Linked Data Maps

A visual entry point for the exploration of datasets

IESD2015 - ***Fabio Valsecchi***, *Matteo Abrate, Clara Bacciu, Maurizio Tesconi, Andrea Marchetti*

Motivation

- Users always ask ***"What is the dataset like?"***
- Linked Data sets are difficult to make sense to **non-experts** of Semantic Web:
  - Content (Data)
  - Structure (Ontologies)
- **Visualizing** or **exploring** LD sets is difficult:
  - Volume
  - Complexity

LD visualization tools

Applications like *LODlive, RelFinder, DBpedia viewer, LOD Visualization*, … feature **some but not all** of the following:

- description of a **single instance**
- exploration of **small groups of instances**
- presentation of a **summary of the whole dataset**

None of them follows ***Shneiderman's Mantra***.

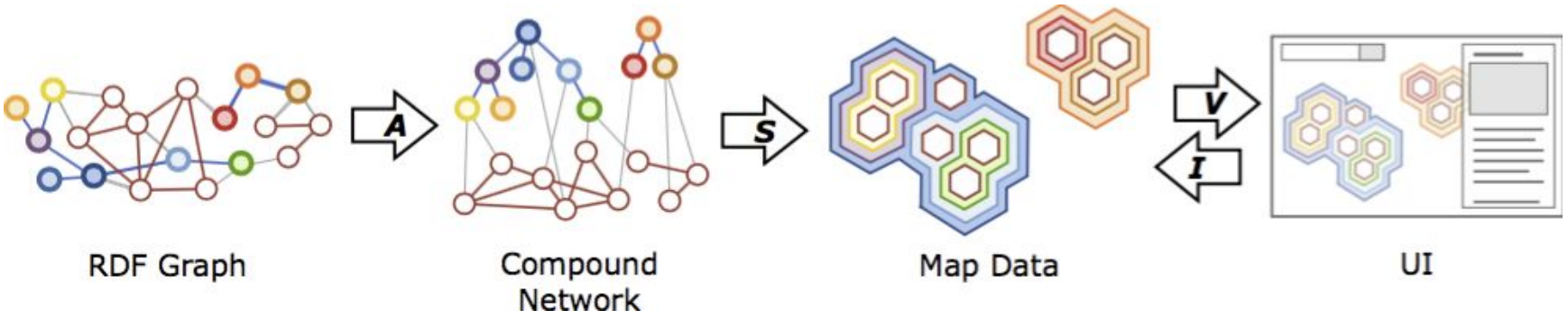
Visual Information-seeking Mantra

"Overview first, zoom and filter, then details on demand."

Lead a user from an overview of the main features of a dataset to its tiniest details.

- Provide an **overview** that acts as an ***entry point*** of the dataset
- Allow to **zoom and filter** for focusing on specific parts of the dataset
- Give **details** on single instances

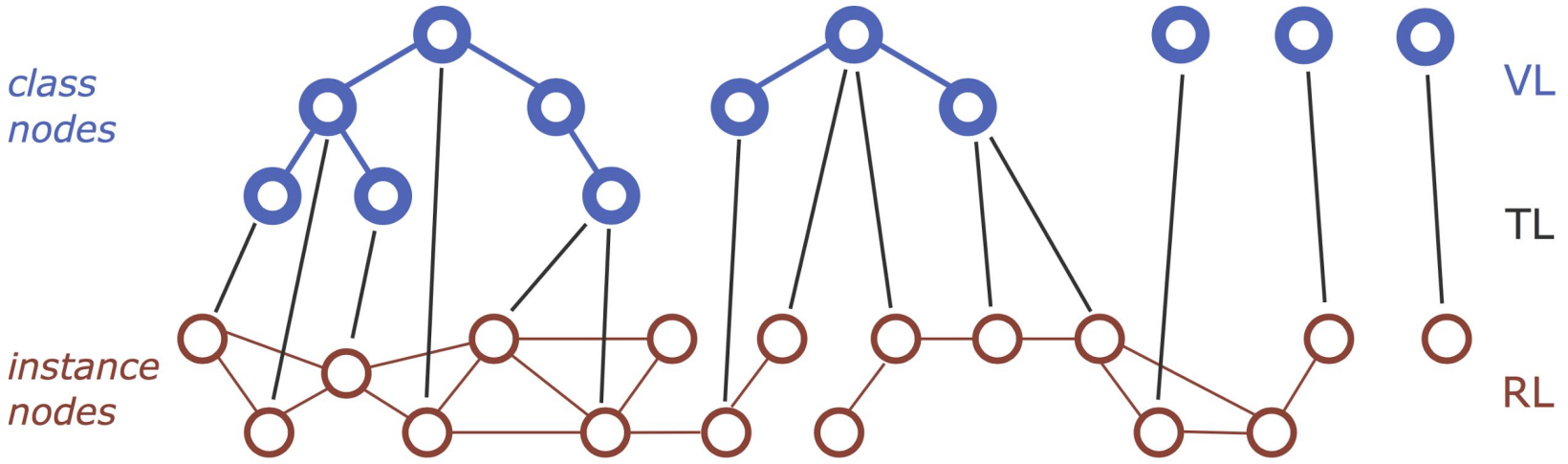
Our approach



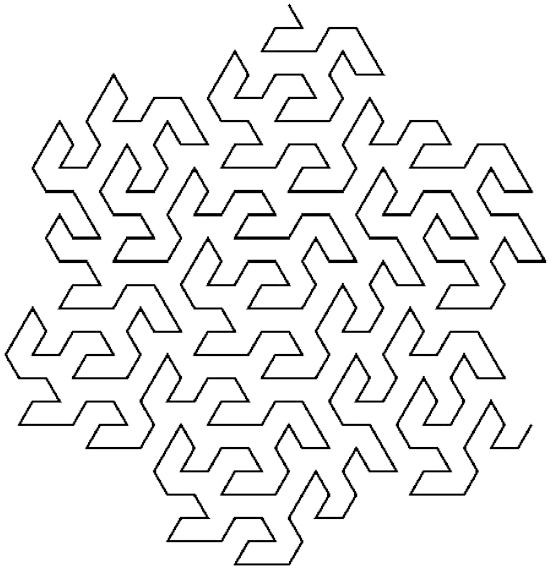
1. Data Abstraction
2. Spatialization Step
3. Visualization
4. Interaction

Data Abstraction

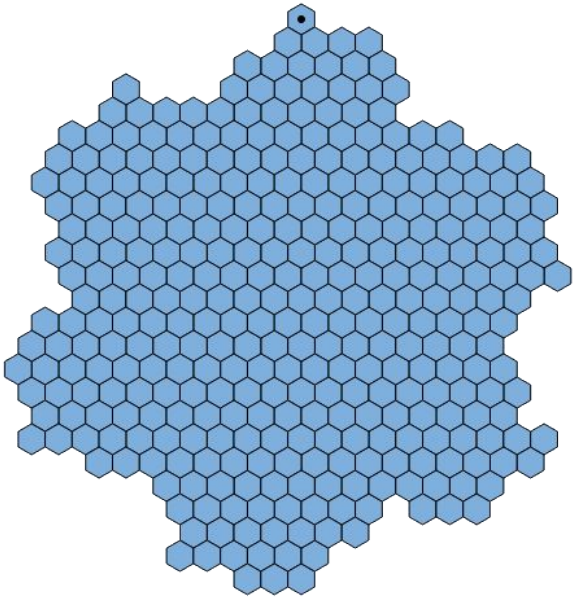
Compound Network: a structure defined by a graph with an associated forest.



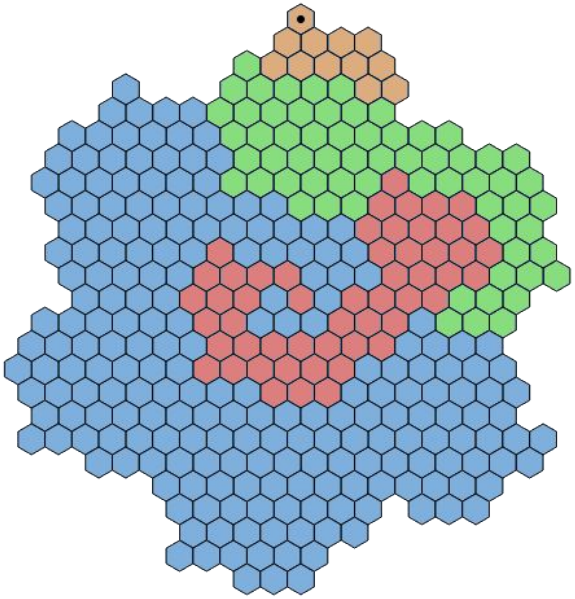
Spatialization step



Gosper **space-filling** curve*



Hexagonal **tiles**



Treemap

*[GosperMap: Using a Gosper Curve for Laying Out Hierarchical Data - Auber, D.]

Use cases

DBpedia*

- **3 billion** RDF triples
- Almost **5 million** instances
- A **hierarchical** ontology composed by **685** classes

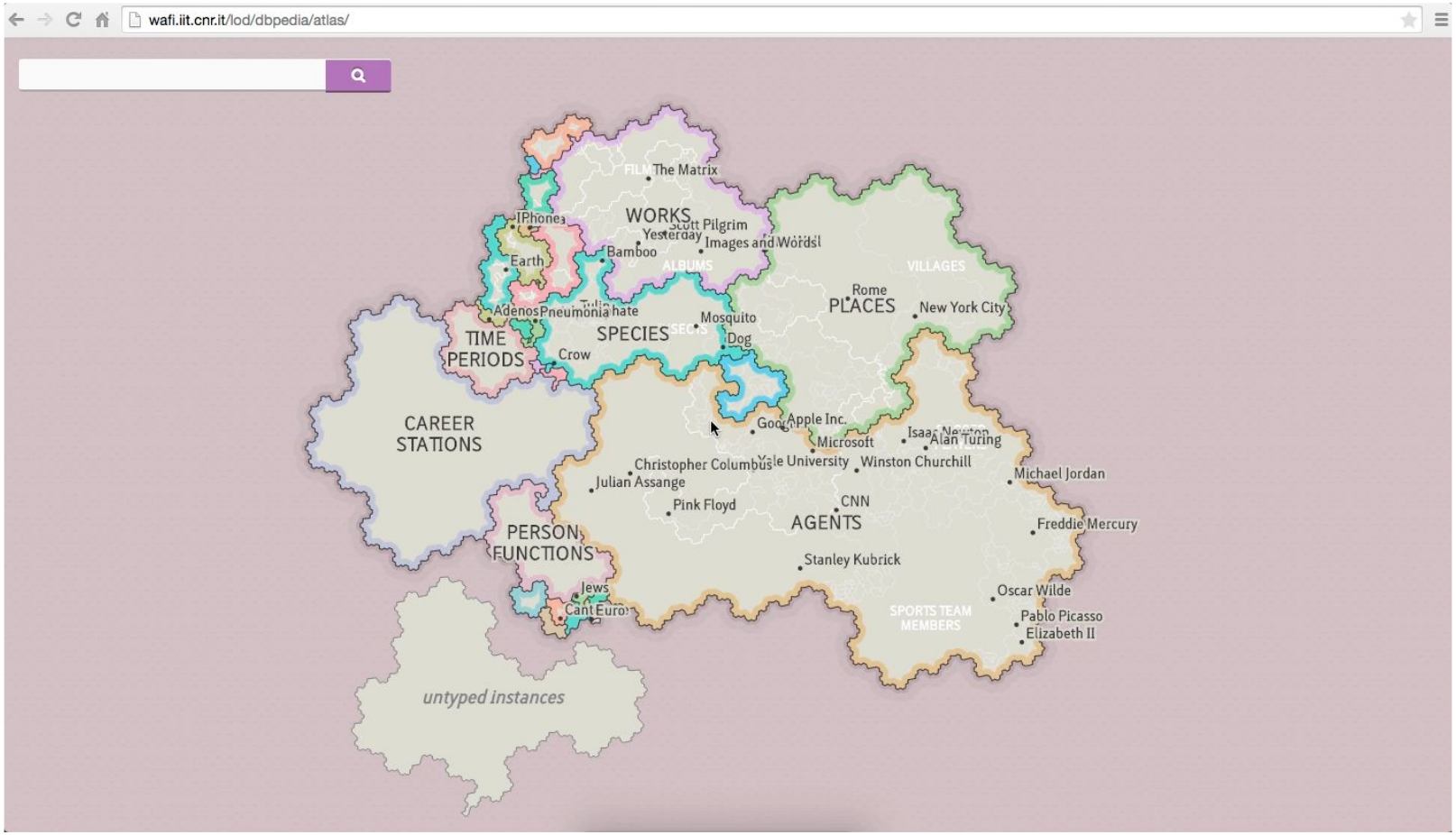
LinkedMDB**

- More than **7 million** RDF triples
- Almost 700 **thousands** instances
- A **flat** ontology comprising **51** classes

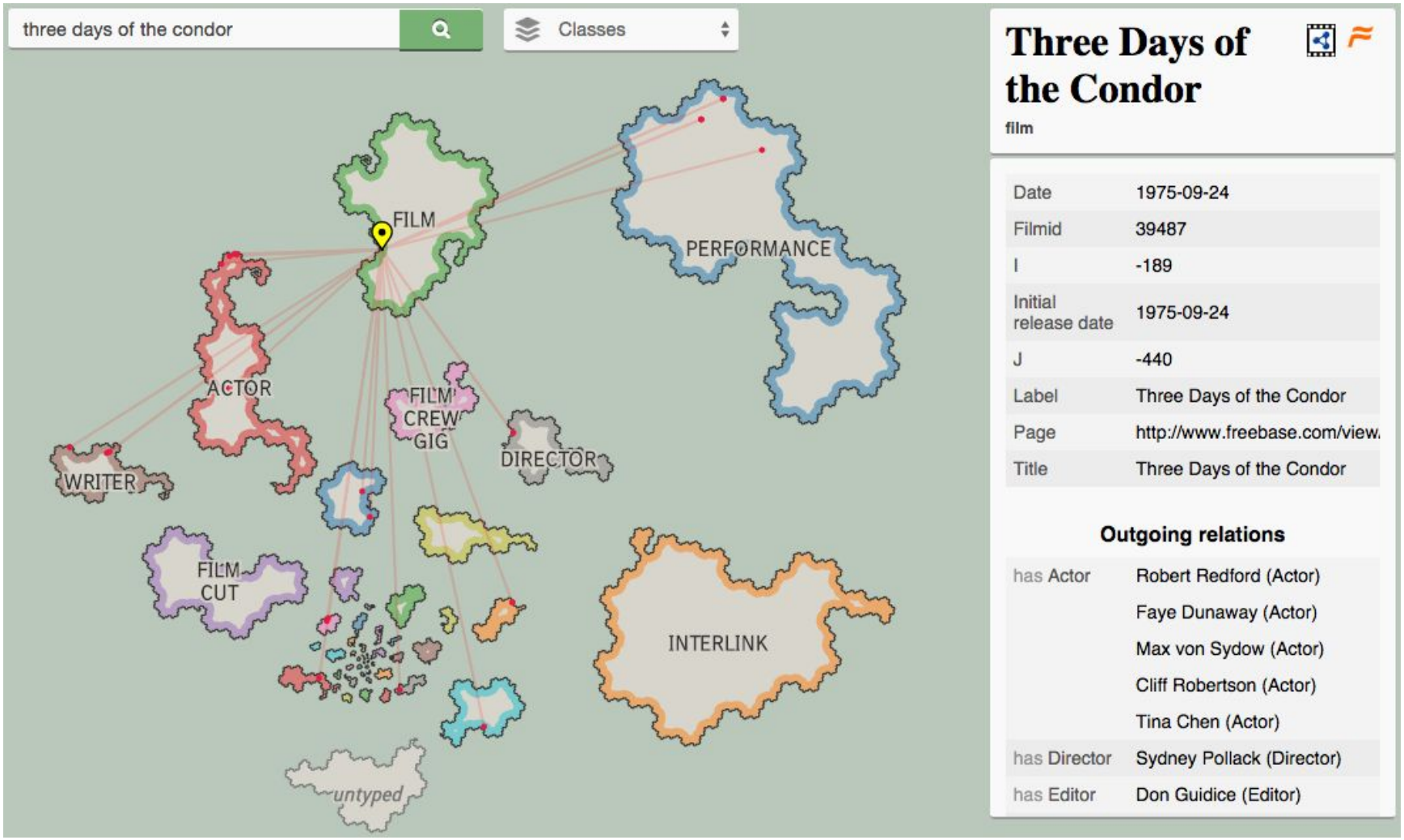
*[DBpedia - A crystallization point for the Web of Data - C. Bizer]

**[Linked Movie Data Base - O. Hassanzadeh]

DBpedia (demo)



LinkedMDB


three days of the condor
Classes
FILM
PERFORMANCE
ACTOR
FILM CREW GIG
DIRECTOR
WRITER
FILM CUT
INTERLINK
untyped
Three Days of the Condor
film
Date 1975-09-24
Filmid 39487
I -189
Initial release date 1975-09-24
J -440
Label Three Days of the Condor
Page http://www.freebase.com/view.
Title Three Days of the Condor
Outgoing relations
has Actor Robert Redford (Actor)
Faye Dunaway (Actor)
Max von Sydow (Actor)
Cliff Robertson (Actor)
Tina Chen (Actor)
has Director Sydney Pollack (Director)
has Editor Don Guidice (Editor)

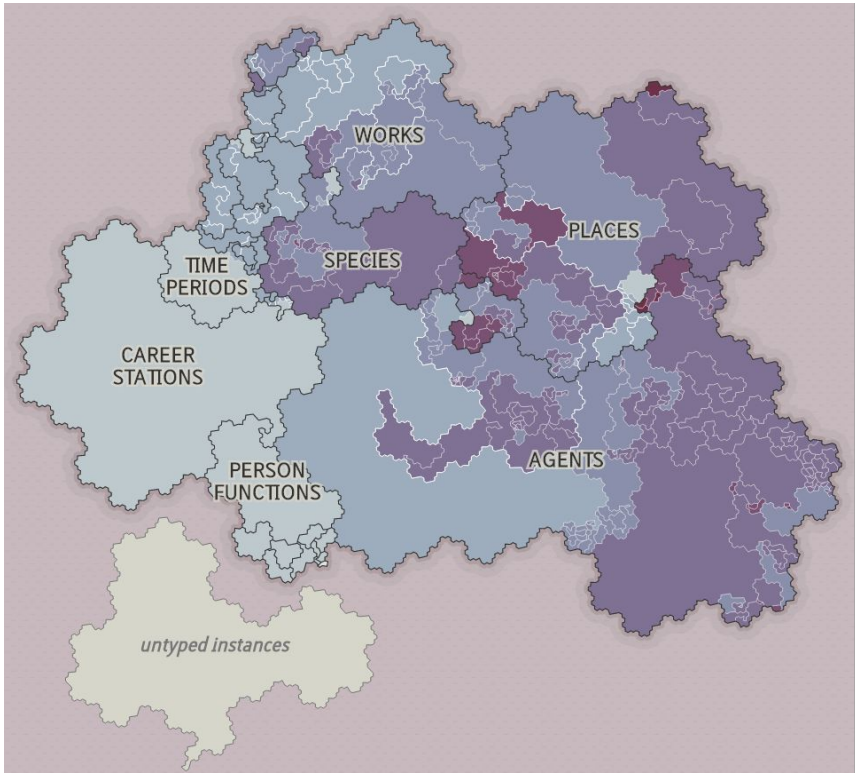
Why a map?

A map can leverage:

- innate **visual perception** abilities
- learned **map-reading** skills

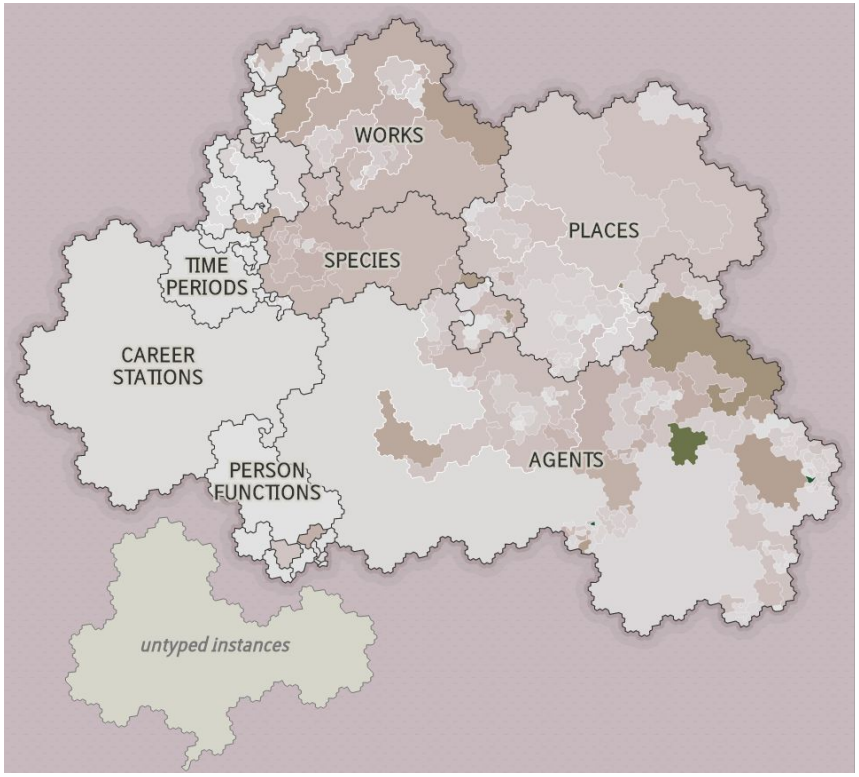
to attain a **high level of efficiency** in communicating features of large scale, complex structures.

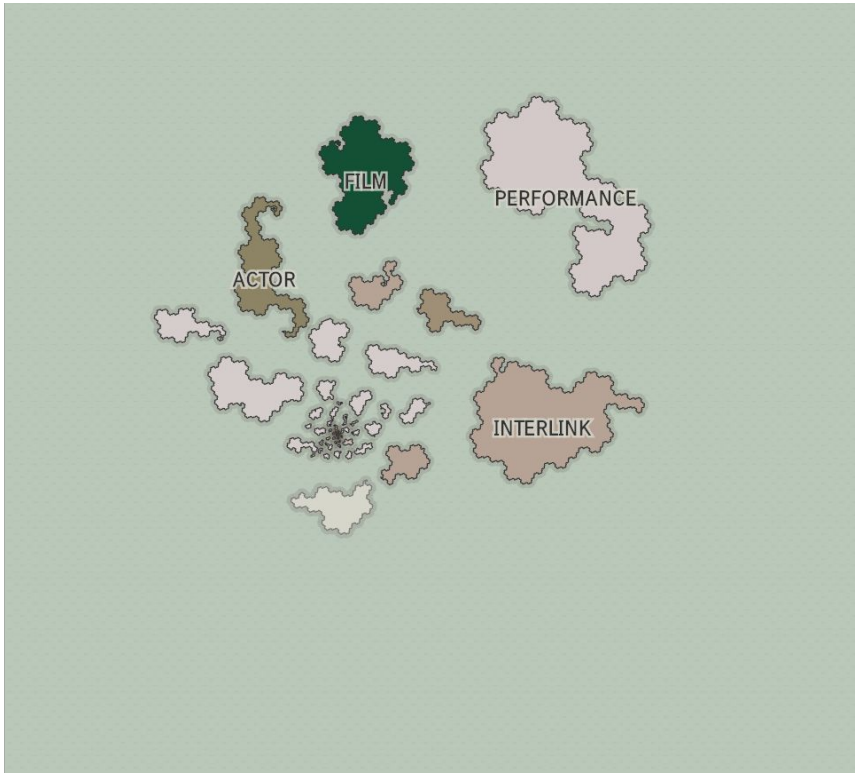
Thematic maps



The depth of the classes in their ontology (the darker, the deeper).

Thematic maps





The density of object properties of each class (the darker, the more dense).

Future Works

- **Similarity:** displace similar instances close together (inside the same region)
- **“Cities”:** implement an automatic system for ranking the importance of instances
- **Level of detail:** as the user zooms in, more content should be shown
- **Additional functionalities:**
  - Advanced search (SPARQL)
  - Path finding features (à la RelFinder)
  - …

Thank you!

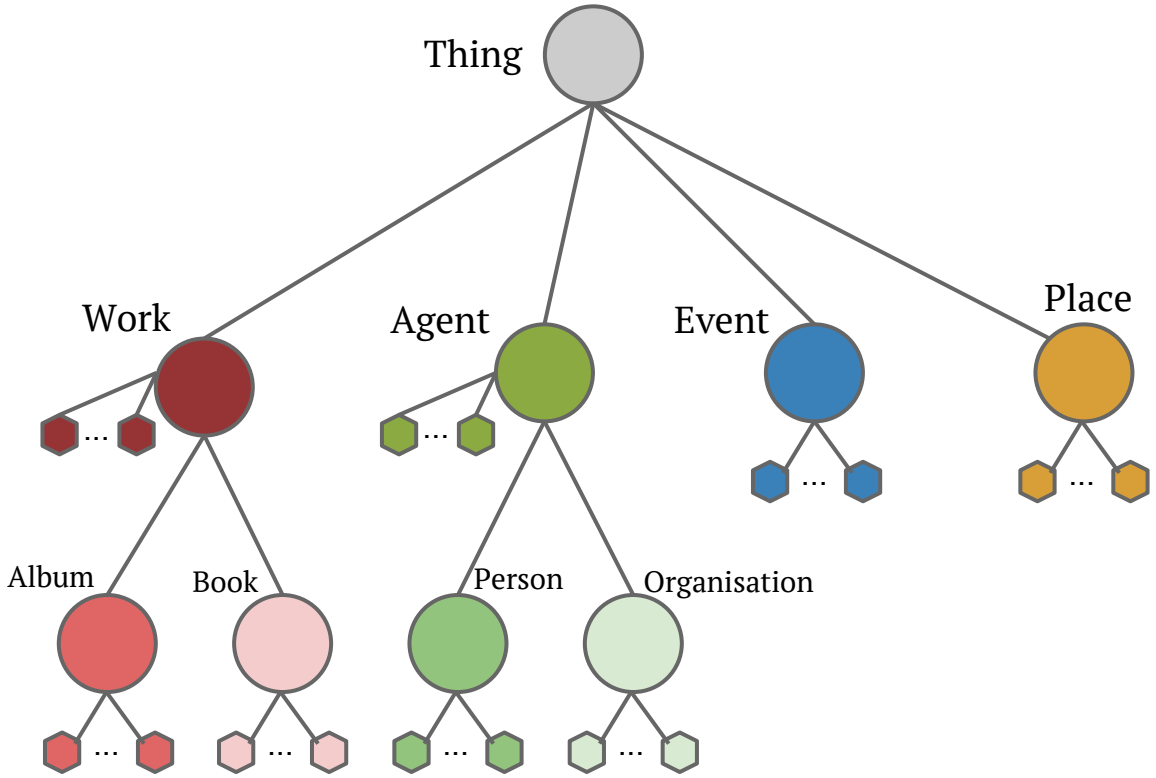
Take a look at our applications:

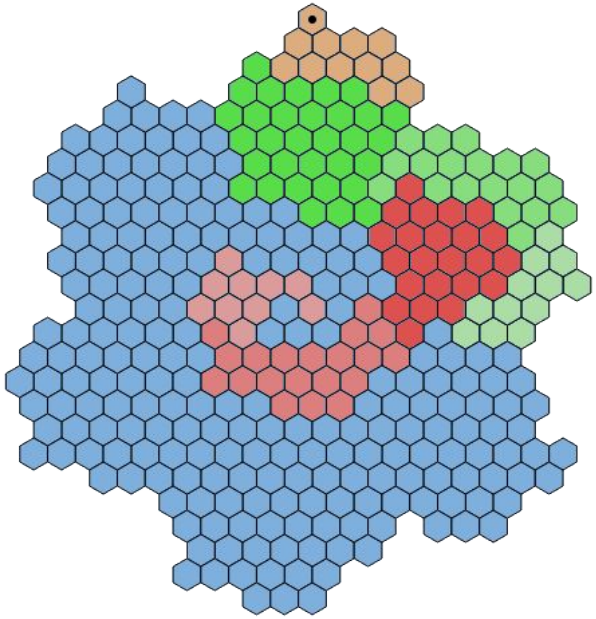
http://wafi.iit.cnr.it/lod/dbpedia/atlas

http://wafi.iit.cnr.it/lod/linkedmdb/atlas

fabio.valsecchi@iit.cnr.it

Spatialization approach





Visualization Pipeline