.

The highest mountain on earth is the Mount Everest

Mountain (semi-structured view)

is a	landform
different from	hill, plain
made of	earth
made of	rock
height	at least 2,500m
samples	Mt. Everest Mt. Kilimanjaro

Mountain

A **mountain** is a large landform that stretches above the surrounding land in a limited area usually in the form of a peak. A mountain is generally steeper than a **hill**.

The highest mountain on earth is the Mount Everest

(unstructured content)

$\sqsubseteq Landform$

$\sqsubseteq \neg Hill \sqcap \neg Plain$

$\sqsubseteq \forall madeOf(Earth \sqcup Rock)$

$\sqsubseteq \exists height. \geq_{2500}$

$Mountain(Mt.Everest)$

$Mountain(Mt.Kilimanjaro)$

(structured content)

Mountain (fully-structured view)

$\sqsubseteq Landform$

$\sqsubseteq \neg Hill \sqcap \neg Plain$

$\sqsubseteq \forall madeOf(Earth \sqcup Rock)$

$\sqsubseteq \exists height. \geq_{2500}$

$Mountain(Mt.Everest)$

$Mountain(Mt.Kilimanjaro)$

Unstructured + Structured content

Chiara Ghidini, **Marco Rospocher**, Luciano Serafini:
Modeling in a Wiki with MoKi: Reference Architecture, Implementation, and Usages Int. J. On Advances in Life Sciences, vol. 4(3&4):111–124 (2012)

Chiara Di Francescomarino, Chiara Ghidini, **Marco Rospocher**:
Evaluating Wiki Collaborative Features in Ontology Authoring. IEEE Trans. Knowl. Data Eng. 26(12): 2997-3011 (2014)

UNIVERSITÀ di VERONA | Dipartimento di LINGUE E LETTERATURE STRANIERE

MoKi

the Modelling WiKi ---

OncoCure

Fond. CARITRO

Edit Plan with Children: Premen-cHR-HER2pos

Description

Description: The diagram represents the recommended treatments for patients in post menopause, with hormone responsive breast cancer and positive cErb2 receptors (Her2 gene over-expressed).
In this diagram, patients are further divided in intermediate and high risk groups, eligible for different treatments, according to conditions on the number of regional lymph nodes with metastasis.

Source

Documents: Protocol Version 2.1, pag 17

Sort key: 170

Plan Attributes

Plan Title: premenopause, certain hormone response, pos

Conditions

Filter Condition: not postmenopause and hormone-responsiver

Setup Condition:

Complete Condition:

Abort Condition:

Suspend Condition:

Reactivate Condition:

Plan Hierarchy View

- Top-level
  - Adjuvant-therapy
    - Adjuvant-phase1
      - Femara
  - Neo-adjuvant-therapy
    - Neo-adjuvant-HER2neg
      - Neo-adjuvant-HER2neg-chemo
        - AC-4
        - Optional-TXT
          - TXT-4
            - TAXOTERE
        - Cape-TXT
      - Cape-TXT
    - Neo-adjuvant-HER2pos
      - NEO-ALLTO
      - Neo-adjuvant-HER2pos-conventional
        - ATH-3
        - TH-3
        - CMFH-4
  - Treatment-of-metastases
  - Treatment-of-relapes
- Bad-stuff

Legend

- Plan with Children with sequential subplans
- Plan with Children with parallel subplans
- Plan with Children with any-order subplans
- Plan with Children with unordered subplans
- Cyclical Plan
- ACT — Plan Activation Plan
  - Green color represents the plan that is to be activated
  - Orange color represents the plan that is to be executed on suspend
  - Red color represents the plan that is to be executed on abort
- Variable Assignment Plan
- ? — Ask Plan
- IF — If Then Else Plan
  - First subplan of this plan is IF condition plan
  - Second subplan of this plan is ELSE condition plan
- User Performed Plan
- Plan yet to be modelled

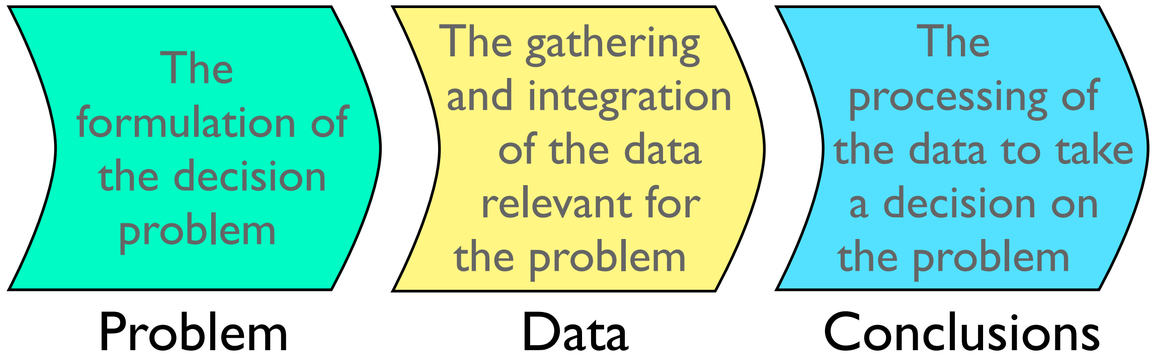
Claudio Eccher, Antonella Ferro, Andreas Seyfang, **Marco Rospocher,** Silvia Miksch:
Modeling Clinical Protocols Using Semantic MediaWiki: The Case of the Oncocure Project. K4HeIP 2008: 42-54 (2008)

Marco Rospocher, Claudio Eccher, Chiara Ghidini, Rakebul Hasan, Andreas Seyfang, Antonella Ferro, Silvia Miksch:
Collaborative Encoding of Asbru Clinical Protocols,Electronic Healthcare (eHealth2010), LNCS, 135–143, (2012)

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE

Ontologies for DSS

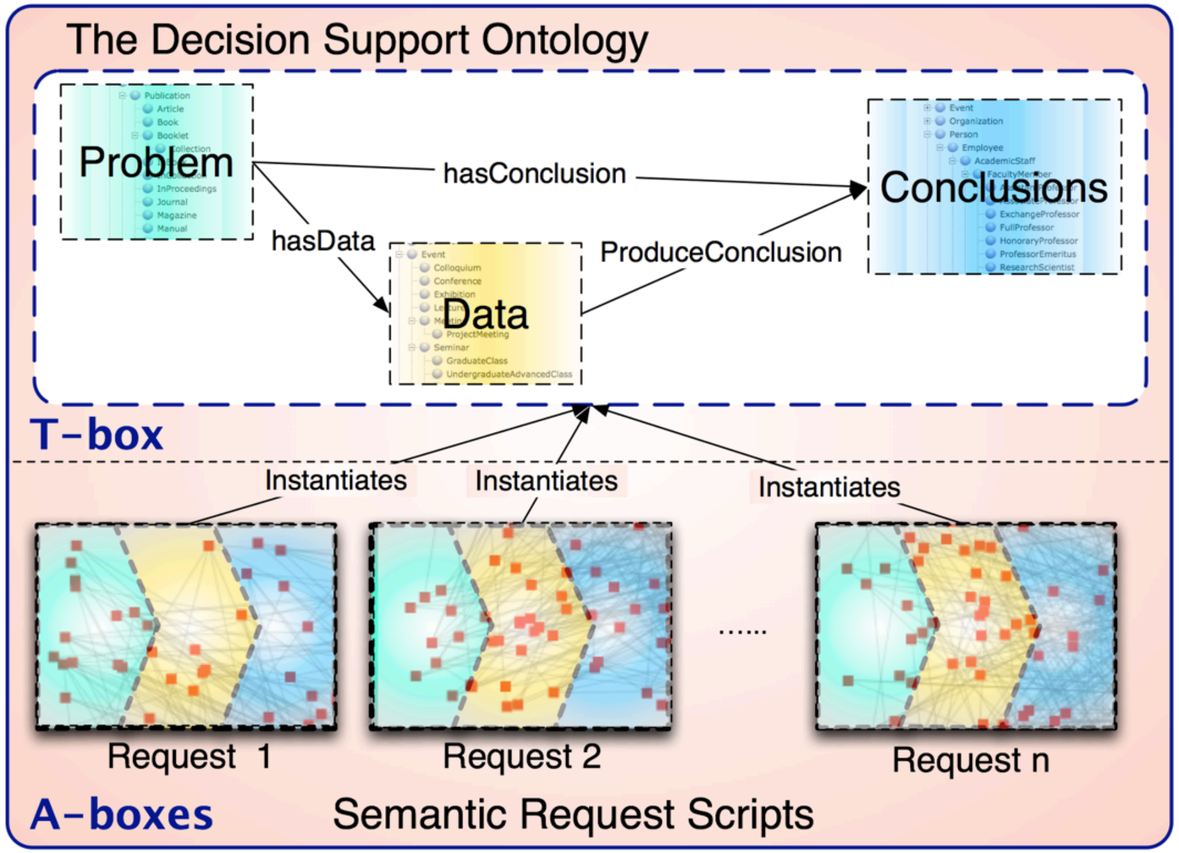
- The decision making process of a Decision Support System (DSS) typically consists of three phases:



- Use-case: a multilingual web-service platform providing personalized environmental information and decision support
  - A pollen allergic person, planning to do some outdoor activities, interested in being notified of potentially harmful environmental conditions
  - A city administrator, to be informed whether the current air quality situation requires some actions to be urgently taken.

Ontologies for DSS

- An ontology-based knowledge base as the main (enhanced) data structure of a DSS
  - T-box:
    - formally represents the content manipulated in the three phases
      - Problem: type of requests, user profile, data to collect
      - Data: data accessed and manipulated by the DSS for the request
      - Conclusions: output produced by the DSS (e.g. warnings/suggestions, data analysis results), decision trace
  - A-box(es):
    - each request submitted to the system corresponds to a single incrementally-built A-Box (a “semantic request script (SRS)”)



Ontologies for DSS

EU-FP7

- A SRS provides a complete “semantic” snapshot of all the information processed and produced by the DSS for a request, with “explanations”
- A (multi-lingual) natural language report can be automatically generated from it
  - especially appreciated by laymen, media corporations, ...

Request Example: a city-admin monitoring the air quality situation in the area she is responsible of

DSS output:

> Situation in the selected area between 08h00 and 20h00 of 07/05/2012. The ozone warning threshold value (240g/m3) was exceeded between 13h00 and 14h00 (247g/m3), the ozone information threshold value (180g/m3) between 12h00 and 13h00 (208g/m3) and between 14h00 and 15h00 (202g/m3). The minimum temperature was 2C and the maximum temperature 17C. The wind was weak (S). There is no data available for carbon monoxide, rain and humidity.
>
> **Ozone warning**: ozone irritates eyes and the mucous membranes of nose and throat. It may also exacerbate allergy symptoms caused by pollen. Persons with respiratory diseases may experience increased coughing and shortness of breath and their functional capacity may weaken. Sensitive groups, like children, asthmatics of all ages and elderly persons suffering from coronary heart disease or chronic obstructive pulmonary disease, may experience symptoms. [...]

Leo Wanner, **Marco Rospocher**, Stefanos Vrochidis, Lasse Johansson, Nadjet B. Agha, Gerard Casamayor, Ari Karppinen, Ioannis Kompatsiaris, Simon Mille, A. Moumtzidou, Luciano Serafini: Ontology-centered environmental information delivery for personalized decision support. Expert Syst. Appl. 42(12): 5032-5046 (2015)

Ontologies for DSS

Edizioni Erickson

ePlanning

TN-FSR

- Individual Education Plan (IEP)

  "*a document that describes integrated and balanced interventions, prepared for students with disabilities in a given period of time, for the purpose of executing the right to education and training*" [Italian Law 104 / 1992]

- A DSS exploiting an ontology, aligned with ICF and ICD10, modelling:
  - functional and cognitive abilities
  - educational goals
  - supporting activities and materials
- Commercially exploited by Edizioni Erickson

SOFIA https://sofia.erickson.it/

Marco Rospocher, Elena Cardillo, Ivan Donadello, Luciano Serafini:
On the Collaborative Development of Application Ontologies: A Practical Case Study with a SME. EKAW 2014: 469-484 (2014)

Sofia Cramerotti, Massimo Turrini, Marco Buccio, Silvia Larentis, **Marco Rospocher**, Luciano Serafini, Elena Cardillo and Ivan Donadello:
ePlanning: an Ontology-based System for Building Individualized Education Plans for Students with Special Educational Needs, MED -- Media Education, vol. 6(1):101–110, Ed. Erickson (2015)

Dipartimento di Eccellenza 2018/2022

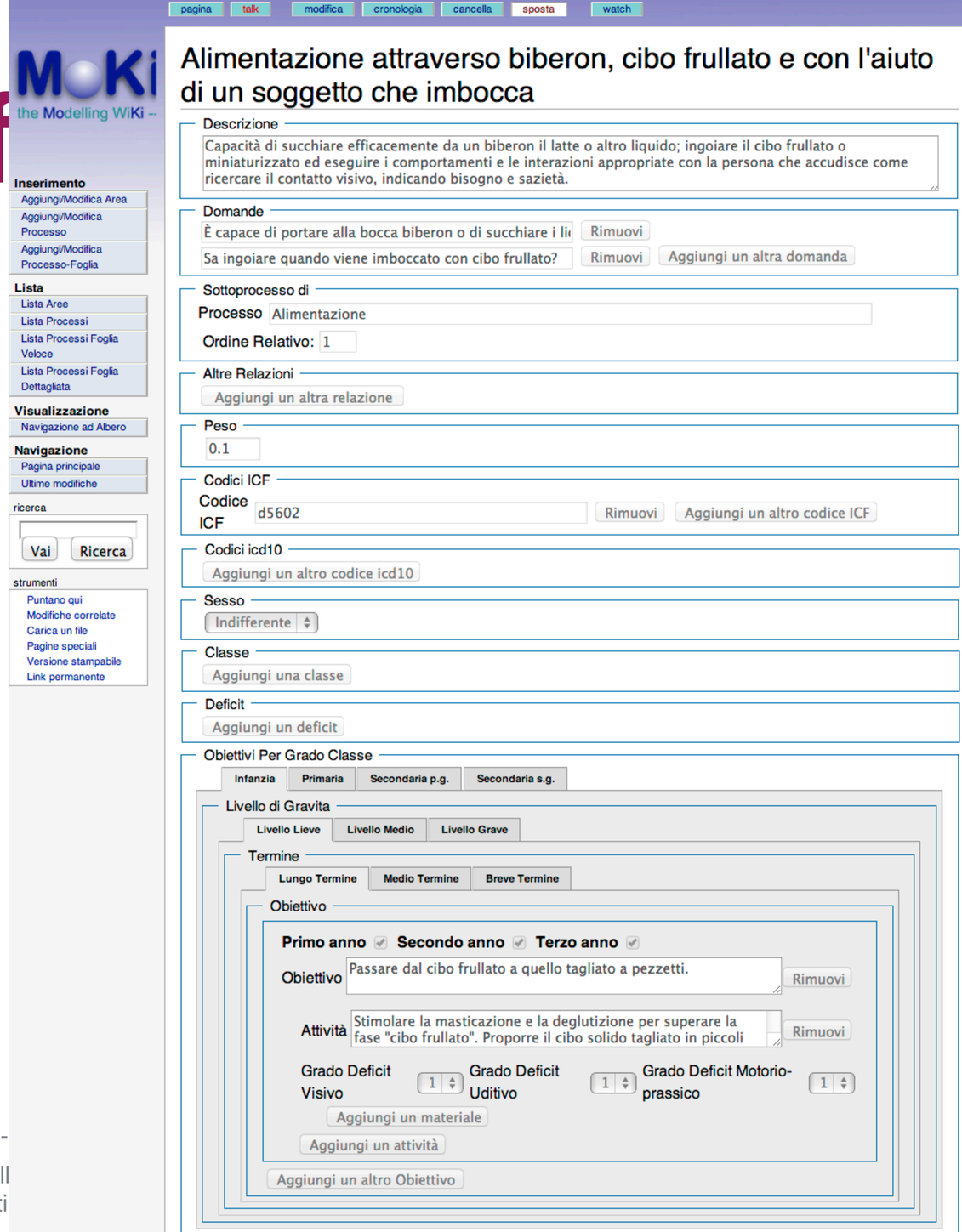
Ontologies f

- Individual Education Plan (IEP)

  "a document that describes integrated and balanced interventions, prepared for students with disabilities in a given period of time, for the purpose of executing the right to education and training" [Italian Law 104 / 1992]

- A DSS exploiting an ontology, aligned with ICF and ICD10, modelling:
  - functional and cognitive abilities
  - educational goals
  - supporting activities and materials
- Commercially exploited by Edizioni Erickson

SOFIA https://sofia.erickson.it/

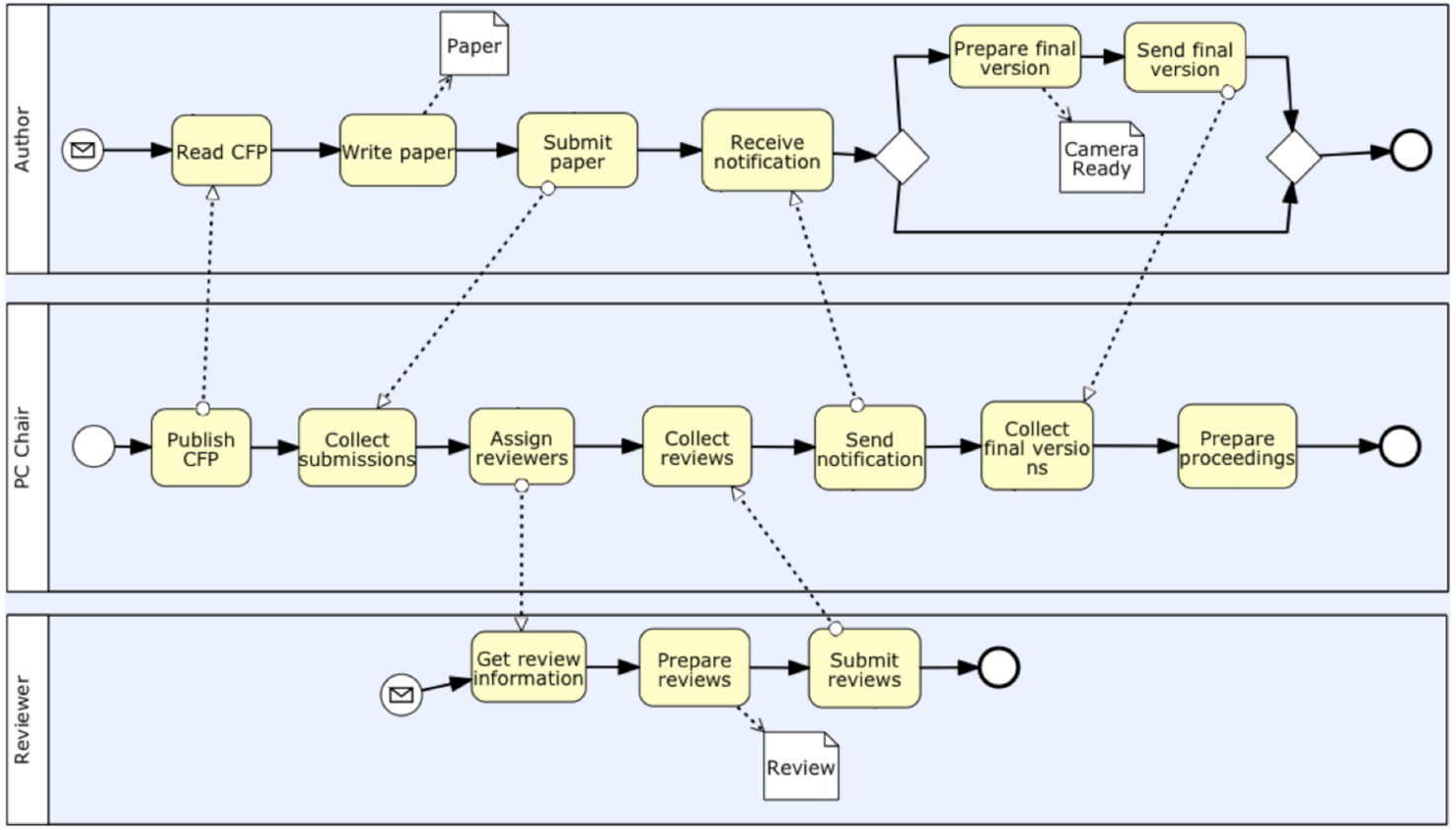


Marco Rospocher, Elena Cardillo, Ivan Donadello, Luciano Serafini:
On the Collaborative Development of Application Ontologies: A Practical Case Study with a SME. EKAW 2014: 469-

Sofia Cramerotti, Massimo Turrini, Marco Buccio, Silvia Larentis, **Marco Rospocher**, Luciano Serafini, Elena Cardill
ePlanning: an Ontology-based System for Building Individualized Education Plans for Students with Special Educati

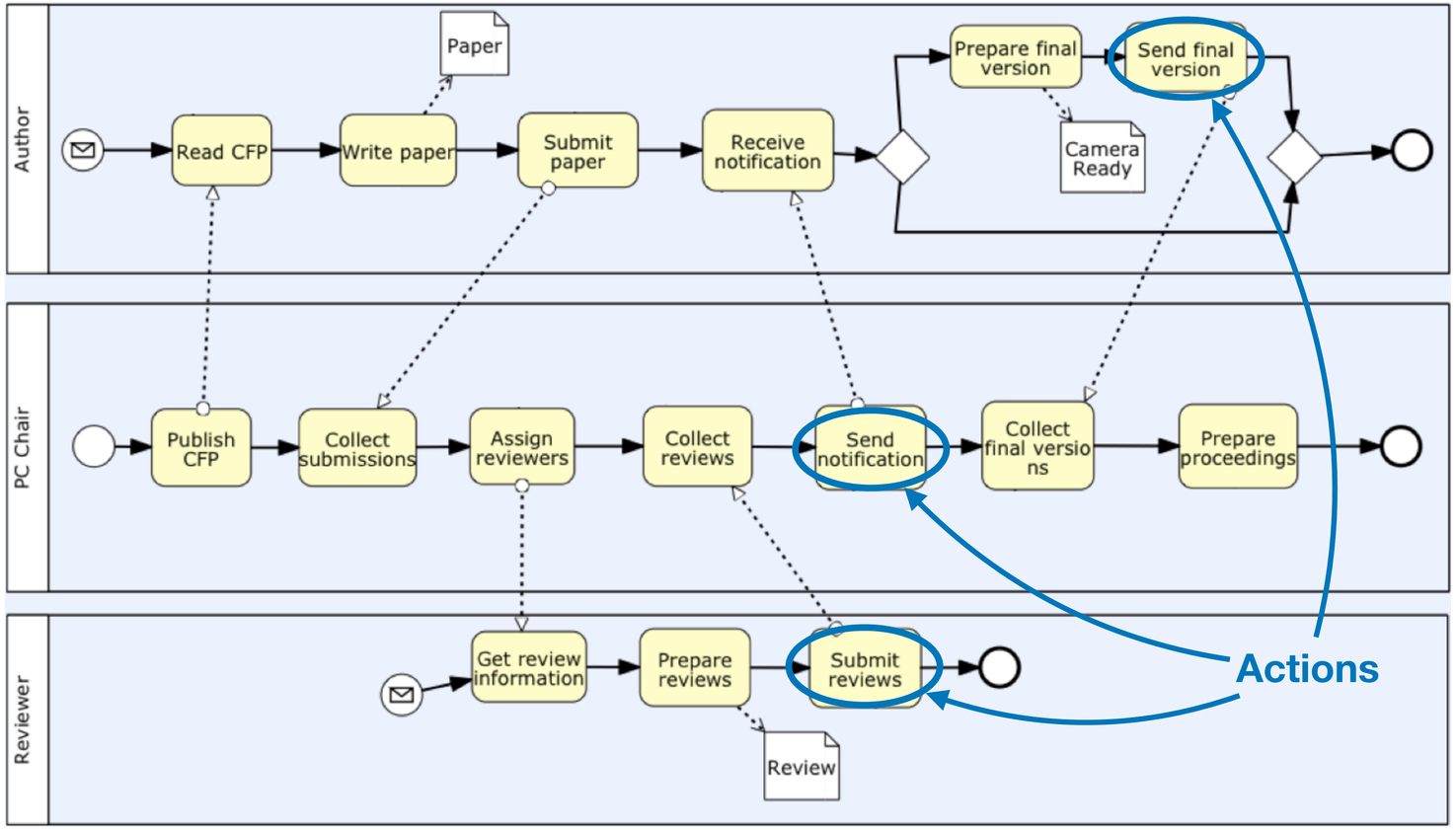
Reasoning on Business Processes

- A lot of semantics is not fully captured by process languages (e.g. BPMN)



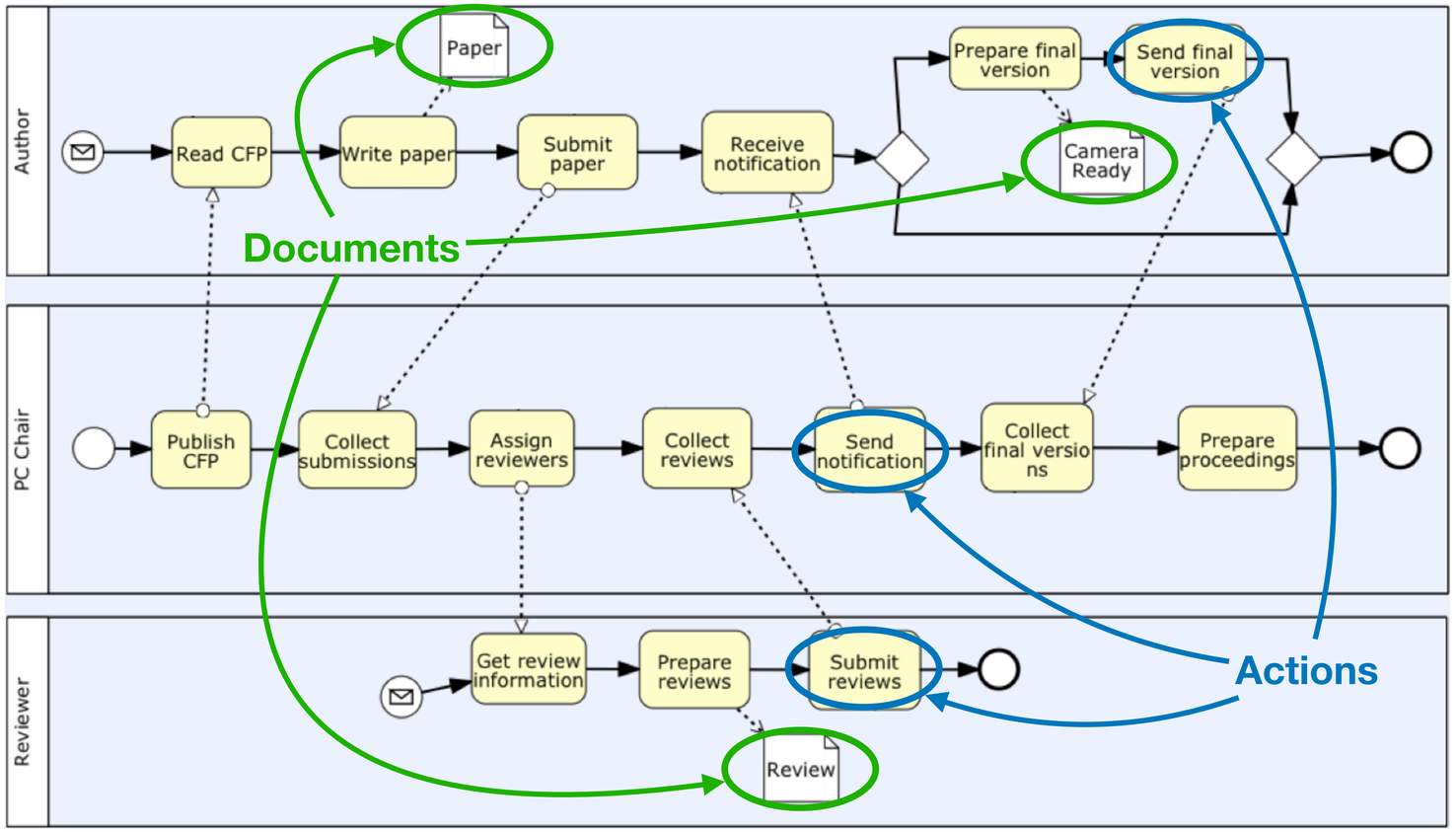
Reasoning on Business Processes

- A lot of semantics is not fully captured by process languages (e.g. BPMN)



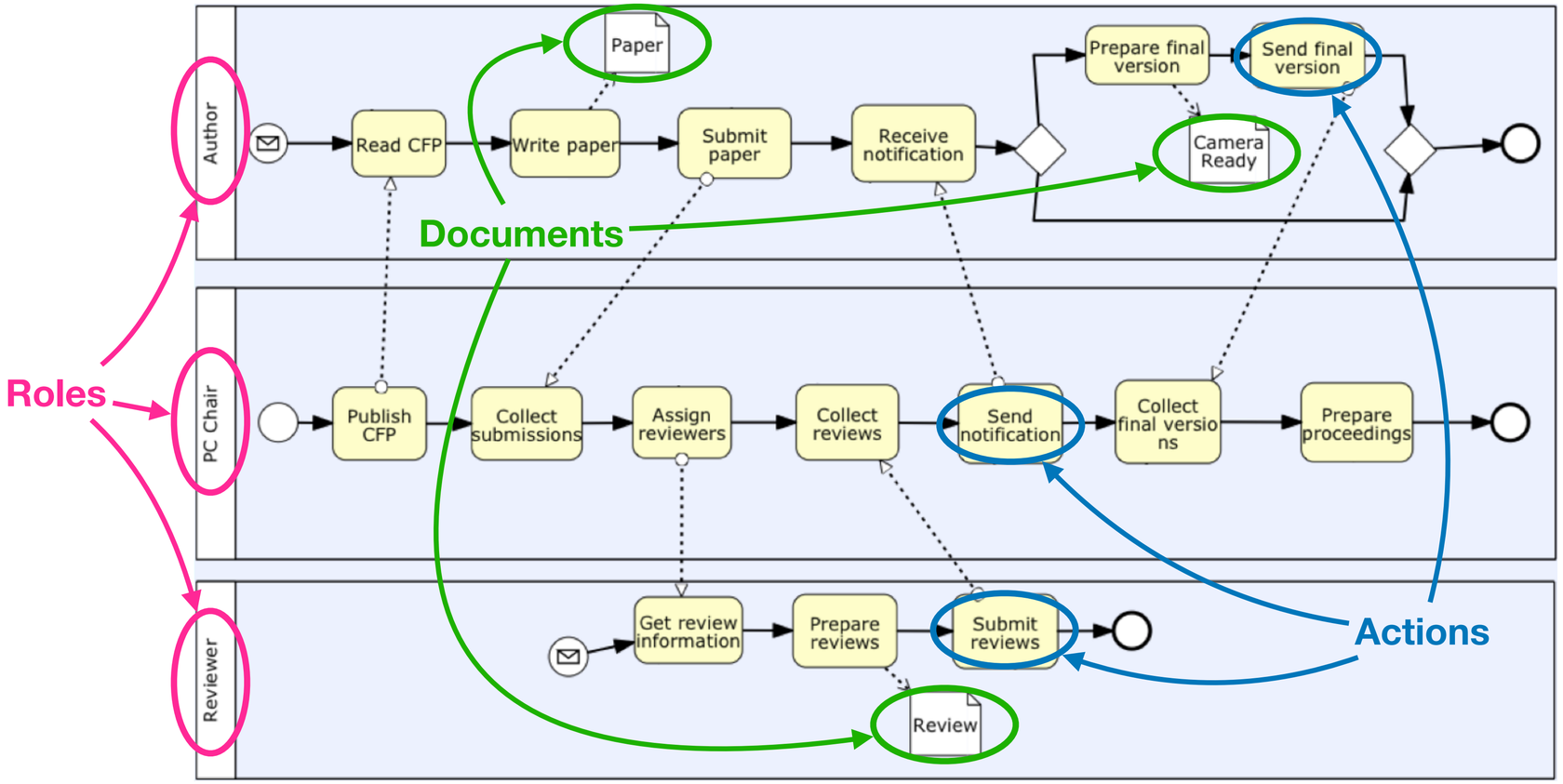
Reasoning on Business Processes

- A lot of semantics is not fully captured by process languages (e.g. BPMN)



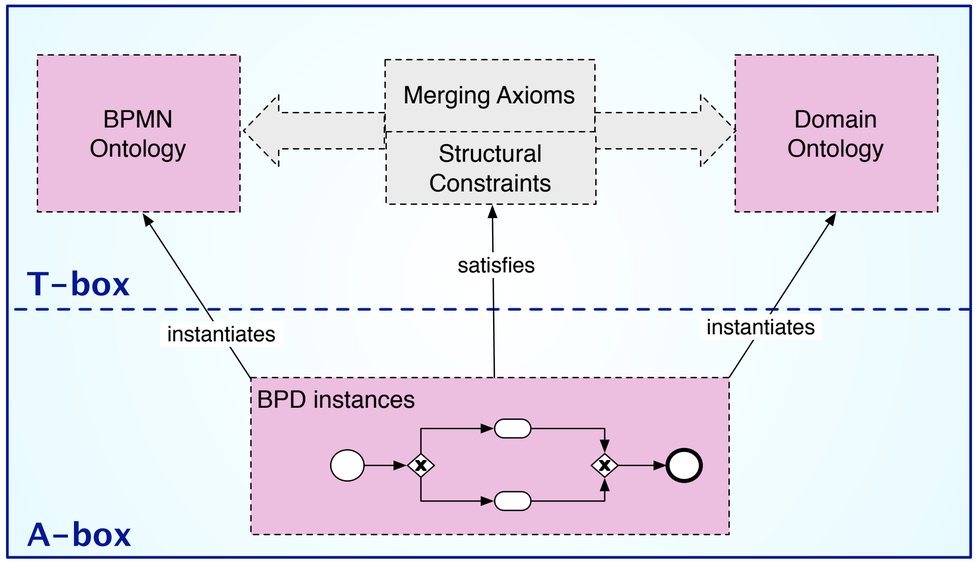
Reasoning on Business Processes

- A lot of semantics is not fully captured by process languages (e.g. BPMN)



Reasoning on Business Processes

- “Semantically annotated” business processes are encoded into an OWL KB



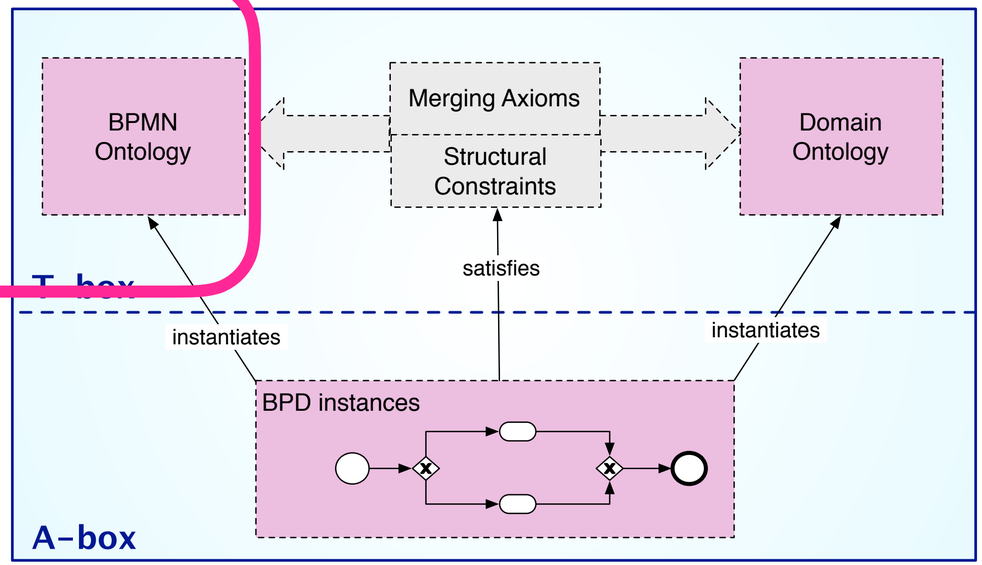
Chiara Di Francescomarino, Chiara Ghidini, **Marco Rospocher**, Luciano Serafini, Paolo Tonella:
Semantically-Aided Business Process Modeling. International Semantic Web Conference 2009: 114-129 (2009)

Chiara Ghidini, Chiara Di Francescomarino, **Marco Rospocher,** Paolo Tonella, Luciano Serafini:
Semantics-Based Aspect-Oriented Management of Exceptional Flows in Business Processes. IEEE Trans. Systems, Man, and Cybernetics, Part C 42(1): 25-37 (2012)

Reasoning on Business Processes

- “Semantically annotated” business processes are encoded into an OWL KB

OWL formalisation of the structural part of BPMN
- the classification of all the elements of the BPMN language
- the formal representation of the attributes and conditions describing how the elements can be combined to obtain a “valid” BPMN business process



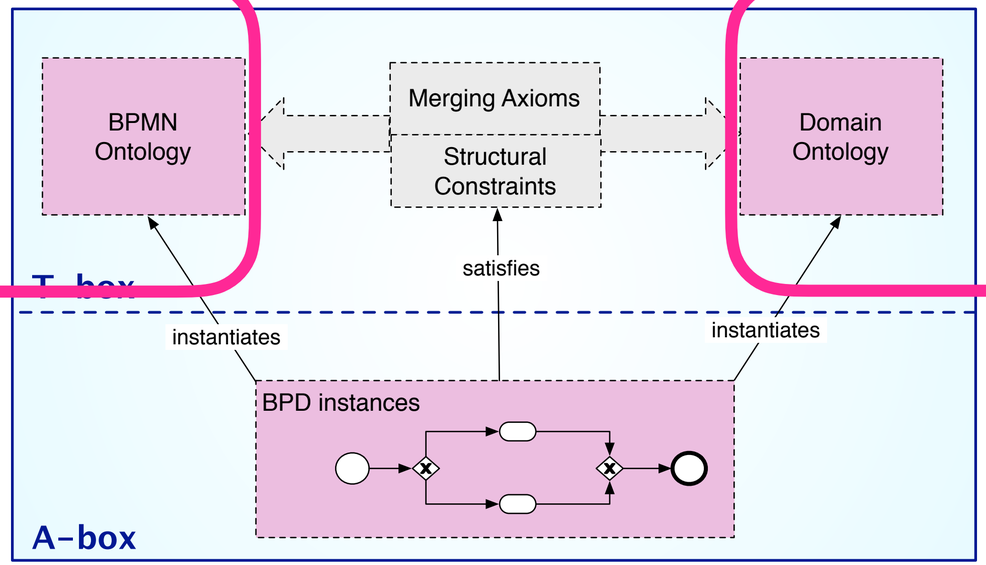
Chiara Di Francescomarino, Chiara Ghidini, **Marco Rospocher**, Luciano Serafini, Paolo Tonella:
Semantically-Aided Business Process Modeling. International Semantic Web Conference 2009: 114-129 (2009)

Chiara Ghidini, Chiara Di Francescomarino, **Marco Rospocher,** Paolo Tonella, Luciano Serafini:
Semantics-Based Aspect-Oriented Management of Exceptional Flows in Business Processes. IEEE Trans. Systems, Man, and Cybernetics, Part C 42(1): 25-37 (2012)

Reasoning on Business Processes

- “Semantically annotated” business processes are encoded into an OWL KB

OWL formalisation of the structural part of BPMN
- the classification of all the elements of the BPMN language
- the formal representation of the attributes and conditions describing how the elements can be combined to obtain a “valid” BPMN business process



Represents the (specific) business domain
Used to annotate the elements of the business process diagram

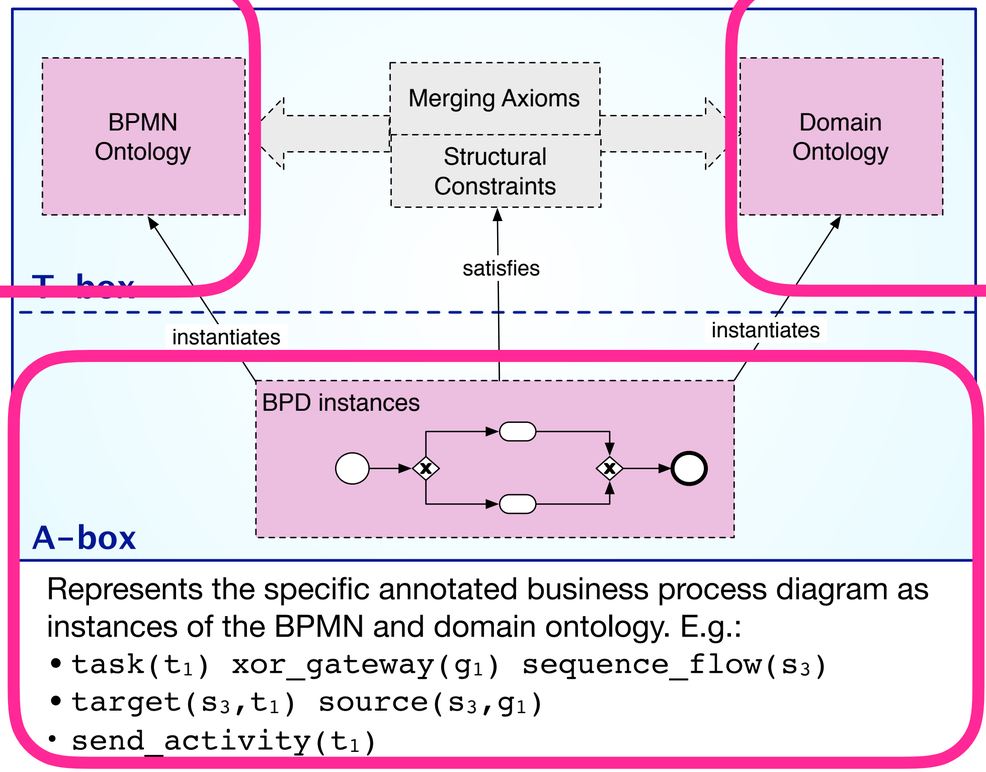
Chiara Di Francescomarino, Chiara Ghidini, **Marco Rospocher**, Luciano Serafini, Paolo Tonella:
Semantically-Aided Business Process Modeling. International Semantic Web Conference 2009: 114-129 (2009)

Chiara Ghidini, Chiara Di Francescomarino, **Marco Rospocher,** Paolo Tonella, Luciano Serafini:
Semantics-Based Aspect-Oriented Management of Exceptional Flows in Business Processes. IEEE Trans. Systems, Man, and Cybernetics, Part C 42(1): 25-37 (2012)

Reasoning on Business Processes

- “Semantically annotated” business processes are encoded into an OWL KB

OWL formalisation of the structural part of BPMN
- the classification of all the elements of the BPMN language
- the formal representation of the attributes and conditions describing how the elements can be combined to obtain a “valid” BPMN business process



Represents the (specific) business domain
Used to annotate the elements of the business process diagram

Represents the specific annotated business process diagram as instances of the BPMN and domain ontology. E.g.:
- $task(t_1)$ $xor\_gateway(g_1)$ $sequence\_flow(s_3)$
- $target(s_3,t_1)$ $source(s_3,g_1)$
- $send\_activity(t_1)$

Chiara Di Francescomarino, Chiara Ghidini, **Marco Rospocher**, Luciano Serafini, Paolo Tonella: Semantically-Aided Business Process Modeling. International Semantic Web Conference 2009: 114-129 (2009)

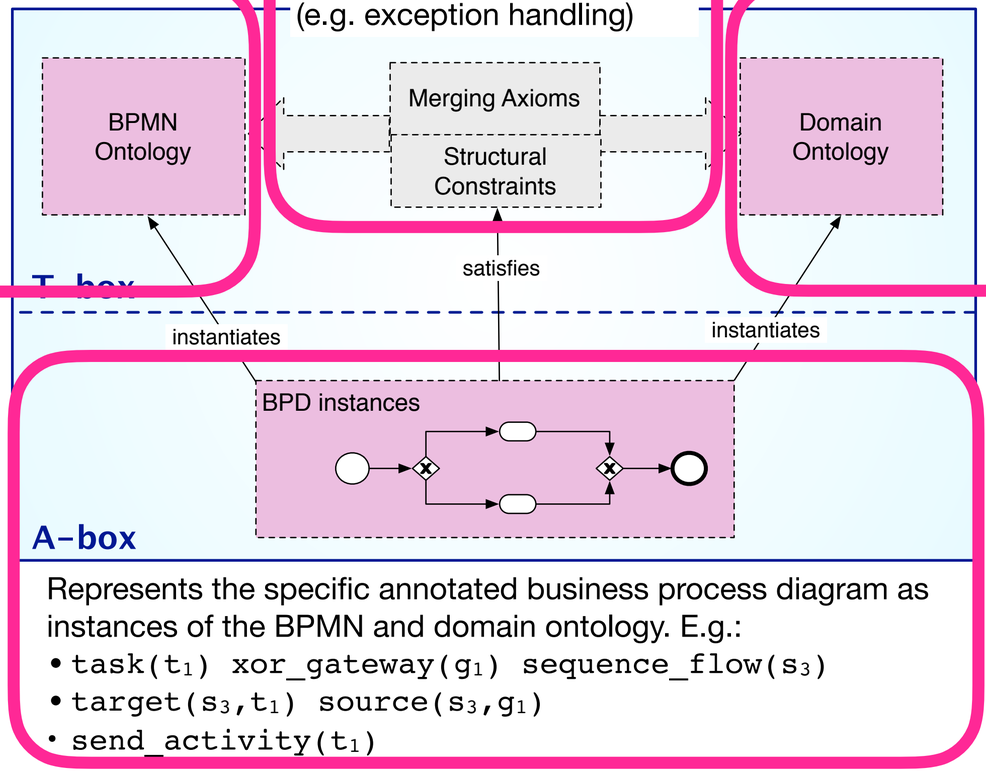
Chiara Ghidini, Chiara Di Francescomarino, **Marco Rospocher,** Paolo Tonella, Luciano Serafini: Semantics-Based Aspect-Oriented Management of Exceptional Flows in Business Processes. IEEE Trans. Systems, Man, and Cybernetics, Part C 42(1): 25-37 (2012)

Reasoning on Business Processes

- “Semantically annotated” business processes are encoded into an OWL KB

OWL formalisation of the structural part of BPMN
- the classification of all the elements of the BPMN language
- the formal representation of the attributes and conditions describing how the elements can be combined to obtain a “valid” BPMN business process

Represent correct/incorrect annotations
Express valid critical patters (e.g. exception handling)



Represents the (specific) business domain
Used to annotate the elements of the business process diagram

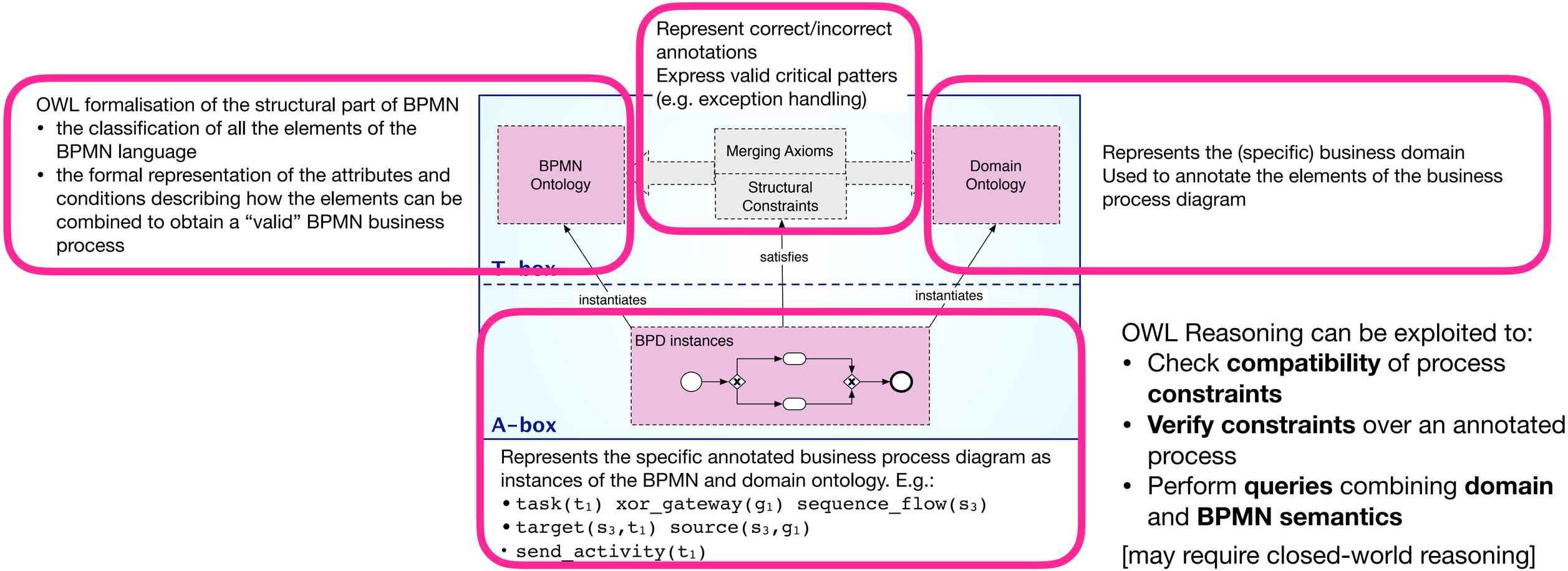
Represents the specific annotated business process diagram as instances of the BPMN and domain ontology. E.g.:
- $task(t_1)$ $xor\_gateway(g_1)$ $sequence\_flow(s_3)$
- $target(s_3,t_1)$ $source(s_3,g_1)$
- $send\_activity(t_1)$

Chiara Di Francescomarino, Chiara Ghidini, **Marco Rospocher**, Luciano Serafini, Paolo Tonella: Semantically-Aided Business Process Modeling. International Semantic Web Conference 2009: 114-129 (2009)

Chiara Ghidini, Chiara Di Francescomarino, **Marco Rospocher,** Paolo Tonella, Luciano Serafini: Semantics-Based Aspect-Oriented Management of Exceptional Flows in Business Processes. IEEE Trans. Systems, Man, and Cybernetics, Part C 42(1): 25-37 (2012)

Reasoning on Business Processes

- "Semantically annotated" business processes are encoded into an OWL KB



OWL Reasoning can be exploited to:

- Check **compatibility** of process **constraints**
- **Verify constraints** over an annotated process
- Perform **queries** combining **domain** and **BPMN semantics**

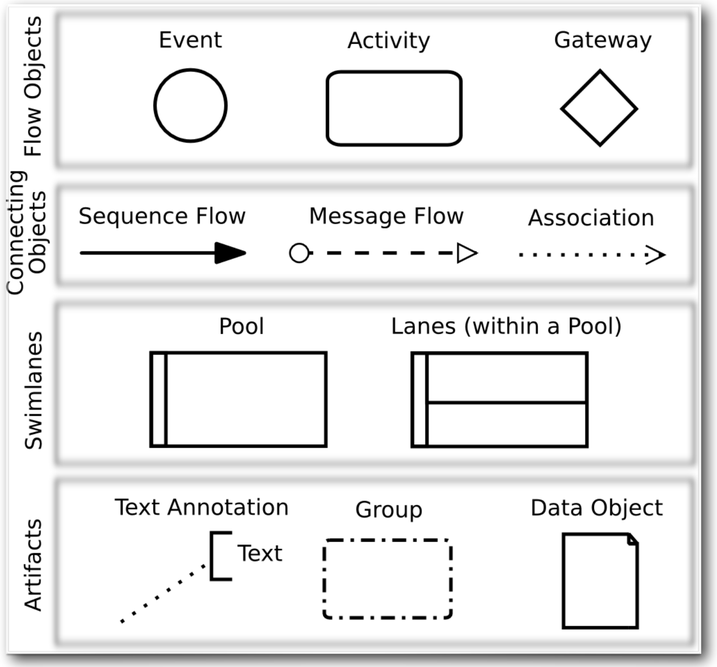
[may require closed-world reasoning]

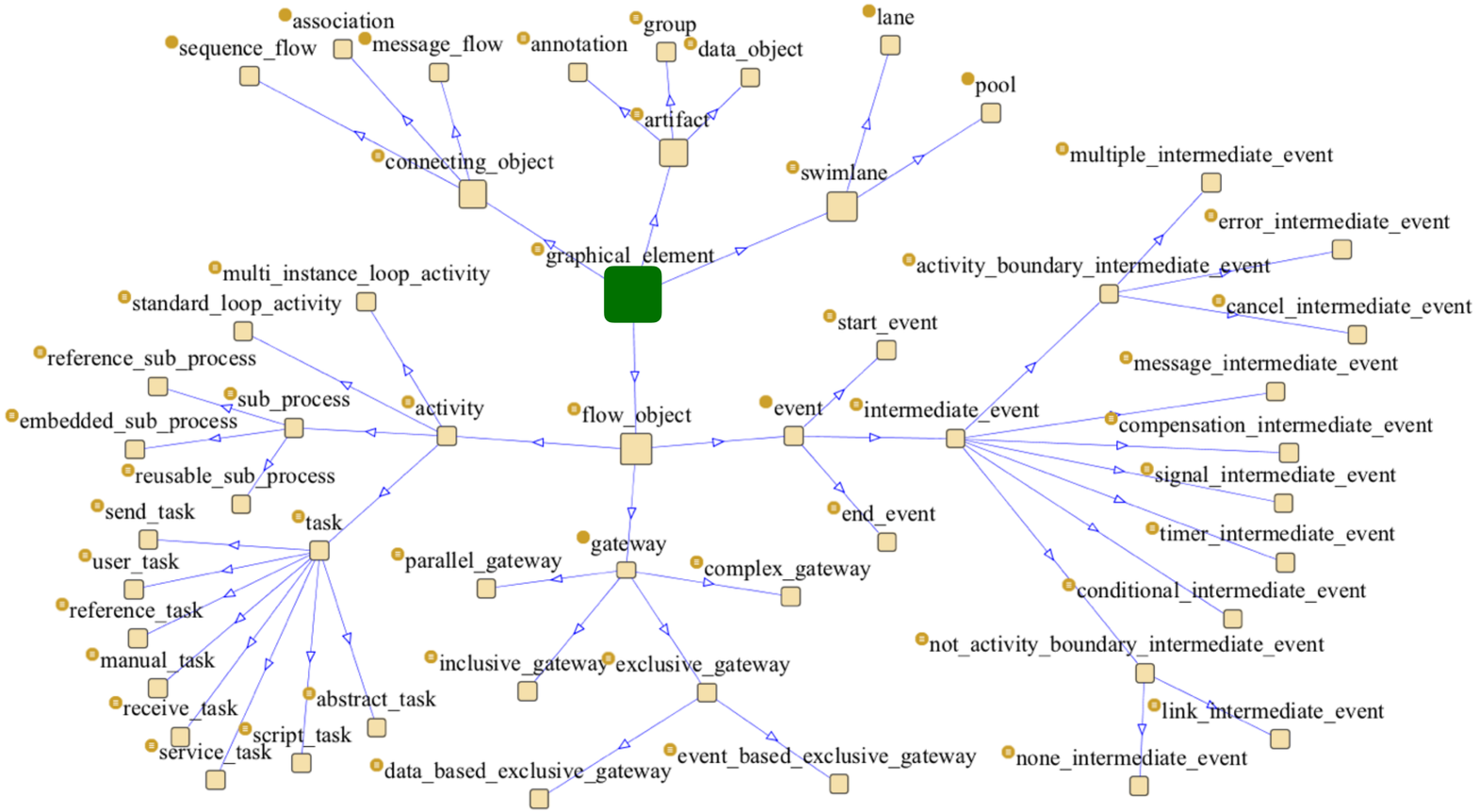
Chiara Di Francescomarino, Chiara Ghidini, **Marco Rospocher**, Luciano Serafini, Paolo Tonella: Semantically-Aided Business Process Modeling. International Semantic Web Conference 2009: 114-129 (2009)

Chiara Ghidini, Chiara Di Francescomarino, **Marco Rospocher,** Paolo Tonella, Luciano Serafini: Semantics-Based Aspect-Oriented Management of Exceptional Flows in Business Processes. IEEE Trans. Systems, Man, and Cybernetics, Part C 42(1): 25-37 (2012)

BMPN ontology

download: https://dkm.fbk.eu/bpmn-ontology

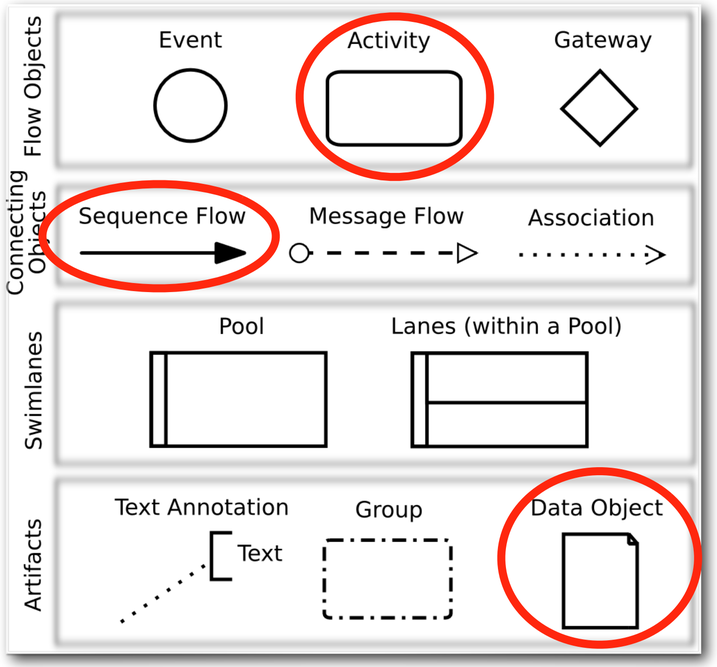


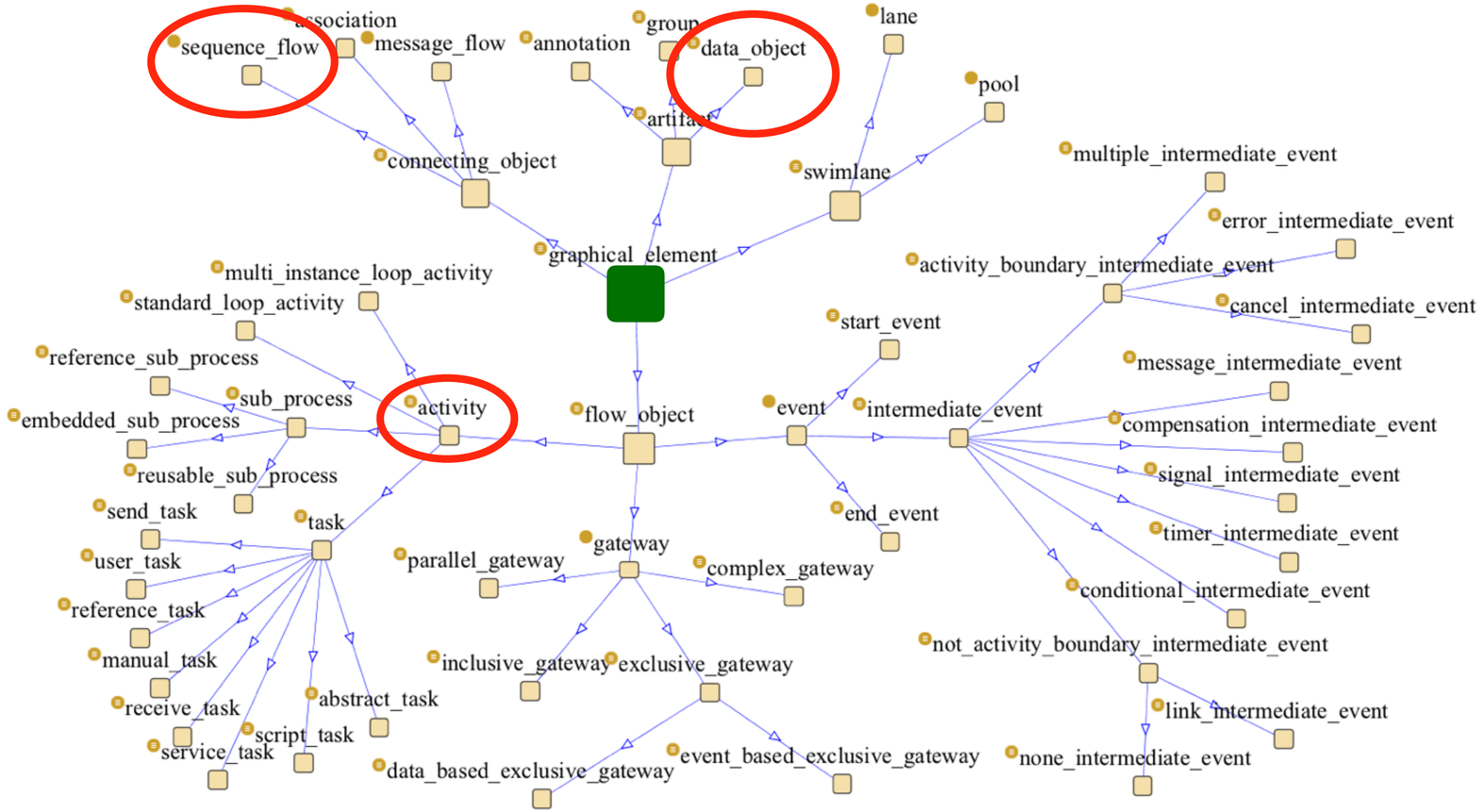


Marco Rospocher, Chiara Ghidini, Luciano Serafini:
An ontology for the Business Process Modelling Notation. FOIS 2014: 133-146 (2014)

BMPN ontology

download: https://dkm.fbk.eu/bpmn-ontology



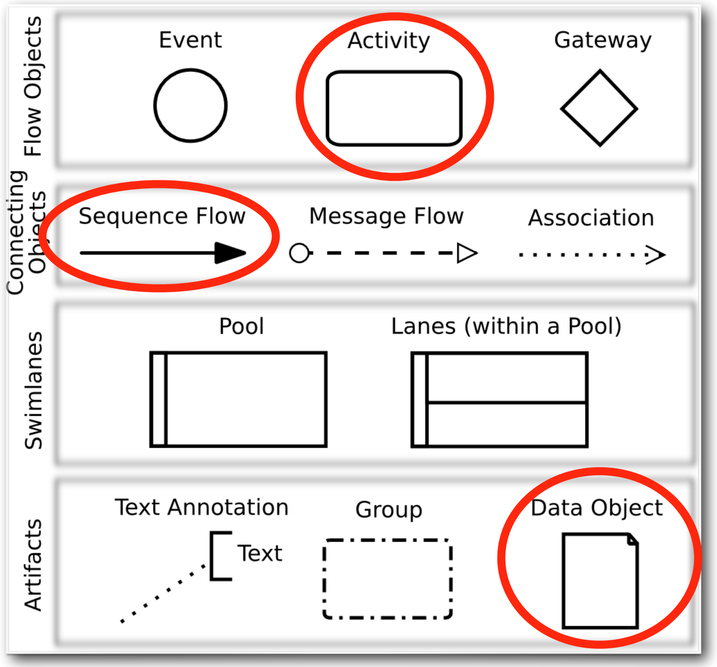


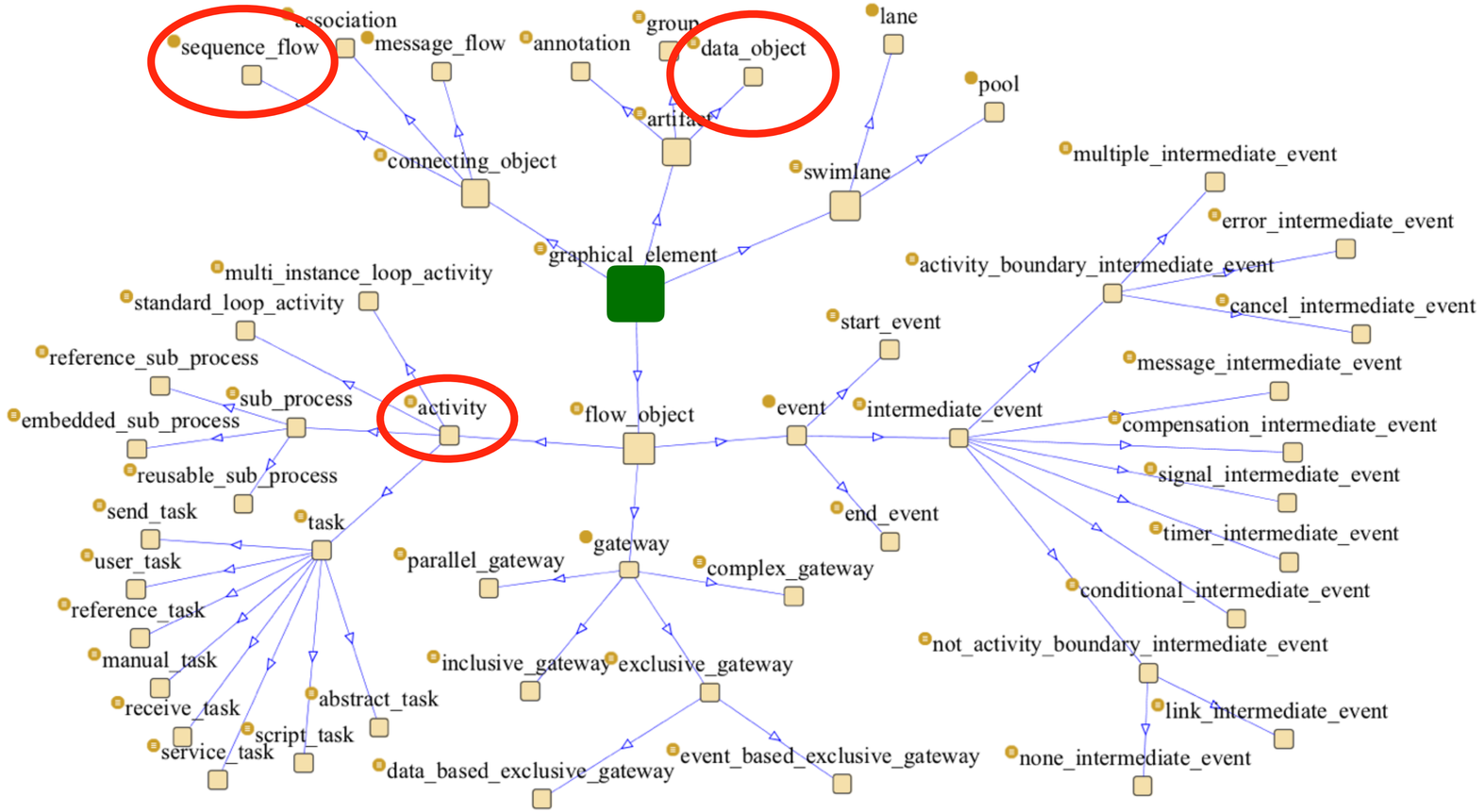
Marco Rospocher, Chiara Ghidini, Luciano Serafini:
An ontology for the Business Process Modelling Notation. FOIS 2014: 133-146 (2014)

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE

BMPN ontology

download: https://dkm.fbk.eu/bpmn-ontology



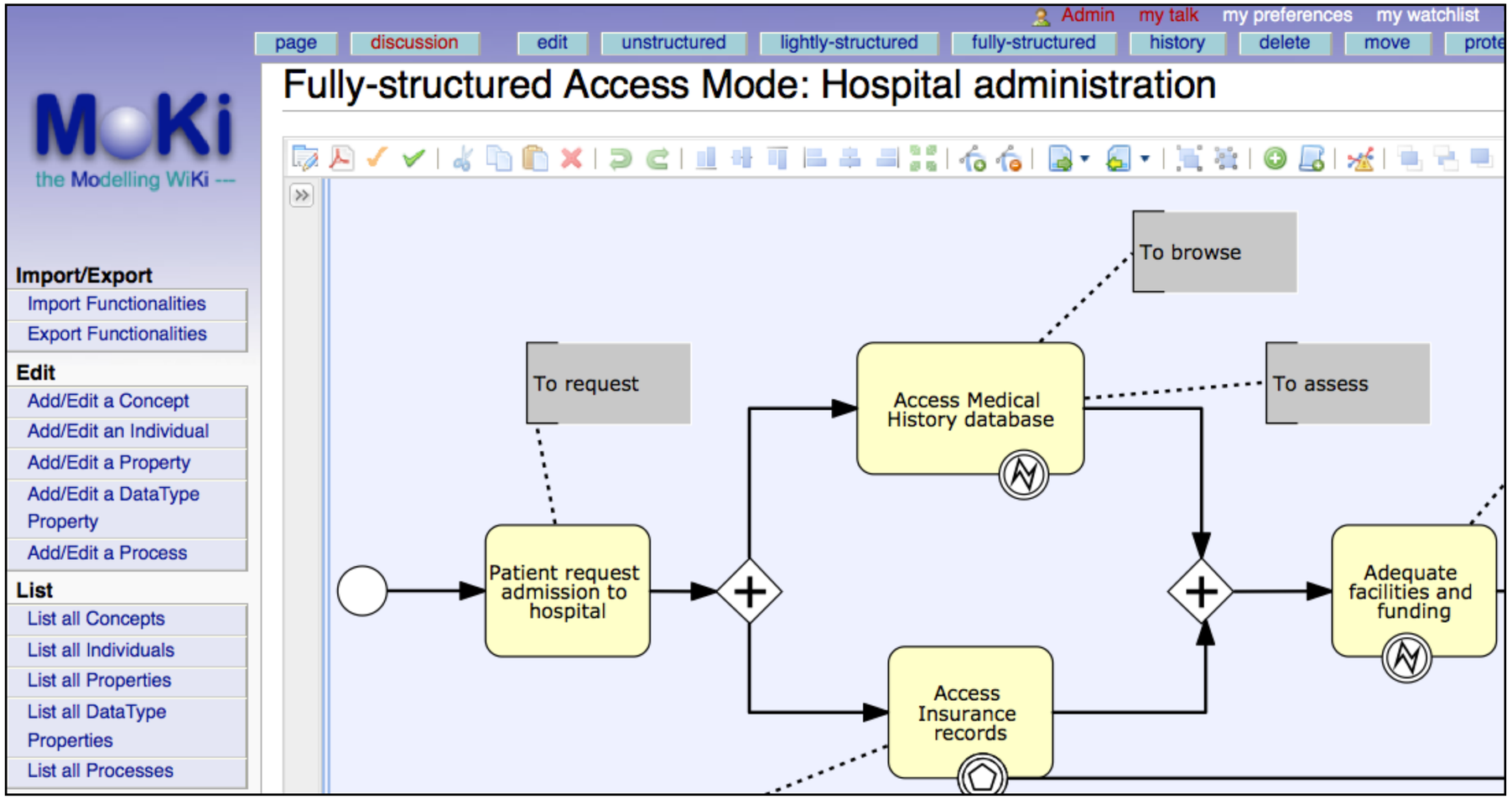


Gates (0-n) : Gate	There MAY be zero or more Gates (except where noted below). Zero Gates are allowed if the Gateway is last object in a Process flow and there are no Start or End Events for the Process. If there are zero or only one incoming Sequence Flow, then there MUST be at least two Gates.

$$Gateway \sqsubseteq (\geq 2)hasSequenceFlowTarget^{-1} \sqcup ((\leq 1)hasSequenceFlowTarget^{-1} \sqcap (\geq 2)hasGatewayGate)$$

Marco Rospocher, Chiara Ghidini, Luciano Serafini:
An ontology for the Business Process Modelling Notation. FOIS 2014: 133-146 (2014)

Collaborative BPMN annotation

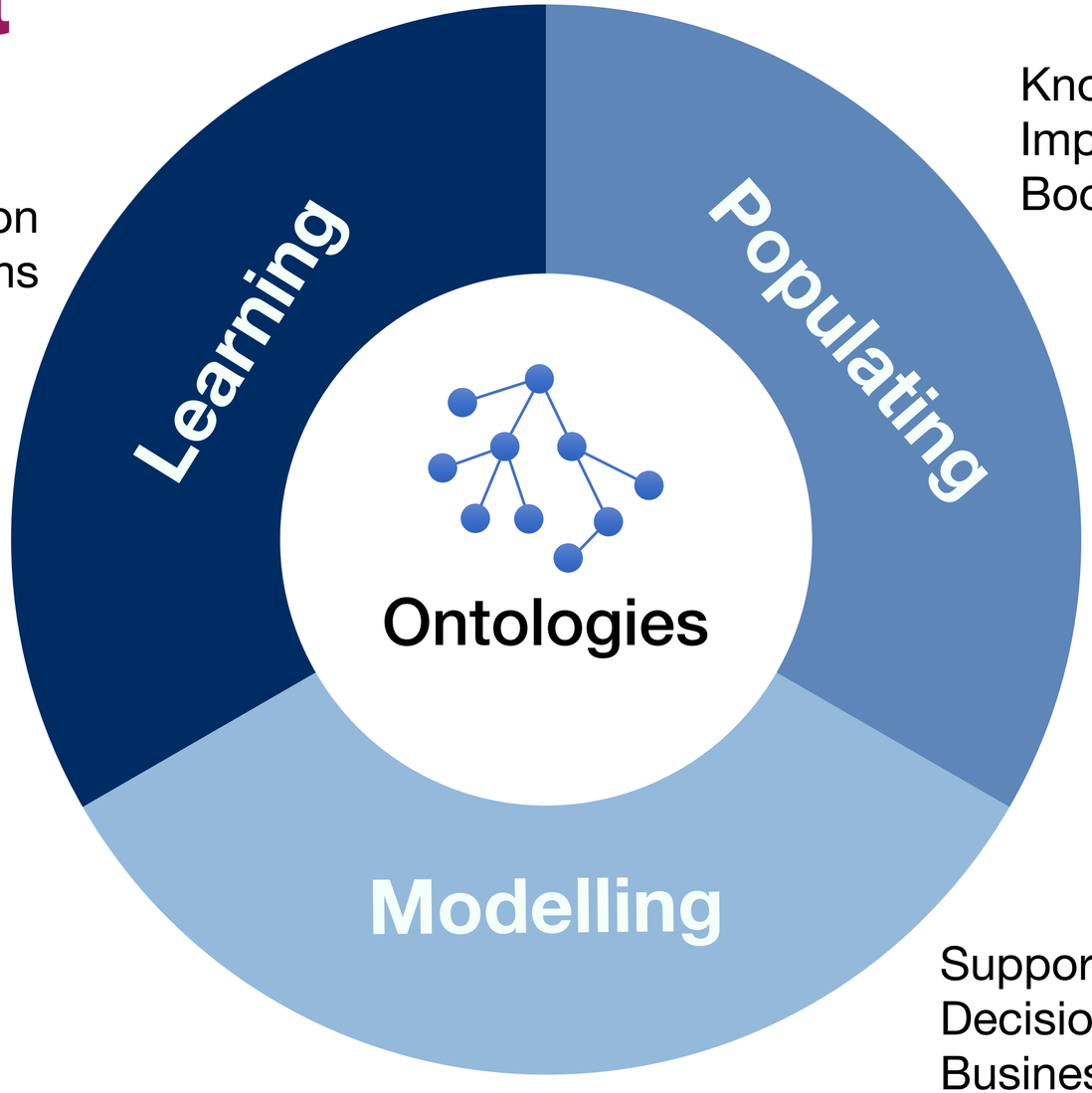


Chiara Di Francescomarino, Chiara Ghidini, **Marco Rospocher,** Luciano Serafini, Paolo Tonella:
A framework for the collaborative specification of semantically annotated business processes. Journal of Software Maintenance 23(4): 261-295 (2011)

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE

Agenda

Terminological Extraction
Learning expressive axioms

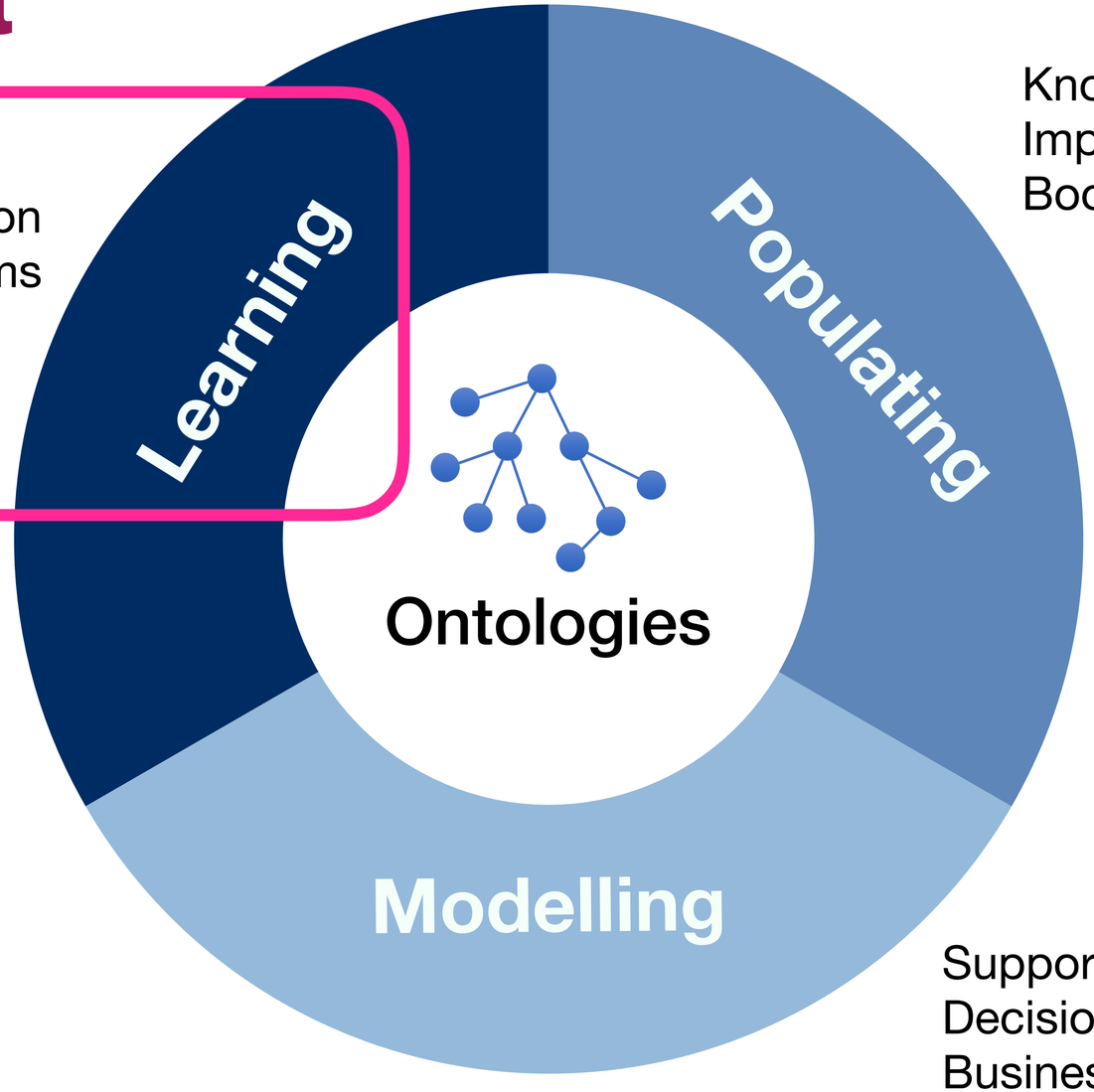


Knowledge Extraction
Improving NLP tasks
Boosting Information Retrieval

Supporting collaborative modelling
Decision Support Systems
Business Processes

Agenda

Terminological Extraction
Learning expressive axioms



Knowledge Extraction
Improving NLP tasks
Boosting Information Retrieval

Supporting collaborative modelling
Decision Support Systems
Business Processes

Concept extraction

PESCaDO
EU-FP7

- Concept extraction from text for ontology learning, extension and validation

Sara Tonelli, **Marco Rospocher**, Emanuele Pianta, Luciano Serafini:
Boosting Collaborative Ontology Building with Key-Concept Extraction. IEEE-ICSC 2011: 316-319 (2011)

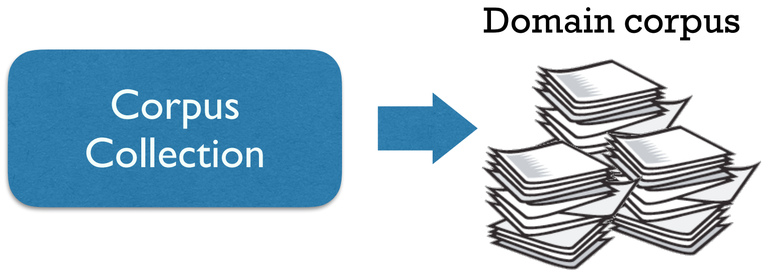
Marco Rospocher, Sara Tonelli, Luciano Serafini, Emanuele Pianta:
Corpus-based terminological evaluation of ontologies. Applied Ontology 7(4): 429-448 (2012)

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE
Dipartimento di Eccellenza 2018/2022

Concept extraction

EU-FP7

- Concept extraction from text for ontology learning, extension and validation



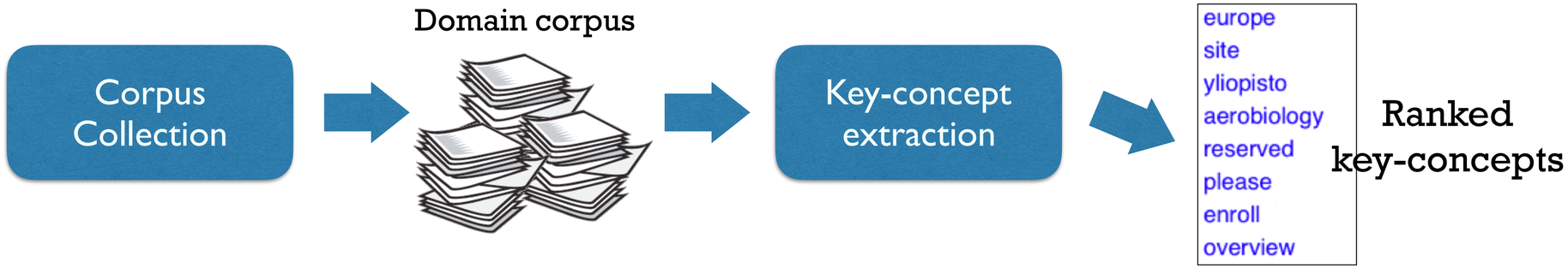
Sara Tonelli, **Marco Rospocher**, Emanuele Pianta, Luciano Serafini:
Boosting Collaborative Ontology Building with Key-Concept Extraction. IEEE-ICSC 2011: 316-319 (2011)

Marco Rospocher, Sara Tonelli, Luciano Serafini, Emanuele Pianta:
Corpus-based terminological evaluation of ontologies. Applied Ontology 7(4): 429-448 (2012)

Concept extraction

EU-FP7

- Concept extraction from text for ontology learning, extension and validation



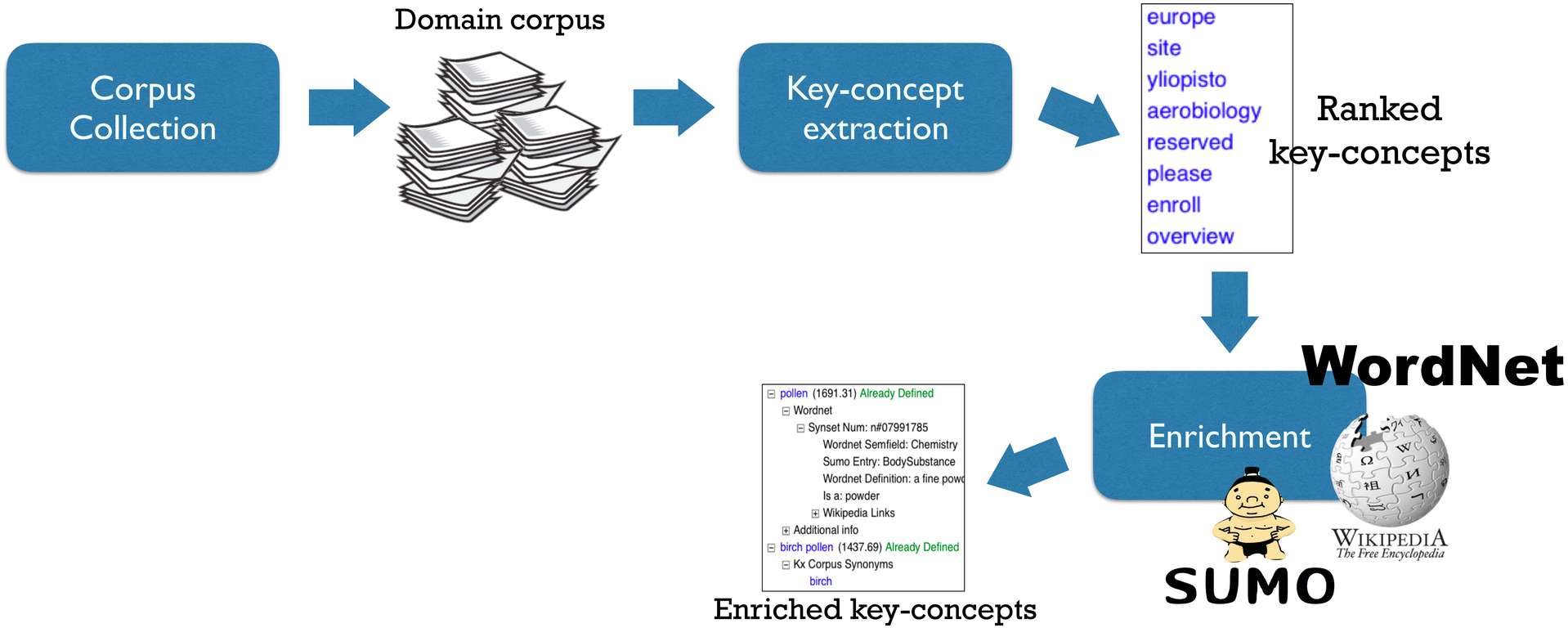
Sara Tonelli, **Marco Rospocher**, Emanuele Pianta, Luciano Serafini:
Boosting Collaborative Ontology Building with Key-Concept Extraction. IEEE-ICSC 2011: 316-319 (2011)

Marco Rospocher, Sara Tonelli, Luciano Serafini, Emanuele Pianta:
Corpus-based terminological evaluation of ontologies. Applied Ontology 7(4): 429-448 (2012)

Concept extraction

EU-FP7

- Concept extraction from text for ontology learning, extension and validation

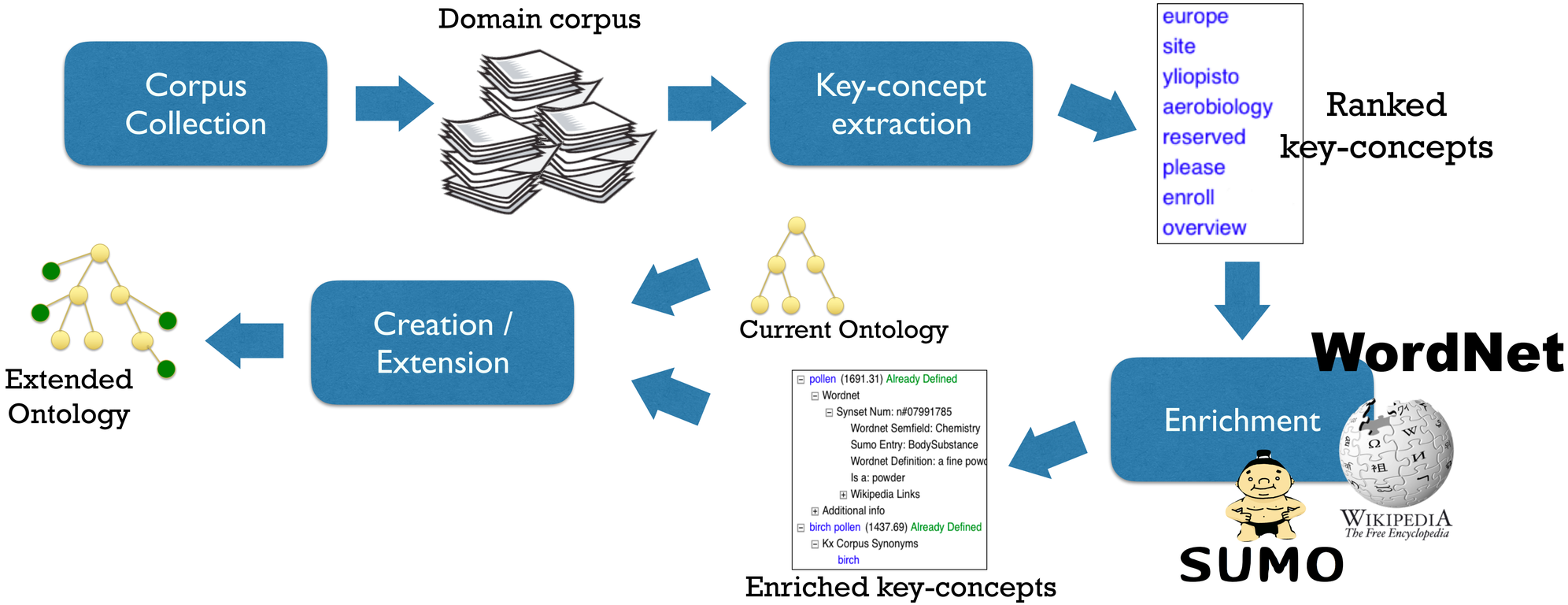


Sara Tonelli, **Marco Rospocher**, Emanuele Pianta, Luciano Serafini:
Boosting Collaborative Ontology Building with Key-Concept Extraction. IEEE-ICSC 2011: 316-319 (2011)

Marco Rospocher, Sara Tonelli, Luciano Serafini, Emanuele Pianta:
Corpus-based terminological evaluation of ontologies. Applied Ontology 7(4): 429-448 (2012)

Concept extraction

- Concept extraction from text for ontology learning, extension and validation



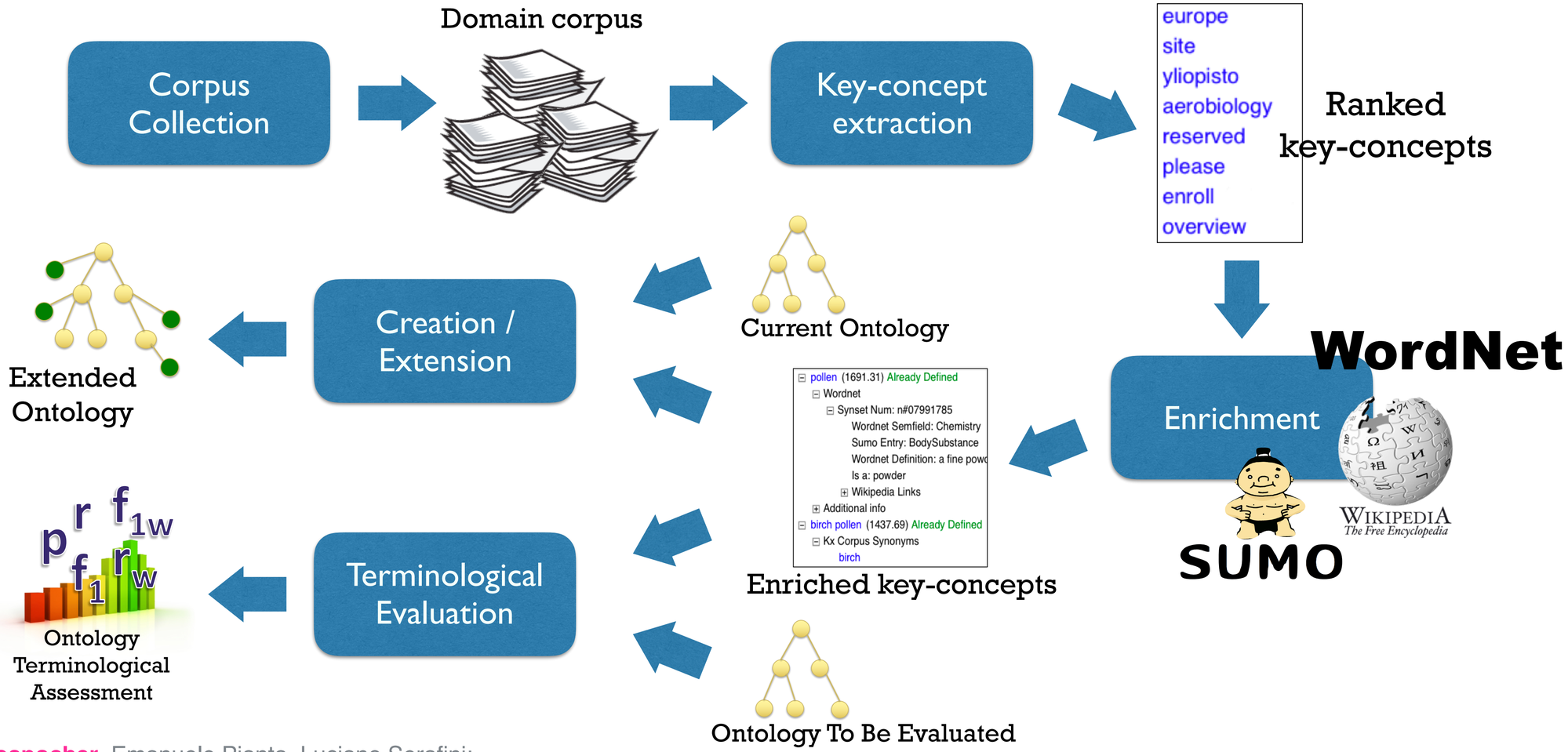
Sara Tonelli, **Marco Rospocher**, Emanuele Pianta, Luciano Serafini:
Boosting Collaborative Ontology Building with Key-Concept Extraction. IEEE-ICSC 2011: 316-319 (2011)

Marco Rospocher, Sara Tonelli, Luciano Serafini, Emanuele Pianta:
Corpus-based terminological evaluation of ontologies. Applied Ontology 7(4): 429-448 (2012)

Concept extraction

EU-FP7

- Concept extraction from text for ontology learning, extension and validation



Sara Tonelli, **Marco Rospocher**, Emanuele Pianta, Luciano Serafini:
Boosting Collaborative Ontology Building with Key-Concept Extraction. IEEE-ICSC 2011: 316-319 (2011)

Marco Rospocher, Sara Tonelli, Luciano Serafini, Emanuele Pianta:
Corpus-based terminological evaluation of ontologies. Applied Ontology 7(4): 429-448 (2012)

Extract new concepts from textual resources

*(Powered by KX - a **K**eyphrase e**X**traction system)*

MoKi
the Modelling WiKi ---

Files

Upload Files

⊞ **Show uploaded files**

Remove all uploaded files

Configure and Run

Re-load Default Settings

Language: english Domain:

Percentage of relevant concepts to return: 15

Take multiword expressions that occur at least:

- either 2 times in a document
- or 5 times in the corpus

Maximum length of multiword expressions: 4

Prefer key-concepts occuring early in the text:

Prefer specific key-phrases: Medium Preference

Extract relevant concepts

Concepts extracted (Ordered by Relevance)	Relevance	100% matching
▶ activity	1.00000	X
▶ attribute	0.88020	
sequence flow	0.71714	X
▶ business process modeling notation	0.70216	
▼ task	0.49418	X
▼ Wordnet		
▼ Synset_#00795720		
Wordnet Definition: any piece of work that is undertaken or attempted		
Is a: work		
Sumo Entry: IntentionalProcess		
▼ Synonyms		
undertaking		
project		
labor		
Hyponims: cinch, breeze, picnic, snap1, duck soup, child's play, pushover, walkover, piece of cake, adventure, escapade, risky venture, dangerous undertaking, assignment, baby, enterprise, endeavor, endeavour, labor of love, labour of love, marathon, endurance contest, no-brainer, proposition, tall order, large order, venture, Manhattan Project		
▶ Wikipedia Links		
▶ mapping	0.48253	
▶ flow	0.47920	
▶ message	0.43927	X
▶ sub process	0.41265	X
▶ gateway	0.39268	X
▶ pool	0.30116	X
message flow	0.27787	X
▶ sequence	0.25790	
▶ expression	0.23461	X

Learning expressive axioms

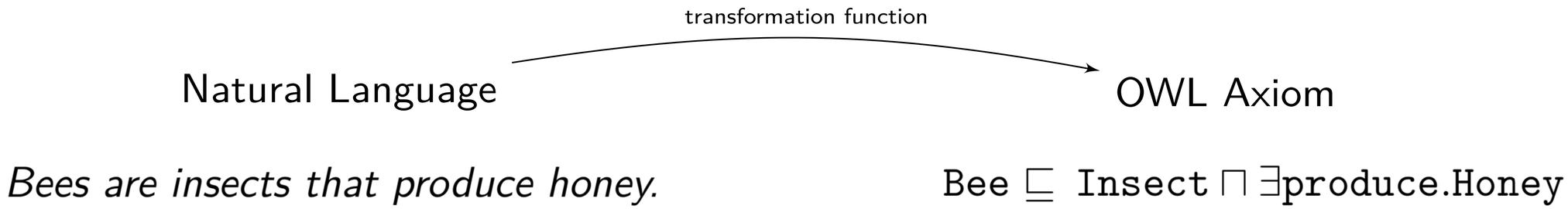
Bees are insects that produce honey.

Learning expressive axioms

Bees are insects that produce honey.

$\texttt{Bee} \sqsubseteq \texttt{Insect} \sqcap \exists \texttt{produce.Honey}$

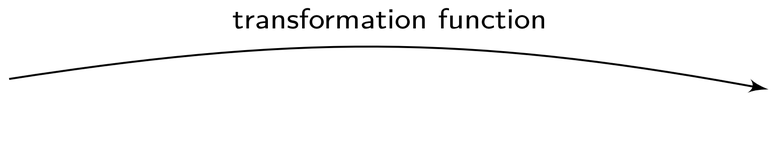
Learning expressive axioms



Learning expressive axioms

All the extralogical symbols come from the sentence.

"*syntactic transformation of natural language definitions into description logic axioms.*" (Völker J., 2008)



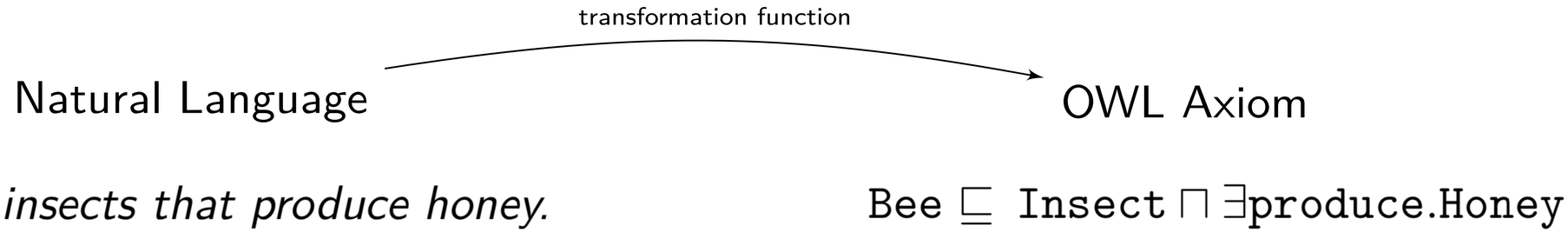
Bees are insects that produce honey.

$\texttt{Bee} \sqsubseteq \texttt{Insect} \sqcap \exists \texttt{produce.Honey}$

Learning expressive axioms

All the extralogical symbols come from the sentence.

"*syntactic transformation of natural language definitions into description logic axioms.*" (Völker J., 2008)



Transforming a sentence into an axiom:

- is it possible to train a machine learning model for this task?
- is it possible to perform the training in a end-to-end fashion?

Tagging & Transducing

A bee is an insect that produces honey.

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE
Dipartimento di Eccellenza 2018/2022

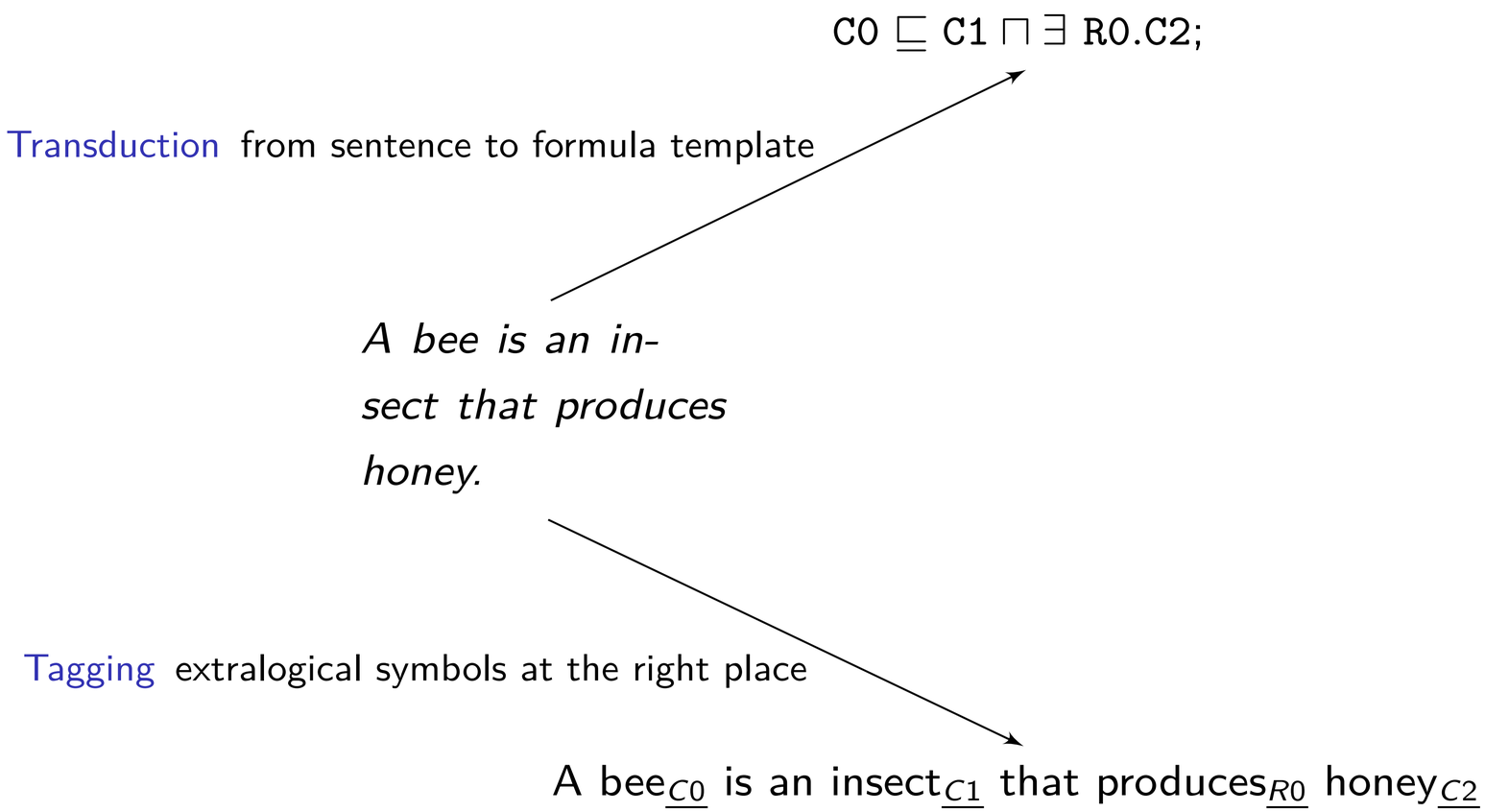
Tagging & Transducing

$\texttt{C0} \sqsubseteq \texttt{C1} \sqcap \exists\ \texttt{R0.C2};$

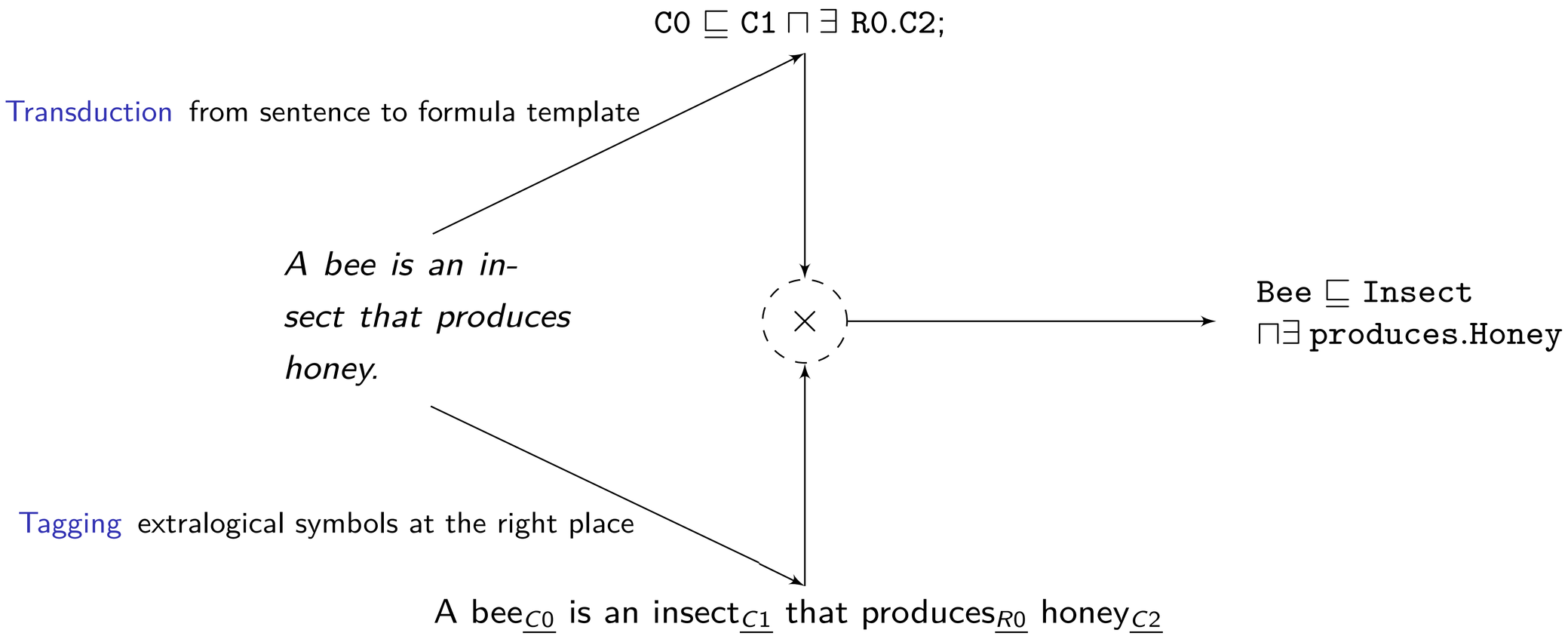
Transduction from sentence to formula template

A bee is an insect that produces honey.

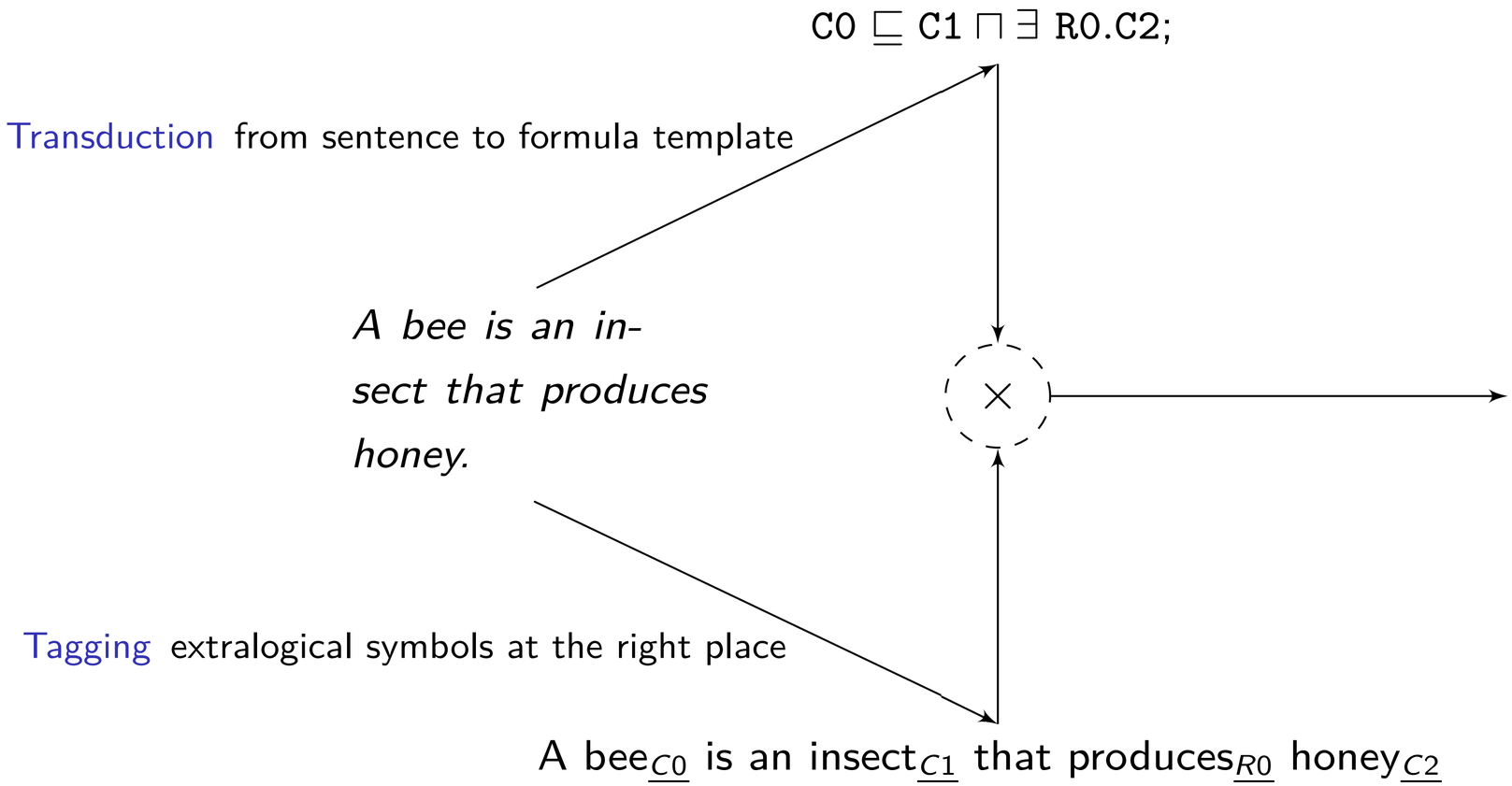
Tagging & Transducing



Tagging & Transducing



Tagging & Transducing



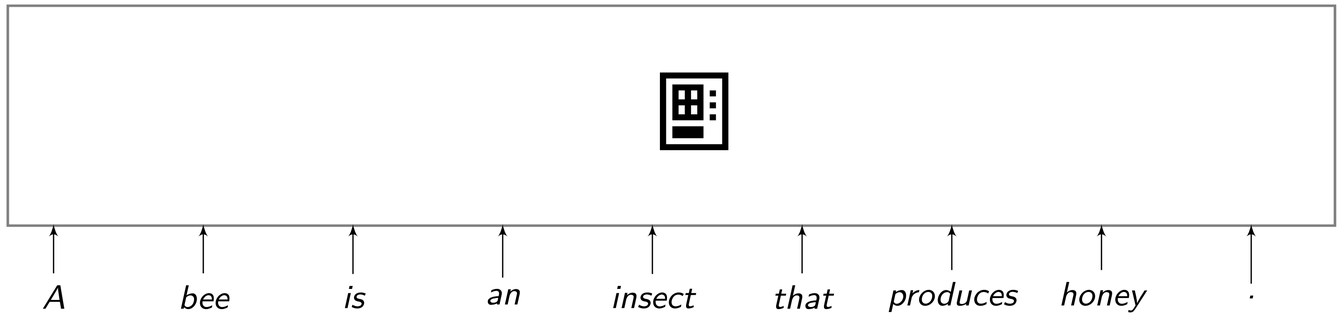
Bee ⊑ Insect
⊓∃ produces.Honey

Implemented as two simple RNNs

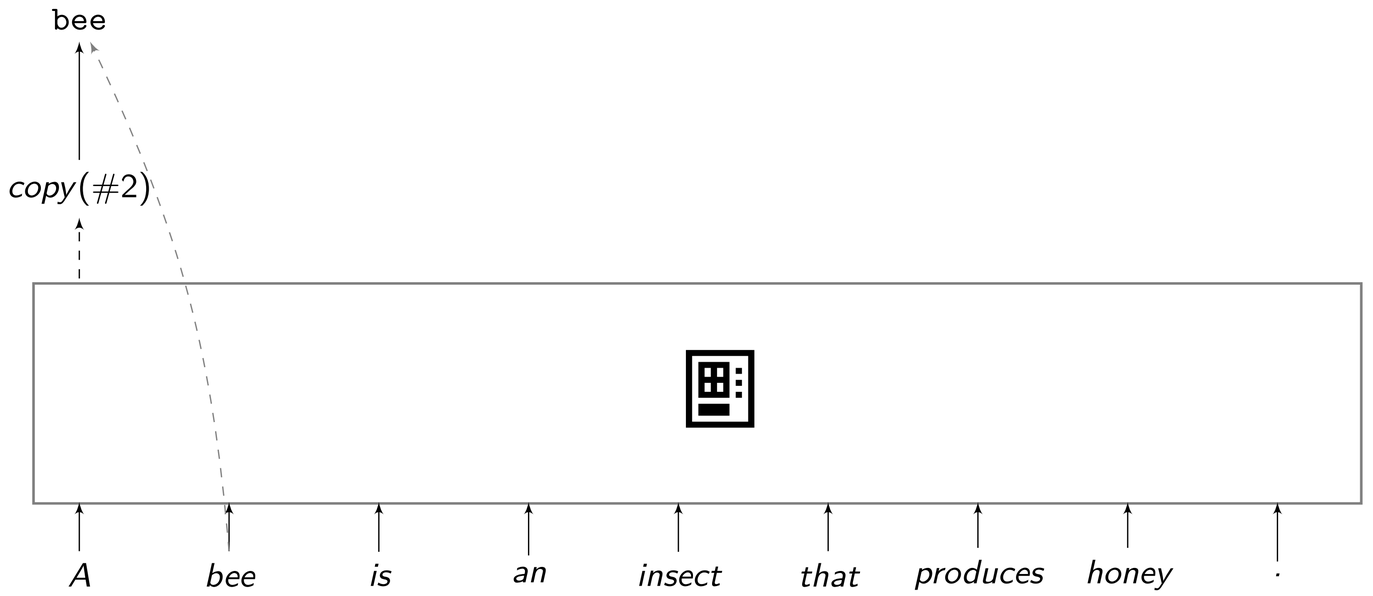
Shows the **feasibility** of the approach, but has some **limitations**:

- Separate training procedures
- Limited number of placeholder → cannot generalise over sentence length

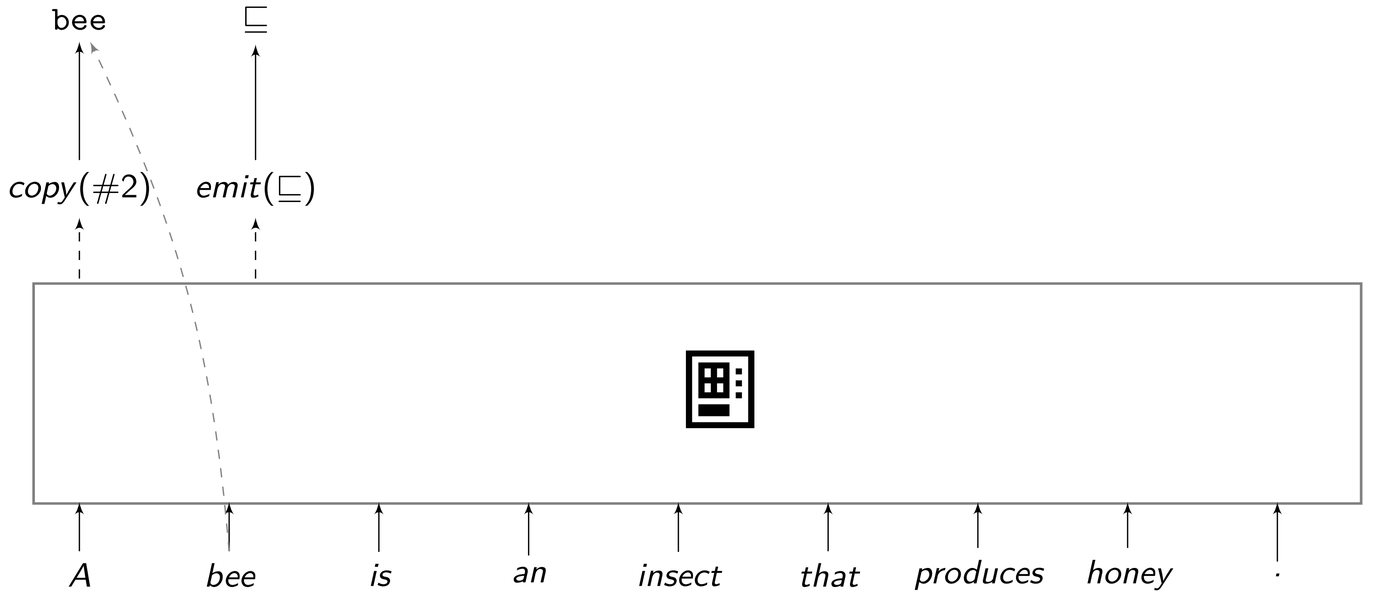
Translating



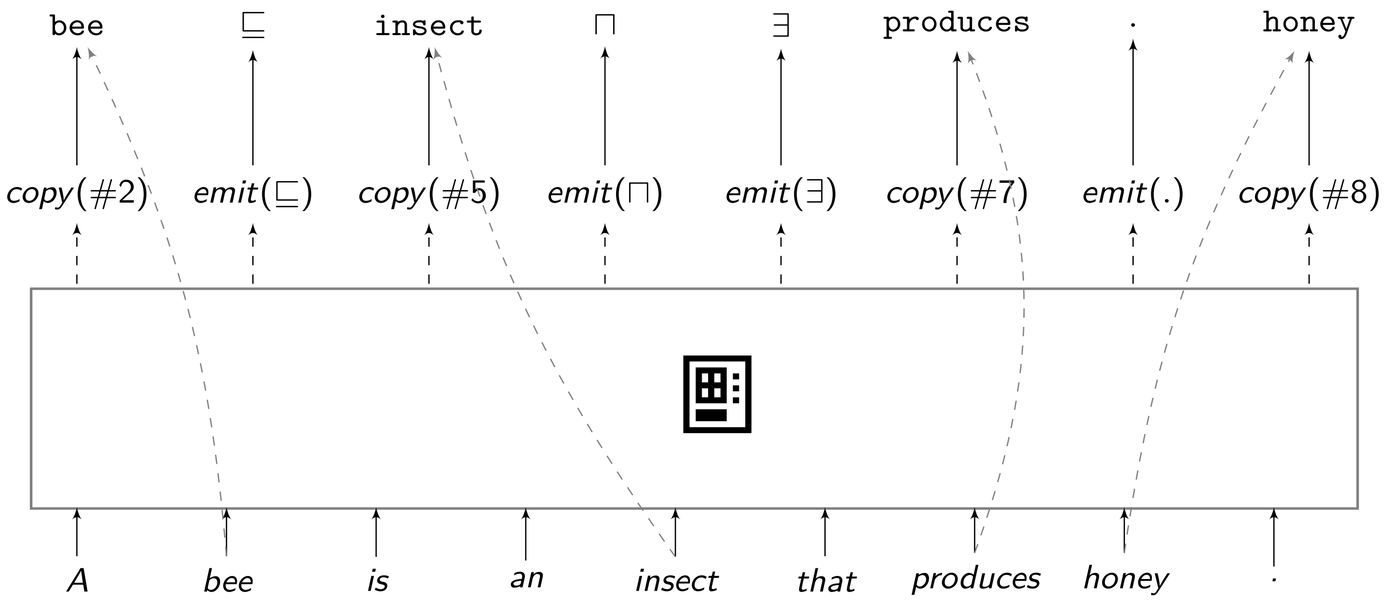
Translating



Translating

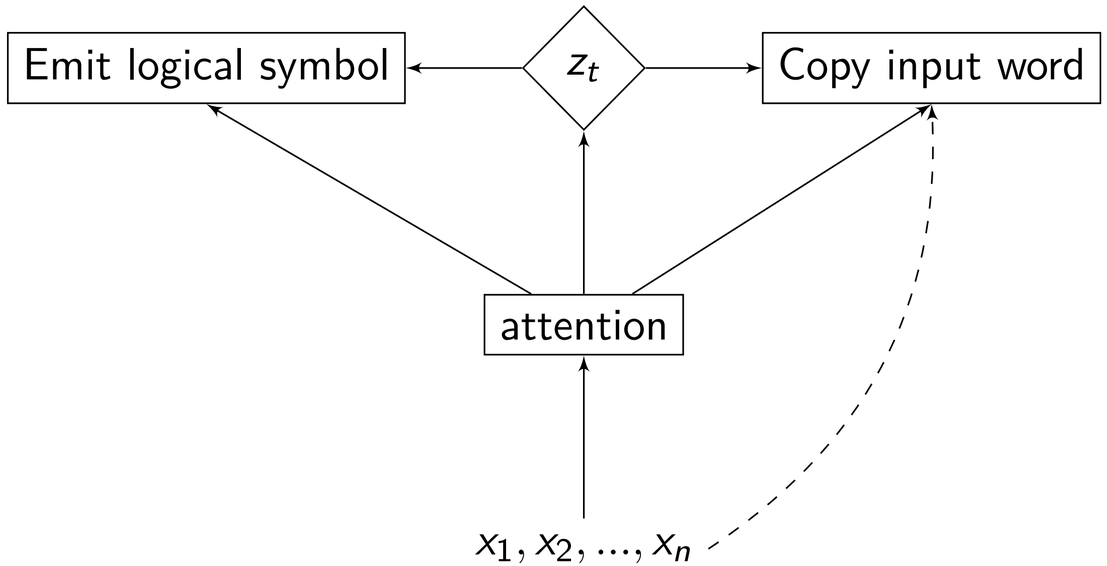


Translating



Quasi-zero vocabulary setting.

Translating

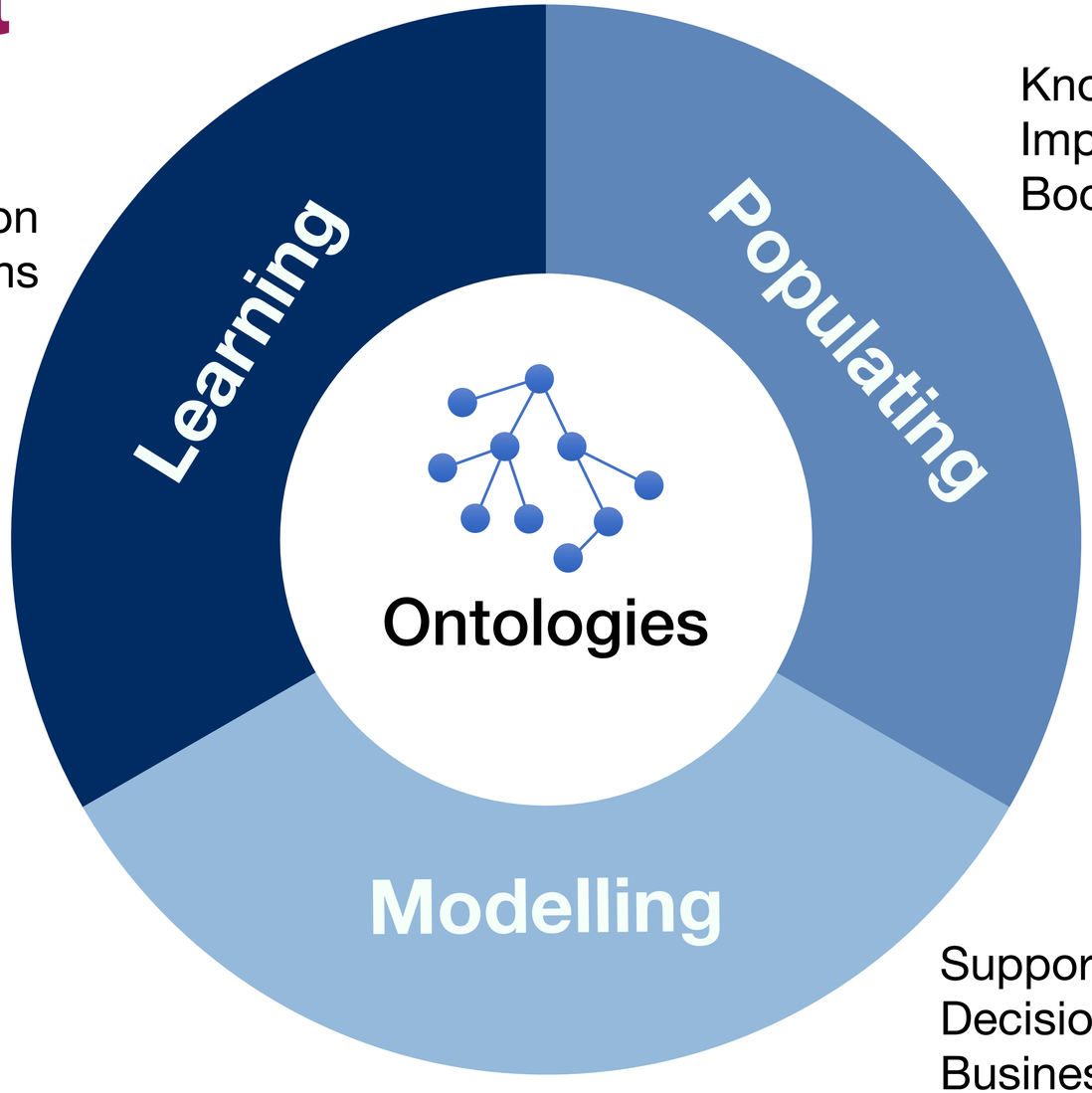


- Generalises well over the syntactic structure of definitions
- Tolerates unknown words

- Besides the NN code, we also developed a training dataset (partly automatically generated) and a reference evaluation dataset
- Code and datasets are available at:
  - https://github.com/dkmfbk/dket/

Agenda

Terminological Extraction
Learning expressive axioms

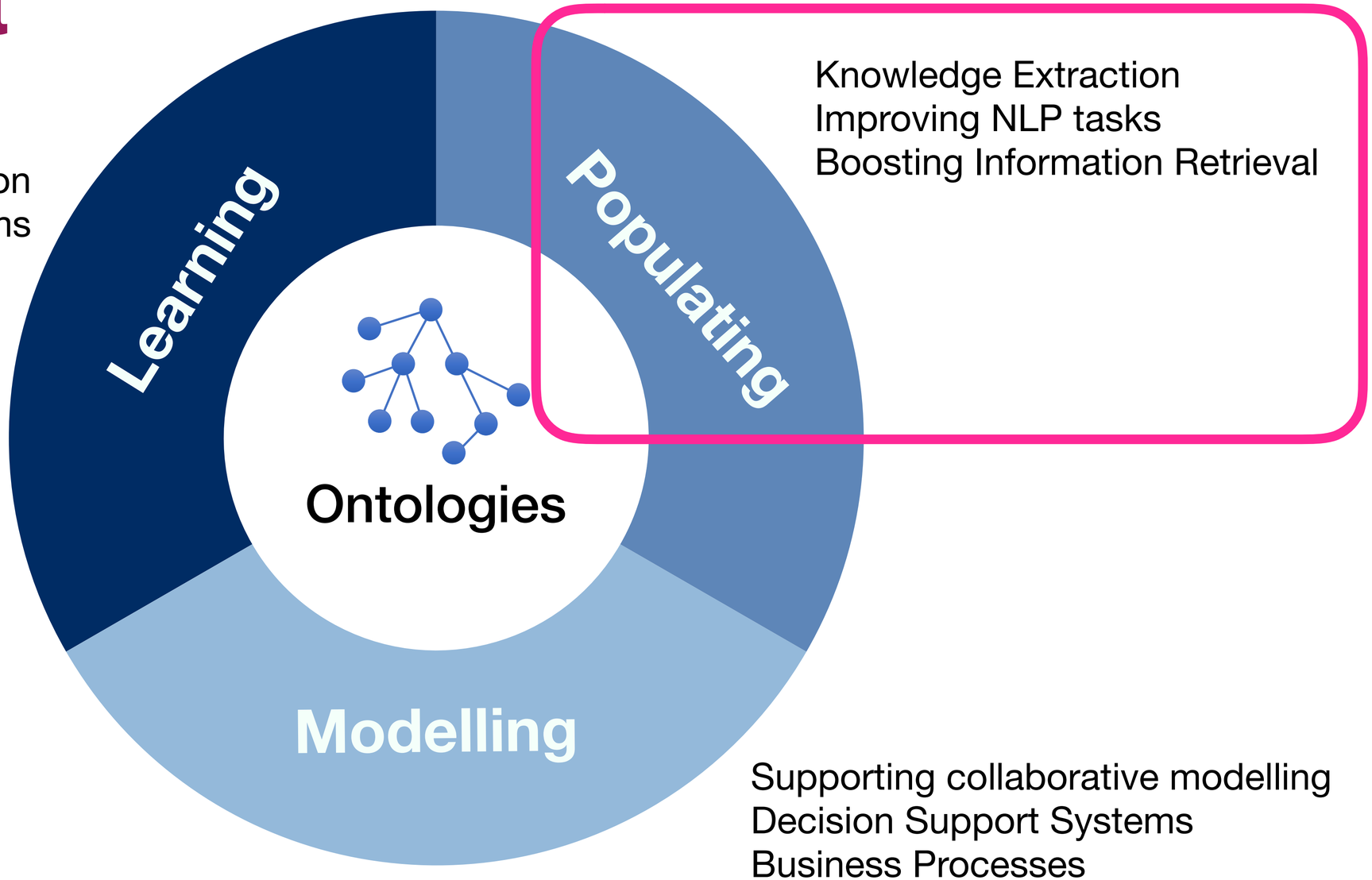


Knowledge Extraction
Improving NLP tasks
Boosting Information Retrieval

Supporting collaborative modelling
Decision Support Systems
Business Processes

Agenda

Terminological Extraction
Learning expressive axioms



Knowledge Extraction
Improving NLP tasks
Boosting Information Retrieval

Supporting collaborative modelling
Decision Support Systems
Business Processes

Event-Centric Knowledge Graph

NewsReader
POST HOC ERGO PROPTER HOC
EU-FP7

Yesterday, Kia has hired Peter Schreyer as chief design officer.

[Newspaper, 2 Aug 2006]

Event-Centric Knowledge Graph

NewsReader
POST HOC ERGO PROPTER HOC
EU-FP7

Yesterday, Kia has hired Peter Schreyer as chief design officer.
[Newspaper, 2 Aug 2006]

hire

Marco Rospocher, Marieke van Erp, Piek Vossen, Antske Fokkens, Itziar Aldabe, German Rigau, Aitor Soroa, Thomas Ploeger, Tessel Bogaard:
Building event-centric knowledge graphs from news. J. Web Semant. 37-38: 132-151 (2016)

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE
Dipartimento di Eccellenza 2018/2022

Event-Centric Knowledge Graph

NewsReader
POST HOC ERGO PROPTER HOC
EU-FP7

Yesterday, Kia has hired Peter Schreyer as chief design officer.
[Newspaper, 2 Aug 2006]

hire
FrameNet:Hiring

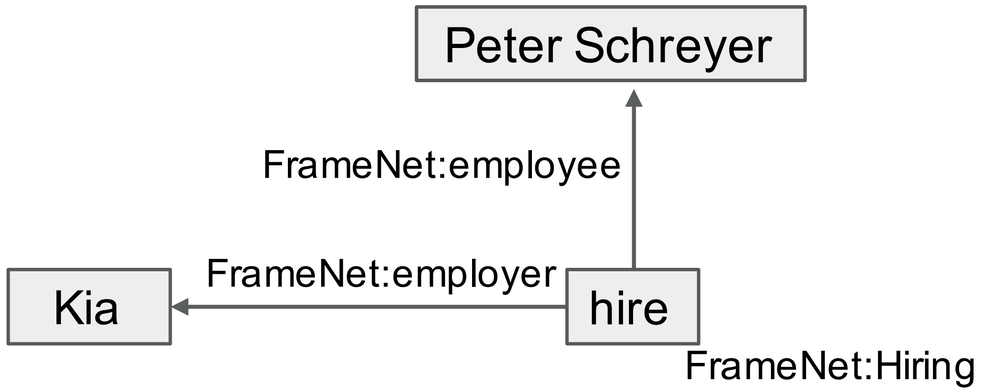
Marco Rospocher, Marieke van Erp, Piek Vossen, Antske Fokkens, Itziar Aldabe, German Rigau, Aitor Soroa, Thomas Ploeger, Tessel Bogaard:
Building event-centric knowledge graphs from news. J. Web Semant. 37-38: 132-151 (2016)

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE
Dipartimento di Eccellenza 2018/2022

Event-Centric Knowledge Graph

NewsReader
POST HOC ERGO PROPTER HOC
EU-FP7

Yesterday, Kia has hired Peter Schreyer as chief design officer.
[Newspaper, 2 Aug 2006]



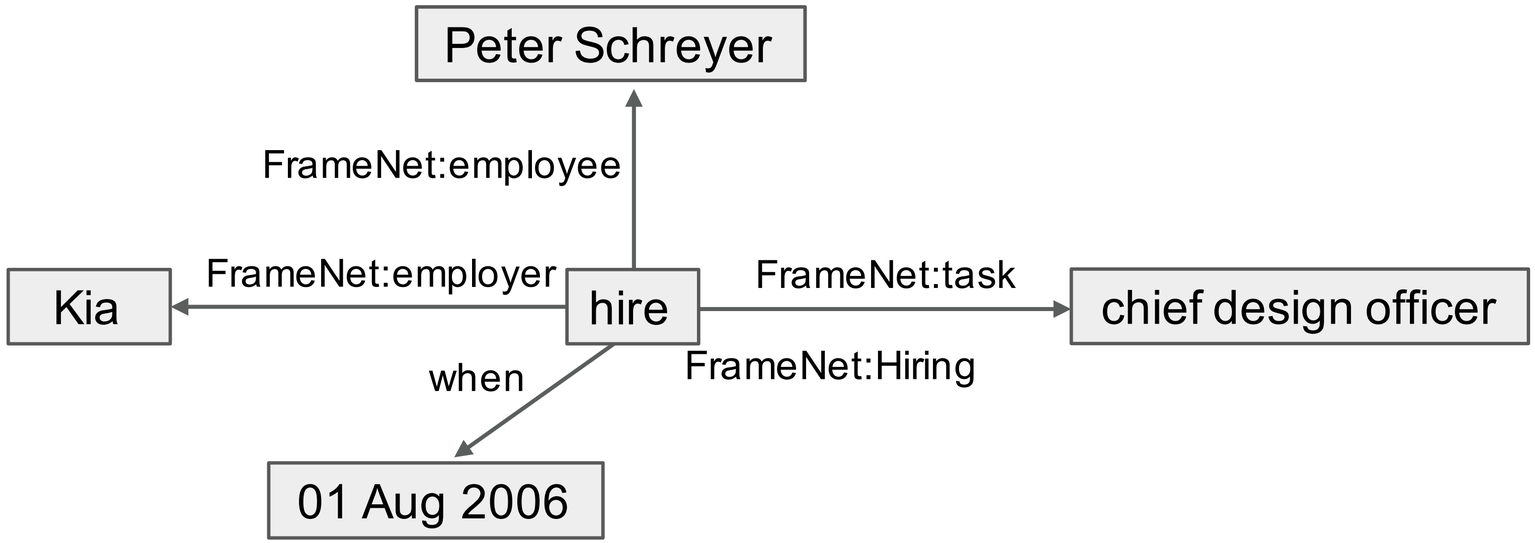
Marco Rospocher, Marieke van Erp, Piek Vossen, Antske Fokkens, Itziar Aldabe, German Rigau, Aitor Soroa, Thomas Ploeger, Tessel Bogaard:
Building event-centric knowledge graphs from news. J. Web Semant. 37-38: 132-151 (2016)

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE

Event-Centric Knowledge Graph

NewsReader
EU-FP7

Yesterday, Kia has hired Peter Schreyer as chief design officer.
[Newspaper, 2 Aug 2006]



Marco Rospocher, Marieke van Erp, Piek Vossen, Antske Fokkens, Itziar Aldabe, German Rigau, Aitor Soroa, Thomas Ploeger, Tessel Bogaard:
Building event-centric knowledge graphs from news. J. Web Semant. 37-38: 132-151 (2016)

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE

Building ECKG

NewsReader
EU-FP7

Yesterday, Kia has hired Peter Schreyer as chief design officer.
[Newspaper, 2 Aug 2006]

NLP Tasks:

Building ECKG

NewsReader
EU-FP7

Yesterday, Kia has hired Peter Schreyer as chief design officer.
[Newspaper, 2 Aug 2006]

Peter Schreyer
Person

Kia
Organization

NLP Tasks:
- Named Entity Recognition and Classification (NERC)

Building ECKG

NewsReader
EU-FP7

Yesterday, Kia has hired Peter Schreyer as chief design officer.
[Newspaper, 2 Aug 2006]

dbpedia:Peter_Schreyer

Peter Schreyer

Person

dbpedia:Kia_Motors

Kia

Organization

NLP Tasks:
- Named Entity Recognition and Classification (NERC)
- Entity Linking (EL)

Building ECKG

NewsReader
EU-FP7

Yesterday, Kia has hired Peter Schreyer as chief design officer.
[Newspaper, 2 Aug 2006]

dbpedia:Peter_Schreyer

Peter Schreyer

Person

dbpedia:Kia_Motors

Kia

Org

rdf:type	owl:Thing
	dbo:Company
	dbo:Agent
	dbo:Organisation
dbo:industry	dbr:Automotive_industry
dbo:location	dbr:Seoul
dbo:numberOfEmployees	33255 (xsd:integer)

NLP Tasks:
- Named Entity Recognition and Clas
- Entity Linking (EL)

Building ECKG

NewsReader
EU-FP7

Yesterday, Kia has hired Peter Schreyer as chief design officer.
[Newspaper, 2 Aug 2006]

dbpedia:Peter_Schreyer

Peter Schreyer

Person

dbpedia:Kia_Motors

Kia

Organization

01 Aug 2006

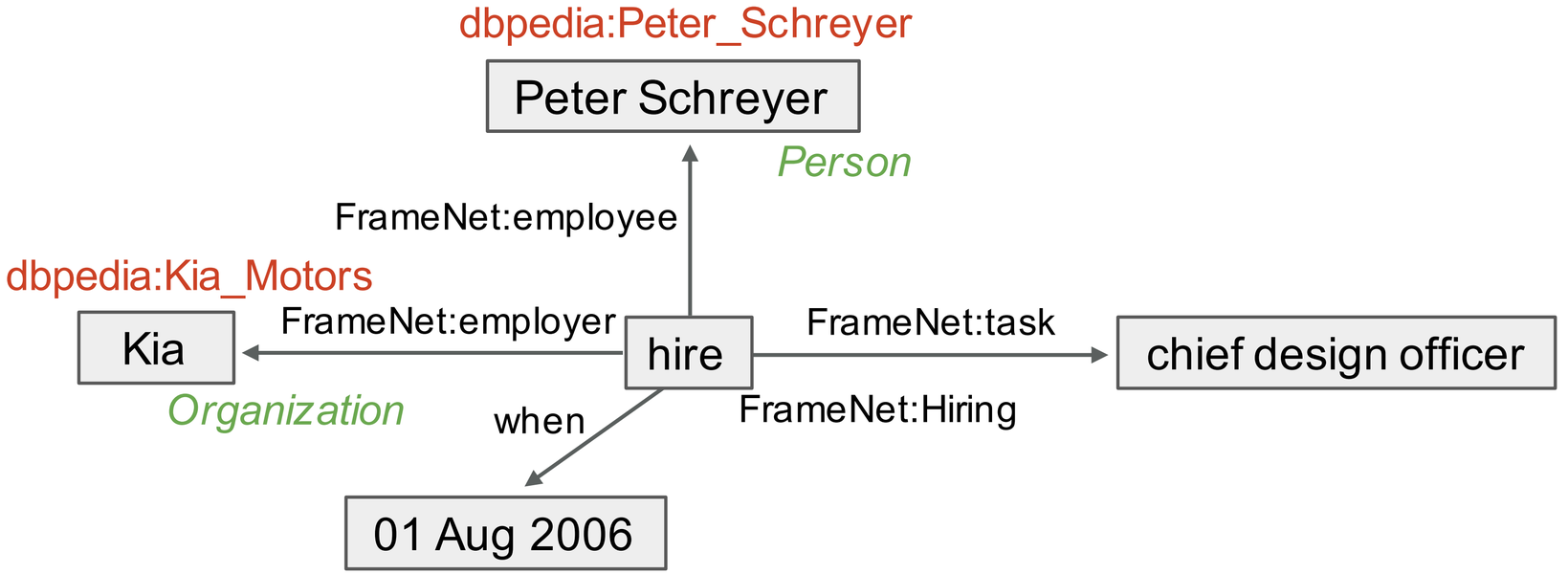
NLP Tasks:
- Named Entity Recognition and Classification (NERC)
- Entity Linking (EL)
- Temporal Expression Recognition and Normalization (TERN)

Building ECKG

EU-FP7

Yesterday, Kia has hired Peter Schreyer as chief design officer.

[Newspaper, 2 Aug 2006]



NLP Tasks:

- Named Entity Recognition and Classification (NERC)
- Entity Linking (EL)
- Temporal Expression Recognition and Normalization (TERN)
- Semantic Role Labeling (SRL)

Building ECKG

EU-FP7

http://www.newsreader-project.eu/

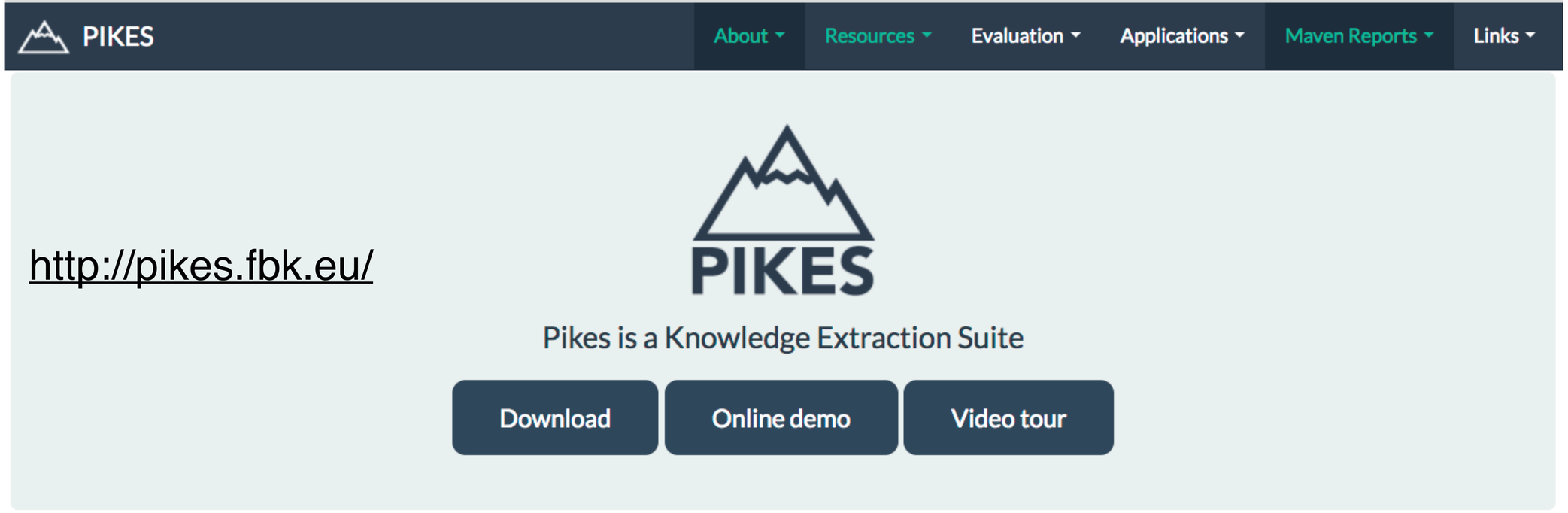
- Also: Factuality, Opinion, Polarity, …
- Languages: EN, IT, ES, NL
- Processed >2M news articles on the crisis in the automotive industry

https://pikes.fbk.eu/

- Fully compliant with Semantic Web / Linked Data best practices
- Languages: EN (+IT coming soon)
- Extremely Efficient (700K words/h)

Piek Vossen, Rodrigo Agerri, Itziar Aldabe, Agata Cybulska, Marieke v. Erp, Antske Fokkens, Egoitz Laparra, A.L. Minard, Alessio P. Aprosio, German Rigau, **Marco Rospocher**, Roxane Segers: NewsReader: Using knowledge resources in a cross-lingual reading machine to generate more knowledge from massive streams of news. Knowl.-Based Syst. 110: 60-85 (2016)

Francesco Corcoglioniti, **Marco Rospocher**, Alessio Palmero Aprosio:
Frame-Based Ontology Population with PIKES. IEEE Trans. Knowl. Data Eng. 28(12): 3261-3275 (2016)



About

PIKES is a Java-based suite that extracts knowledge from textual resources. The tool implements a rule-based strategy that reinterprets the output of semantic role labelling (SRL) tools in light of other linguistic analyses, such as dependency parsing or co-reference resolution, thus properly capturing and formalizing in RDF important linguistic aspects such as argument nominalization, frame-frame relations, and group entities.

Features

- Argument nominalization using SRL
- Frame-frame relations extractions
- Entity grouping exploiting linking and co-reference
- Extensible and replaceable NLP pipeline
- Interlinked three-layer representation model exposed as RDF
- Instance RDF triples annotated with detailed information of the mentions (via named graph)
- REST API service included, built on top of Grizzly
- Based on Java 8 and RDFpro

News

- 2018-05-25 Paper on PSL4EA accepted at ISWC2018! Additional material available here.
- 2018-04-16 Paper on JPARK accepted at IJCAI-ECAI-18! Additional material available here.
- 2017-10-27 Added instructions to use PIKES for batch processing documents.
- 2017-04-18 Added documentation in download page.
- 2016-09-01 Video recording of the ESWC2016 presentation available on videolectures.net
- 2016-08-18 Paper accepted at IEEE TKDE journal!
- 2016-03-18 Added Using PIKES for Information Retrieval section
- 2016-02-23 Paper on using PIKES for Information Retrieval accepted at ESWC 2016 conference!

PIKES

Pikes is a Knowledge Extraction suite

Video tutorial | Jump to example

Write a text in English and press the blue button.

G. W. Bush and Bono are very strong supporters of the fight of HIV in Africa. Their March 2002 meeting resulted in a 5 billion dollar aid.

Pick a sample text:

[Select one]

Submit

PIKES

Pikes is a Knowledge Extraction suite

Video tutorial | Jump to example

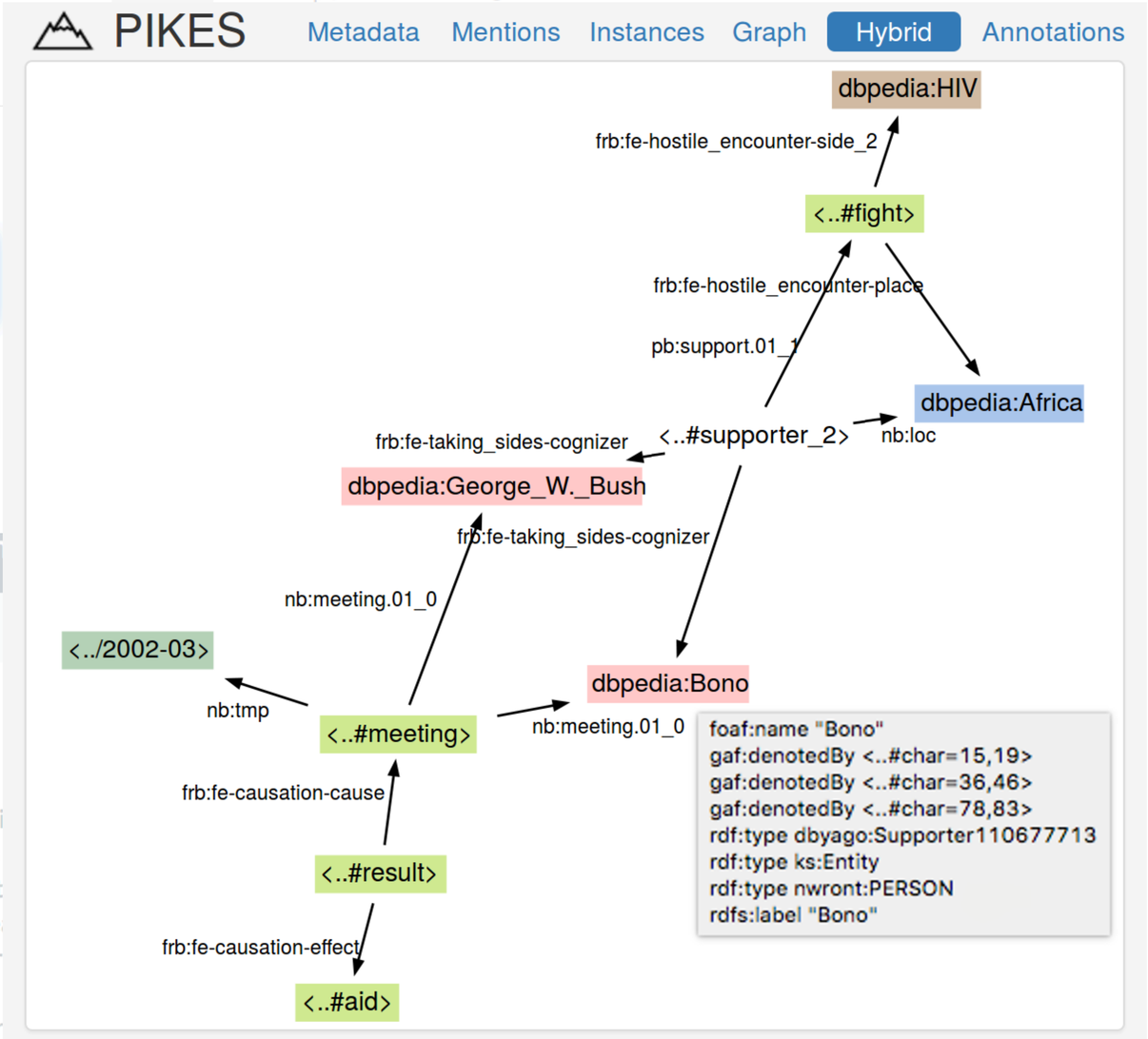
Write a text in English and press the blue button.

G. W. Bush and Bono are very strong supporters of the fight of HIV in Africa. Their March 2002 meeting resulted in a 5 billion dollar aid.

Pick a sample text:

[Select one]

Submit



PIKES

Pikes is a Knowledge Extraction suite

Video tutorial | Jump to example

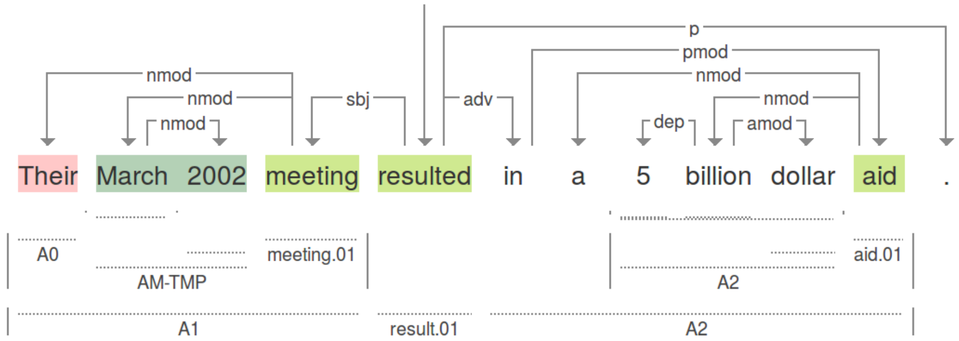
Write a text in English and press the blue button.

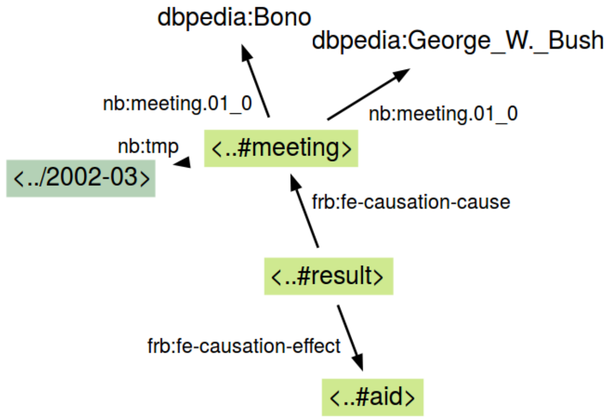
G. W. Bush and Bono are very strong supporters of the fight of HIV in Africa. Their March 2002 meeting resulted in a 5 billion dollar aid.

PIKES Metadata Mentions Instances Graph Hybrid Annotations

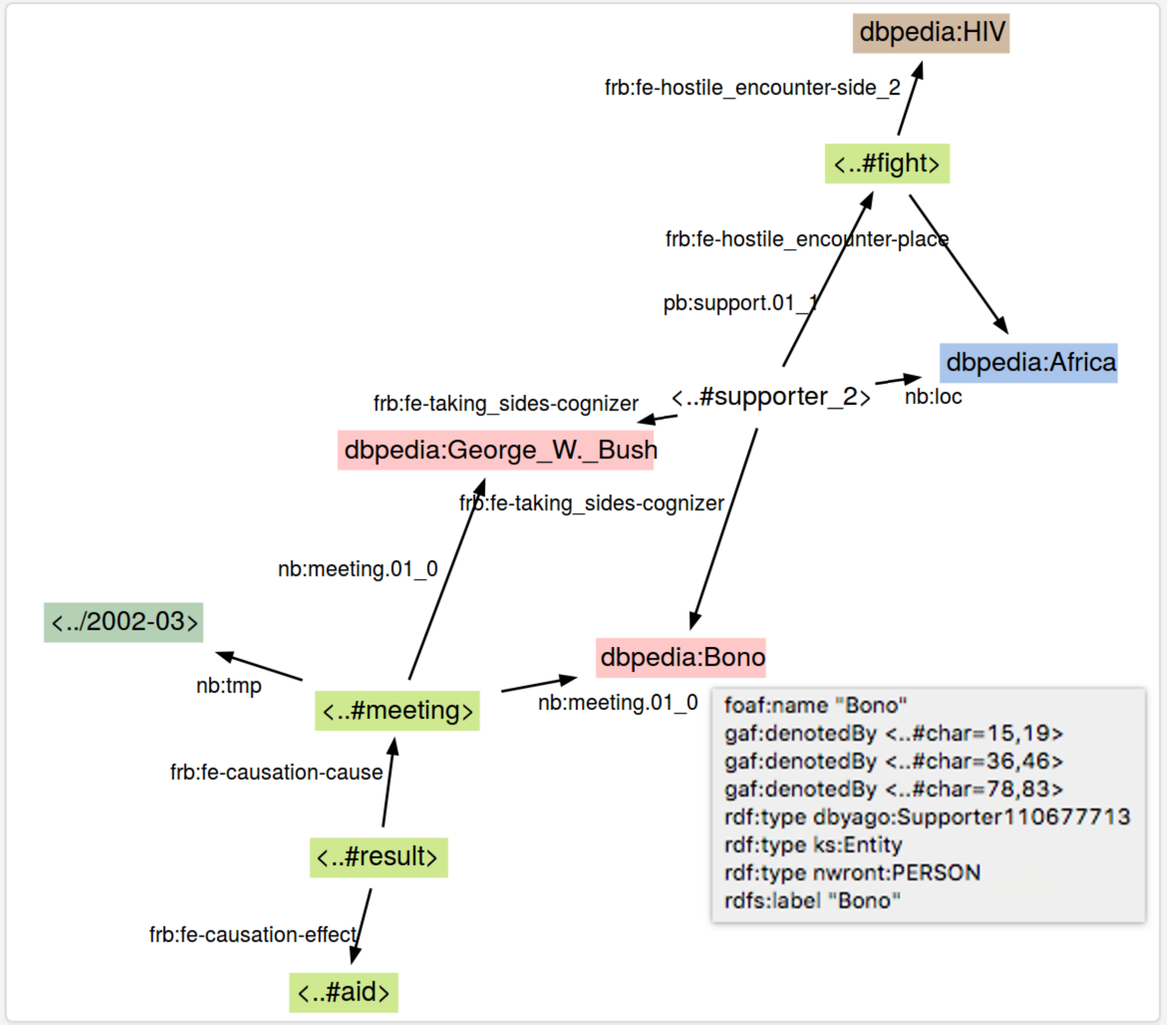
S1 G. W. Bush and Bono are very strong supporters of the fight of HIV in Africa.

S2





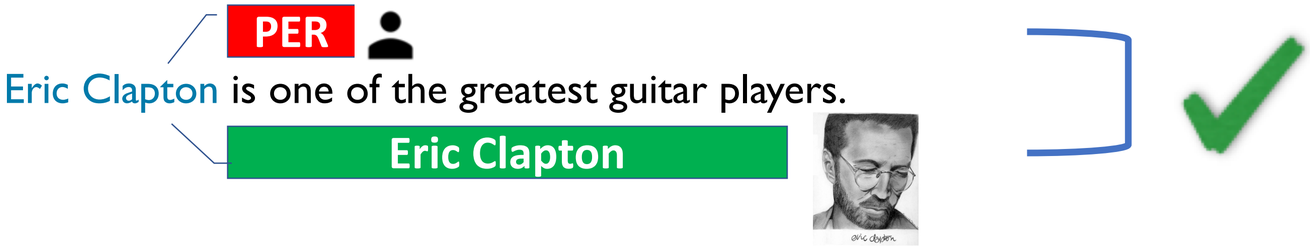
PIKES Metadata Mentions Instances Graph Hybrid Annotations



Improving NLP via Ontologies

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE

Improving NLP via Ontologies



Improving NLP via Ontologies

PER

Eric Clapton is one of the greatest guitar players.

Eric Clapton

PER

Mr. Washington was runner-up at Wimbledon in 1996.

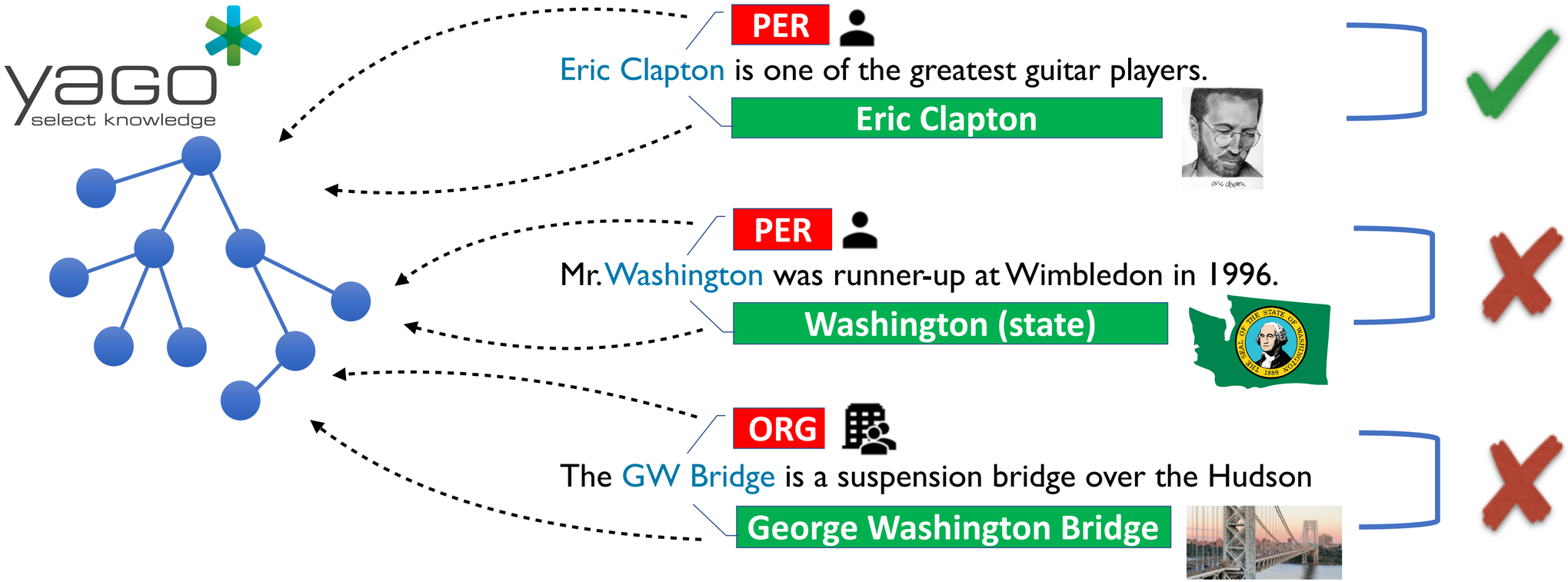
Washington (state)

ORG

The GW Bridge is a suspension bridge over the Hudson

George Washington Bridge

Improving NLP via Ontologies



Improving NLP via Ontologies

… $token_1$ $token_2$ $token_3$ $token_4$ $token_5$ $token_6$ ….

Marco Rospocher, Francesco Corcoglioniti:
Joint Posterior Revision of NLP Annotations via Ontological Knowledge. IJCAI 2018: 4316-4322

Marco Rospocher:
An Ontology-Driven Probabilistic Soft Logic Approach to Improve NLP Entity Annotations. International Semantic Web Conference (1) 2018: 144-161

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE
Dipartimento di Eccellenza 2018/2022

Improving NLP via Ontologies

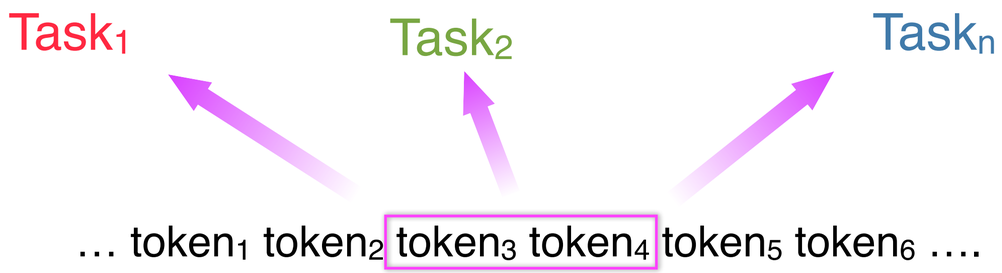
… $token_1$ $token_2$ $token_3$ $token_4$ $token_5$ $token_6$ ….

Marco Rospocher, Francesco Corcoglioniti:
Joint Posterior Revision of NLP Annotations via Ontological Knowledge. IJCAI 2018: 4316-4322

Marco Rospocher:
An Ontology-Driven Probabilistic Soft Logic Approach to Improve NLP Entity Annotations. International Semantic Web Conference (1) 2018: 144-161

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE
Dipartimento di Eccellenza 2018/2022

Improving NLP via Ontologies

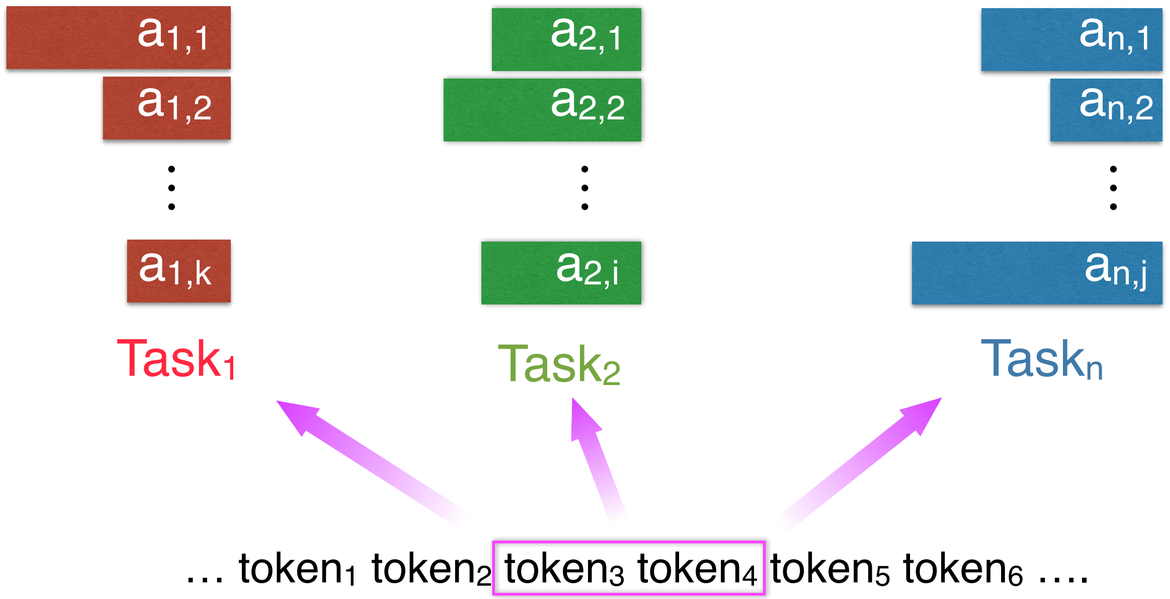


Marco Rospocher, Francesco Corcoglioniti:
Joint Posterior Revision of NLP Annotations via Ontological Knowledge. IJCAI 2018: 4316-4322

Marco Rospocher:
An Ontology-Driven Probabilistic Soft Logic Approach to Improve NLP Entity Annotations. International Semantic Web Conference (1) 2018: 144-161

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE
Dipartimento di Eccellenza 2018/2022

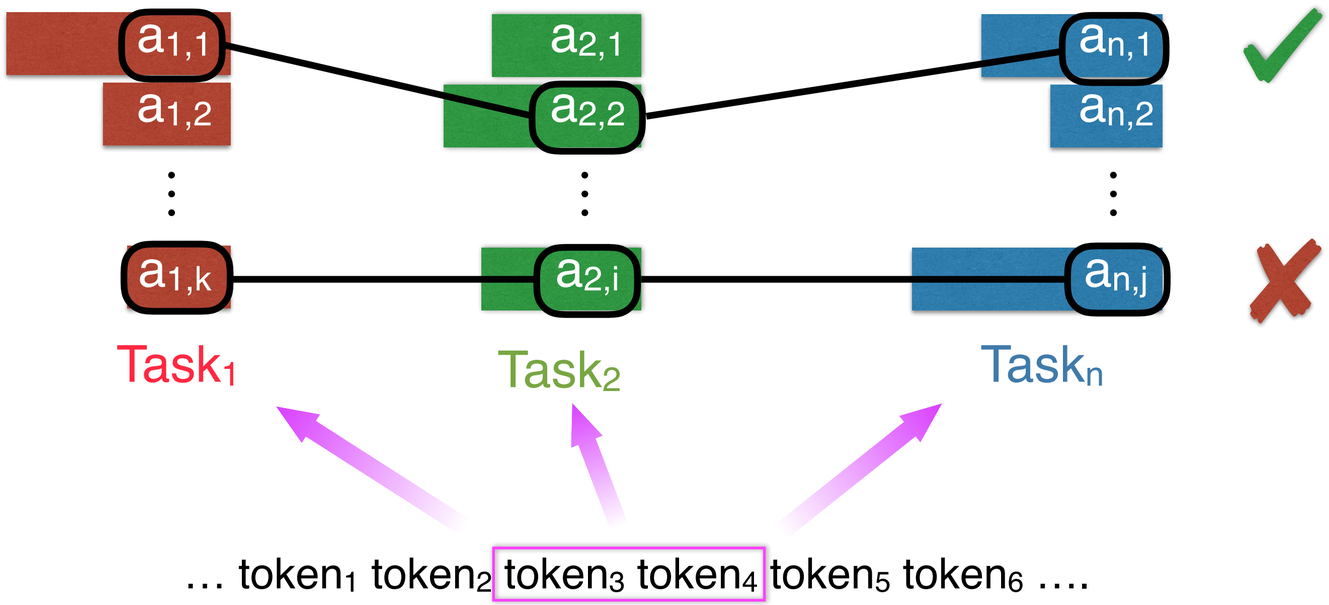
Improving NLP via Ontologies



Marco Rospocher, Francesco Corcoglioniti:
Joint Posterior Revision of NLP Annotations via Ontological Knowledge. IJCAI 2018: 4316-4322

Marco Rospocher:
An Ontology-Driven Probabilistic Soft Logic Approach to Improve NLP Entity Annotations. International Semantic Web Conference (1) 2018: 144-161

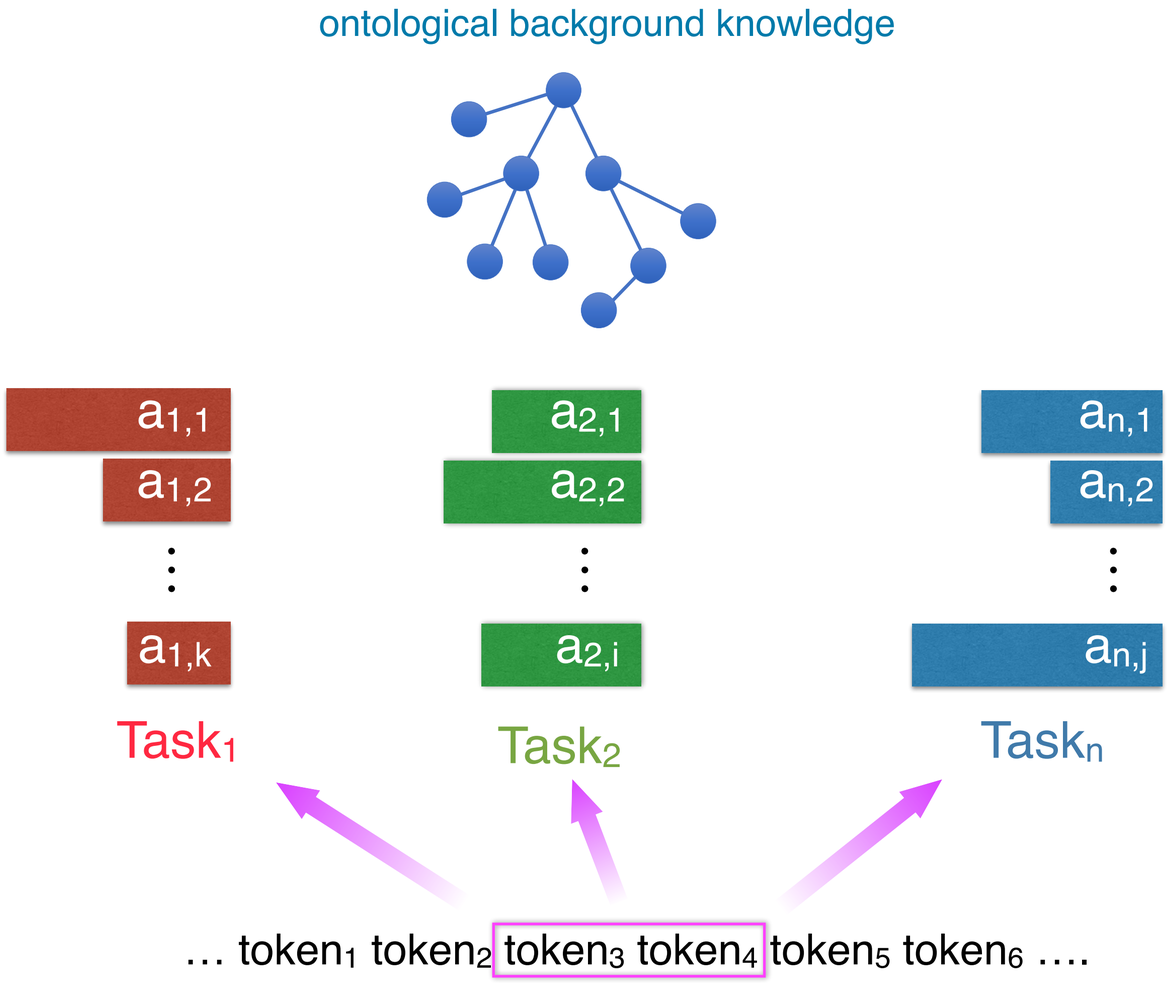
Improving NLP via Ontologies



Marco Rospocher, Francesco Corcoglioniti:
Joint Posterior Revision of NLP Annotations via Ontological Knowledge. IJCAI 2018: 4316-4322

Marco Rospocher:
An Ontology-Driven Probabilistic Soft Logic Approach to Improve NLP Entity Annotations. International Semantic Web Conference (1) 2018: 144-161

Improving NLP via Ontologies

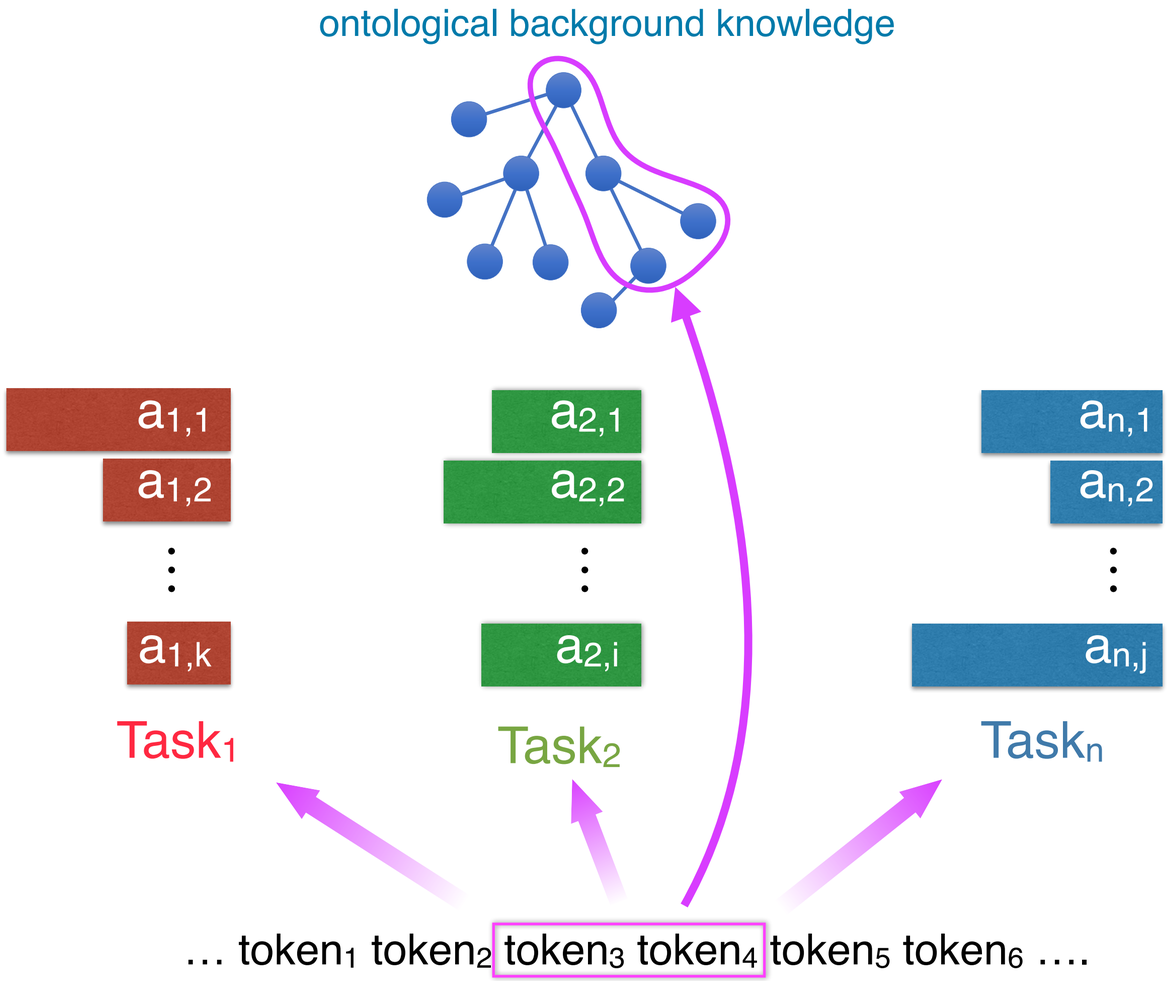


Marco Rospocher, Francesco Corcoglioniti:
Joint Posterior Revision of NLP Annotations via Ontological Knowledge. IJCAI 2018: 4316-4322

Marco Rospocher:
An Ontology-Driven Probabilistic Soft Logic Approach to Improve NLP Entity Annotations. International Semantic Web Conference (1) 2018: 144-161

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE

Improving NLP via Ontologies

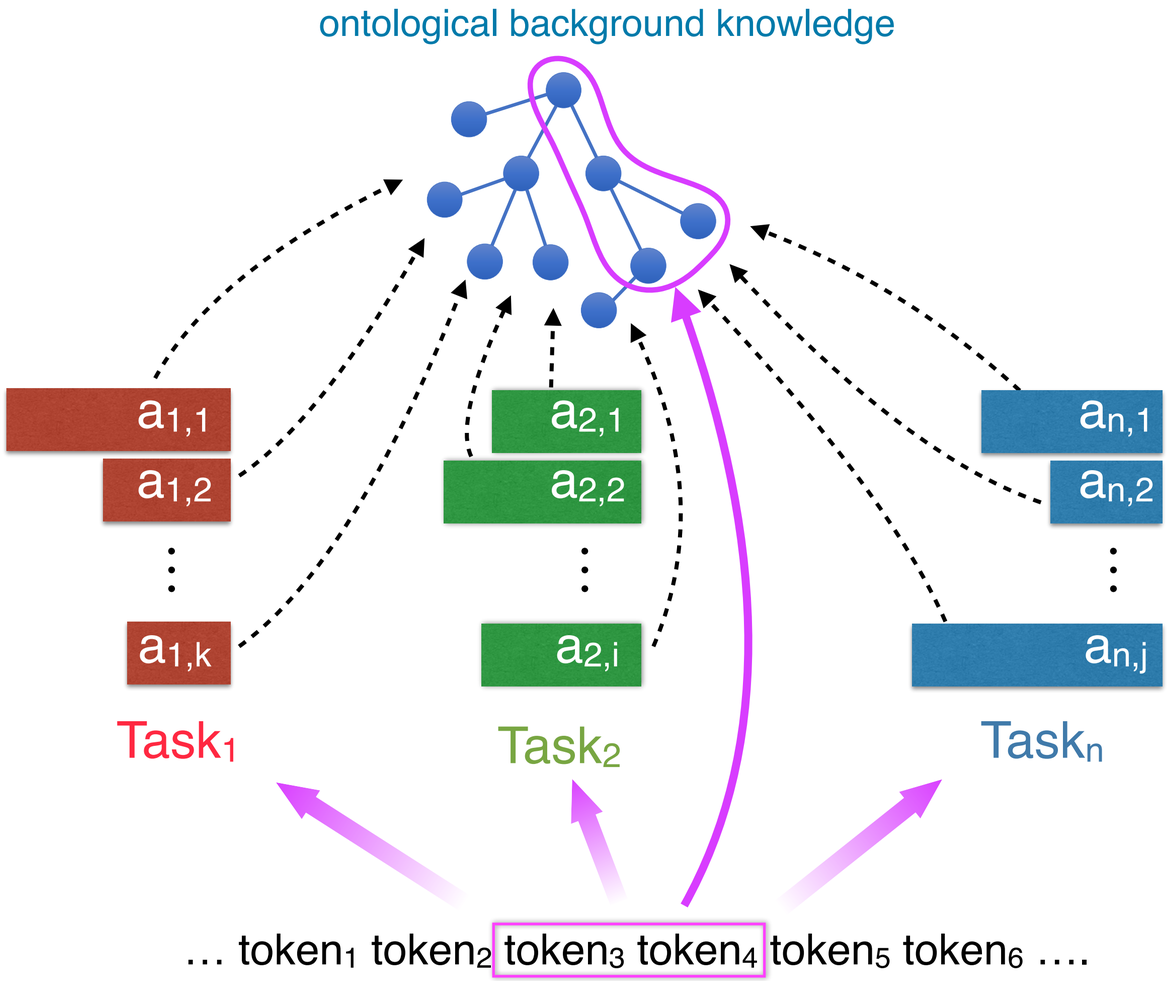


Marco Rospocher, Francesco Corcoglioniti:
Joint Posterior Revision of NLP Annotations via Ontological Knowledge. IJCAI 2018: 4316-4322

Marco Rospocher:
An Ontology-Driven Probabilistic Soft Logic Approach to Improve NLP Entity Annotations. International Semantic Web Conference (1) 2018: 144-161

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE
Dipartimento di Eccellenza 2018/2022

Improving NLP via Ontologies

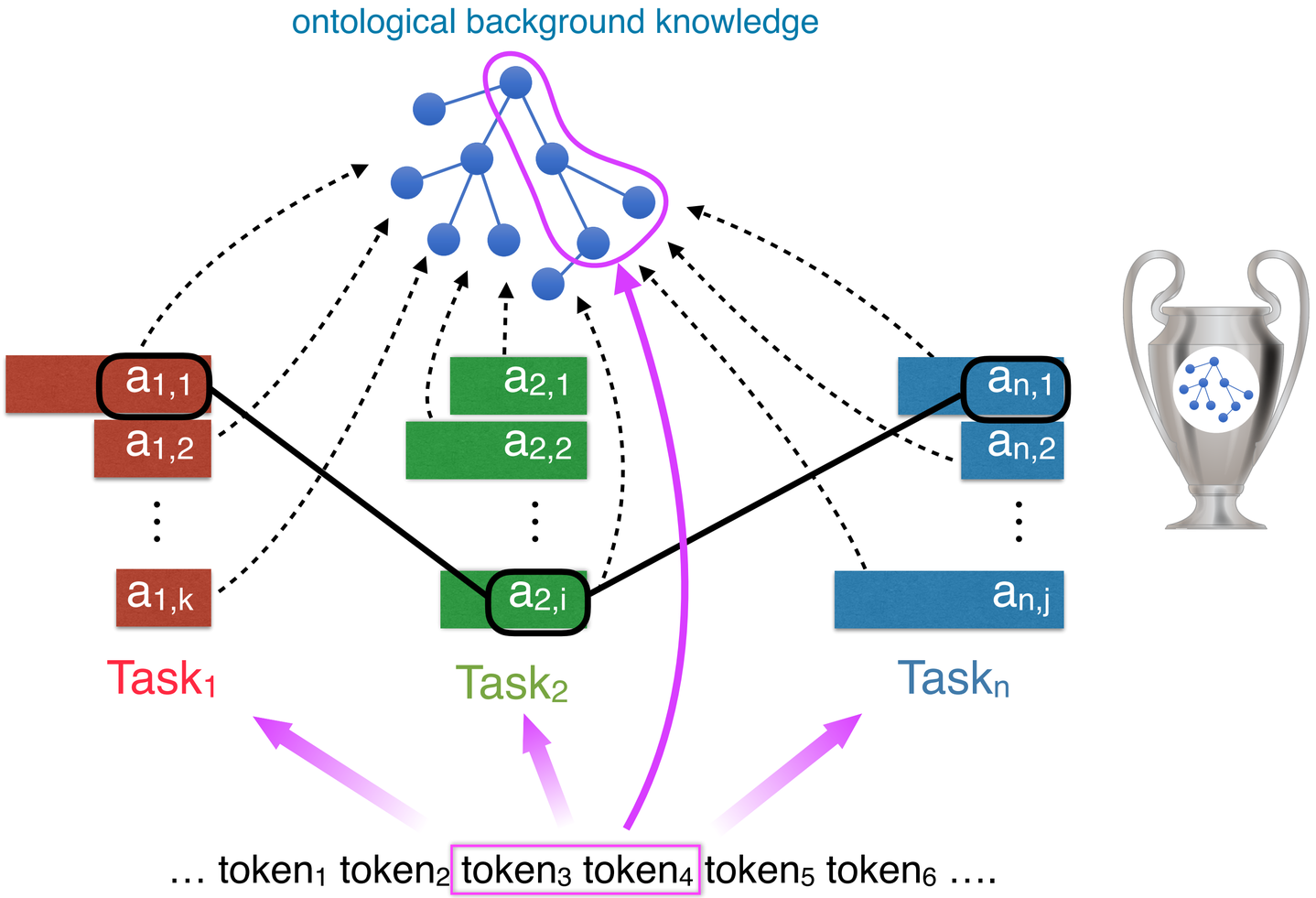


Marco Rospocher, Francesco Corcoglioniti:
Joint Posterior Revision of NLP Annotations via Ontological Knowledge. IJCAI 2018: 4316-4322

Marco Rospocher:
An Ontology-Driven Probabilistic Soft Logic Approach to Improve NLP Entity Annotations. International Semantic Web Conference (1) 2018: 144-161

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE
Dipartimento di Eccellenza 2018/2022

Improving NLP via Ontologies

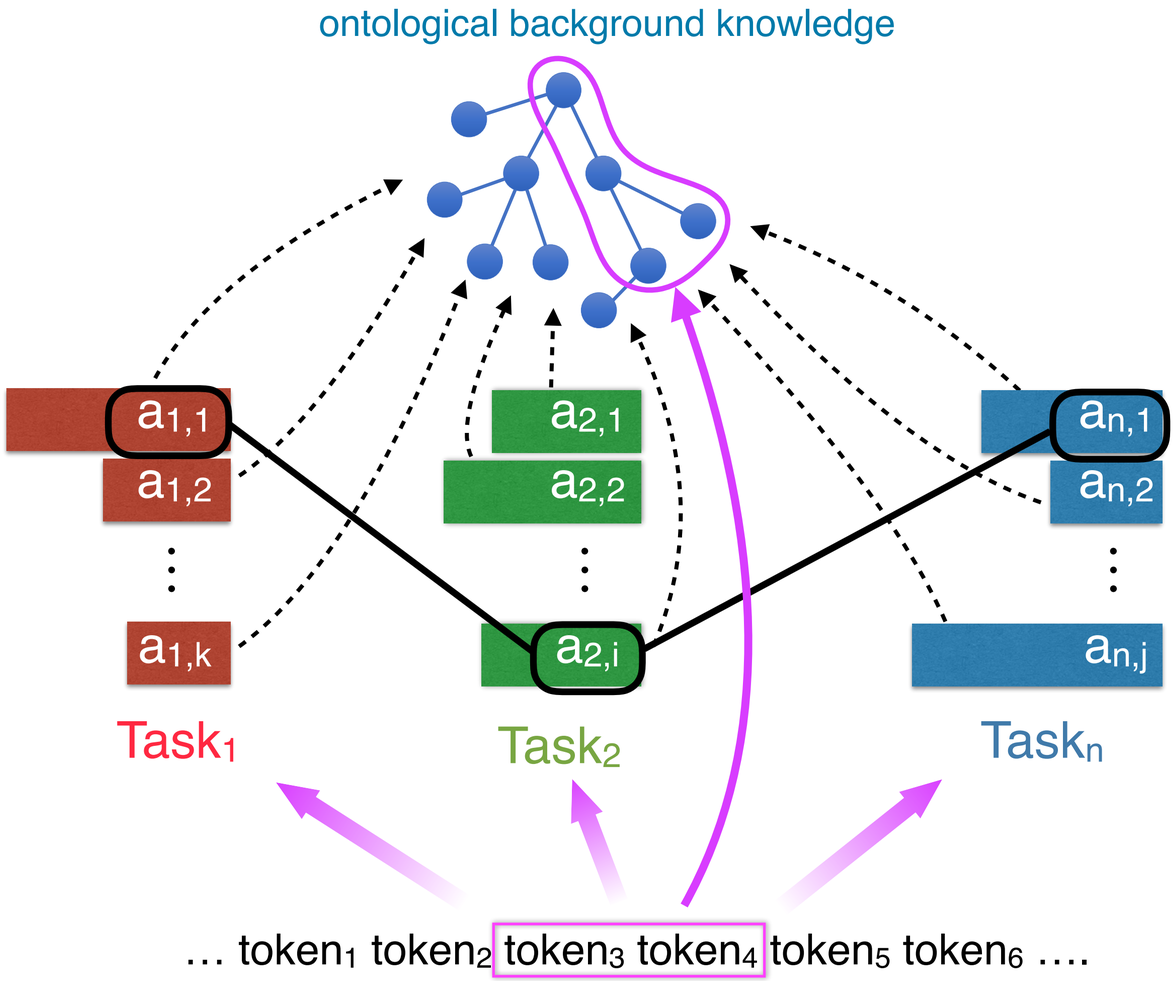


Marco Rospocher, Francesco Corcoglioniti:
Joint Posterior Revision of NLP Annotations via Ontological Knowledge. IJCAI 2018: 4316-4322

Marco Rospocher:
An Ontology-Driven Probabilistic Soft Logic Approach to Improve NLP Entity Annotations. International Semantic Web Conference (1) 2018: 144-161

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE

Improving NLP via Ontologies



JPARK

- Bayesian model that estimates the posterior probability of the combinations of NLP annotations
- https://pikes.fbk.eu/jpark.html

PSL4EA

- Probabilistic Soft Logic (PSL) model that computes the best combination of annotations via Most Probable Explanation (MPE) inference
- https://pikes.fbk.eu/psl4ea.html

Marco Rospocher, Francesco Corcoglioniti:
Joint Posterior Revision of NLP Annotations via Ontological Knowledge. IJCAI 2018: 4316-4322

Marco Rospocher:
An Ontology-Driven Probabilistic Soft Logic Approach to Improve NLP Entity Annotations. International Semantic Web Conference (1) 2018: 144-161

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE

KnowledgeStore

https://knowledgestore.fbk.eu/

NewsReader

EU-FP7

KnowledgeStore UI | Lookup | SPARQL query | Reports

```
SELECT DISTINCT ?event ?event_label ?year ?month ?day
WHERE {
    ?event a sem:Event, eso:JoiningAnOrganization .
    ?event rdfs:label ?event_label .
    ?event eso:employment-employer dbpedia:Kia_Motors .
    ?event sem:hasTime ?time .
    ?time owltime:inDateTime ?time_owl .
    ?time_owl owltime:year ?year; owltime:month ?month; owltime:day ?day .
}
ORDER BY ?year ?month ?day
```

Timeout 600 s | Display results | Download as... | example query

6 results in 10064 ms show / hide query panel

event	event_label	year	month	day
<..#ev44>	hire	2005	9	18
<..#ev4>	hire	2006	8	1
<..#ev31>	hiring	2006	8	1
<..#ev8>	hire	2010	1	13
<..#ev2>	hire	2010	1	14
<..#ev8>	hire	2010	1	14

KnowledgeStore

https://knowledgestore.fbk.eu/

NewsReader

EU-FP7

KnowledgeStore UI | Lookup | SPARQL query | Reports

```
SELECT DISTINCT ?event ?event_label ?year ?month ?day
WHERE {
    ?event a sem:Event, eso:JoiningAnOrganization .
    ?event rdfs:label ?event_label .
    ?event eso:employment-employer dbpedia:Kia_Motors .
    ?event sem:hasTime ?time .
    ?time owltime:inDateTime ?time_owl .
    ?time_owl owltime:year ?year; owltime:month ?month; owltime:day ?day .
}
ORDER BY ?year ?month ?day
```

Show me all the hiring events where KIA was involved as employer

Timeout 600 s | Display results | Download as... | example query

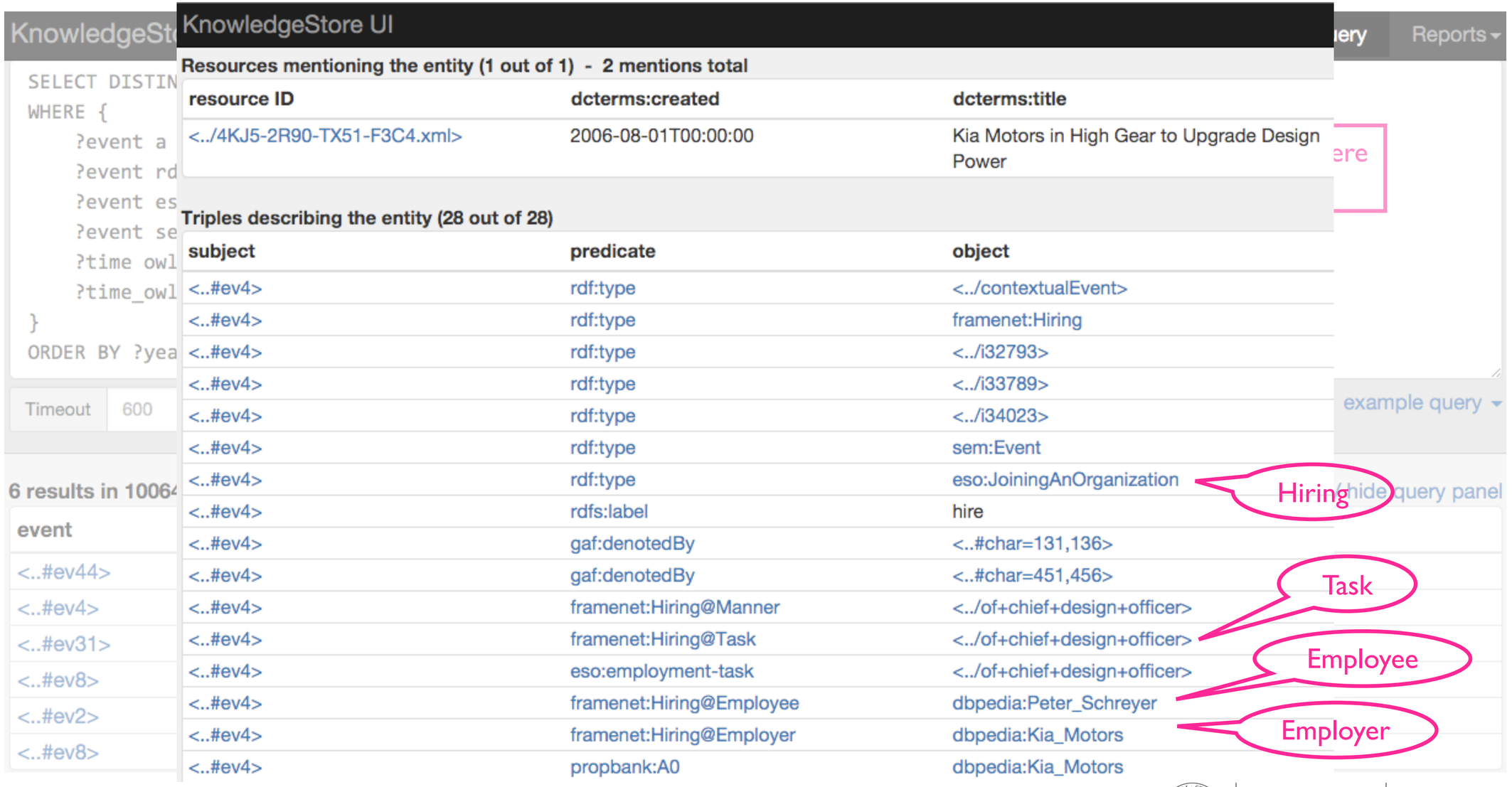
6 results in 10064 ms show / hide query panel

event	event_label	year	month	day
<..#ev44>	hire	2005	9	18
<..#ev4>	hire	2006	8	1
<..#ev31>	hiring	2006	8	1
<..#ev8>	hire	2010	1	13
<..#ev2>	hire	2010	1	14
<..#ev8>	hire	2010	1	14

KnowledgeStore

https://knowledgestore.fbk.eu/

EU-FP7



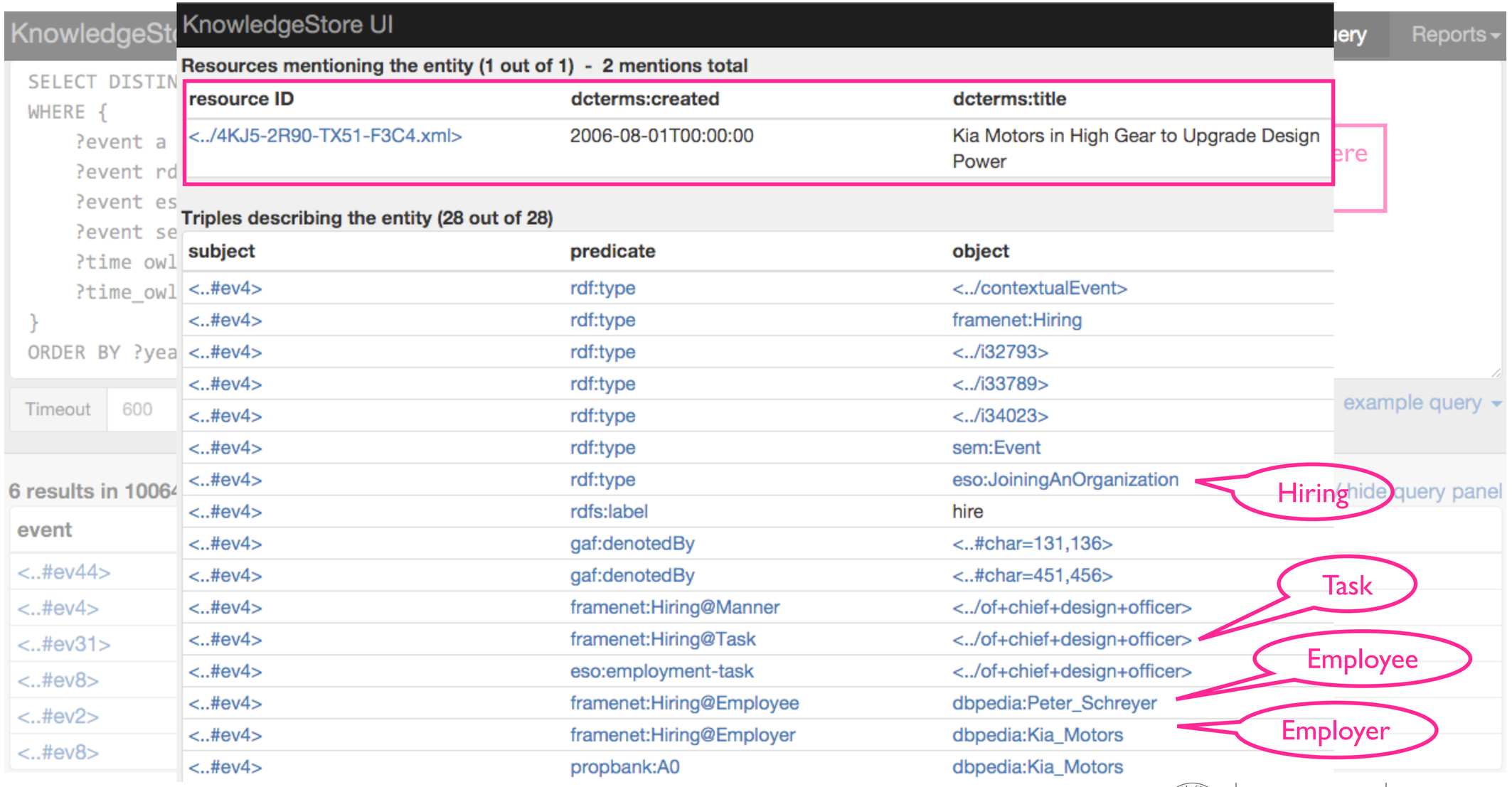
Francesco Corcoglioniti, **Marco Rospocher**, Roldano Cattoni, Bernardo Magnini, Luciano Serafini:
The KnowledgeStore: A Storage Framework for Interlinking Unstructured and Structured Knowledge. Int. J. Semantic Web Inf. Syst. 11(2): 1-35 (2015)

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE

KnowledgeStore

https://knowledgestore.fbk.eu/

EU-FP7



Francesco Corcoglioniti, **Marco Rospocher**, Roldano Cattoni, Bernardo Magnini, Luciano Serafini:
The KnowledgeStore: A Storage Framework for Interlinking Unstructured and Structured Knowledge. Int. J. Semantic Web Inf. Syst. 11(2): 1-35 (2015)

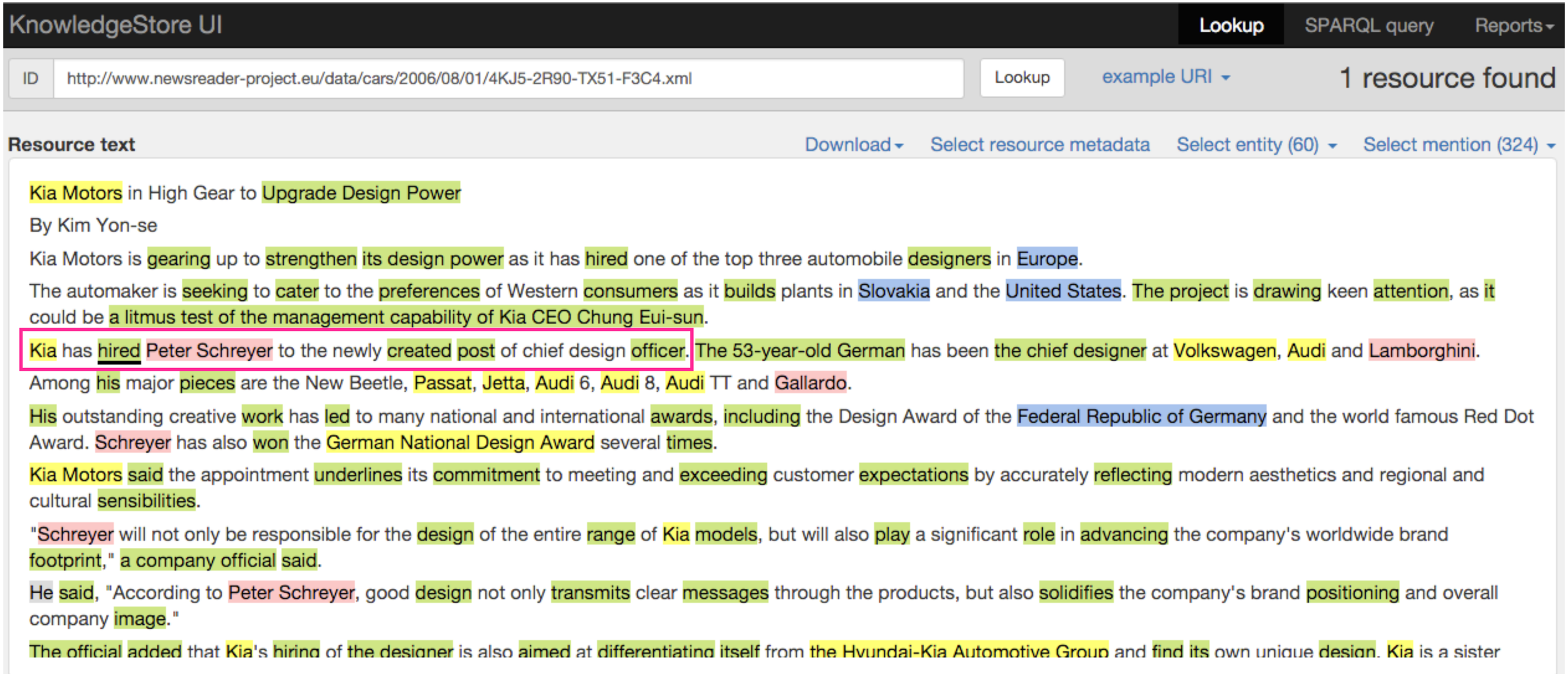
UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE

KnowledgeStore

https://knowledgestore.fbk.eu/

NewsReader

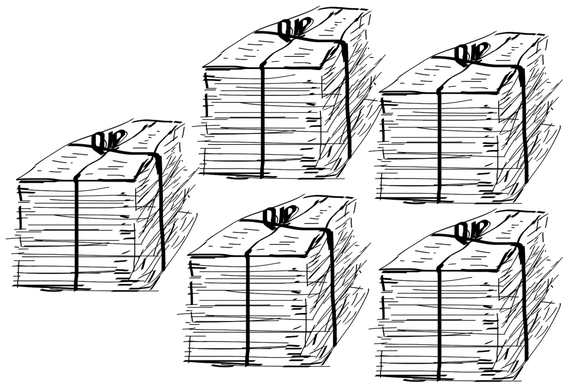
EU-FP7



Francesco Corcoglioniti, **Marco Rospocher**, Roldano Cattoni, Bernardo Magnini, Luciano Serafini:
The KnowledgeStore: A Storage Framework for Interlinking Unstructured and Structured Knowledge. Int. J. Semantic Web Inf. Syst. 11(2): 1-35 (2015)

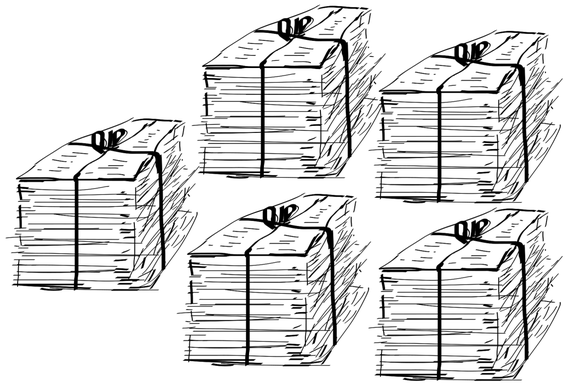
UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE
Dipartimento di Eccellenza 2018/2022

Ontology Population for IR



Ontology Population for IR

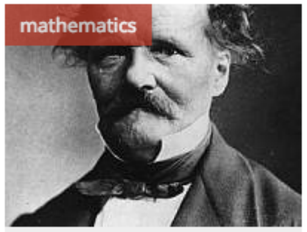
astronomers influenced by Gauss



UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE
Dipartimento di Eccellenza 2018/2022

Ontology Population for IR

astronomers influenced by Gauss



Ernst Kummer and his Achievements in Mathematics

29. January 201 ld Sack

On January 29, 1 hematician Ernst Eduard Kummer was born. One of his ons is the introduction of ideal numbers, which are defined as al subgroup of a ring fundamental theorem of arithmetic to complex discovered the fourth order surface based on...

2

Heinrich Olbers and the Olbers' Paradox

11. October 2014 Harald Sack

Heinrich Olbers (17 ctober 11, 1758, German physician and astronomer Heinri thias Olbers was born. Besides his discovery of coments planets, Olbers is best known for his new method to calculate the velocity of falling stars. Maybe you have also heard of the famous Olbers' paradox, which asks...

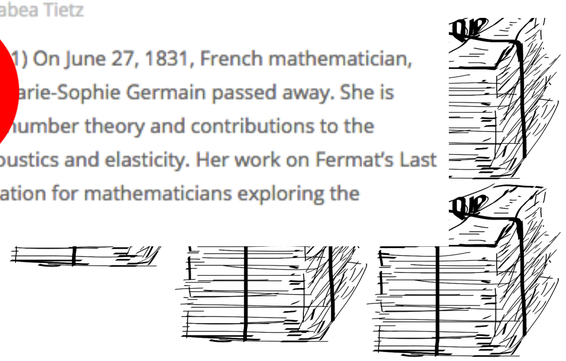
1

Sophie Germain and the Chladni Ex

27. June 2014 0 Tabea Tietz

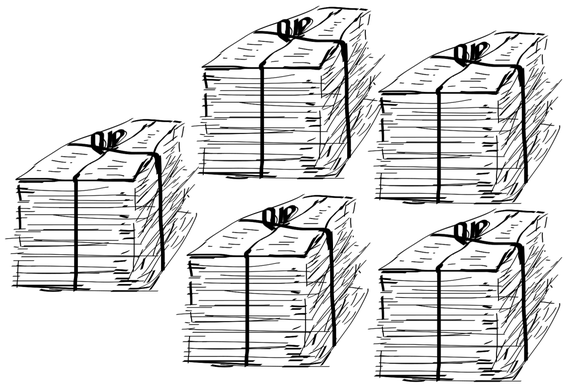
Sophie Germ 1) On June 27, 1831, French mathematician, physicist, ar arie-Sophie Germain passed away. She is best known umber theory and contributions to the applied mathem acoustics and elasticity. Her work on Fermat's Last Theorem provided a foundation for mathematicians exploring the subject...

3



Ontology Population for IR

astronomers influenced by Gauss



PIKES + DBpedia + YAGO (select knowledge) = Better Retrieval!

KE4IR

- Enriched Vector Space Model (VSM)
- Evaluated on several datasets (e.g., WES, TREC 6-7-8-9-2001)
- https://pikes.fbk.eu/ke4ir.html

Marco Rospocher, Francesco Corcoglioniti, Mauro Dragoni:
Boosting Document Retrieval with Knowledge Extraction and Linked Data. Semantic Web 10(4): 753-778 (2019)

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE

Where am I heading to?

Beyond just reading… Understanding!

Yesterday, Kia has hired Peter Schreyer as chief design officer.

[Newspaper, 2 Aug 2006]

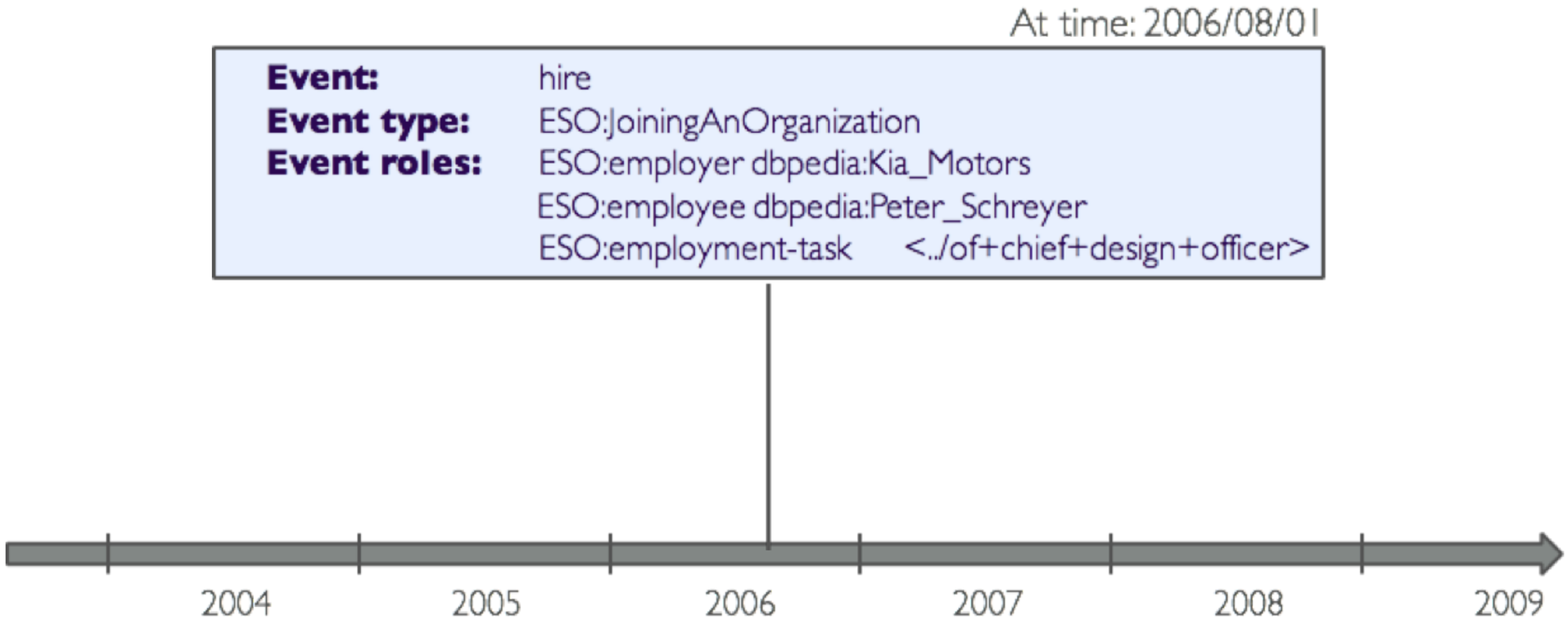
Roxane Segers, Piek Vossen, **Marco Rospocher**, Luciano Serafini, Egoitz Laparra and German Rigau:
ESO: a Frame based Ontology for Events and Implied Situations. In Proceedings of the MAPLEX 2015 Workshop (2015)

Francesco Corcoglioniti, **Marco Rospocher**, Michele Mostarda, Marco Amadori:
Processing billions of RDF triples on a single machine using streaming and sorting. ACM-SAC 2015: 368-375 (2015)

Beyond just reading… Understanding!

Yesterday, Kia has hired Peter Schreyer as chief design officer.

[Newspaper, 2 Aug 2006]



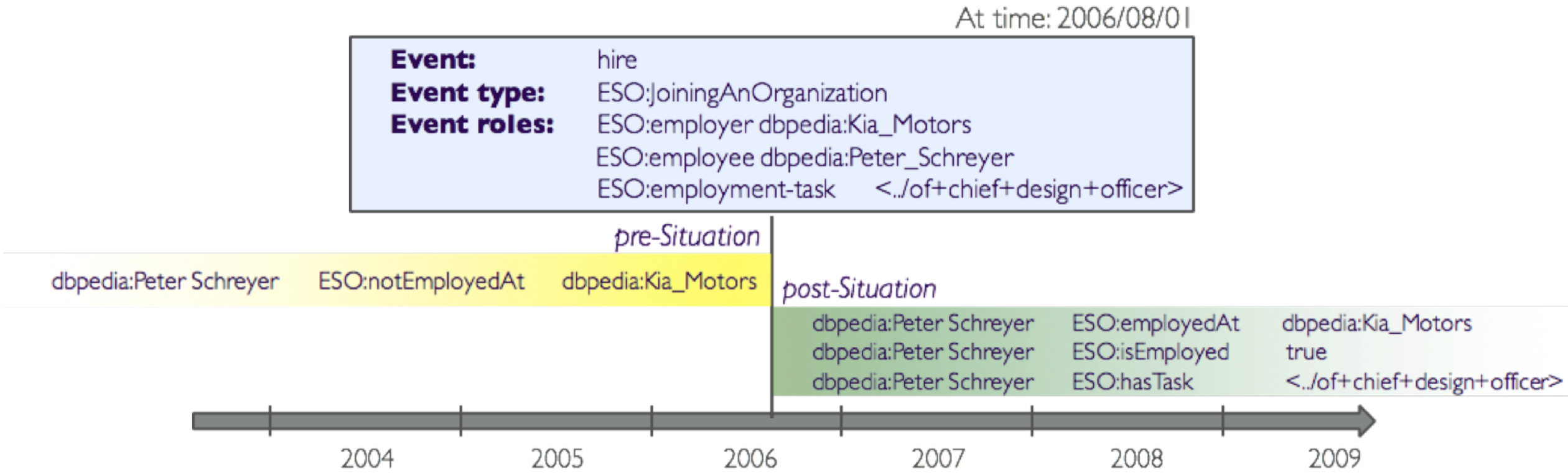
Roxane Segers, Piek Vossen, **Marco Rospocher**, Luciano Serafini, Egoitz Laparra and German Rigau:
ESO: a Frame based Ontology for Events and Implied Situations. In Proceedings of the MAPLEX 2015 Workshop (2015)

Francesco Corcoglioniti, **Marco Rospocher**, Michele Mostarda, Marco Amadori:
Processing billions of RDF triples on a single machine using streaming and sorting. ACM-SAC 2015: 368-375 (2015)

Beyond just reading… Understanding!

Yesterday, Kia has hired Peter Schreyer as chief design officer.

[Newspaper, 2 Aug 2006]



Roxane Segers, Piek Vossen, **Marco Rospocher**, Luciano Serafini, Egoitz Laparra and German Rigau:
ESO: a Frame based Ontology for Events and Implied Situations. In Proceedings of the MAPLEX 2015 Workshop (2015)

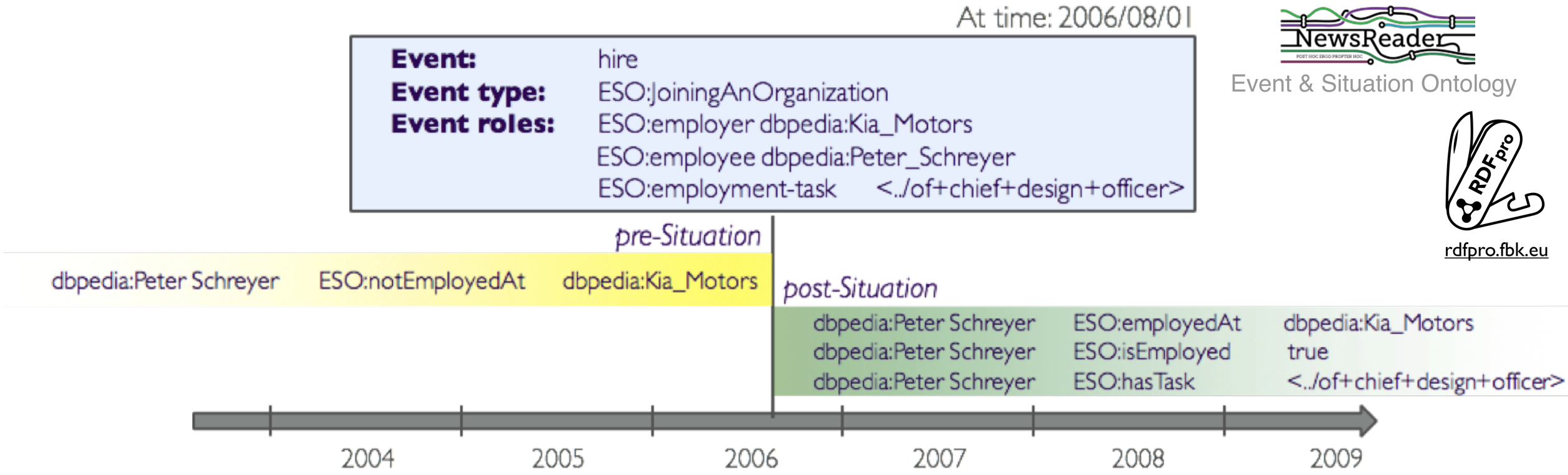
Francesco Corcoglioniti, **Marco Rospocher**, Michele Mostarda, Marco Amadori:
Processing billions of RDF triples on a single machine using streaming and sorting. ACM-SAC 2015: 368-375 (2015)

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE

Beyond just reading… Understanding!

Yesterday, Kia has hired Peter Schreyer as chief design officer.

[Newspaper, 2 Aug 2006]



NewsReader

Event & Situation Ontology

RDFpro

rdfpro.fbk.eu

Roxane Segers, Piek Vossen, **Marco Rospocher**, Luciano Serafini, Egoitz Laparra and German Rigau:
ESO: a Frame based Ontology for Events and Implied Situations. In Proceedings of the MAPLEX 2015 Workshop (2015)

Francesco Corcoglioniti, **Marco Rospocher**, Michele Mostarda, Marco Amadori:
Processing billions of RDF triples on a single machine using streaming and sorting. ACM-SAC 2015: 368-375 (2015)

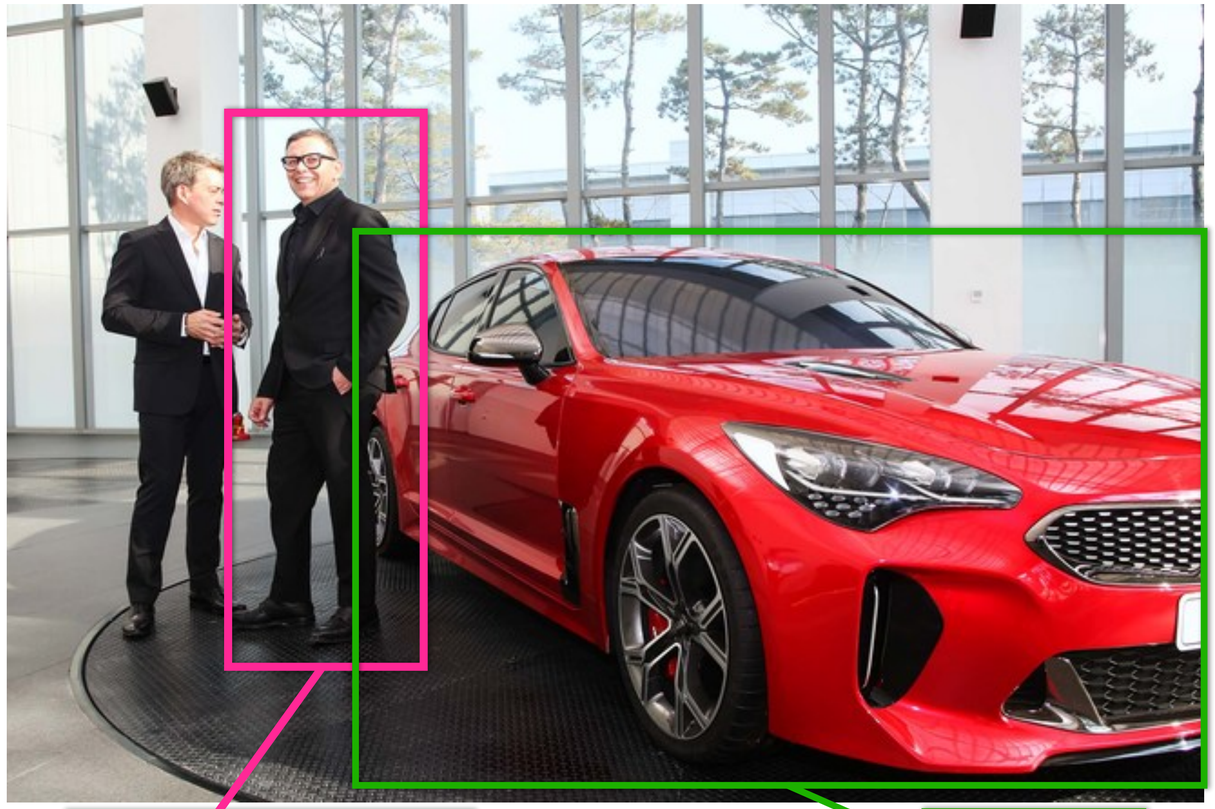
UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE

Beyond just text… Multi-modalities!

Peter Schreyer showcasing a new KIA car

Shahi Dost, Luciano Serafini, **Marco Rospocher**, Lamberto Ballan, Alessandro Sperduti: VTKEL: A resource for Visual-Textual-Knowledge Entity Linking, Accepted at ACM-SAC2020 Dataset. https://doi.org/10.6084/m9.figshare.7882781.v5

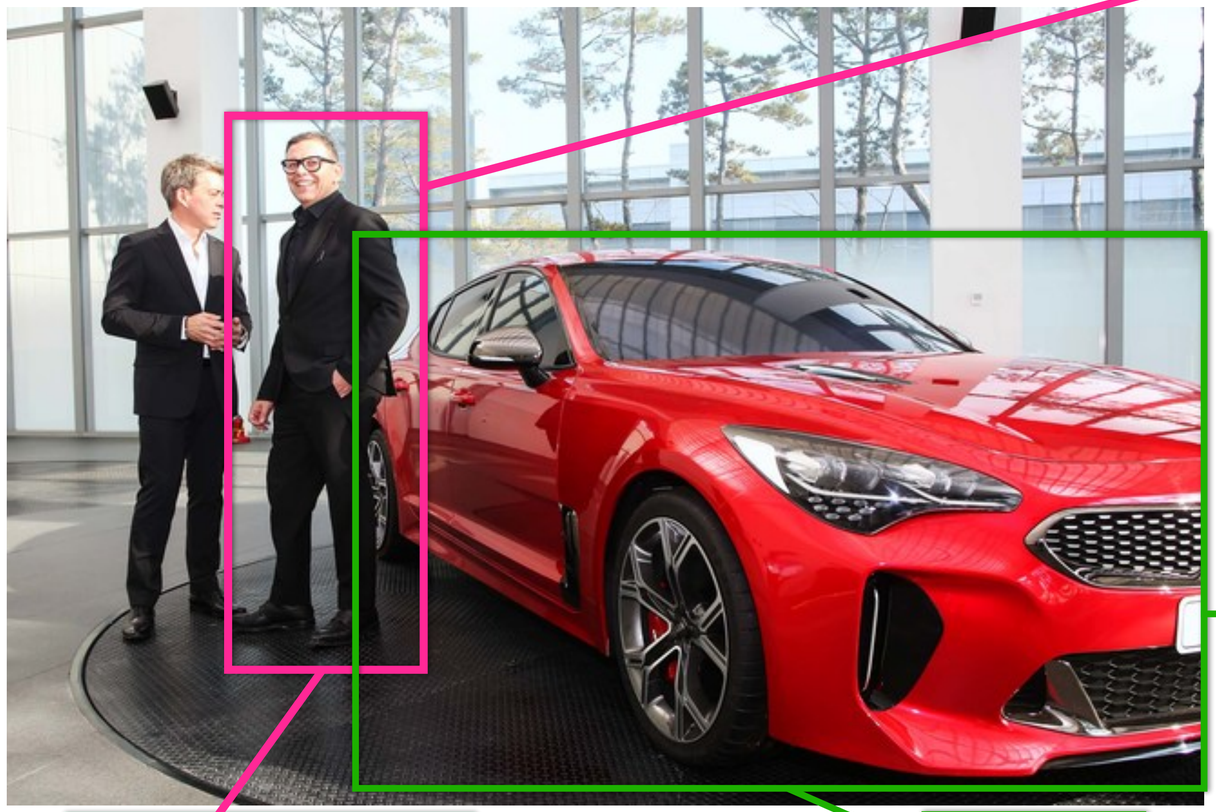
Beyond just text… Multi-modalities!



Shahi Dost, Luciano Serafini, **Marco Rospocher**, Lamberto Ballan, Alessandro Sperduti:
VTKEL: A resource for Visual-Textual-Knowledge Entity Linking, Accepted at ACM-SAC2020
Dataset. https://doi.org/10.6084/m9.figshare.7882781.v5

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE

Beyond just text… Multi-modalities!



Peter Schreyer showcasing a new KIA car

dbpedia:Peter_Schreyer

dbo:abstract	Peter Schreyer (born 1953) is a German automobile designer for Hyundai Motor and Kia Motors, widely known for his design contributions to the Audi TT. He has been the chief design officer at Kia Motors since 2006 and [..]. (en)
dbo:birthDate	1953-1-1
dbo:birthPlace	dbr:Bavaria dbr:West_Germany dbr:Bad_Reichenhall
dbo:education	dbr:Munich_University_of_Applied_Sciences dbr:Royal_College_of_Art
dbo:nationality	dbr:Germany

dbpedia:Kia_Motors

dbo:abstract	▪ Kia Motor Corporation (Hangul: 기아자동차; hanja: 起亞自動車, IPA: [ki.a]) (stylized as KIA), headquartered in Seoul, is South Korea's second-largest automobile manufacturer, following the Hyundai Motor Company, with sales of over 3.3 million vehicles in 2015.[..] (en)
dbo:location	▪ dbr:Seoul
dbo:equity	▪ 1.99E10
dbo:industry	▪ dbr:Automotive_industry

Shahi Dost, Luciano Serafini, **Marco Rospocher**, Lamberto Ballan, Alessandro Sperduti: VTKEL: A resource for Visual-Textual-Knowledge Entity Linking, Accepted at ACM-SAC2020 Dataset. https://doi.org/10.6084/m9.figshare.7882781.v5

Metadata Quality in DH archives

- “Automatize” the Bruce and Hillman [“Metadata in Practice,” ALA Editions, 2004] framework for metadata curation
  - metadata: completeness, accuracy, logical consistency,…

- Use Case:

JPARK
pikes.fbk.eu/jpark

KE4IR
pikes.fbk.eu/ke4ir

BPMN Ontology
dkm.fbk.eu/bpmn-ontology

PREMON
premon.fbk.eu

PESCaDO
bit.ly/pescado-onto

NewsReader
Event & Situation Ontology
github.com/newsreader/eso

PSL4EA
pikes.fbk.eu/psl4ea

PIKES
pikes.fbk.eu

Marco Rospocher

http://marcorospocher.com/ @ marco.rospocher@univr.it @marcorospocher

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE
Dipartimento di Eccellenza 2018/2022 MIUR

RDFpro
rdfpro.fbk.eu

KnowledgeStore
knowledgestore.fbk.eu

MoKi
the Modelling WiKi ---
moki.fbk.eu

TEX-OWL
github.com/dkmfbk/TexOwl

JPARK
pikes.fbk.eu/jpark

KE4IR
pikes.fbk.eu/ke4ir

BPMN Ontology
dkm.fbk.eu/bpmn-ontology

PREMON
premon.fbk.eu

PESCaDO
bit.ly/pescado-onto

NewsReader
Event & Situation Ontology
github.com/newsreader/eso

PSL4EA
pikes.fbk.eu/psl4ea

PIKES
pikes.fbk.eu

Marco Rospocher

http://marcorospocher.com/ @ marco.rospocher@univr.it @marcorospocher

UNIVERSITÀ di VERONA
Dipartimento di LINGUE E LETTERATURE STRANIERE
Dipartimento di Eccellenza 2018/2022 MIUR

Slide:
http://bit.ly/2019DIuniVR

RDFpro
rdfpro.fbk.eu

KnowledgeStore
knowledgestore.fbk.eu

MoKi
the Modelling WiKi ---
moki.fbk.eu

TEX-OWL
github.com/dkmfbk/TexOwl