ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

HYPERMODELEX

European Research Council
Established by the European Commission

Co-funded by
the European Union

Hybrid Classification of European Legislation using Sustainable Development Goals

Michele Corazza, Franco M. T. Gatti, Salvatore Sapienza, Monica Palmirani

CIRSFID-ALMA AI, University of Bologna, Italy

Introduction

- In deliberative institutions there is a need to detect where a specific policy is implemented in legal acts;
- This detection process should be as fine-grained as possible, to allow the institutions to assess the extent of the legislative efforts;
- In this context, the Sustainable Development Goals (SDG) provide are world-wide policies whose implementation should be monitored in European legislative documents;
- We aim to use a hybrid AI approach to classify EU legislative documents with their related SDG in an unsupervised manner.

Sustainable Development Goals

- Part of the 2030 United Nations Agenda for Sustainable Development;
- Adopted by all United Nation member states in 2015;
- Composed of 17 goals, whose aim is to "peace and prosperity for people and the planet";
- Each goal is composed of one or more finer-grained targets, which are used to assess progress;

Goal	Targets
1 End poverty in all its forms everywhere	1.1 By 2030, eradicate extreme poverty for all people everywhere, currently measured as people living on less than $1.25 a day
	1.2 By 2030, reduce at least by half the proportion of men, women and children of all ages living in poverty in all its dimensions according to national definitions

Research Questions

RQ1: which parts (Recitals, definition, main body, etc.) of the legislative document are connected with the targets of the SDGs taxonomy?

RQ2: what is the evolution of SDGs classification over time?

Akoma Ntoso

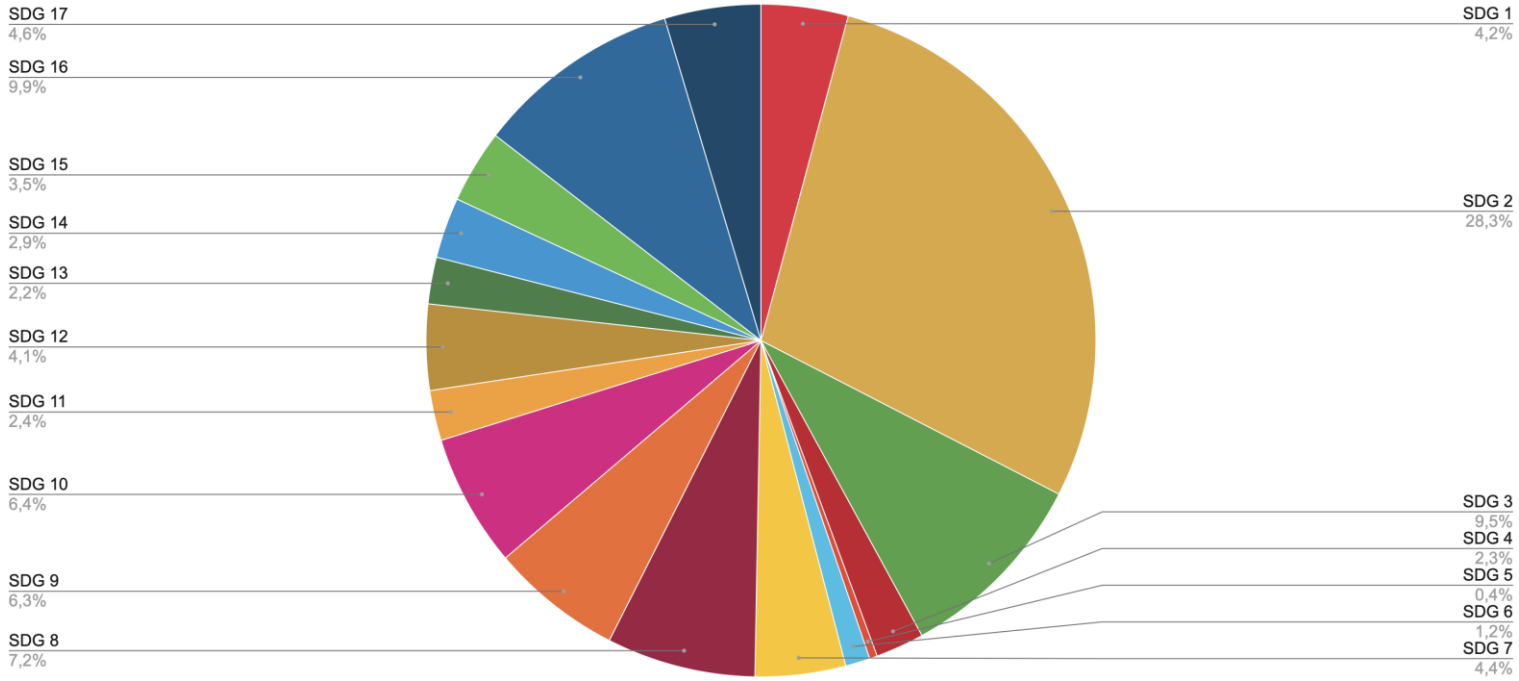
- An XML standard for legal documents (both legislative and judicial);
- Used by a multitude of institutions around the world;
- It allows us to leverage information about the structure (articles, commas, lists, etc) of legislative documents
- It encodes information about **normative references**;
- It also allows us to consider the temporal aspects of the law.

AKOMA NTOSO

Architecture for Knowledge-Oriented Management of African Normative Texts using Open Standards and Ontologies

Dataset

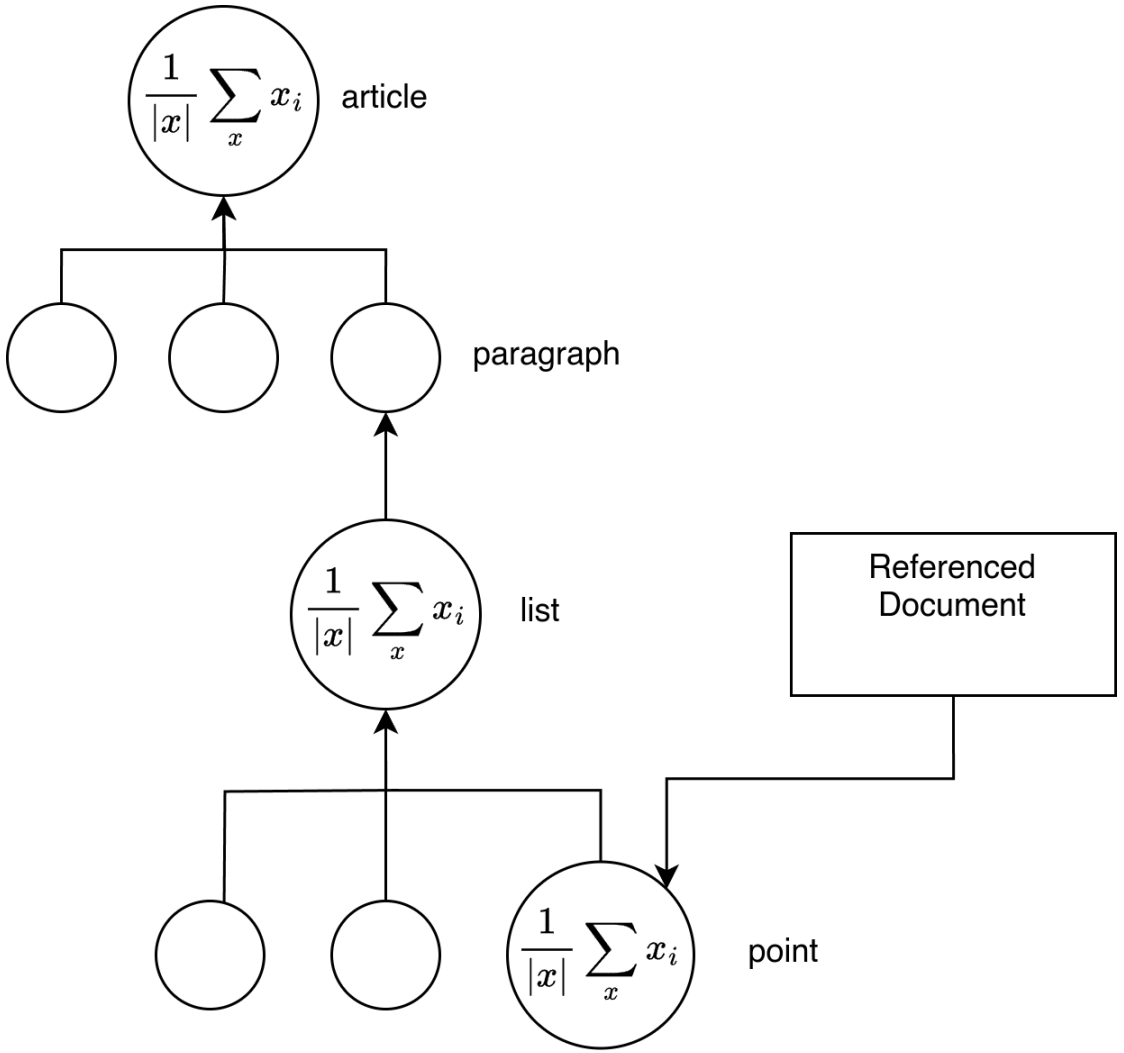
- 3846 European legislative documents converted in Akoma Ntoso XML;
- Documents from the Juncker Commission's mandate (2015-2019)
- They were annotated at the document level by the Joint Research Centre of the EU commission with their related SDG targets.



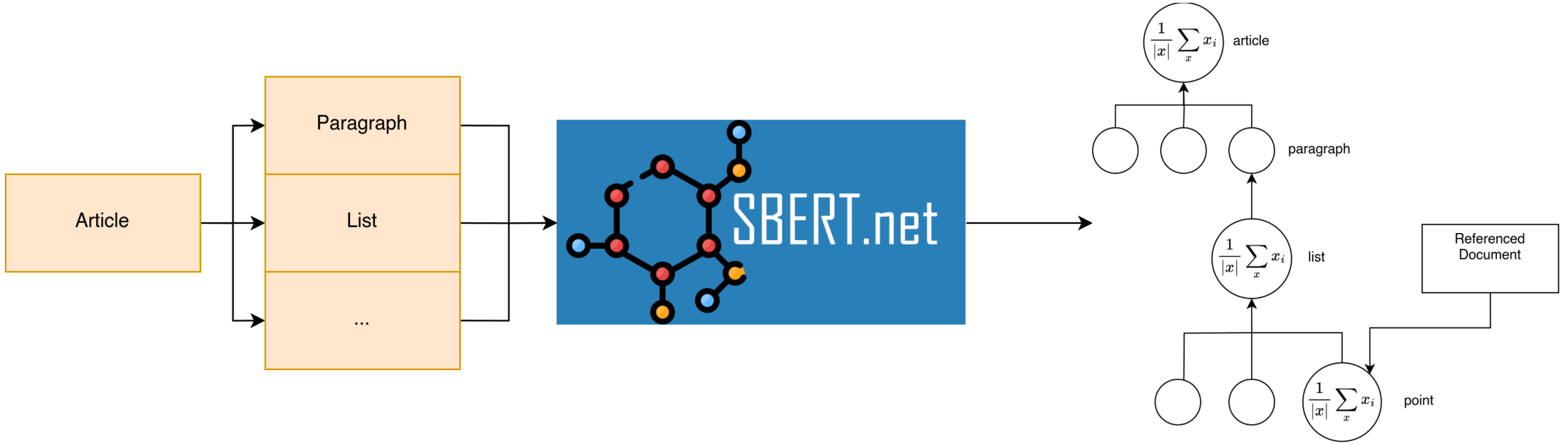
Sentence Transformers

- Models which have been trained using an approach analogous to **siamese neural networks**
- They produce a **vector representation for sentences** (embedding) which can be compared with other vectors using a metric (cosine similarity)
- We use them to compare legislative documents with the description of SDG targets.

Vector representation for document parts: tree strategy



Method

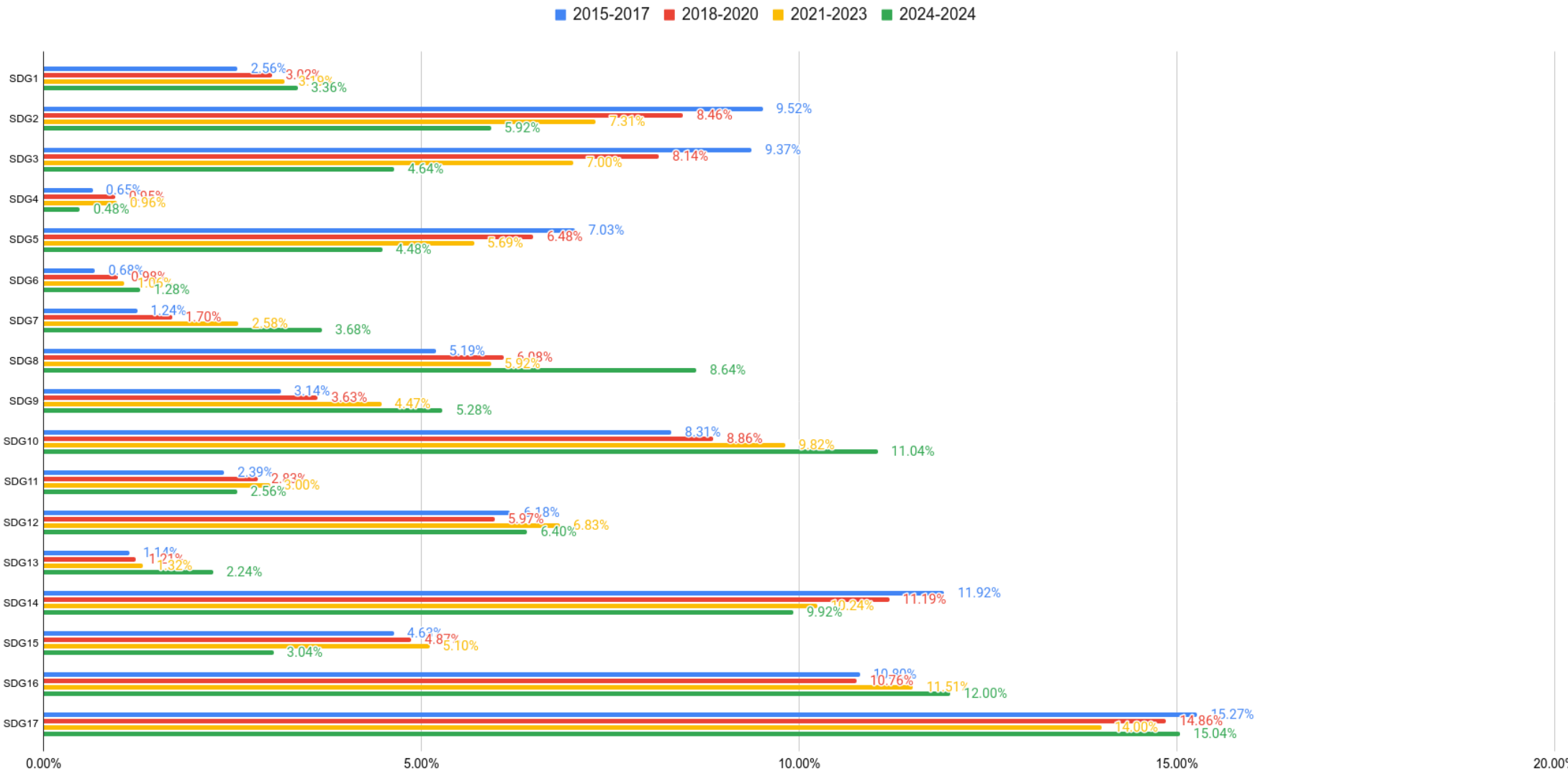


Document level results

Strategy	Average	Precision	Recall	F1 Score
All articles	Macro	**0.11**	0.16	0.07
	Weighted	**0.37**	0.22	0.14
Random (All articles)	Macro	0.03 ± 0.0003	0.33 ± 0.008	0.06 ± 0.0004
	Weighted	0.12 ± 0.001	0.41 ± 0.003	0.17 ± 0.001
First four	Macro	0.09	0.09	0.04
	Weighted	0.29	0.12	0.06
Random (First four)	Macro	0.02 ± 0.0006	0.08 ± 0.007	0.03 ± 0.0008
	Weighted	0.09 ± 0.003	0.1 ± 0.003	0.08 ± 0.002
Recitals + first four	Macro	0.09	0.36	**0.10**
	Weighted	0.27	0.44	**0.22**
Random (Recitals + first four)	Macro	0.03 ± 0.0001	$\mathbf{0.52} \pm 0.008$	0.05 ± 0.0002
	Weighted	0.10 ± 0.0008	$\mathbf{0.65} \pm 0.003$	0.16 ± 0.001

Table 1: Precision, recall and F1 score for the four strategies, obtained from Macro and Weighted averages over individual classes. In bold, the best values for each metric. Underlined, the higher metric when comparing each strategy with its baseline. For each baseline we report the means and standard deviations of the metrics over 100 runs.

Monitoring SDGs through time



ALMA MATER STUDIORUM
UNIVERSITÀ DI BOLOGNA

Conclusions

RQ1: the first four articles and the recitals are better suited to detect SDG targets according to our experiments;

RQ2: steady decrease in SDG02 (Zero Hunger) related legislation from 2018-2020 to 2024. SDG13 (Climate Action) exhibits a very recent increase which almost doubles its percentage. EU legislation related to SDG03 (Good Health and Well-being) decreases during the 2021-2023 period in comparison to 2015-2017.

Our work produces a finer-grained traceability of the SDGs policies in EU legislation allowing the legislator to detect the articles where the association is weakest. This tool could be used during drafting to propose better legal definitions and to improve the implementation of the SDGs.