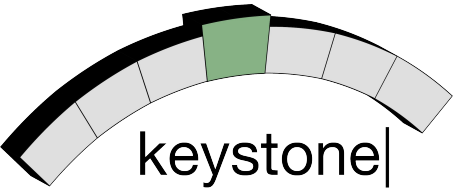
keystone

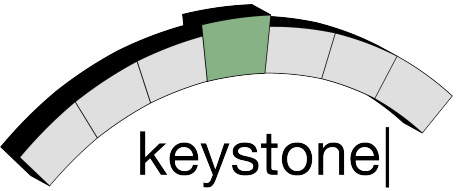
# Group 3

Ana Mestrovic
Andrea Calì
Elton Domnori
John Breslin
Jose Viqueira
Martin Ugarte
Slavko Žitnik
Velislava Stoykova



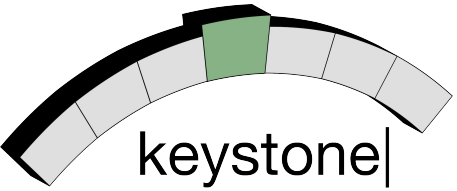

# The model: graph based

- We use a **directed** graph
- The data/ontology representation is **flat** (from the logical point of view)
- Document terms and ontology terms are nodes
- Arcs are labelled with a **relation type** (e.g. generalisation, similarity, geo proximity)
- A **layer** is a subgraph identified by a set of arc labels (relation types) and nodes specific to that layer




# The model (cont'd)

- Nodes corresponding to doc terms carry a score; they are common to all layers

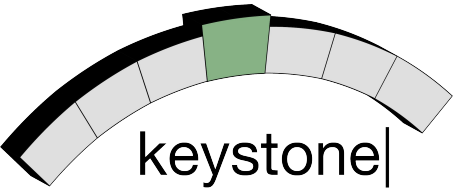



# General overview

**Input:** set of keywords

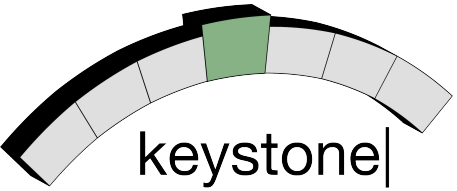
**Data:** documents; ontology; vocabulary

**Output:** ranked list of documents



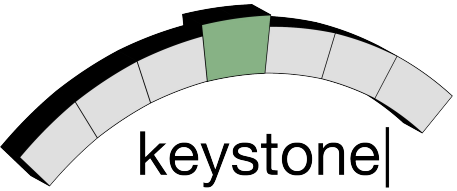

# Components

- Expand query
- Apply layers, one by one
- Build a graph
- Use graph to query the document set
- Rank the results




# Algorithm

- For each layer:
  - For each keyword, represent the keyword in terms of the base nodes (as a vector of N elements)
  - Based on layer type (relationships between leaf nodes, proximity, etc.), comput scores of vector elements
  - Create a vector with N elements that represents the query
- This weighted vector (per layer) will be used to retrieve documents later
- Vectors will be weighted and combined based on queries, user profiles

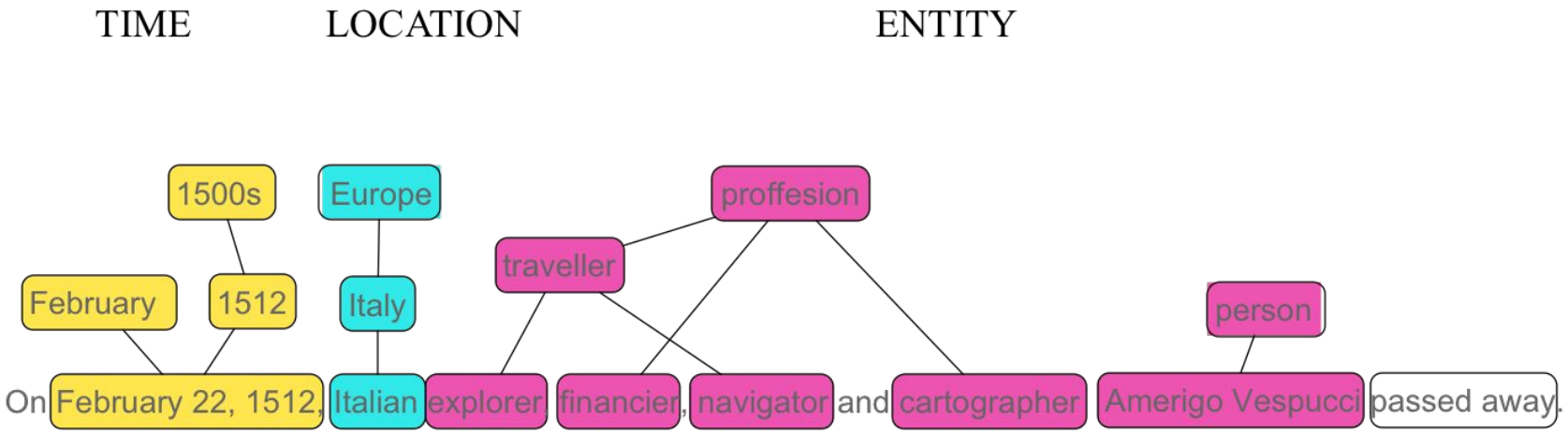



# A toy example

http://blog.yovisto.com/amerigo-vespucci-and-the-new-world/

*“On February 22, 1512, Italian explorer, financier, navigator and cartographer Amerigo Vespucci passed away. He first demonstrated that Brazil and the West Indies did not represent Asia’s eastern outskirts.”*

# Layers based on example



keystone

# Functional architecture

Query keywords

*{February, 1512, Amerigo, Italian, explorer, ...}*

Ontology

Query expander

Layer applier and graph search

Results ranker

Ranked documents

# Mapping team member expertise

| Name | Step 1: Access to datasets/sources | Step 2: Solutions to query (methods) | Step 3: Evaluation |
|---|---|---|---|
| Ana Mestrovic | √ | | |
| Andrea Calì | √ | | √ (maybe) |
| Elton Domnori | | √ | |
| John Breslin | √ | | |
| Jose Viqueira | √ | | |
| Martin Ugarte | | √ | |
| Slavko Žitnik | | √ | √ |
| Velislava Stoykova | √ | | |