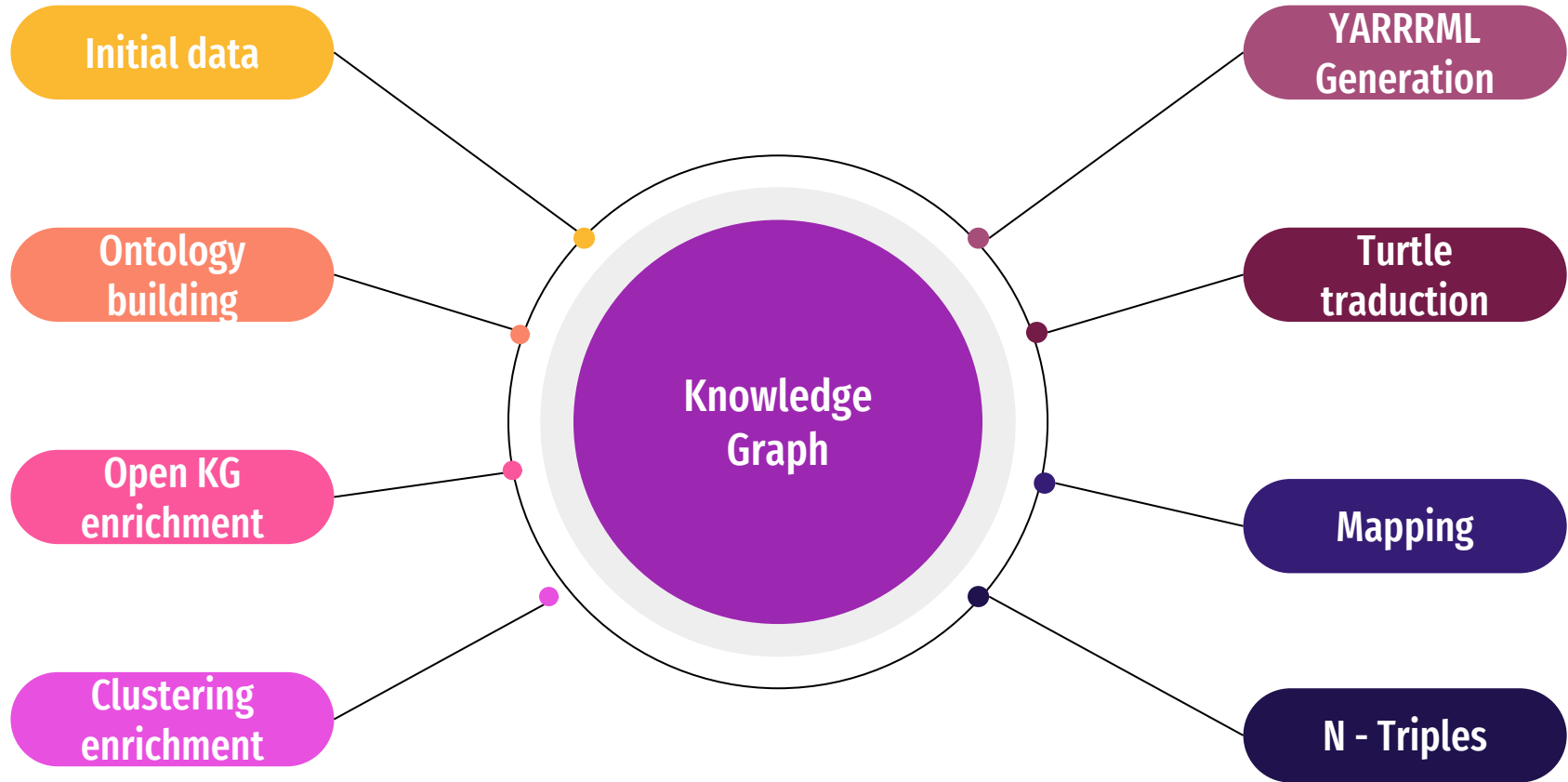


# Research Knowledge Graph

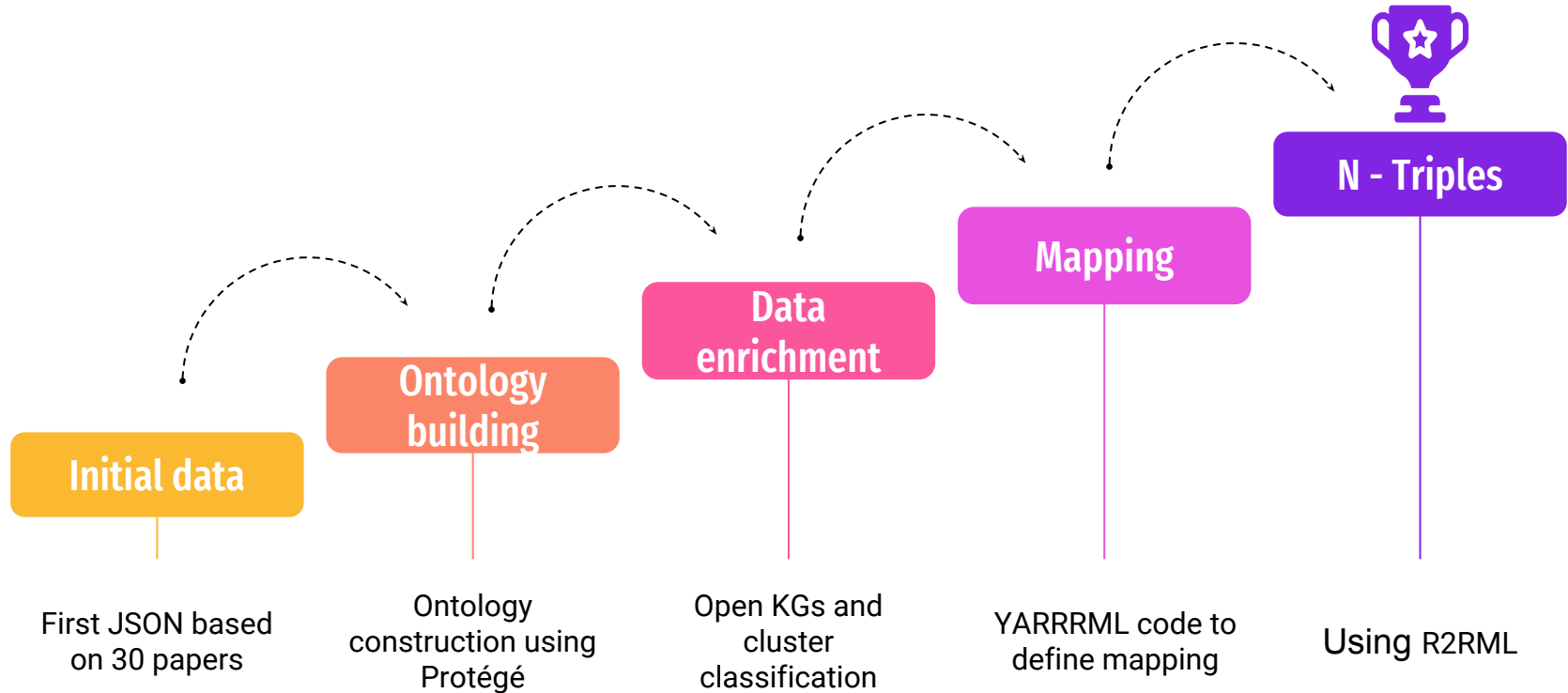
Artificial Intelligence And Open Science  
In Research Software Engineering

Miguel Casado Pina  
Gonzalo Solórzano Calvo  
Álvaro Pérez Cáceres  
Eduardo Marquina García

# Introduction



# Steps Followed



# Data Source

## Raw papers

30 papers from PubMed repository.

## KeyWords extraction

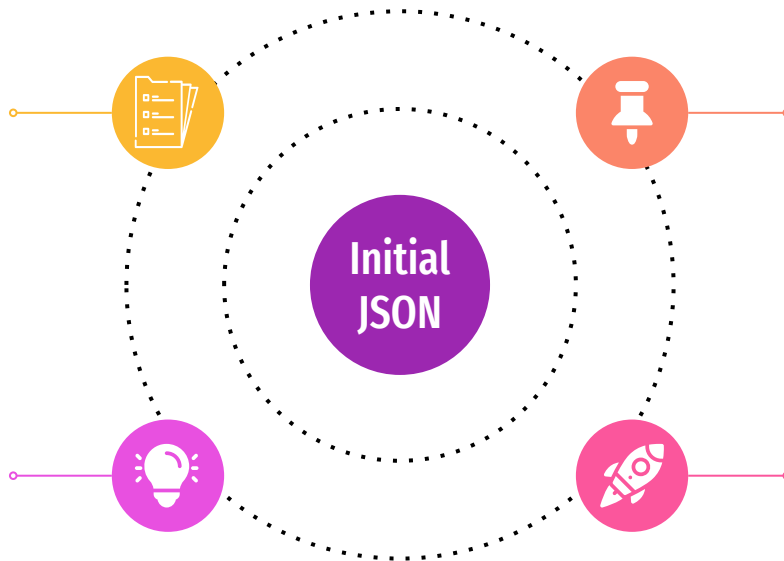
One Keywords list per paper

## Data extraction

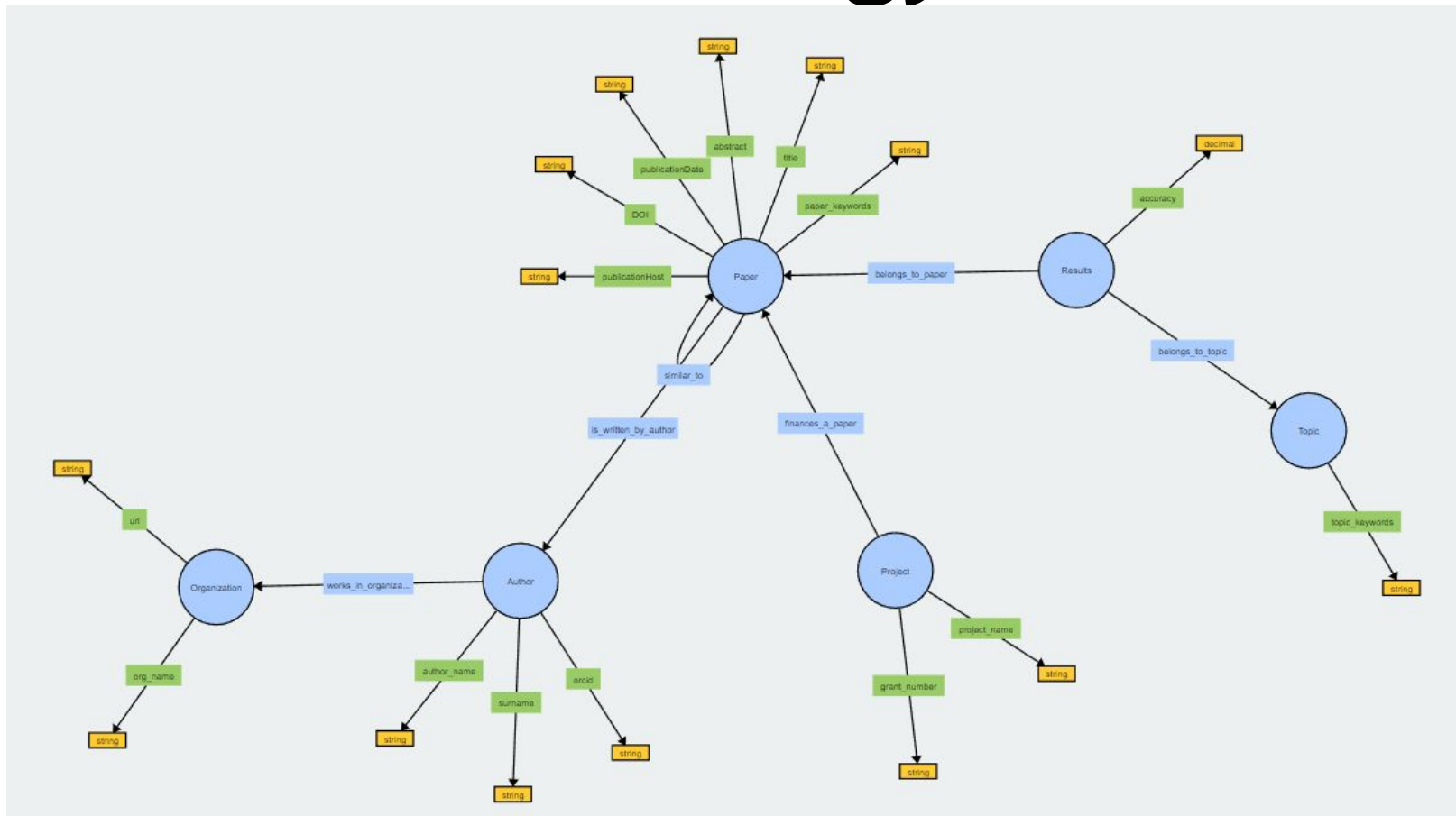
Using Grobid

## Initial Labels

- Id Paper
- Title
- Abstract
- Author
- Keyword
- DOI



# Ontology

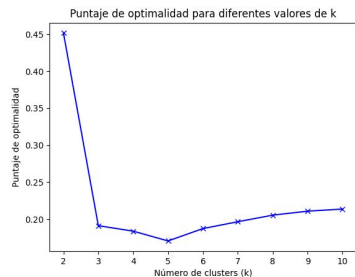


# Data Enrichment (KG)

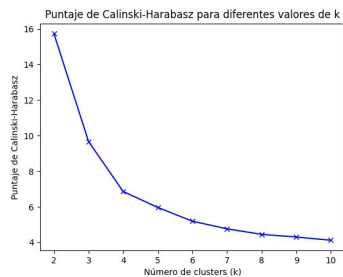


# Data Enrichment (Clustering)

## K-Means



## Calinski-Harabasz



## LatentDirichletAllocation

2 clusters

2 topics

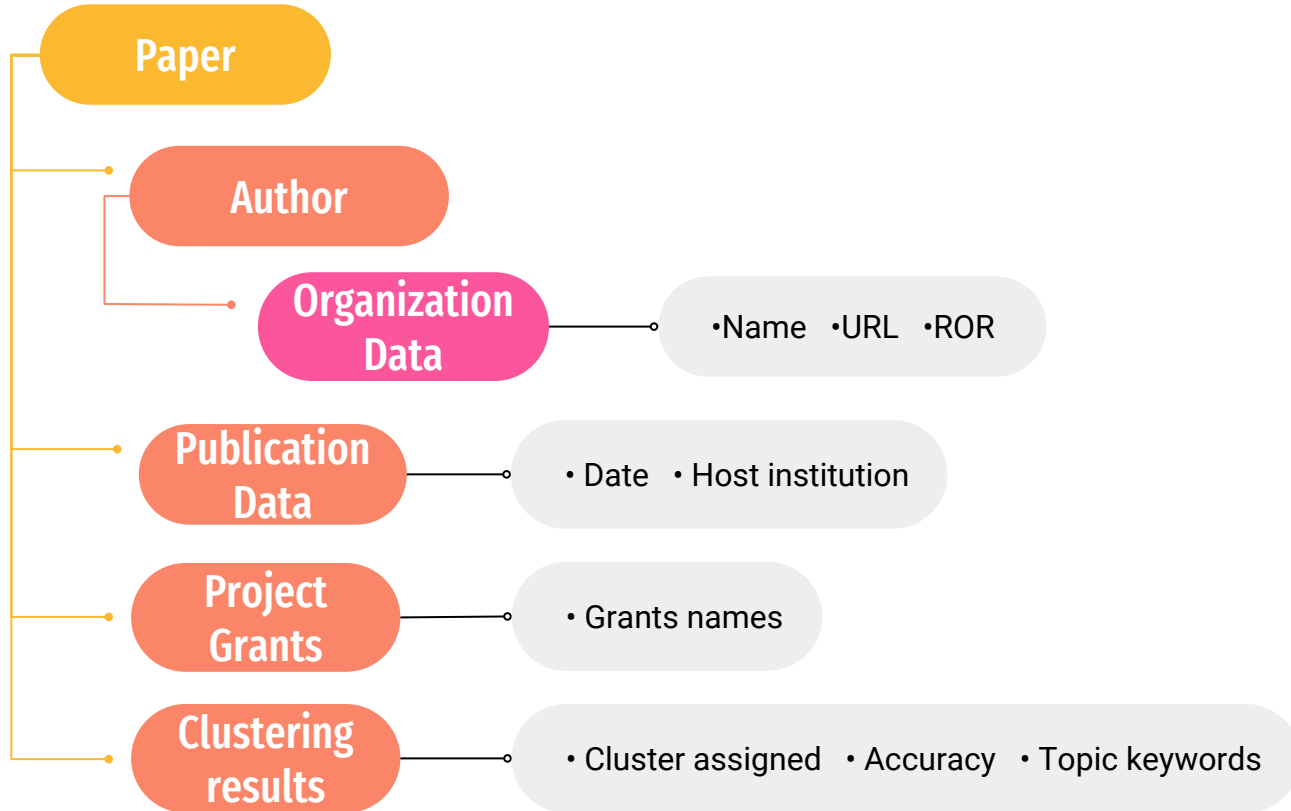
Precision of belonging to a  
cluster (accuracy)

## Metrics

Silhouette coefficient:  
0.2562

Davies-Bouldin index:  
1.5295

# Data Enrichment





# Mapping & N-Triples



**YARRRML**

Definition of the mapping structure



**R2RML**

Mapping tool



**N-TRIPLES**

List of entities



**APACHE-JENA-FUSEKI**

SPARQL query testing tool

# Questions & Requests

