

# Annotation lexico-grammaticale du grec/latin

[https://github.com/OdysseusPolymetis/phi  
lo\\_num\\_grenoble\\_24](https://github.com/OdysseusPolymetis/phi_lo_num_grenoble_24)

# Qu'est-ce que "pré-traiter" ?

# La lemmatisation

C'est un retour du mot à son entrée lexicale commune (« forme canonique » enregistrée dans les dictionnaires de la langue), c'est-à-dire sans flexions.

Exemple :

<b>Quo</b>	<b>usque</b>	<b>tandem</b>	<b>abutere</b>	<b>Catilina</b>
Qui	usque	tandem	abutor	Catilina

# L'étiquetage de nature

On appelle ça "posttagging". Cela consiste à rechercher la nature des mots.

Exemple :

<b>Quo</b>	<b>usque</b>	<b>tandem</b>	<b>abutere</b>	<b>Catilina</b>
Qui	usque	tandem	abutor	Catilina
PRON	ADP	ADV	VERB	NOUN

# Autres tâches (variées)

Entre autres : la NER

Exemple :

<b>Quo</b>	<b>usque</b>	<b>tandem</b>	<b>abutere</b>	<b>Catilina</b>
Qui	usque	tandem	abutor	Catilina
PRON	ADP	ADV	VERB	NOUN
				PERS

# Pourquoi traiter les textes

[https://github.com/OdysseusPolymetis/philosophy\\_num\\_grenoble\\_24/blob/main/why\\_nlp.ipynb](https://github.com/OdysseusPolymetis/philosophy_num_grenoble_24/blob/main/why_nlp.ipynb)

# Différents outils à comparer

# Outil 1 : pie-extended

[https://github.com/OdysseusPolymetis/phi  
lo\\_num\\_grenoble\\_24/blob/main/pie\\_for\\_co  
mparison.ipynb](https://github.com/OdysseusPolymetis/phi_lo_num_grenoble_24/blob/main/pie_for_comparison.ipynb)

## Outil 2 : cltk/stanza

[https://github.com/OdysseusPolymetis/phi  
lo\\_num\\_grenoble\\_24/blob/main/cltk\\_for\\_c  
omparison.ipynb](https://github.com/OdysseusPolymetis/phi_lo_num_grenoble_24/blob/main/cltk_for_comparison.ipynb)

[https://github.com/OdysseusPolymetis/phi  
lo\\_num\\_grenoble\\_24/blob/main/stanza\\_fo  
r\\_comparison.ipynb](https://github.com/OdysseusPolymetis/phi_lo_num_grenoble_24/blob/main/stanza_for_comparison.ipynb)

## Outil 3 : spacy

[https://github.com/OdysseusPolymetis/phi  
lo\\_num\\_grenoble\\_24/blob/main/spacy\\_for  
\\_comparison.ipynb](https://github.com/OdysseusPolymetis/phi_lo_num_grenoble_24/blob/main/spacy_for_comparison.ipynb)

## Outil 4 : BERT

[https://github.com/OdysseusPolymetis/phi  
lo\\_num\\_grenoble\\_24/blob/main/ancient\\_gr  
eek\\_bert\\_for\\_comparison.ipynb](https://github.com/OdysseusPolymetis/phi_lo_num_grenoble_24/blob/main/ancient_greek_bert_for_comparison.ipynb)

**En conclusion**

Grenoble -- 18 janvier 2024

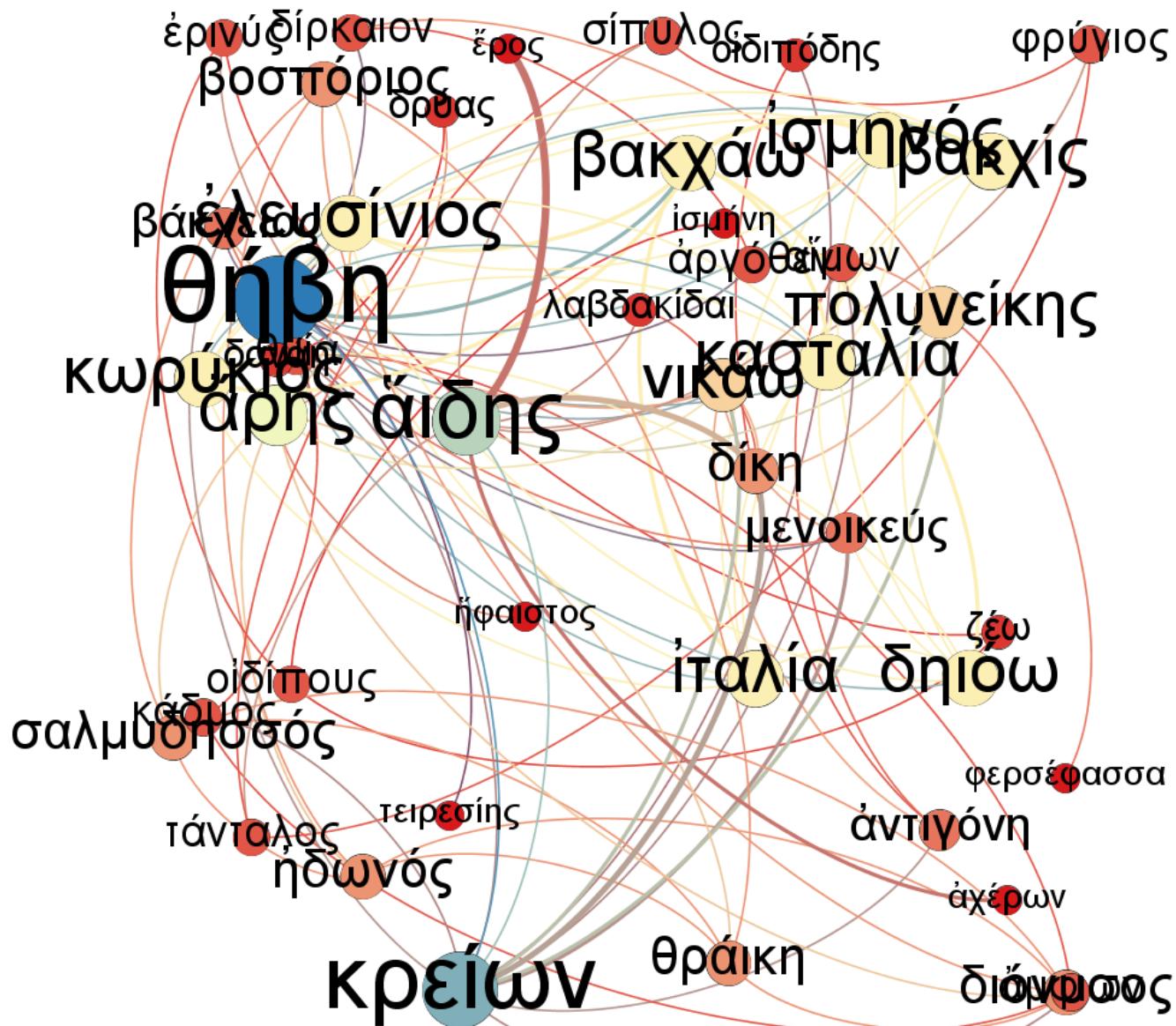
**Je vous remercie.**

[marianne.reboul@ens-lyon.fr](mailto:marianne.reboul@ens-lyon.fr)

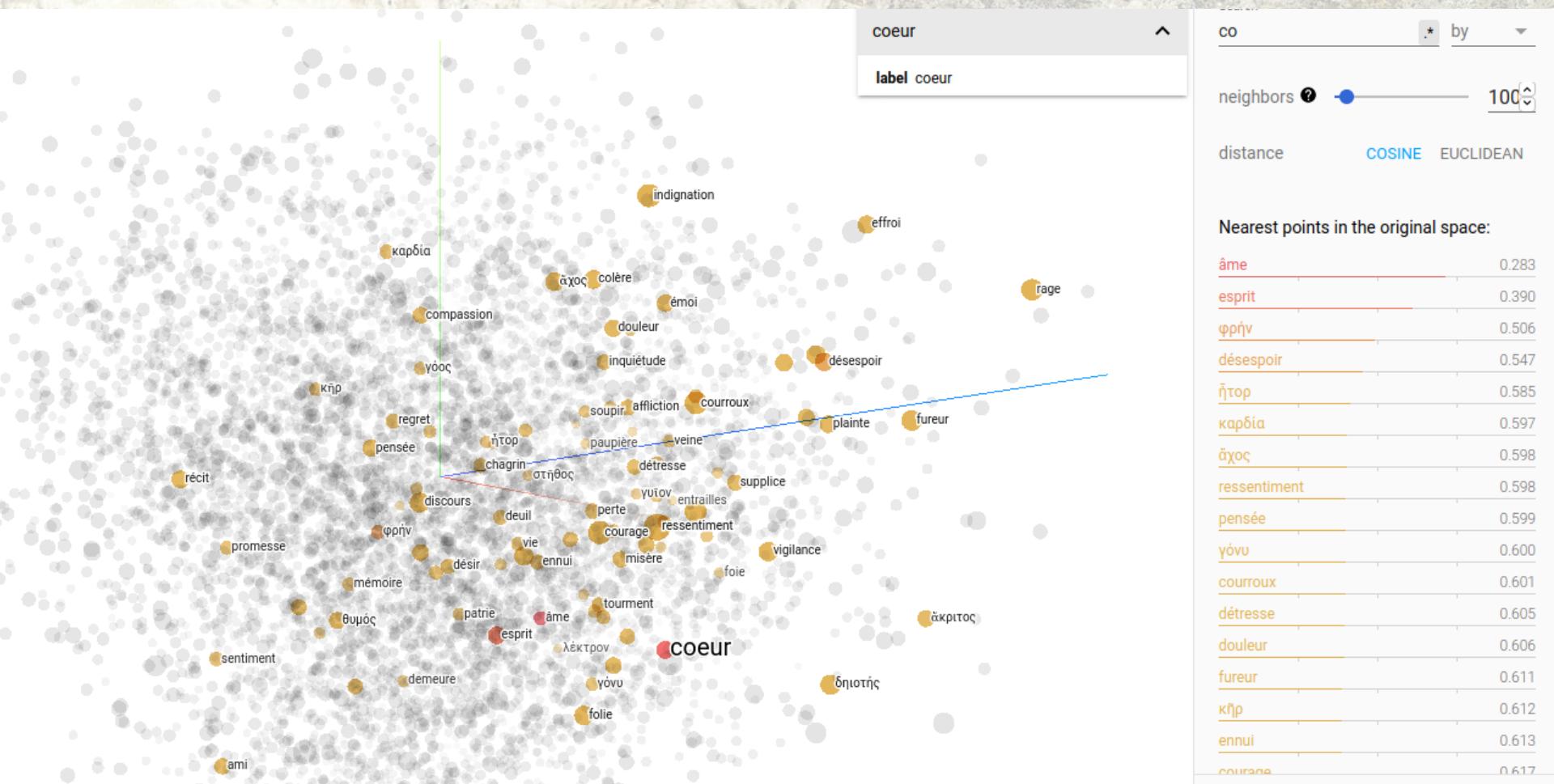
# Références (et remerciements)

- deucalion
- pie-extended
- stanza
- ancient-greek-bert
- spacy
- udpipe
- voyanttools
- cltk

# Addenda :Quelques exemples de visualisations récentes

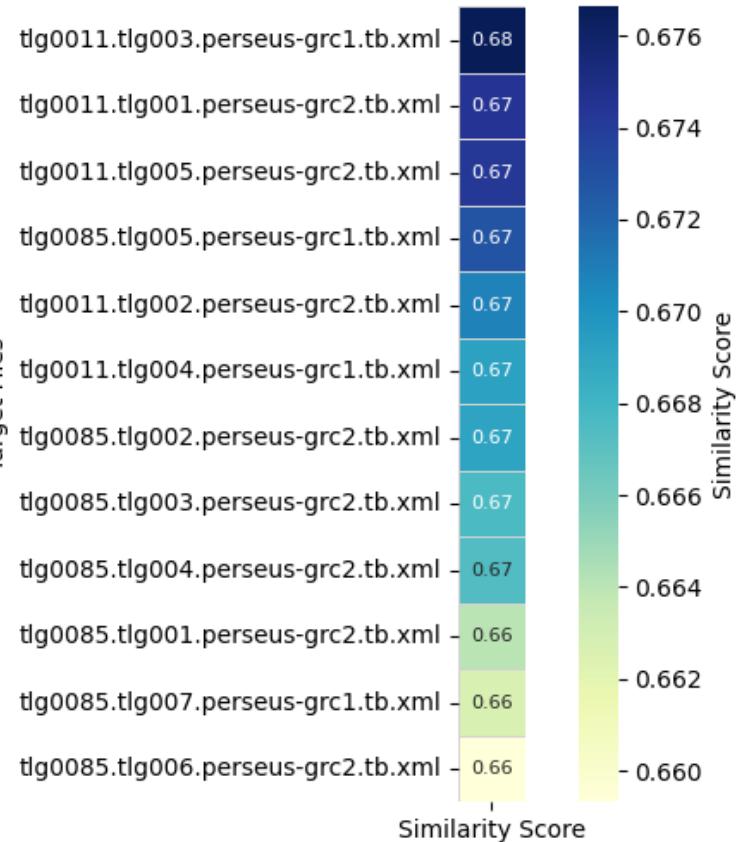


# Grenoble -- 18 janvier 2024



# Similarité globale entre Homère et les tragiques

Similarity Heatmap for tlg0012.tlg001.perseus-grc1.tb.xml



Similarity Heatmap for tlg0012.tlg002.perseus-grc1.tb.xml

