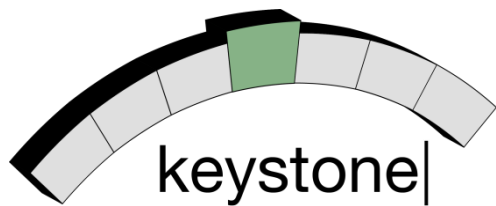


Hackathon

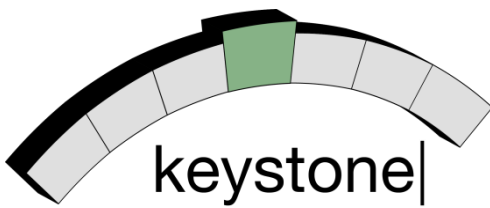
Group 2 led by Mauro Dragoni



The members of the group

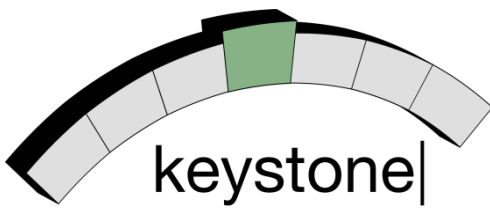
- ▶ A table mapping team members' expertise in the aspects related to the solution to the query

| Name | Step 1: access to the data sources | Step 2: solution to the query | Step 3: evaluation |
|----------------------|---|---|--------------------|
| Mauro Dragoni | Concept indexing Text annotation | Document retrieval guided by semantics | |
| Antonio Fariña | Compact RDF representation RDF Indexing | | |
| Bernard Espinasse | text indexing annotation | | |



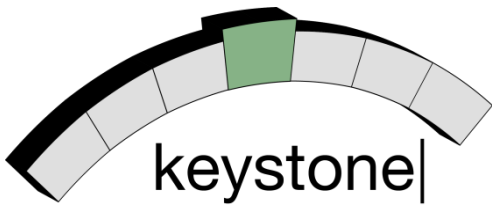
The members of the group

| Name | Step 1: access to the data sources | Step 2: solution to the query | Step 3: evaluation |
|----------------|--|-------------------------------|--|
| Dragana Dudic | Ontology learning, ontology population | NER | |
| Atanas Hristov | High performance computing | big data analytics | |
| Hussain Mahdi | Text Categorization, Text Indexing, keyword indexing | | real-time response of system, subjective user satisfaction |

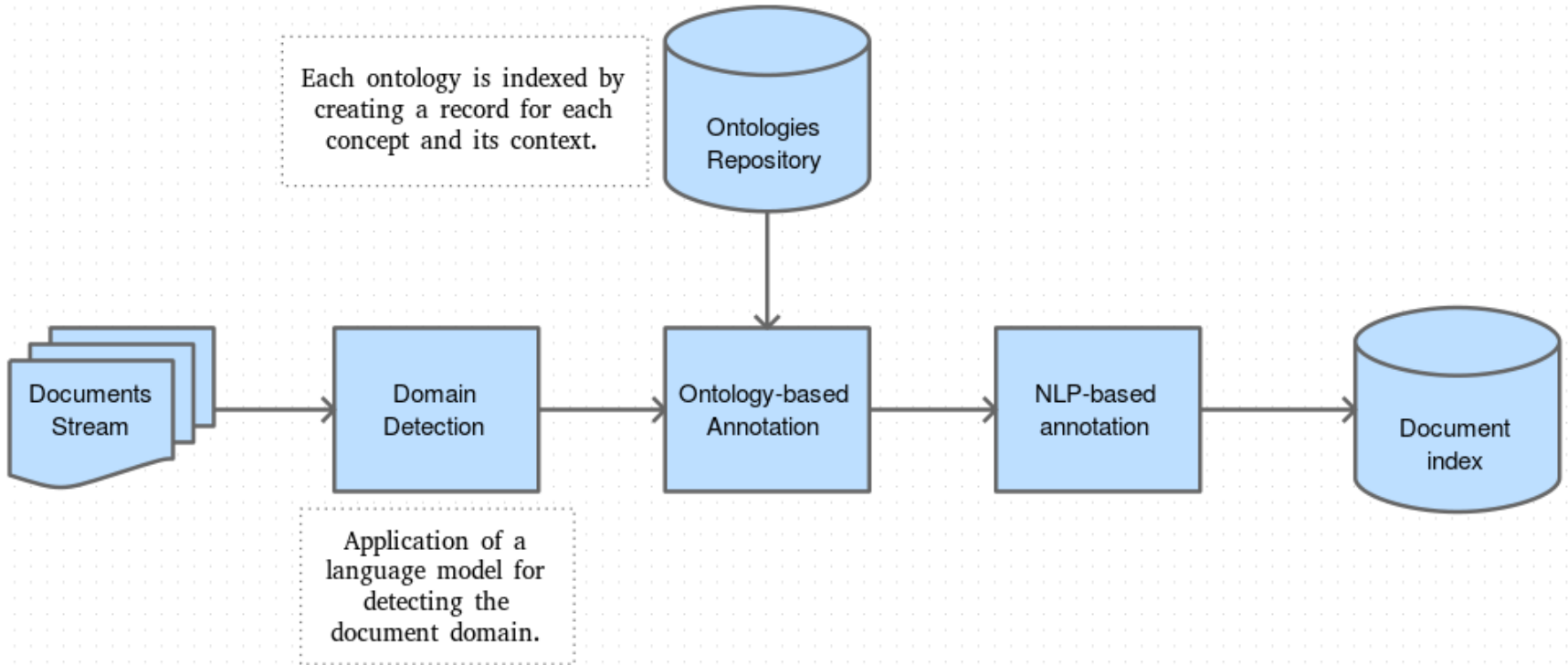


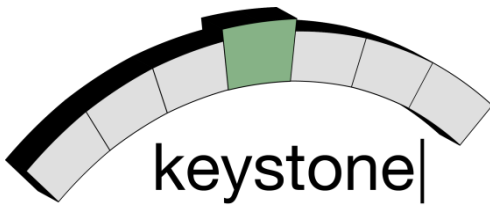
The members of the group

| Name | Step 1: access to the data sources | Step 2: solution to the query | Step 3: evaluation |
|--------------------|---|--|--------------------|
| Claudia Ifrim | text indexing | active learning algorithms | |
| Wouter Addink | specific for biodiversity: TDWG Ontology, DwC and ABCD | ES Indexing | |
| Raquel Trillo_Lado | Semantic relatedness measures and Word Sense Disambiguation Ontology Matching Ontology Population | Translation of Keyword queries to SQL and SPARQL queries | |



Functional Architecture-Backend





Challenges and Issues

▶ Datasets:

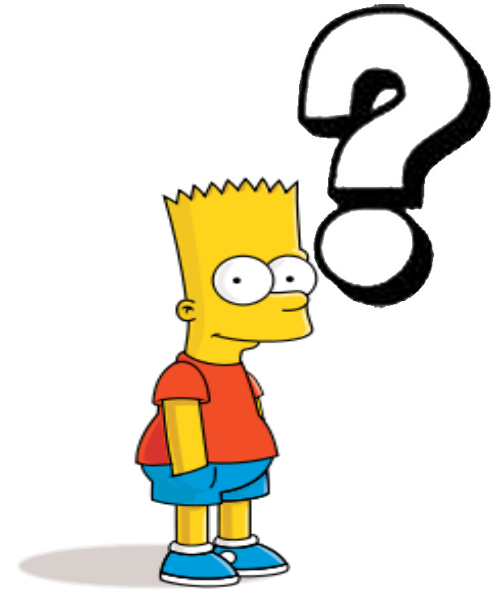
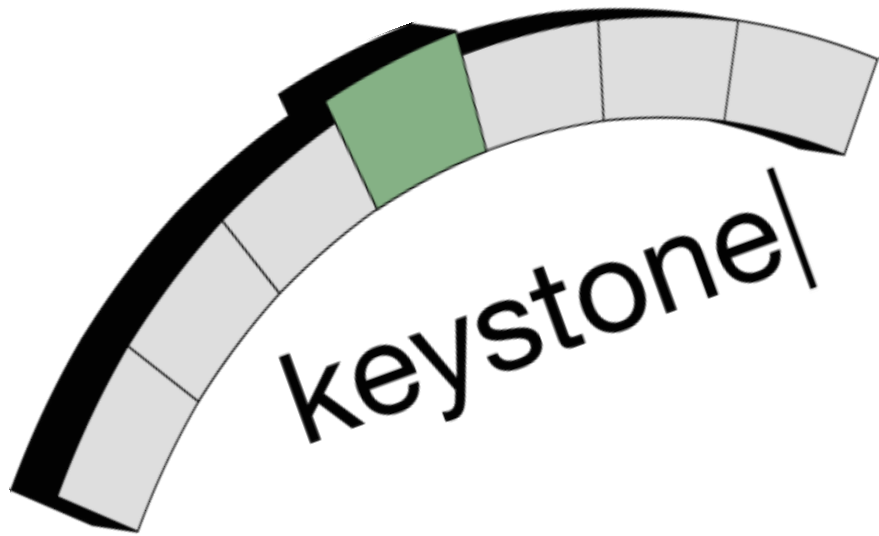
- ▶ There are many possible domains but there are not domain ontologies for all of them
- ▶ Some ontologies can be very poor. Measurements to evaluate the quality of ontologies

▶ Query Methods:

- ▶ Translating the user's Keyword query into a SPARQL query if the annotations of the documents are stored in RDF.
- ▶ The same query for different users can have a different translation according to user's preferences.

▶ Evaluation:

- ▶ In the medical domain:
 - **BioAsq.org**: *A challenge on large-scale biomedical semantic indexing and question answering*
 - *MuchMore collection (muchmore.dfki.de)*



Thank you very much!

Group 2 led by Mauro Dragoni