

UNIOR NLP@COFE Datathon 2022

Task 1 - Keyword Extraction from Proposals

Gennaro Nolano, Maria Pia di Buono, Johanna Monti
UNIOR NLP Research Group - University of Naples "L'Orientale"
{gnolano, mpdibuono, jmonti}@unior.it

Proposed workflow

Multilingual proposals

Automatic KeyPhrase
Extraction

MultiWord
Expression
Discovery

External
Knowledge
Injection

Graph-Based
Candidate
Representation

Multilingual
Keywords

Automatic Keyphrase Extraction

Proposal id: 56234

Proposal title: Make the carbon pricing system global

Proposal body: Compliance and implementation of carbon prices is a solution to mitigate pollution levels. Europe buys products worldwide thus contributing with its demand to the CO2 foreign production. The carbon pricing system cannot be sustainable if it's not global, and the EU needs to do more to make polluters pay.



*carbon prices: 0.14,
products: 0.11,
pollution levels: 0.10,
europe: 0.09,
implementation: 0.09,
solution: 0.08,
compliance: 0.07,
demand: 0.05,
co2 foreign production: 0.05,
carbon pricing system: 0.04*

Multiword Expression Discovery

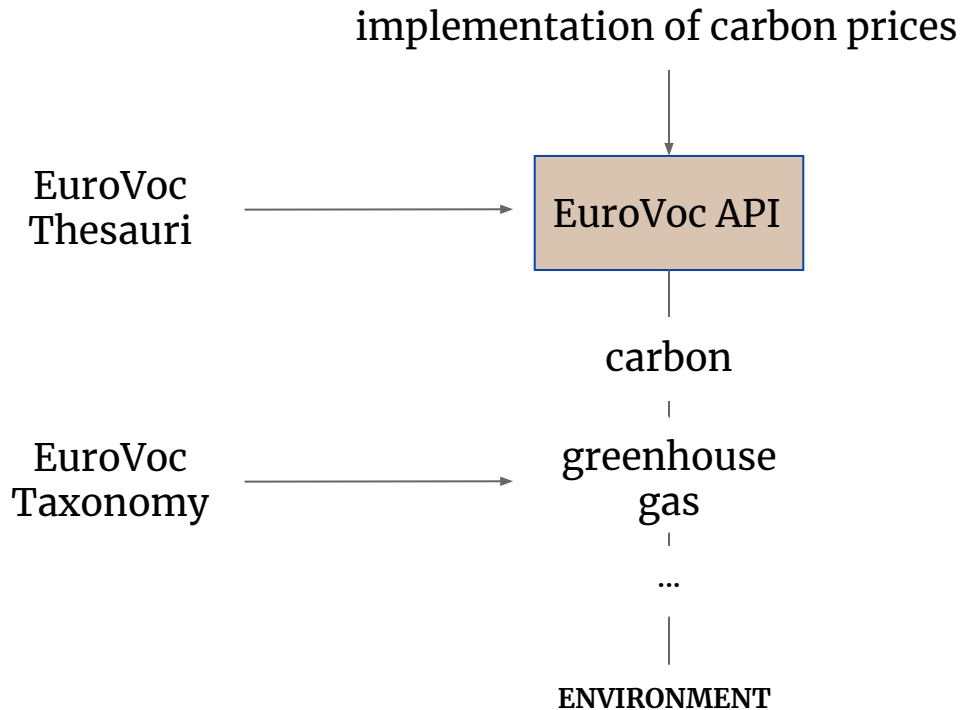
carbon prices: 0.14,
products: 0.11,
pollution levels: 0.10,
europe: 0.09,
implementation: 0.09,
...

Compliance and implementation of
carbon prices is a solution to mitigate
pollution levels

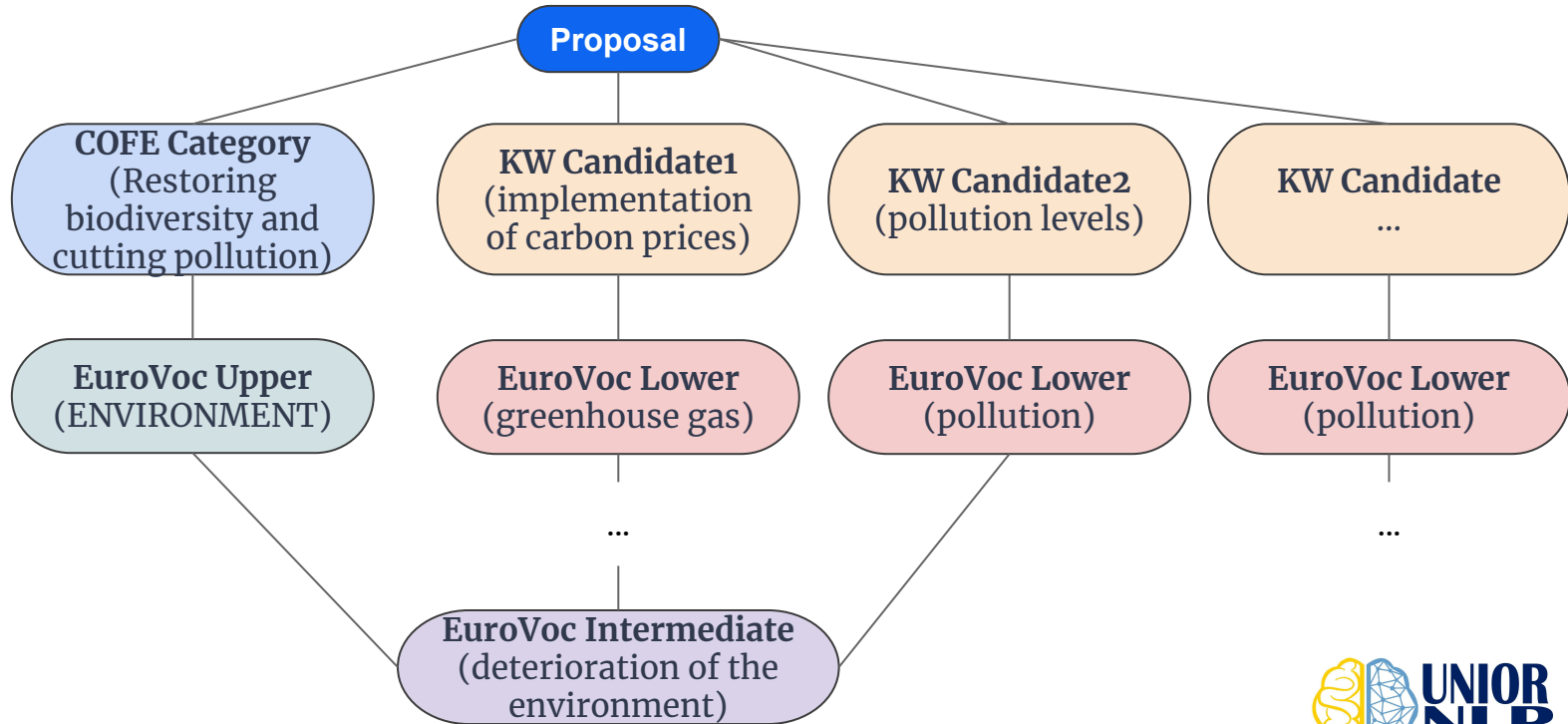
implementation of carbon prices: $(0.14 + 0.09)/2 = 0.115$

Distance	MWE Pattern
0	w1 w2
1	w1 preposition w2 w1 adjective w2
2	w1 preposition determiner w2 w1 adjective preposition w2
3	w1 adjective preposition determiner w2

External Knowledge Injection



Graph-Based Candidate Representation



Graph-Based Candidate Re-Ranking

$$\text{Finale Score} = e^{\text{mpr}} + e^{\text{bc}}$$

mpr = Multipartite Ranking score

bc = betweenness centrality

Thank you all for your attention