

# Tutorial on Semantic Schema Discovery: principles, methods and future research directions Part 5

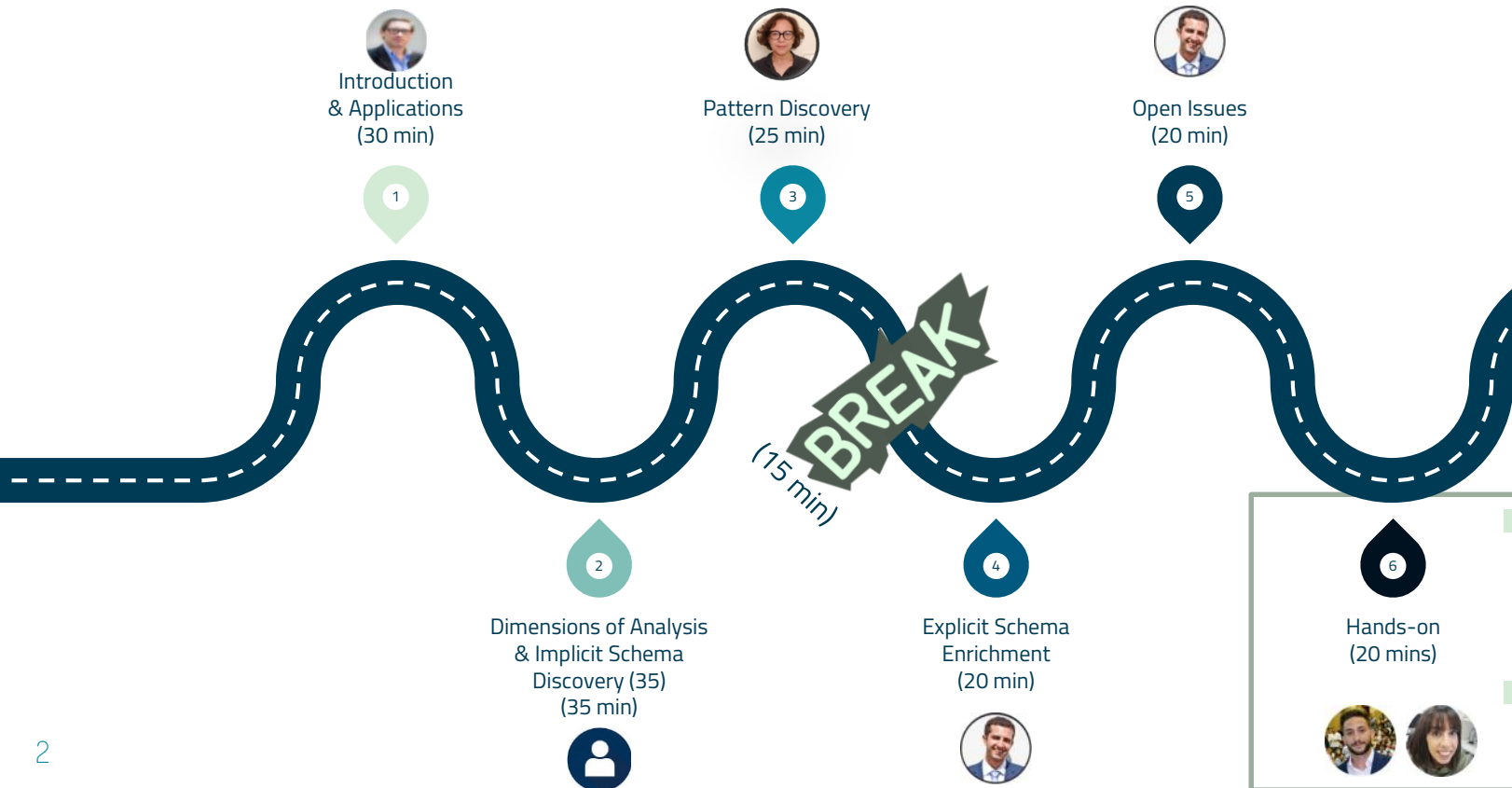
Kenza Kellou-Menouer, Nikolaos Kardoulakis, Georgia  
Troullinou, Zoubida Kedad, Dimitris Plexousakis,  
Haridimos Kondylakis



Équipes Traitement  
de l'Information  
et Systèmes

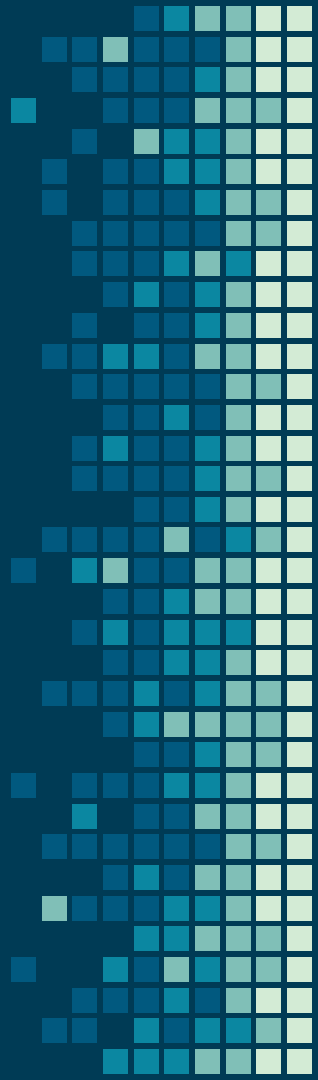


# ROADMAP

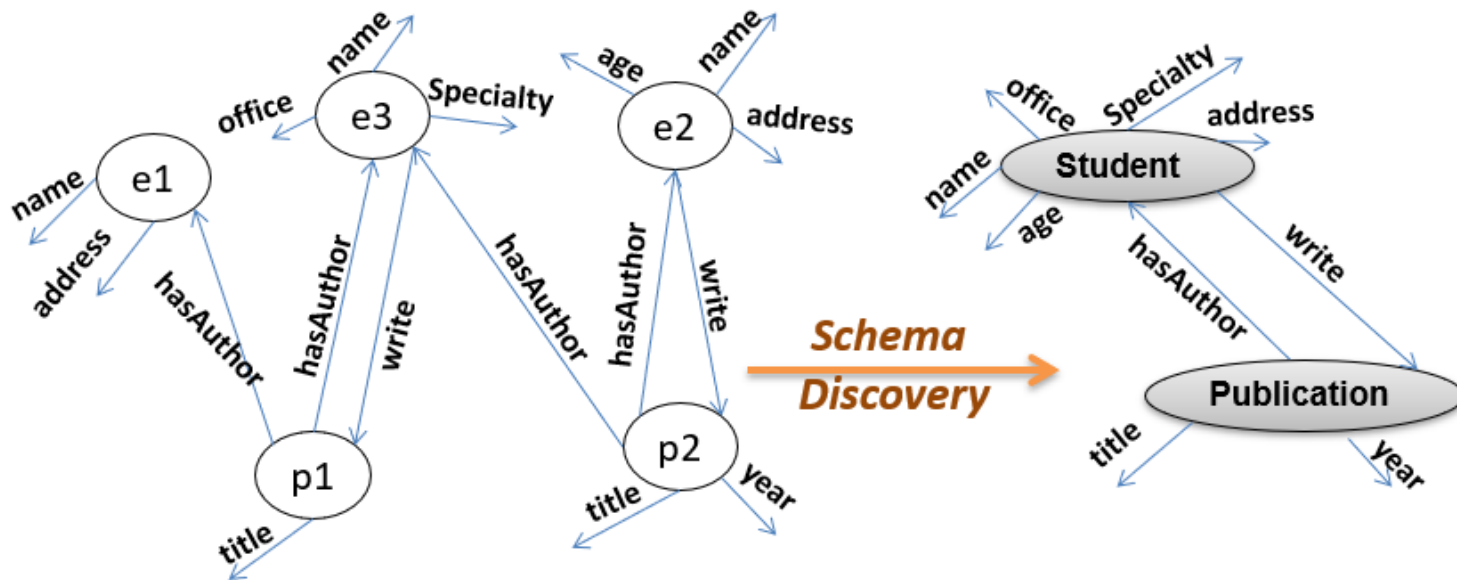




# Hands-on



# Example of Schema Discovery

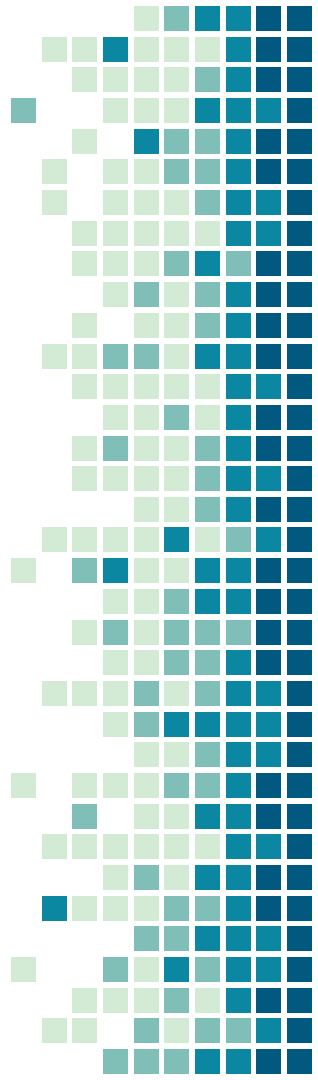
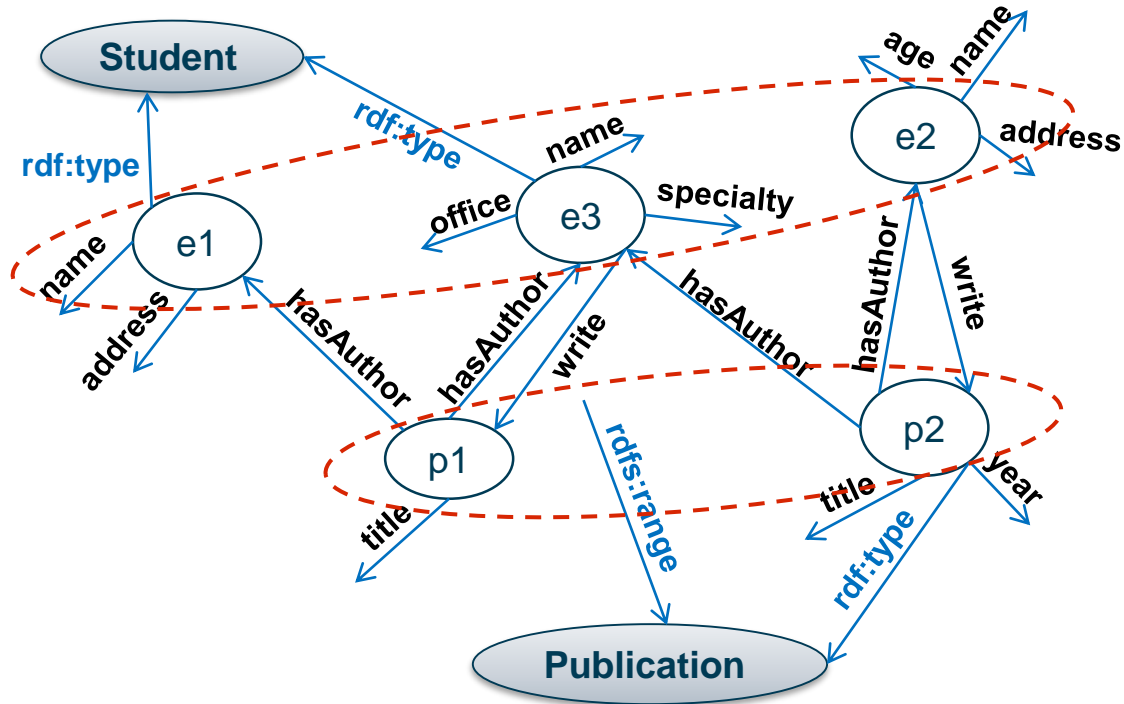


# Schema Discovery Approaches

- Implicit Schema Discovery
- Explicit Schema Enrichment
- Structural Patterns Discovery of the Instances

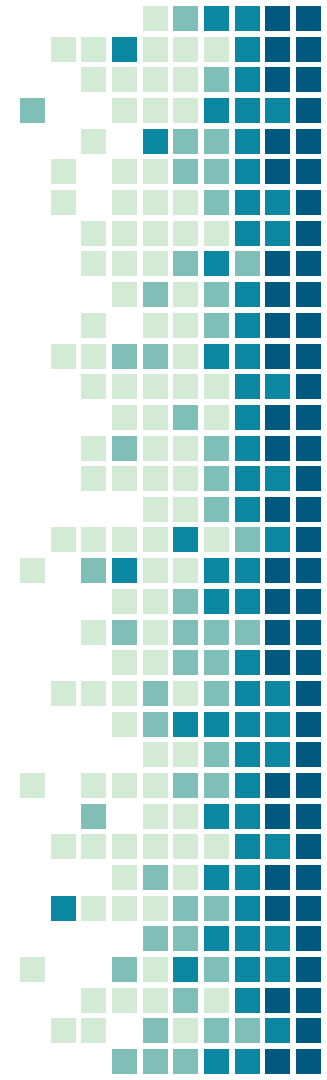


# Implicit Schema Discovery

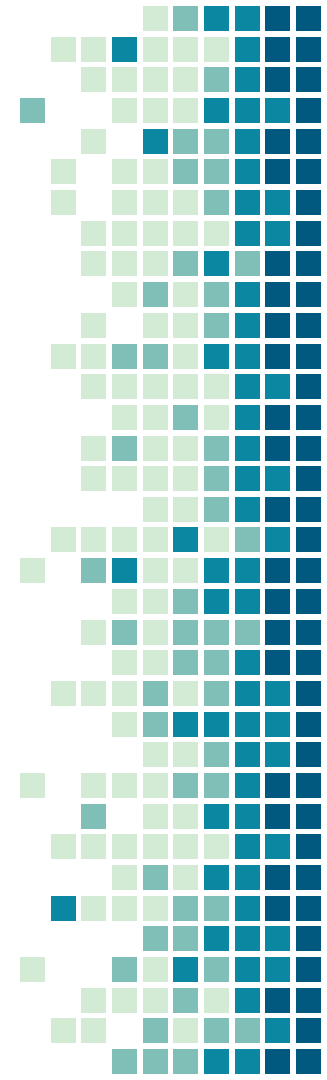
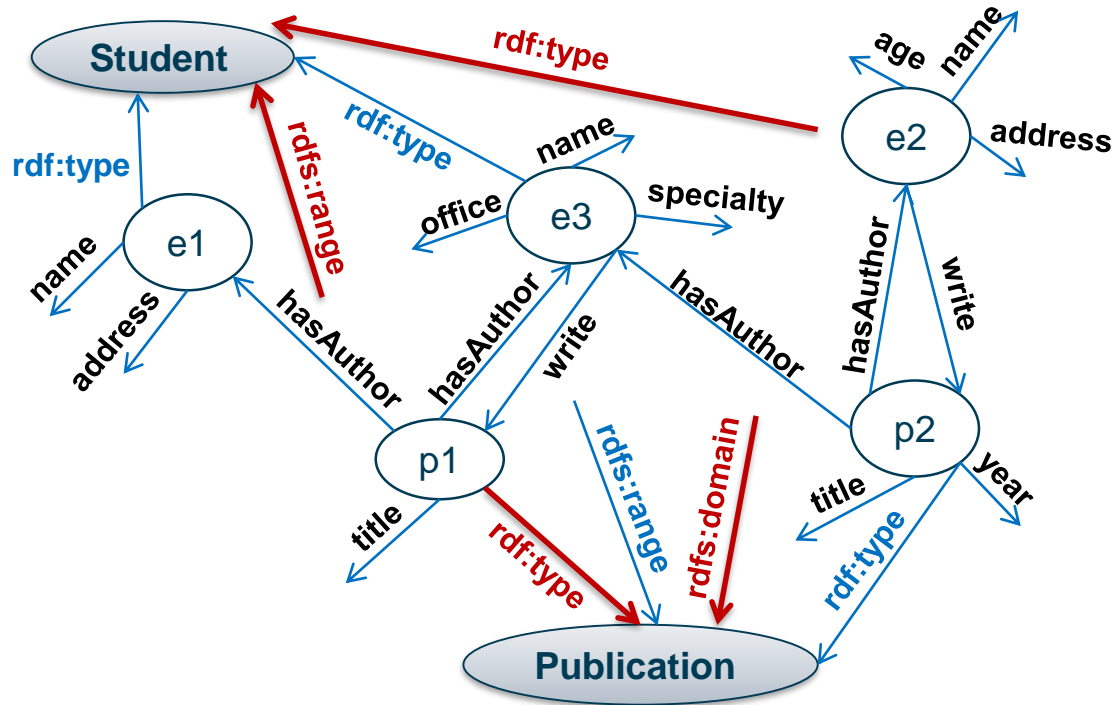


# Schema Discovery Approaches

- Implicit Schema Discovery
- **Explicit Schema Enrichment**
- Structural Patterns Discovery of the Instances



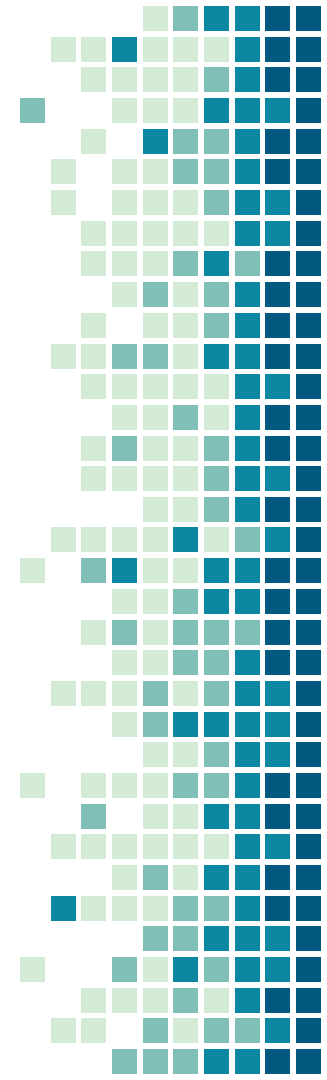
# Explicit Schema Discovery





# Schema Discovery Approaches

- Implicit Schema Discovery
  - Explicit Schema Enrichment
  - Structural Patterns Discovery of the Instances
- } **Hybrid Approach**

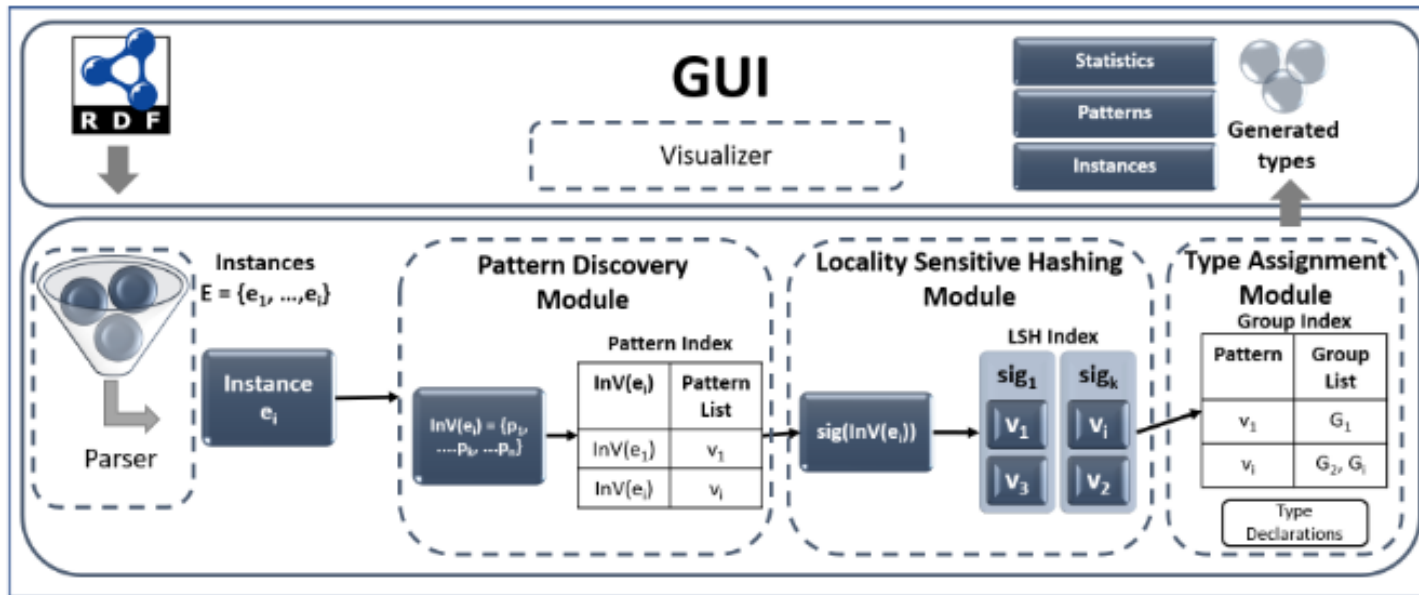


# HInT: Hybrid and Incremental Type Discovery



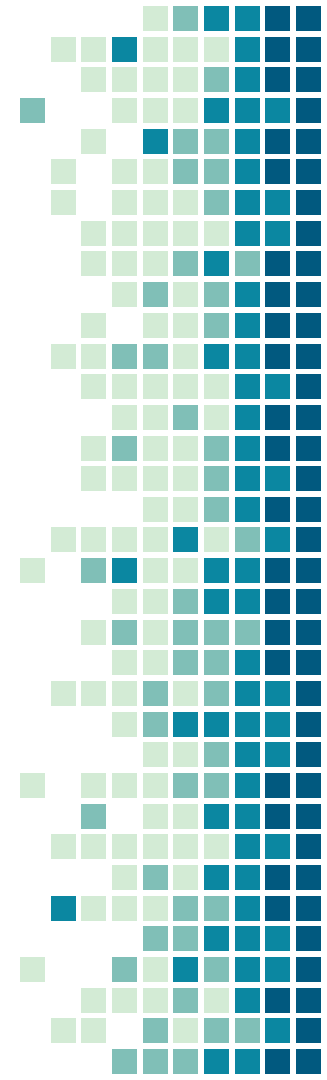
- Hybrid type discovery
  - Implicit: does not require existence of typing information
  - Explicit: exploits typing information if available
- Incremental
- Scalable

# HInT: Hybrid and Incremental Type Discovery



# Schema Discovery Approaches

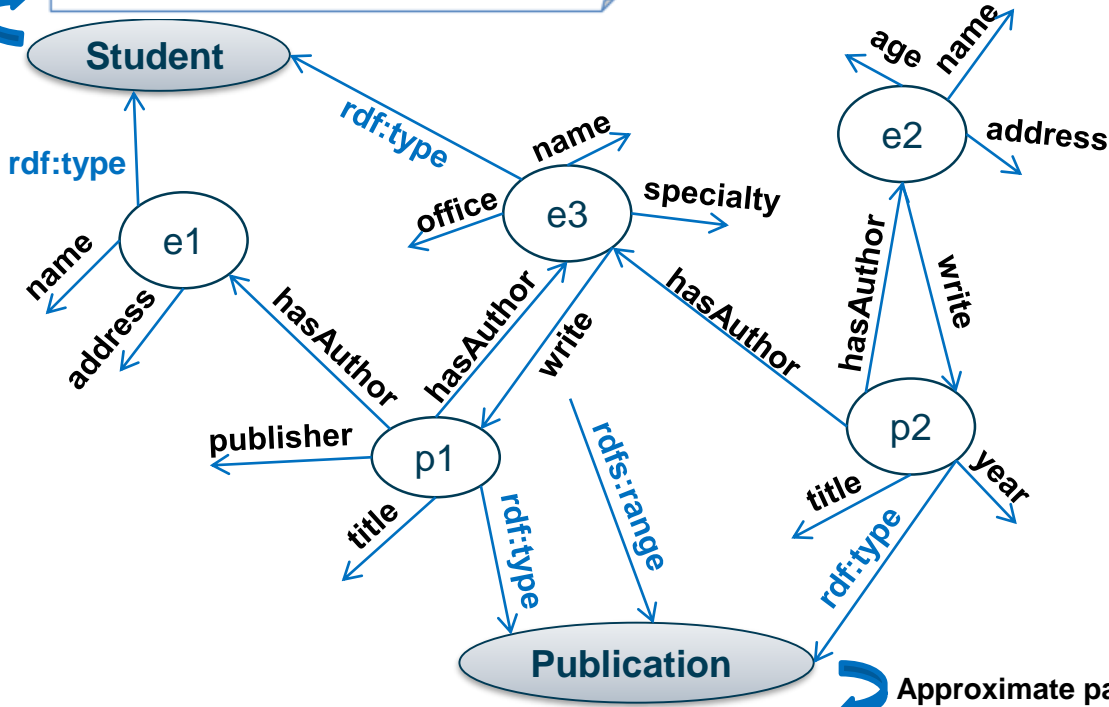
- Implicit Schema Discovery
  - Explicit Schema Enrichment
- } Hybrid Approach
- Structural Patterns Discovery of the Instances: Set of Properties
    - Exact Patterns
    - Approximate Patterns



# Structural Pattern Discovery

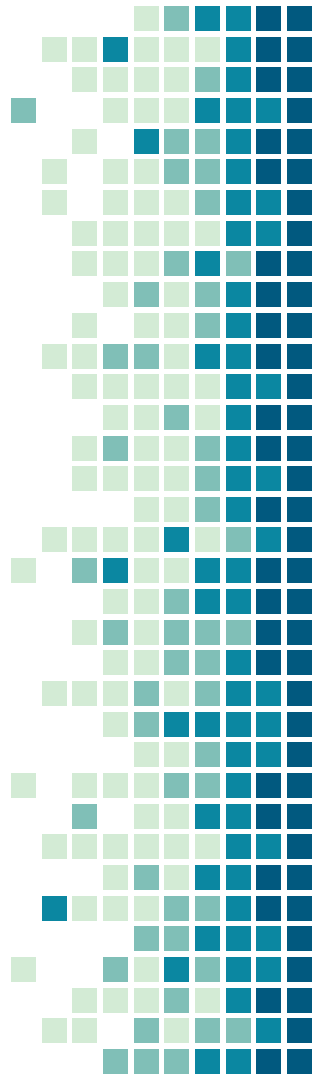
V1 = { $\vec{\text{name}}$ ,  $\vec{\text{address}}$ ,  $\overleftarrow{\text{hasAuthor}}$ }  
V2 = { $\vec{\text{name}}$ ,  $\vec{\text{office}}$ ,  $\vec{\text{specialty}}$ ,  $\overleftarrow{\text{hasAuthor}}$ ,  $\vec{\text{write}}$ }

Exact patterns



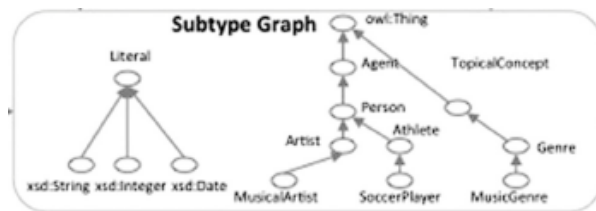
Approximate patterns

V' = { $\vec{\text{title}}$ ,  $\vec{\text{write}}$ ,  $\overleftarrow{\text{hasAuthor}}$ ,  $\vec{\text{year(?)}}$ ,  $\vec{\text{publisher(?)}}$ }

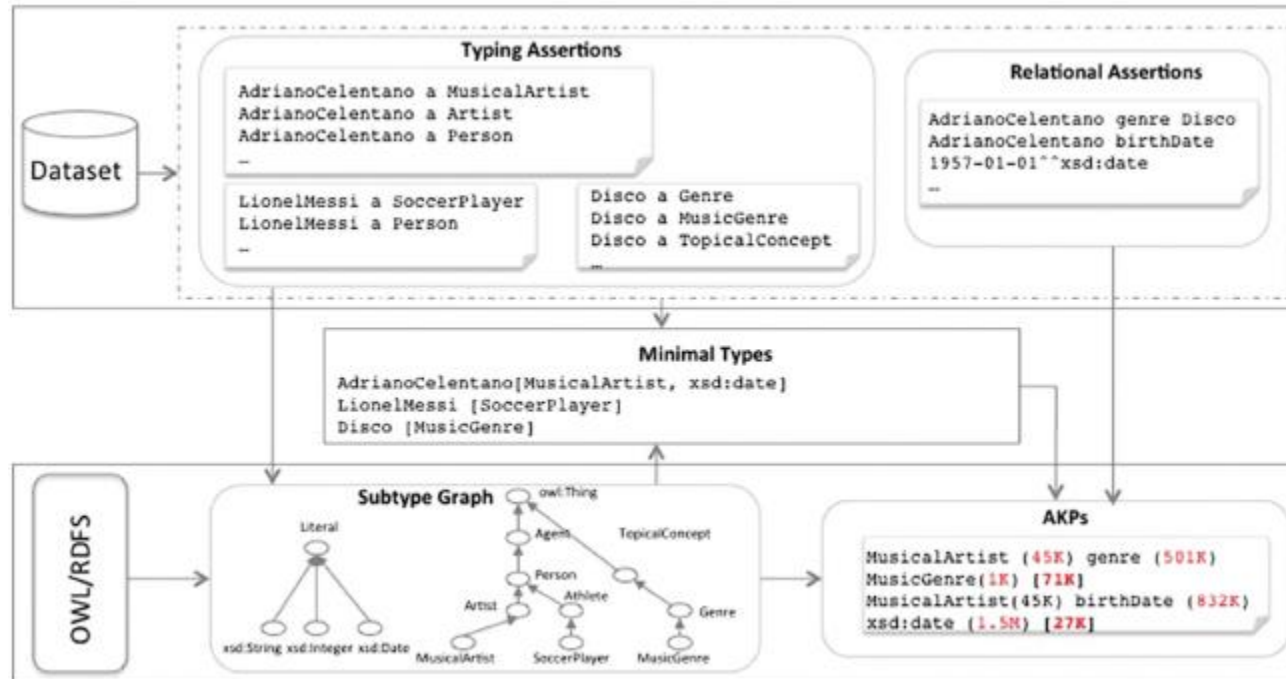


# ABSTAT: Ontology-Driven Linked Data Summaries with Pattern Minimalization

- Ontology-driven Linked data Summarization Model
- Represents a Pattern as a triple (C, P,D) such that C and D are types and P is a property
- Produces a base of *minimal type patterns*
  - i.e., a subset of the patterns such that every other pattern can be derived using a subtype graph
- Produces statistics



# ABSTAT: Ontology-Driven Linked Data Summaries with Pattern Minimalization



# THANKS!

Any questions?

You can find us at

[https://users.ics.forth.gr/~kondylak/  
iswc\\_2022\\_tutorial/](https://users.ics.forth.gr/~kondylak/iswc_2022_tutorial/)

