Politecnico di Bari – Istituto Poligrafico Zecca dello Stato

Navigating the Legal Landscape: Developing Italy's Official Legal Knowledge Graph for Enhanced Legislative and Public Services

Vito Walter Anelli, Eros Brienza, Marco Recupero, Francesco Greco, Andrea De Maria, Tommaso Di Noia and Eugenio Di Sciascio

POLIGRAFICO
E ZECCA
DELLO STATO
ITALIANO

Politecnico
di Bari

GOALS

· · · · · · · · · ·

Revolutionizing Policy Management
with Cutting-Edge AI Tools

01 Enable a Deep Learning-based classification function for "Normattiva", the Italian web portal of current legislation.

02 Create a Knowledge Graph of the current legislation, enabling Content Negotiation, SPARQL queries, and Graph exploration considering the impact of a law on previous legislation.

03 Create a search tool based on Transformer embeddings, to retrieve paragraphs with high "semantic" similarity.

PAST

Handmade law issuing process
Several human actors are involved

Handmade law classification
A shortlist of experts has to analyse a large corpus of legal documents

Human-driven search
There are no public tools for exploring the current legislation

Support for Akoma Ntoso XML schemas
IPZS has begun to represent laws using the Akoma Ntoso XML schemas

TODAY

Explorable Knowledge Graph
LOD paves the way to query the Knowledge Graph and explore the graph.

5-STAR Linked Open Data
The entire Italian legal corpus is now in the LOD standard. Italy must publish public document in the Open Data standard (EU directive).

Deep Learning-based Similarity and Classification tools
AI-powered tools to support the daily life of legislators, judges, lawyers, and citizens.

Domain-specific stakeholders

Who pushes for Normattiva AI-based tools?

1

Vincenzo

Lawyer

Including judges and the other legislation-related professions

2

Gianluca

Legislator

Including consultants and other professions that support the legislation process

3

Ylenia

Normattiva editor

Why do we need AI-powered tools?
Context Analysis

187.000

Laws enacted since 1861

Proclamation of the Unification of Italy

111.000

Active laws

Including national legislation only

647

Decrees, laws and regulations

Enacted in the last parliamentary term

Law-Classification AI-powered tool
Model and Results

82%
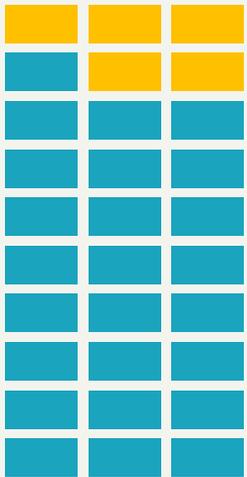
- Extreme *Multi-label Text Classification.*
- NLP pre-processing
- **FastText**: efficient model designed by Meta

12700
Dataset

Laws considered for building the models

400
Labels

Number of classes for the Classification System

2300
Different predictions

Number of classes that do not correspond to the ones selected by humans

Law-Classification AI-powered tool
Strategy and System's modules
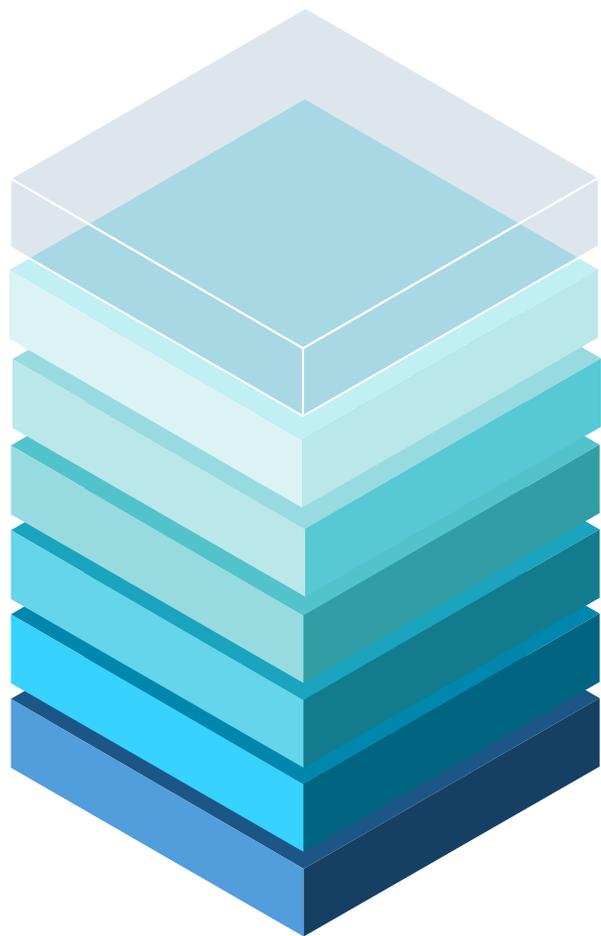
- **Batch law loader:** loading of the entire Italian legal corpus

- **Data processing:** NLP pre-processing (lowering – stemming – tokenization), filtering, domain-specific transformations

- **Data filtering:** stratification, splitting, domain-specific filtering

- **Classification model loader:** DL model loading and saving utilities
- **trainer:** model training
- **exposer:** model serving

- **Report generator:** performance analysis and logs

Italian Legislation Knowledge Graph
Ontologies and standards for representing Knowledge in the legal landscape

```xml
<?xml version="1.0" encoding="UTF-8"?><akomaNtoso xmlns="http://docs.oasis-open.org/legaldocml/ns/akn/3.0"
xmlns:eli="http://data.europa.eu/eli/ontology#"
xmlns:fo="urn:oasis:names:tc:opendocument:xmlns:xsl-fo-compatible:1.0"

. . .

</akomaNtoso>
```
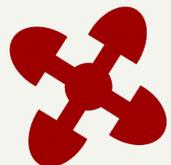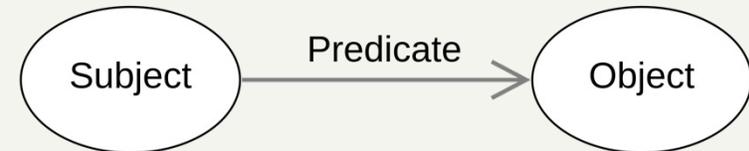

Akoma Ntoso
XML for parliamentary, legislative & judiciary documents

Italian Legislation Knowledge Graph

The system analyzes the normative corpus and construct an RDF graph then stored in the triplestore.
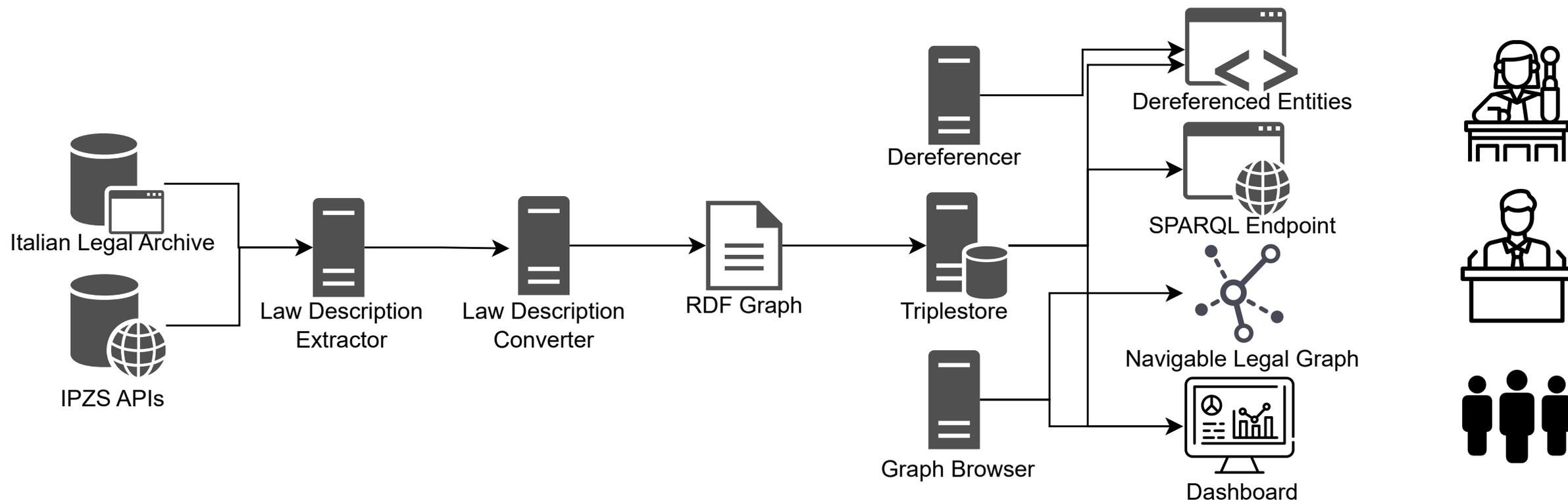
Italian Legislation Knowledge Graph
The dereferenced description of an active law

090G0294
http://localhost:8080/lodview/akn/it/act/legge/stato/1990-08-07/241

|	
label	090G0294
<http://localhost:8080/lodview/akn/it/**passive_reference**>	/akn/it/camera-it:parlamento:scheda.progetto.legge:camera;10.legislatura;1913 @it
	/akn/it/senato-it:parl:ddl:senato;10.legislatura;2226 @it
<http://localhost:8080/lodview/akn/it/**approval_date**>	1990-08-07
<http://localhost:8080/lodview/akn/it/**local_id**>	090G0294
<http://localhost:8080/lodview/akn/it/**preface**>	LEGGE 07 agosto 1990 241 Nuove norme in materia di procedimento amministrativo e di diritto di accesso ai documenti amministrativi. Entrata in vigore della legge: 2-9-1990 @it
<http://localhost:8080/lodview/akn/it/**publication_date**>	1990-08-18
<http://localhost:8080/lodview/akn/it/**title**>	Nuove norme in materia di procedimento amministrativo e di diritto di accesso ai documenti amministrativi. @it
<http://localhost:8080/lodview/akn/it/**based_on**>	<http://localhost:8080/lodview/akn/it/act/legge/stato/1990-08-07/241> ↳ 090G0294

Italian Legislation Knowledge Graph
## A queryable and explorable Knowledge Graph
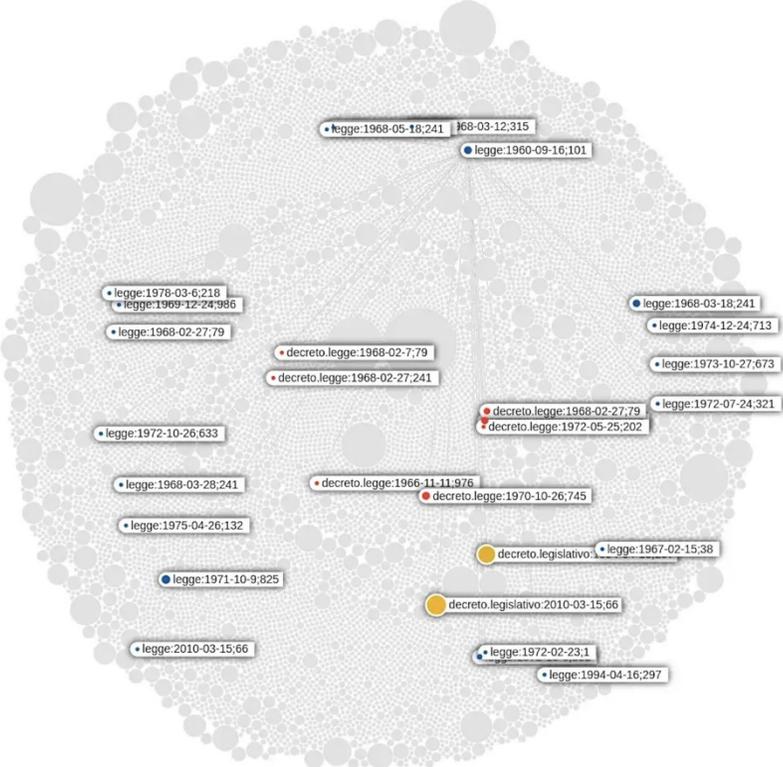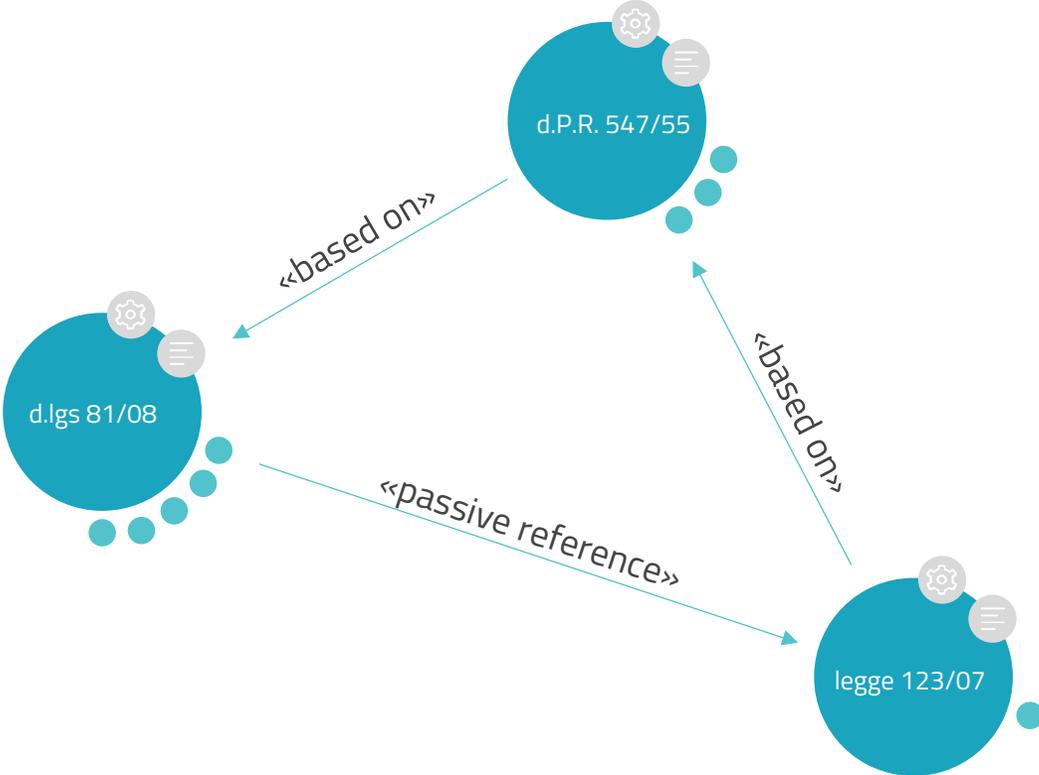
Italian Legislation Knowledge Graph

The system analyzes the normative corpus and construct an RDF graph then stored in the triplestore.
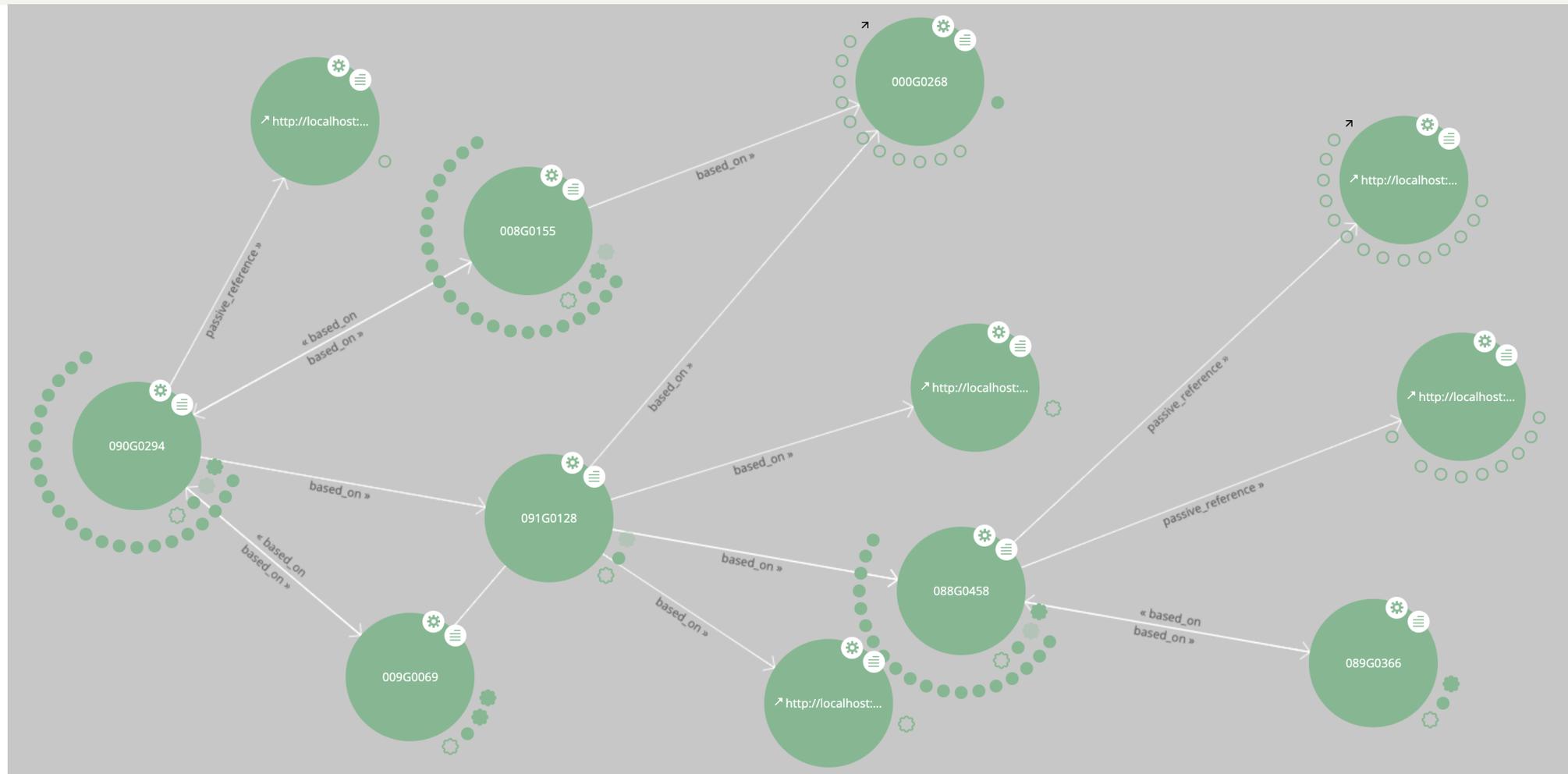
Italian Legislation Knowledge Graph
A pleasant side effect: 5-star Linked Open Data

5* STAR OPEN DATA
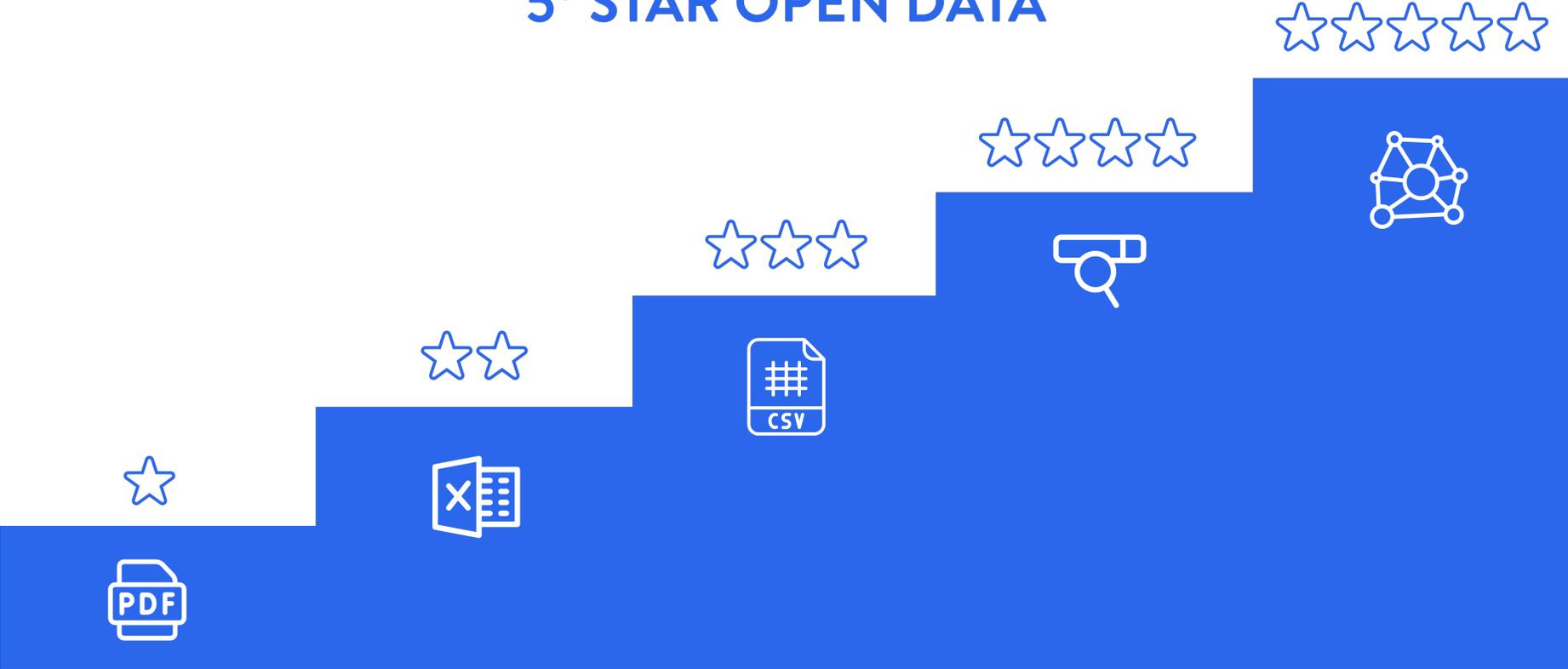
«Semantic» Similarity Search tool
Strategy and System's modules
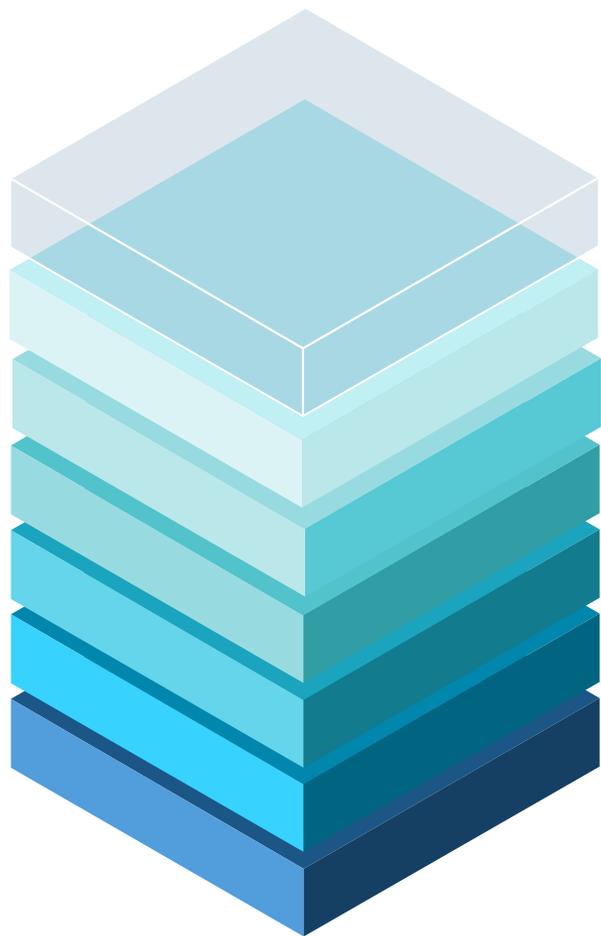
- **Batch law loader:** loading of the entire Italian legal corpus

- **Data processing:** NLP pre-processing (lowering – tokenization), filtering, domain-specific transformations

- **Data encoding:** Encoding of the corpus with a Transformer-like architecture (i.e. paraphrase-multilingual-mpnet-base-v2)

- **Similarity model loader:** DL model loading and saving utilities
- **trainer:** model encoding of the new law draft
- **exposer:** model serving

«Semantic» Similarity Search tool
An easy-to-use interface

Ricerca atti attraverso similarità semantica

Decadi considerate:

Seleziona tutto   Deseleziona tutto

☑ **2020**   ☑ **2010**   ☑ **2000**   ☑ **1990**

☑ **1980**   ☑ **1970**   ☑ **1960**   ☑ **1950**

☑ **1940**   ☑ **1930**   ☑ **1920**   ☑ **1910**

☑ **1900**   ☑ **1890**   ☑ **1880**   ☑ **1870**

☑ **1860**

Numero di documenti da restituire:

```
5
```

Testo da ricercare:

```
La circolazione navale militare è incrementata.
```

Invia

«Semantic» Similarity Search tool
Results

Frase	Atto	Similarità	Num	Testo
La circolazione navale militare è incrementata.	/akn/it/act/decreto_legislativo/stato/1997-04-30/165	0.8004570007324219	56	per i militari dell'esercito e dell'aeronautica il servizio reso a bordo di navi militari e quello reso sulla costa in tempo di guerra e' aumentato della meta'.
La circolazione navale militare è incrementata.	/akn/it/act/decreto_del_presidente_della_repubblica/stato/1973-12-29/1092	0.8004570007324219	69	per i militari dell'esercito e dell'aeronautica il servizio reso a bordo di navi militari e quello reso sulla costa in tempo di guerra e' aumentato della meta'.
La circolazione navale militare è incrementata.	/akn/it/act/decreto_legislativo/stato/2010-03-15/66	0.7834142446517944	6722	4. per i militari dell'esercito e dell'aeronautica il servizio reso a bordo di navi militari e quello reso sulla costa in tempo di guerra e' aumentato della meta'».
La circolazione navale militare è incrementata.	/akn/it/act/decreto_legislativo/stato/1997-04-30/165	0.7562923431396484	52	- il servizio prestato dai militari della marina a bordo di navi in armamento o in riserva e' aumentato di un terzo; lo stesso aumento si applica per il servizio prestato da detti militari sulla costa in tempo di guerra.
La circolazione navale militare è incrementata.	/akn/it/act/decreto_legislativo/stato/2010-03-15/66	0.751998782157898	6721	3. il servizio prestato a bordo delle navi in armamento o in riserva dai militari addetti alle macchine e' aumentato di due quinti.

Conclusion and Future Work

Next step: June 2023

JUNE 2023

Publicly available Prototypes

The three prototypes will be integrated with the Normattiva platform soon.

FUTURE DIRECTIONS

Improve Search and Conversational abilities

The project follow-up is considering new research paths that include the exploitation of the state-of-the-art Information Retrieval techniques, and the integration of the extracted knowledge in a Conversational Agent for Question-Answering.

Politecnico di Bari – Istituto Poligrafico Zecca dello Stato

Navigating the Legal Landscape: Developing Italy's Official Legal Knowledge Graph for Enhanced Legislative and Public Services

Vito Walter Anelli, Eros Brienza, Marco Recupero, Francesco Greco, Andrea De Maria, Tommaso Di Noia and Eugenio Di Sciascio