



HYPERMODELEX



HyperModeLex applications

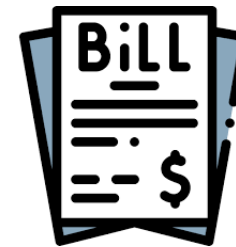
Michele Corazza – Researcher at
CIRSIFD-ALMA AI, Department of
Legal Studies, University of Bologna,
Italy

ECAI 2025 – University of Bologna
Day 1 - October 25, 2025

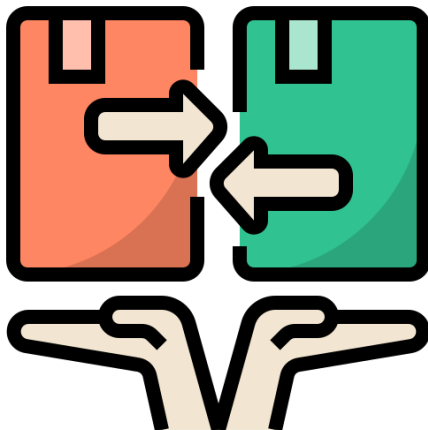


Index of unconstitutionality of bill

- In 2024, 44,5% of the judgments of the Italian Constitutional Court were about the constitutional illegitimacy of law. It is a problem of democracy, dialogue between institutions, division of powers.
- Can we prevent this situation?



- Bill of the Chamber of Deputies



- Italian legislation (in force/not in force)
- Constitutional Courts - all
- Abstract of the Constitutional Courts
- Ratio decidendi annotation made by the judges
- Classification case-law references (positive or negative)

The factors tested for the unconstitutionality index

The similarities are computed using our hierarchical embeddings and cosine similarity

1. spj = similarity of «bill» vs. «judgments»
2. scn = similarity of «bill commission description» vs. «ratio decidendi» annotated by the judges
3. spm = similarity of «bill» vs. «abstract of the case-law provided by the Constitutional Court
4. normative references of abrogated law are considered

The factors tested for the unconstitutionality index

$$R(b) = \frac{\sum_{h=1}^{10} s_h(b) I_h(b)}{\sum_{i=1}^{10} s_i(b)}$$

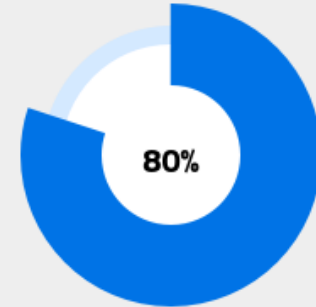
- The risk of unconstitutionality $R(b)$ is computed as the product between the similarities of all the bills with the 10 most similar judgments and an indicator function that indicates that the h -th judgment contains an unconstitutionality judgment

Under evaluation in the Chamber of Deputies of Italy

Ricerca:

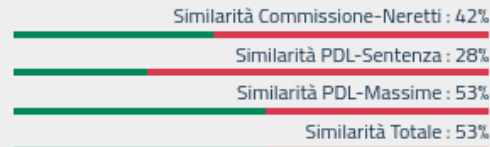
/akn/it/bill/
propostadilegge/2024-
01-16/1651/!main

Valore di predizione:



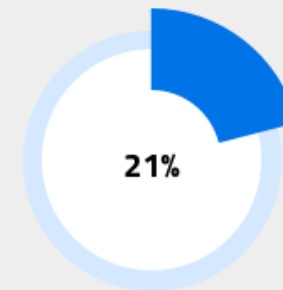
Sentenza N. 35 Anno 2017

FRBRExpressionuri: /akn/it/judgment/sentenza/
cortecostituzionale/2017-02-09/35/!main



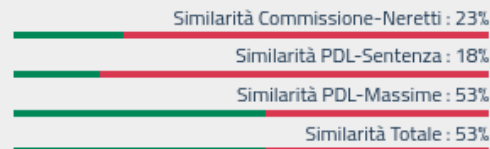
[Riferimenti normativi abrogati](#)

Riferimenti normativi abrogati:

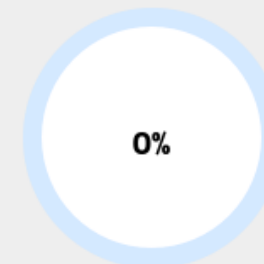


Sentenza N. 86 Anno 2009

FRBRExpressionuri: /akn/it/judgment/sentenza/
cortecostituzionale/2009-03-27/86/!main



Riferimenti normativi abrogati:



Evaluation

Similarity Criterion	Top 10 Recall	Top 5 Recall
Articles-Paragraphs	0	0
Commission-Ratio decidendi	0	0
Articles-Abstract	0.51	0.31

Table 4. Results of the experiments with 10 bills

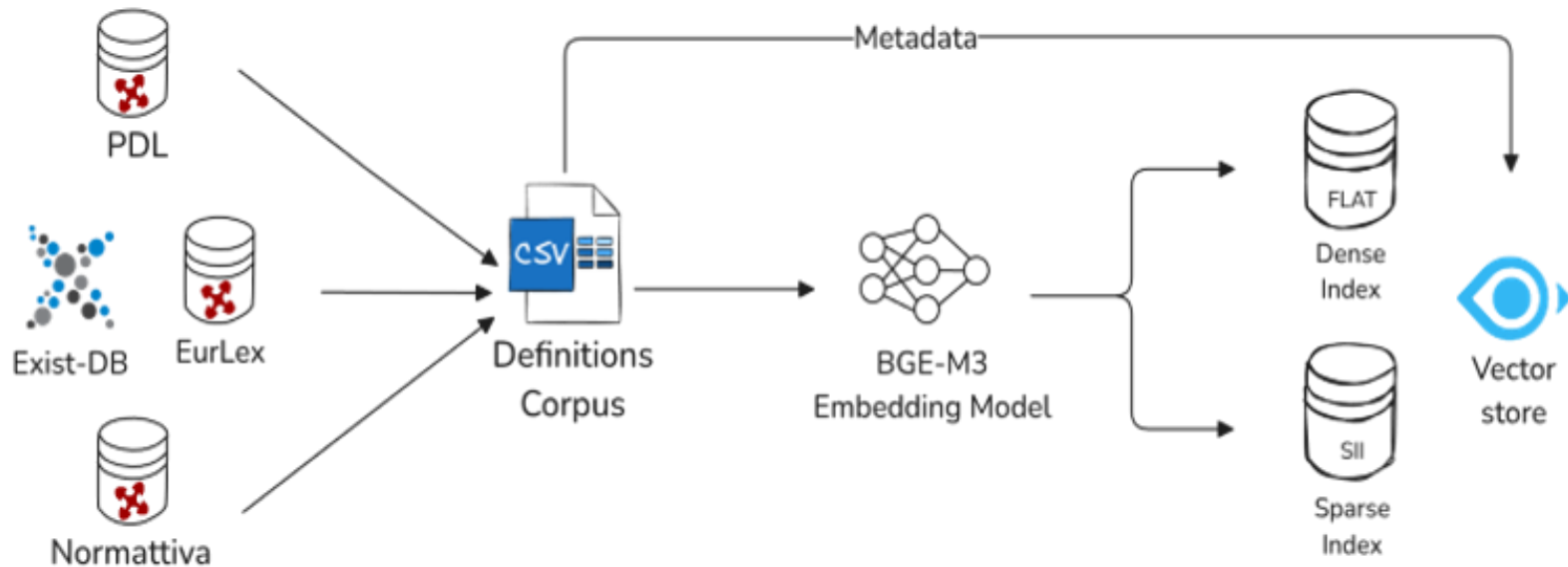
Agentic retrieval and generation of definitions

One of the tasks that are performed during legislative drafting is the creation of **definitions**.

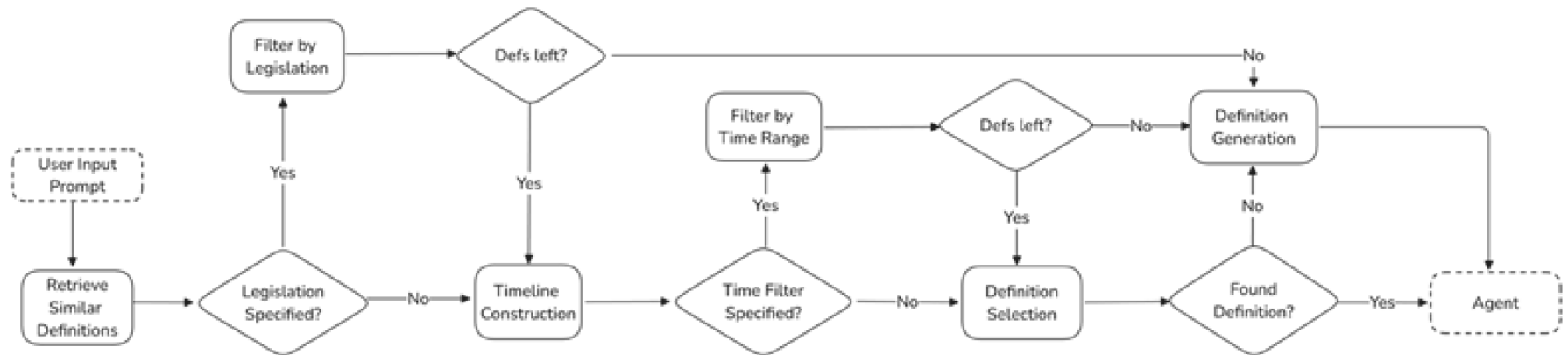
These definitions are used to clearly establish terminology that is present in the laws, and in this context, our goals are:

1. Identifying the appropriate term according to the domain of the act, avoiding ambiguity
2. Retrieving existing legislative definitions of the same term, to prevent conflicts
3. Drafting an appropriate definition when none exists

Embeddings for legislative definitions



Pipeline



Quantitative Evaluation

LLM used: LLama 3.3 70B-Instruct

Definition generation evaluation

Dataset: LexDrafter¹ (1330 entries)

Model	BLEU				ROUGE			BERTScore			BLEURT
	1	2	3	4	1	2	L	P	R	F1	
DefAgent-IT	0.34	0.26	0.22	0.19	0.35	0.21	0.31	0.85	0.87	0.86	0.51
DefAgent-EU	0.26	0.16	0.12	0.09	0.35	0.19	0.31	0.78	0.81	0.79	0.49
LD-Vicuna	0.25	0.13	0.07	0.04	–	–	–	0.83	0.81	0.82	0.47
LD-LLama-2	0.28	0.15	0.09	0.05	–	–	–	0.83	0.81	0.82	0.47

Table 1. Summary of the automated generation metrics scores for our models (DefAgent-IT and DefAgent-EU for Italian and European documents) and for the two models from LexDrafter (LD-Vicuna and LD-LLama-2).

¹ Chouhan, Ashish, and Michael Gertz. 2024. “LexDrafter: Terminology Drafting for Legislative Documents Using Retrieval Augmented Generation.”

Evaluation: Qualitative Evaluation

Metric	IT	EU
Accuracy	3.57	3.97
Contextual Appropriateness	4.32	4.28
Completeness	3.94	4.01
Consistency	4.23	4.19
Timeliness	4.06	4.41
References accuracy	2.25	2.16

Table 2. Average scores for each dimension

Metric	IT	EU
Hallucination Rate	0.10	0.08
Irrelevance Rate	0.10	0.04

Table 3. Hallucination and irrelevance rates

Issues:

- References too broad, out of scope (e.g., references to other domains), not in force, hallucinated, redundant or recursive
- abstract or overly legalistic formulations

Conclusions

- In order to mitigate the risks of the application of LLMs to the legal domain, we can apply a hybrid approach
- The Akoma Ntoso standard allows us to leverage its rich information and to use them with sophisticated NLP/AI models (structure of the documents, definitions, references, jurisdiction, etc)
- By enriching statistical models with symbolic information, we reduce the risk of hallucinations and ensure that the models consider the relevant (i.e. in force) documents while also considering the temporal aspects of the law.