

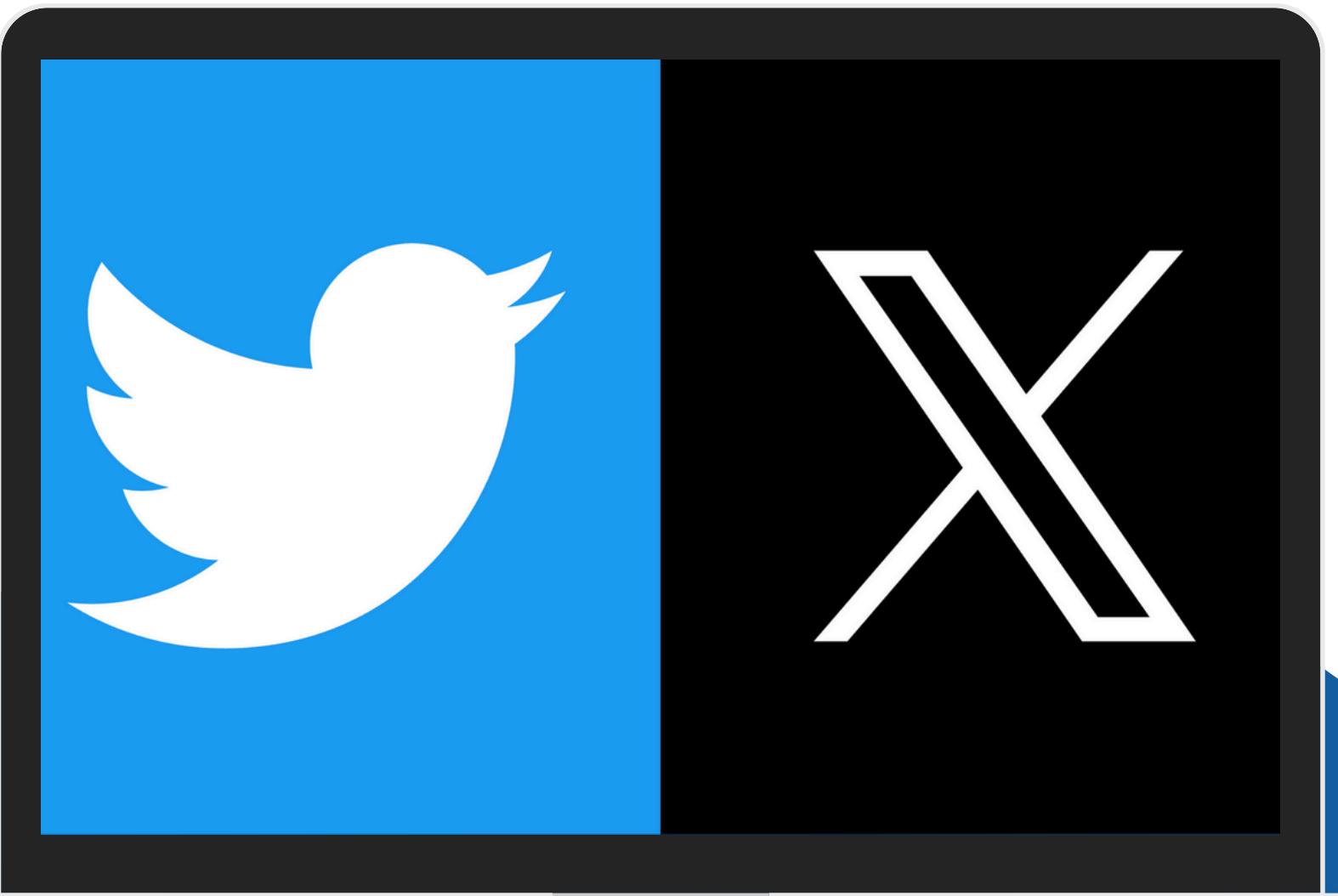


Controverses de Tweets

Réalisé par Jonathan DUCKES, Audric GIRONDIN, Maéva
MAÏO, Houria SAYAH, Alya ZOUZOU



janssen
Horizon



INTRODUCTION

- **Controverse**: un échange d'opinions divergentes mettant en lumière des points de vue contradictoires
- **Interventions non médicamenteuses (INM)** : mesures préventives et stratégies de soins non liées aux médicaments

Comment les visualisations de données peuvent contribuer à une meilleure compréhension des facteurs de controverses sur Twitter ?

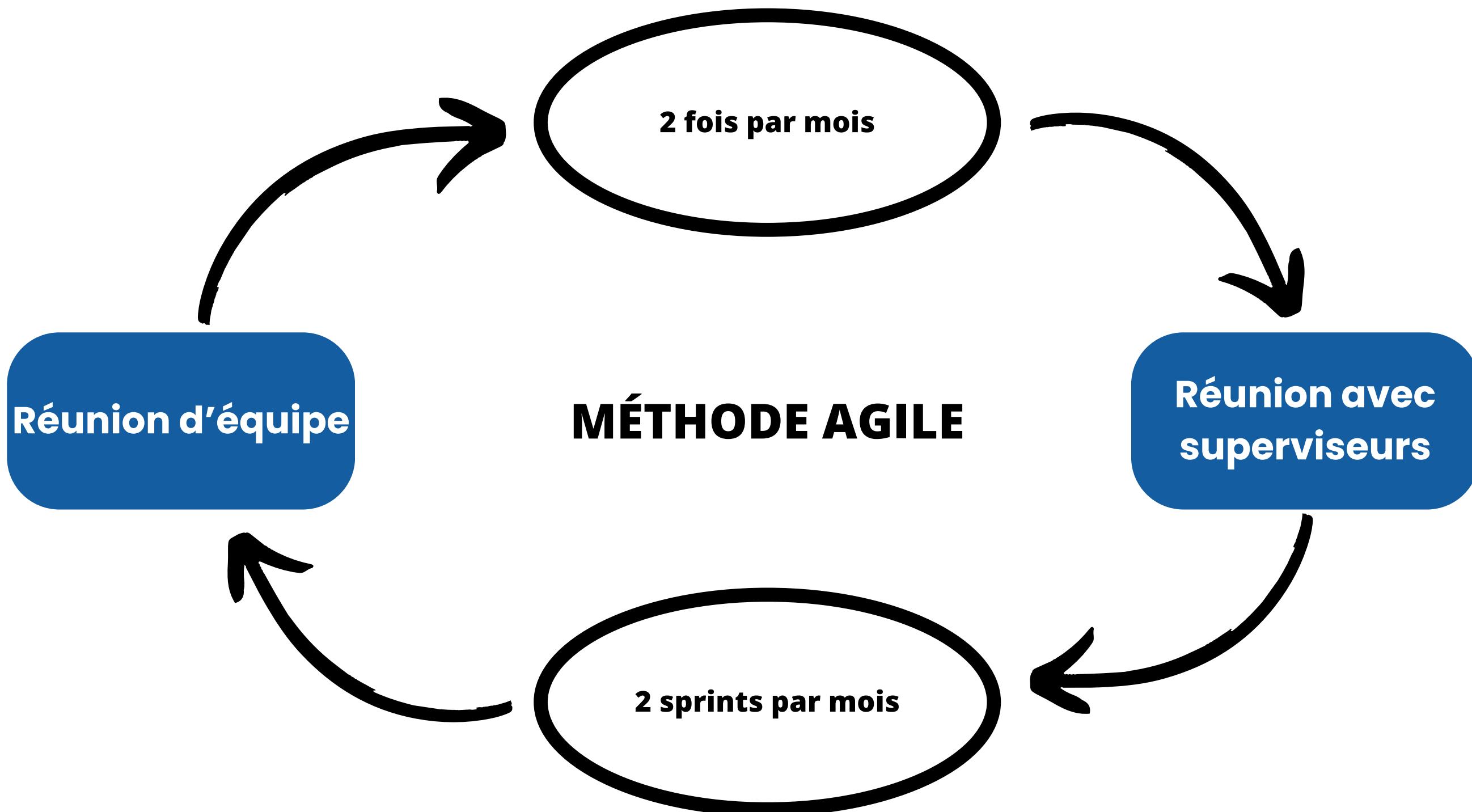


Sommaire

▶ Gestion de projet	3
▶ Traitement de données	4
▶ Analyses et visualisations des tweets	5
▶ Indice de controverse	12
▶ Conclusion	14
▶ Perspectives	16



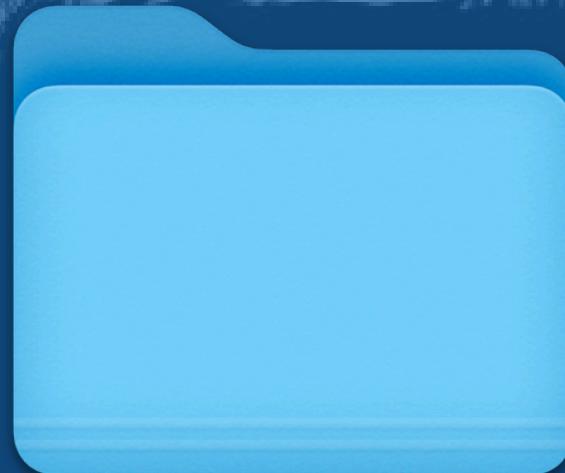
Gestion de projet





**Quelles sont nos
données ?**

Traitement de données



Données brutes

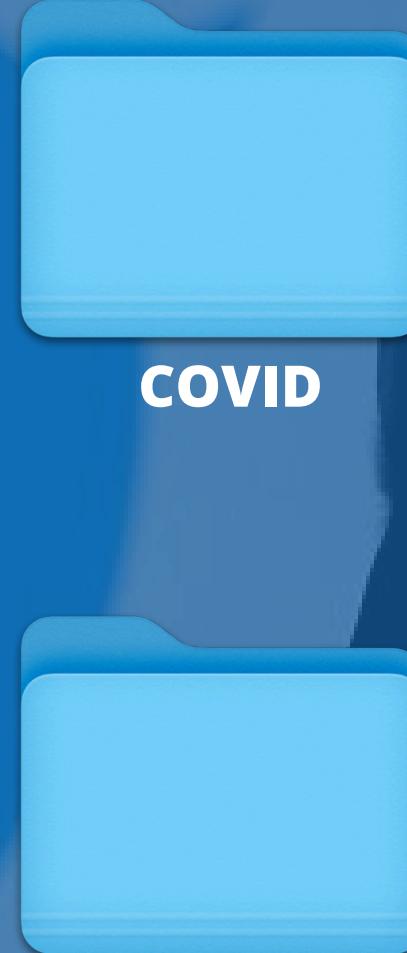
Nombre de tweets : 5 486 603

Taille : 23 G



Traitement de données

- Récupérer
- Trier et organiser
- Script filtre Python
- Traitement et analyse des tweets
- Préparation pour visualisation



COVID

INM

Dimensions : 249 456, 17

Likes : 7 209 886

Retweets : 2 048 663

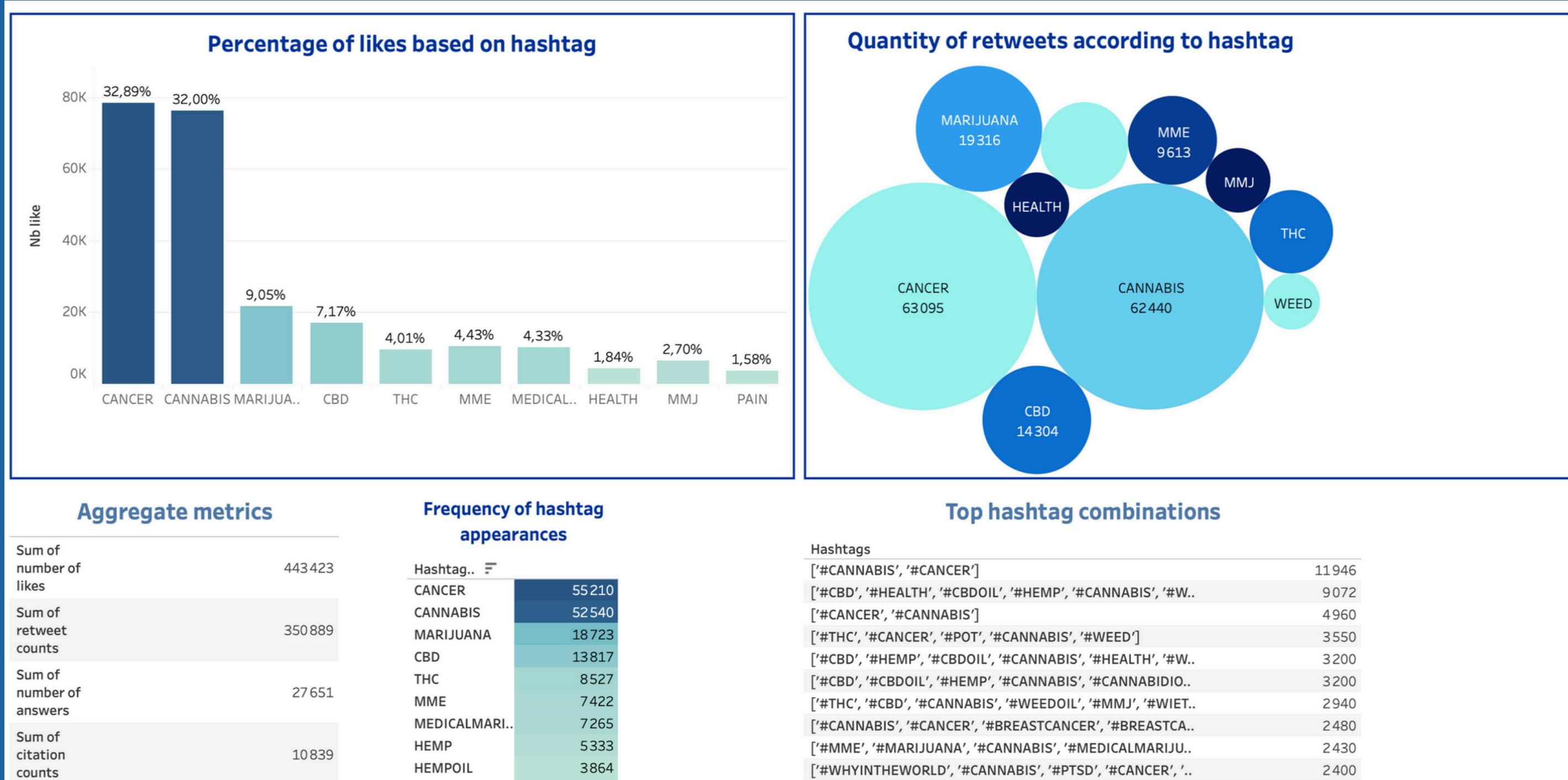
Réponses : 623 221

Citations : 280 448



**Comment
représenter
nos données ?**

Exploration visuelle et traitement des tweets



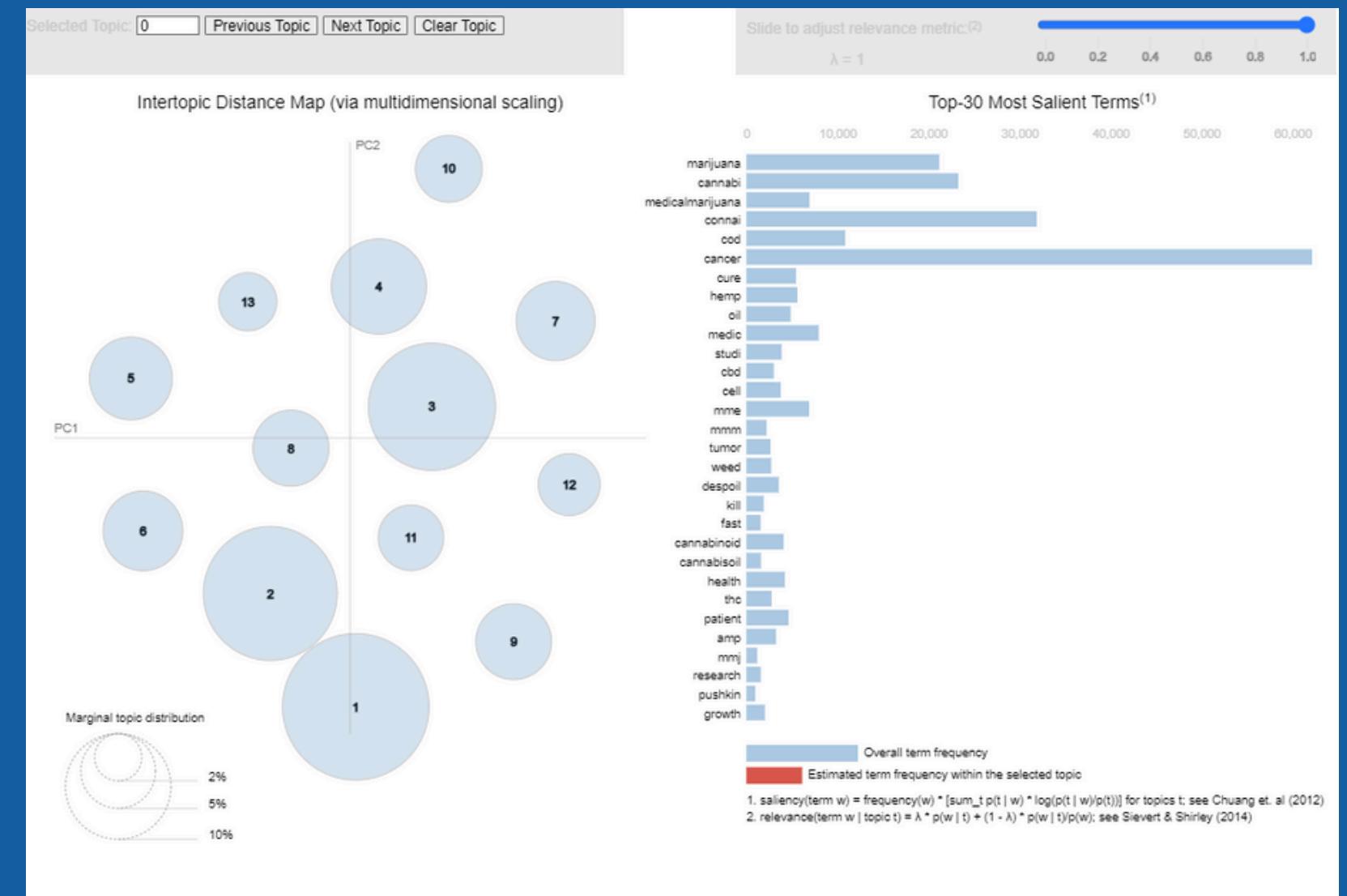


**Quelles sont les
caractéristiques des
textes qualifiés de
controversés ?**

Topic Modeling - LDA

Label	Top Words
Treatment with Cannabis	[cancer, cannabi, connai, cell, marijuana, cure, studi, treatment, cannabinoid, kill]
Diet and Intermittent Fasting	[cancer, cannabi, breast, connai, breastcanc, analysi, extens, v, sport, wast]
Health Benefits of Cannabis	[cancer, health, daili, connai, cannabi, latest, benefit, cbd, hemp, ill]
CBD and Pain Management	[cancer, connai, cbd, hemp, boil, health, anxieti, cannabi, cod, epilepsi]
Health and CBD Oil Usage	[cancer, connai, cb, cod, hemp, list, mmm, e, wietoli, weedoil]
THC and Tumor Inhibition	[cancer, connai, mme, tumor, cod, growth, cannabinoid, cannabi, thc, marijuana]
Research on Cancer Prevention	[cancer, connai, cannabi, amp, pain, fight, effect, growth, patient, cell]
Epilepsy and Medical Marijuana	[cancer, marijuana, connai, cannabi, cod, medic, mme, medicalmarijuana, despoil, help]
Medical Marijuana Usage	[cancer, connai, cannabi, marijuana, leukaemia, medicalmarijuana, mmm, health, medicin, cod]
Health Risks and Marijuana	[cancer, connai, marijuana, medicalmarijuana, pushkin, hemp, donald, cod, weed, amp]
Cannabis as a Cure	[cancer, cannabi, connai, marijuana, stop, eat, cure, plant, food, amp]
Lifestyle and Cancer Support	[cancer, connai, cannabi, health, fast, diet, sport, cure, port, amp]
THC and Cervical Cancer Treatment	[cancer, cannabi, studi, would, new, connai, great, suggest, cervic, disappear]

Chaney & Blei. Visualizing topic models. In Proceedings of the International AAAI Conference on Web and Social Media (2012 Vol. 6, No. 1, pp. 419-422).

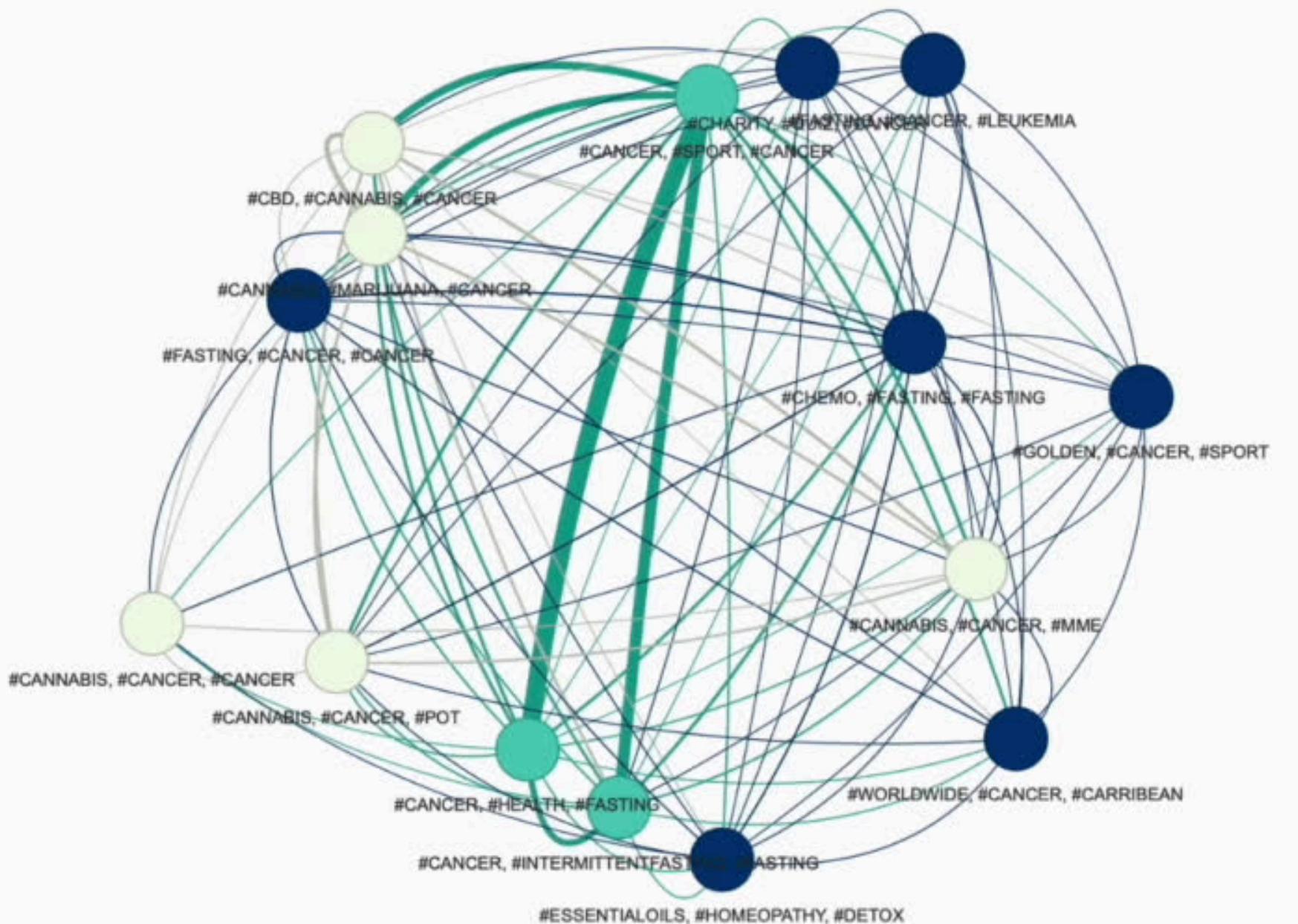




**Comment expliquer les
interactions
entre les utilisateurs au
sein de communautés ?**

Détection de communautés dans Graphique de Cluster de hashtags

- Clusters
- Co-occurrences
- Communauté



Que et al. Scalable Community Detection with the Louvain Algorithm, In Proceedings of the 2015 IEEE International Parallel and Distributed Processing Symposium, Hyderabad, India, 2015, 28-37



**Existe-t-il un lien
entre les sentiments
et la controverse ?**

Sentiment Analysis with



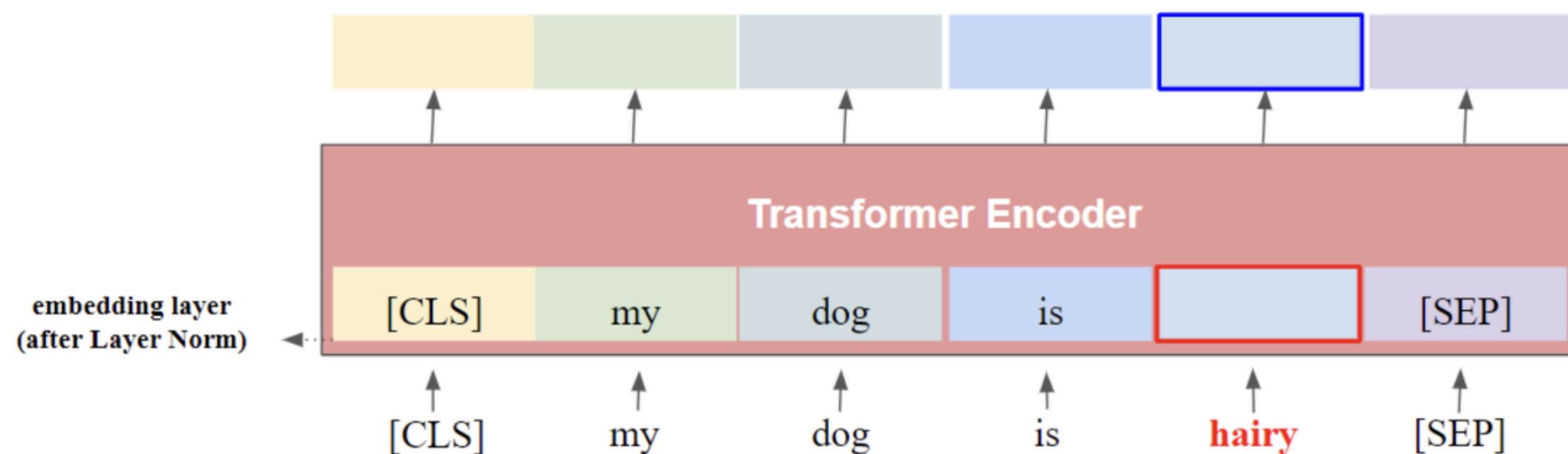
Hugging Face

RoBERTa

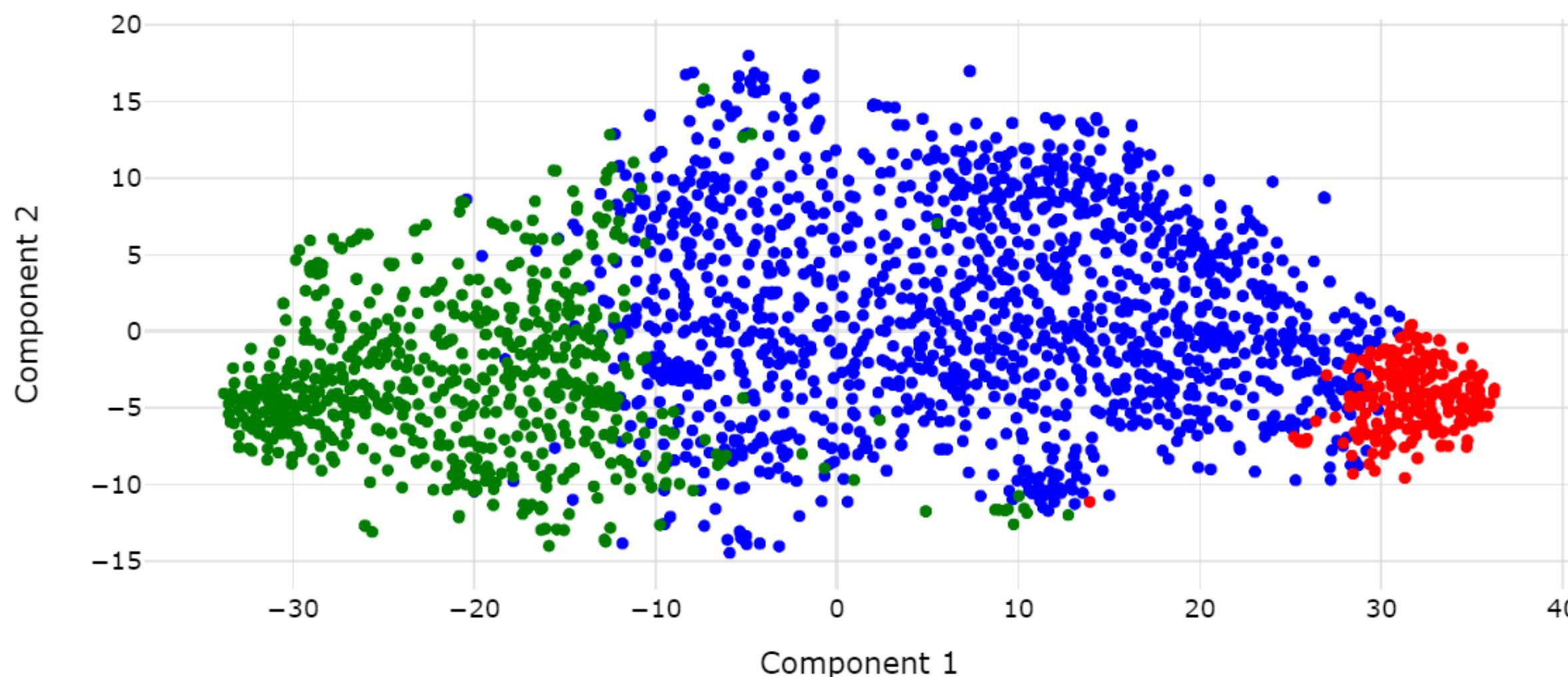
- Compréhension Bidirectionnelle
- Modélisation du langage masqué (MLM)
- Prédiction de la phrase suivante (NSP)

How ?

Meta



t-SNE Visualization of Feature Spaces for Sentiment Analysis



NEGATIVE

- While #Cancer is big business for some, it's a business that only exists because the #Cure is illegal. #Cannabis...
- It doesn't matter to me if you smoke it or take drops, Just think we are healthier, saving the environment, and tax payer money, Fk big pharma. #cannabis #cbdoil #hempoil #stress #ibs #cancer #anxiety #organic #facialcream #painrub #sleep #arthritis #fibromialgia #insomnia

POSITIVE

- We found that almost 25%-35% of the #mice were completely cured of #cancer after administering #Cannabis.



- Medical Marijuana Helped Save This Child From #Cancer <https://t.co/My5dO4ksKR> #marijuana #cannabis #MME #Cancer http



**Peut-on classifier les
tweets selon
l'opinion exprimée ?**

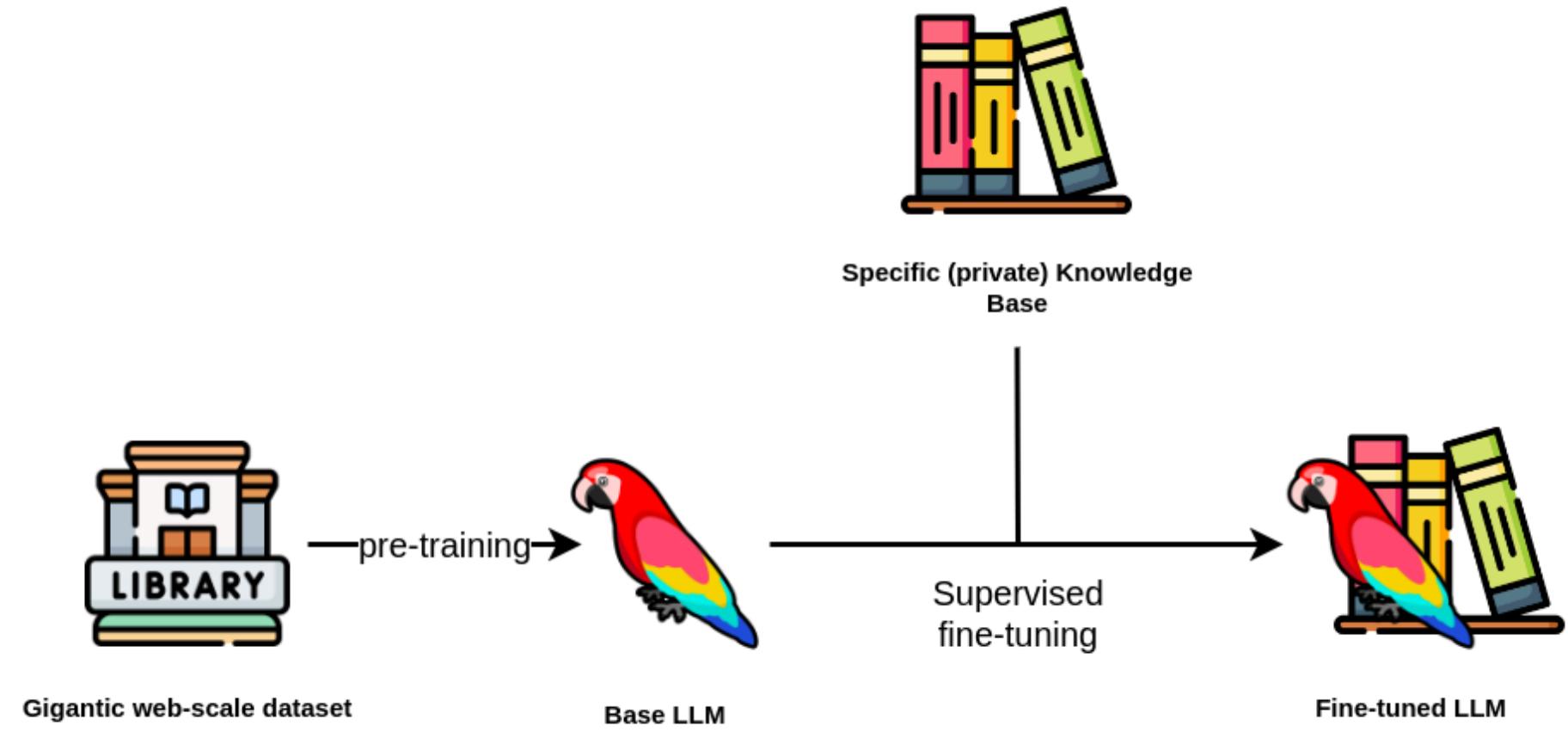
Classification d'opinions



Hugging Face

Le Modèle Pré-entraîné : google-bert/bert-base-uncased

- Base d'entraînement : **80% (450 tweets)**
- Base de test : **20% (50 tweets)**
- **3 époques**
- Labélisation : **FOR, AGAINST, NEUTRAL**



**Déséquilibre de classes
AGAINST MINORITAIRE !**

FOR : Cancer patient: 'I had to become an outlaw or die' <https://t.co/iKhdt4MDR> via @user #MME #Cancer #Cannabis http

AGAINST : New research suggests that #cannabis extract can have a "dramatic effect" on brain #cancer. http #breakthetaboo

Classification d'opinions



Hugging Face

Nos résultats

Accuracy Globale : 86%

Précision	Valeur
FOR	0.971
AGAINST	0.500
NEUTRAL	0.643

Problème pour la classe
AGAINST - MINORITAIRE



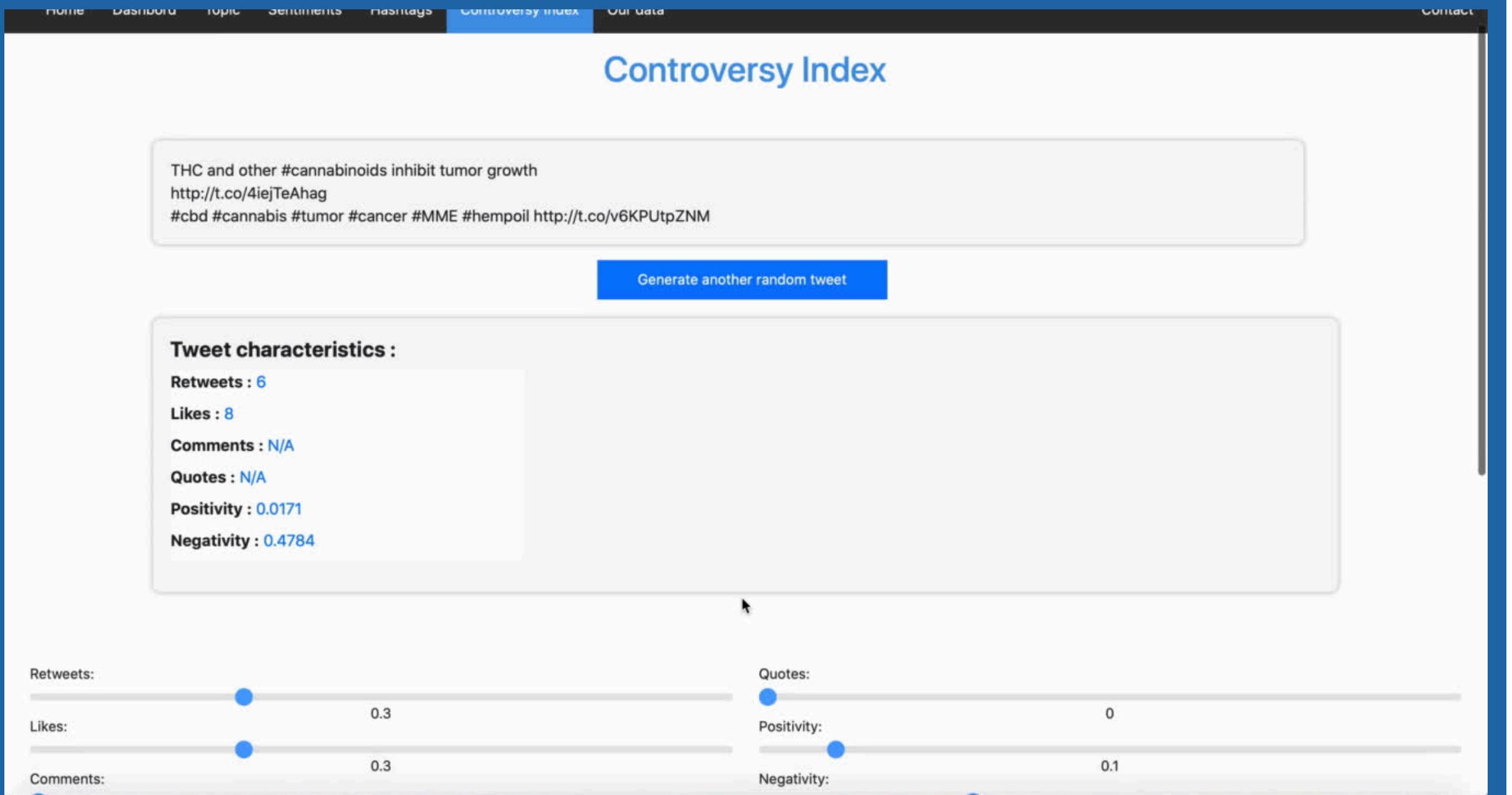
1. DATA AUGMENTATION
2. LOSS pour gérer le déséquilibre de classes
 - Weighted Categorical Cross-Entropy Loss
 - Focal Loss



**Comment
quantifier la
controverse ?**

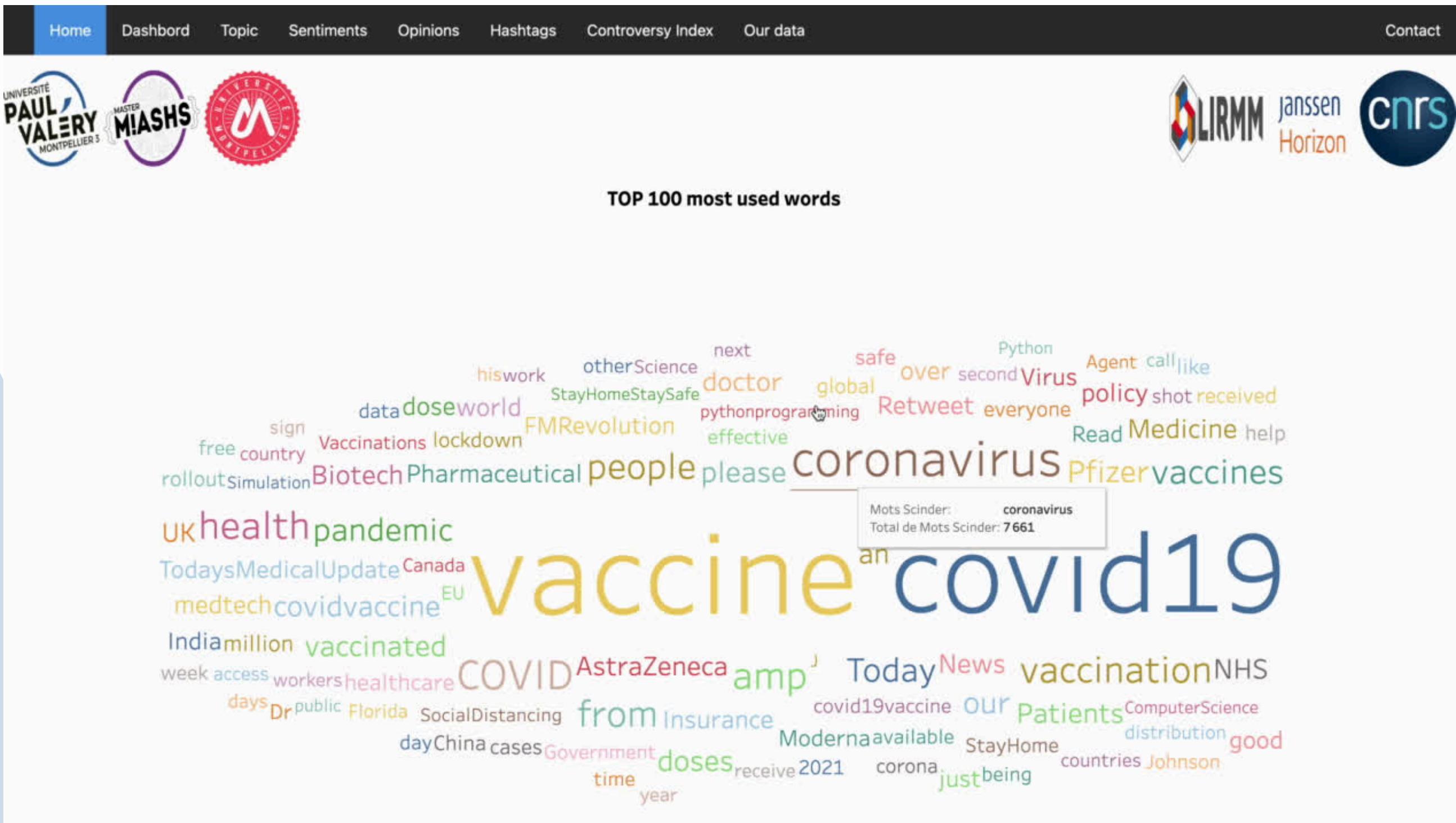
Indice de controverse

- Métrique essentielle
- Plusieurs facteurs pris en compte
- Seuils choisis par l'utilisateur





**Où trouver
nos
recherches ?**



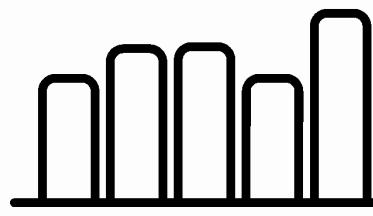
Regroupement des analyses

- Un onglet par axe d'analyse
- Accès à nos données

CONCLUSION

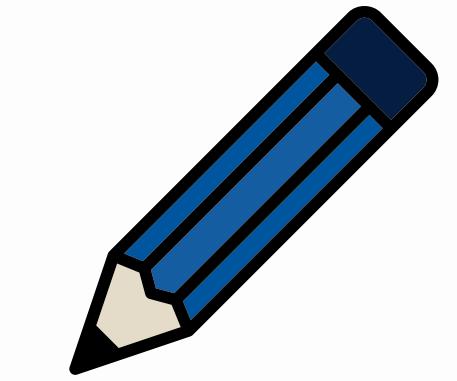
- Diversité des sujets de santé abordés
 - Emotions
 - Sujets de discussion
 - Clusters des communautés
- } Contributions importantes aux débats en ligne

Difficultés



Quantité de données

- Temps d'attente long : exécution des scripts, chargement des visualisations, calcul...



Labellisation

- Quantité limitée de données labellisées

Perspectives

- 01 Labellisation binaire sur la controverse
- 02 Choix meilleure loss fine-tuning





MERCI !

Sans oublier nos superviseurs :
Sandra BRINGAY et Maximilien SERVAJEAN

