

L'analyse de la série Chernobyl par l'IA

Pierre-Loïc Bayart

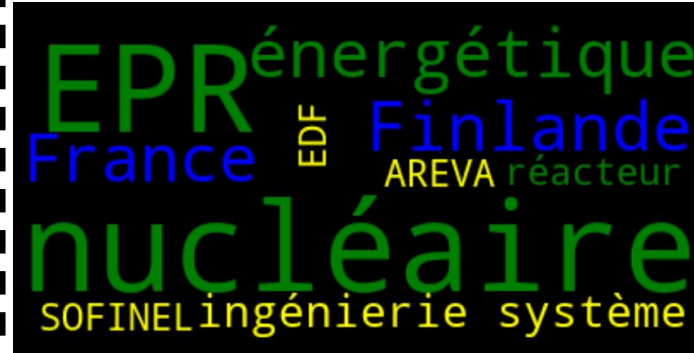
[illegible][illegible]

PYCON.FR
BORDEAUX 2019
du 31 octobre au 3 novembre 2019

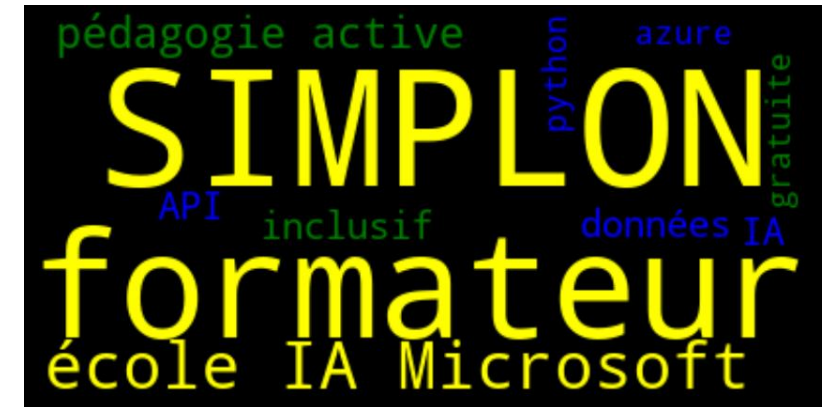
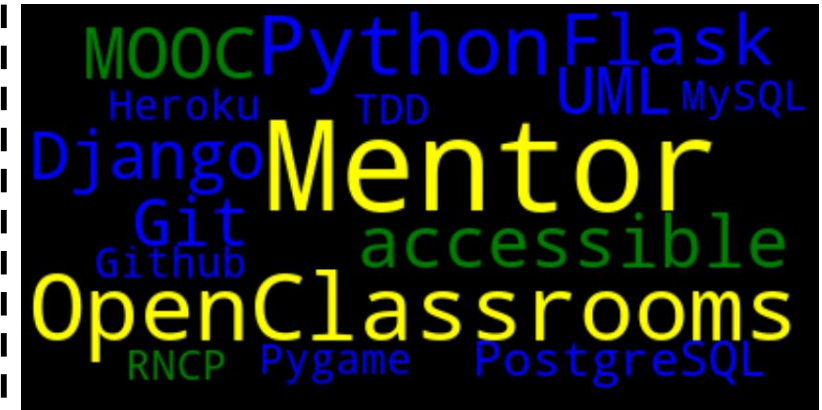
Pourquoi ce sujet de conférence?



2010

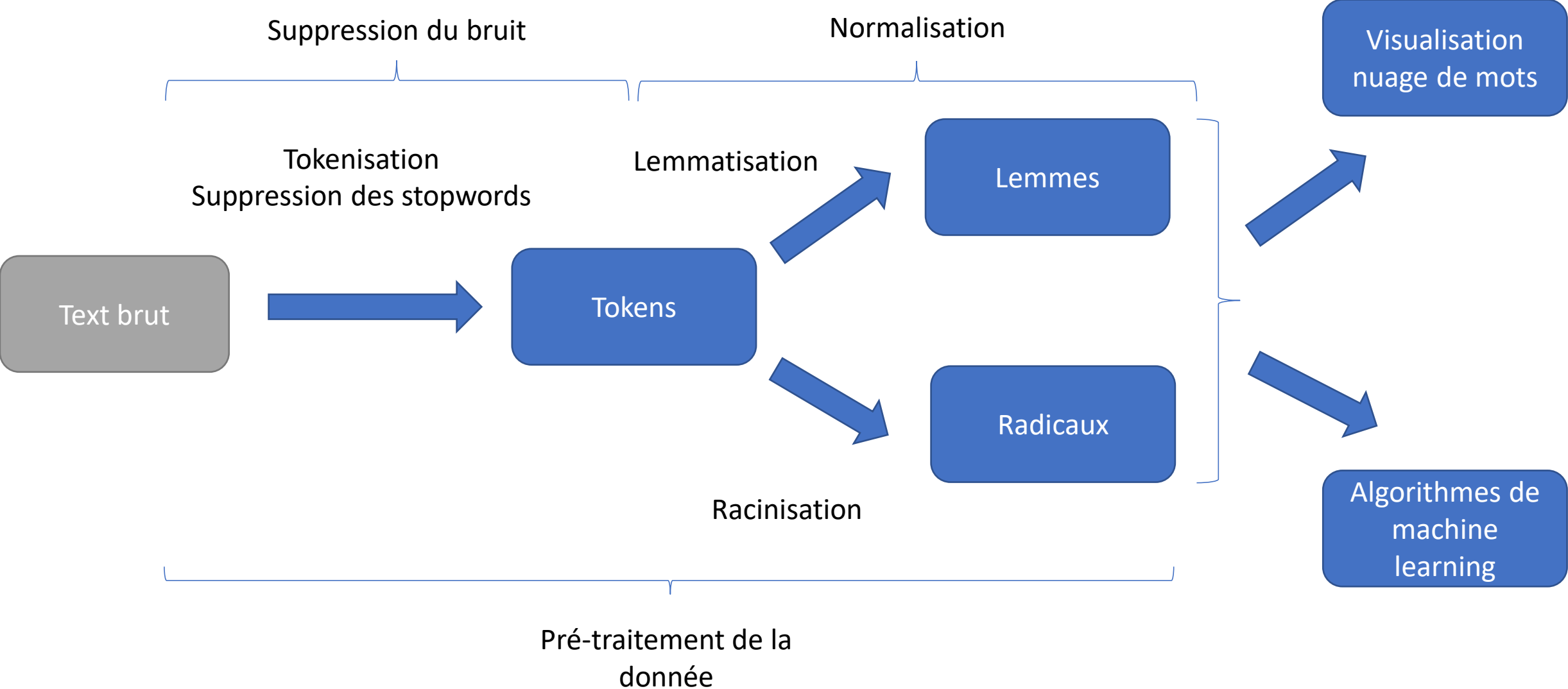


2018



La structure du code d'analyse

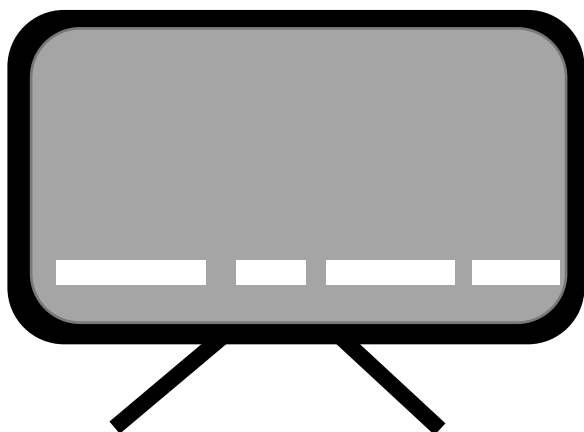
Les étapes classiques d'une analyse du langage naturel



Collecte des données

Analyse du langage naturel

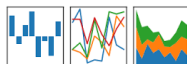
Visualisation et analyse



Fichiers SRT

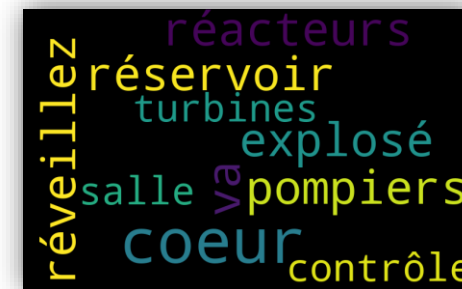
re, pandas

pandas
 $y_{it} = \beta^T x_{it} + \mu_i + \epsilon_{it}$



DataFrame

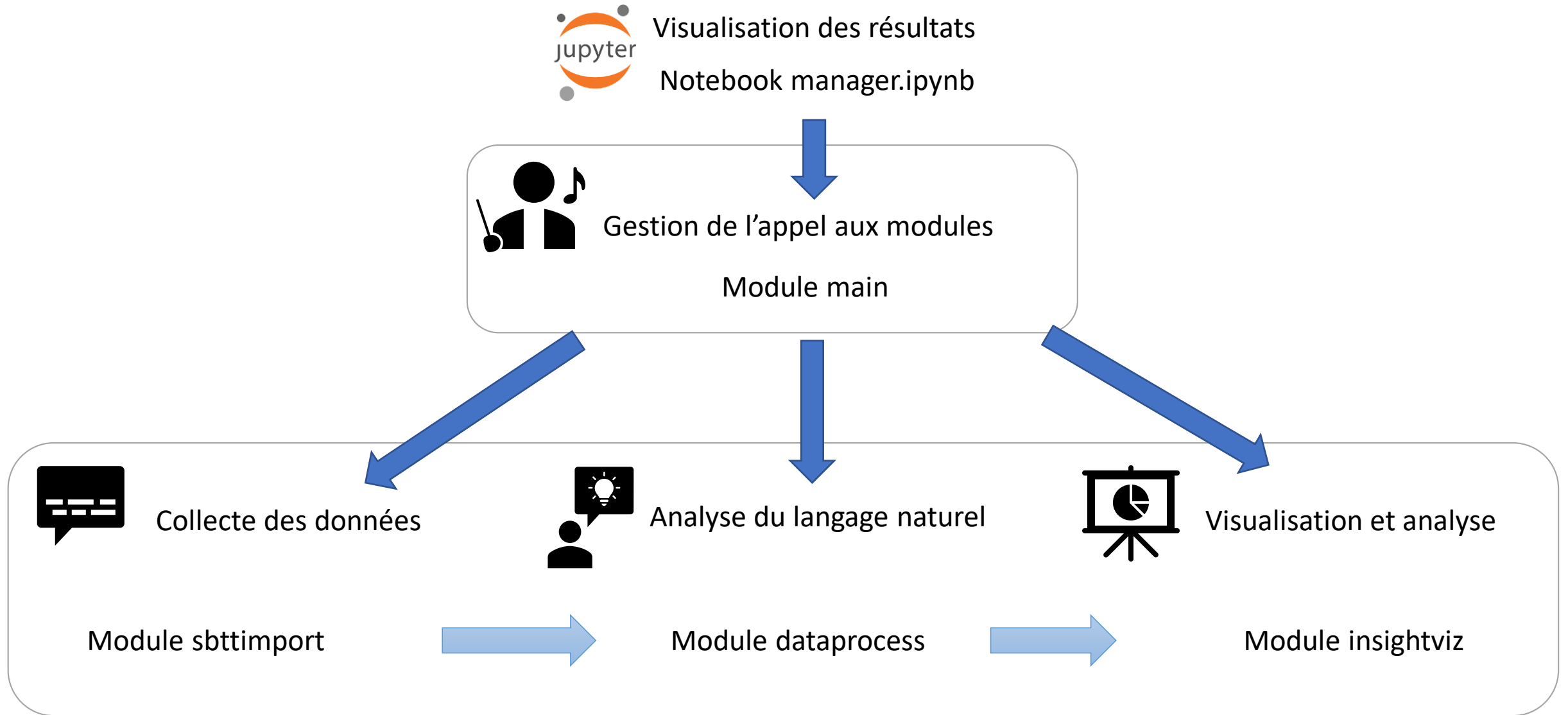
Pandas, sklearn, nltk



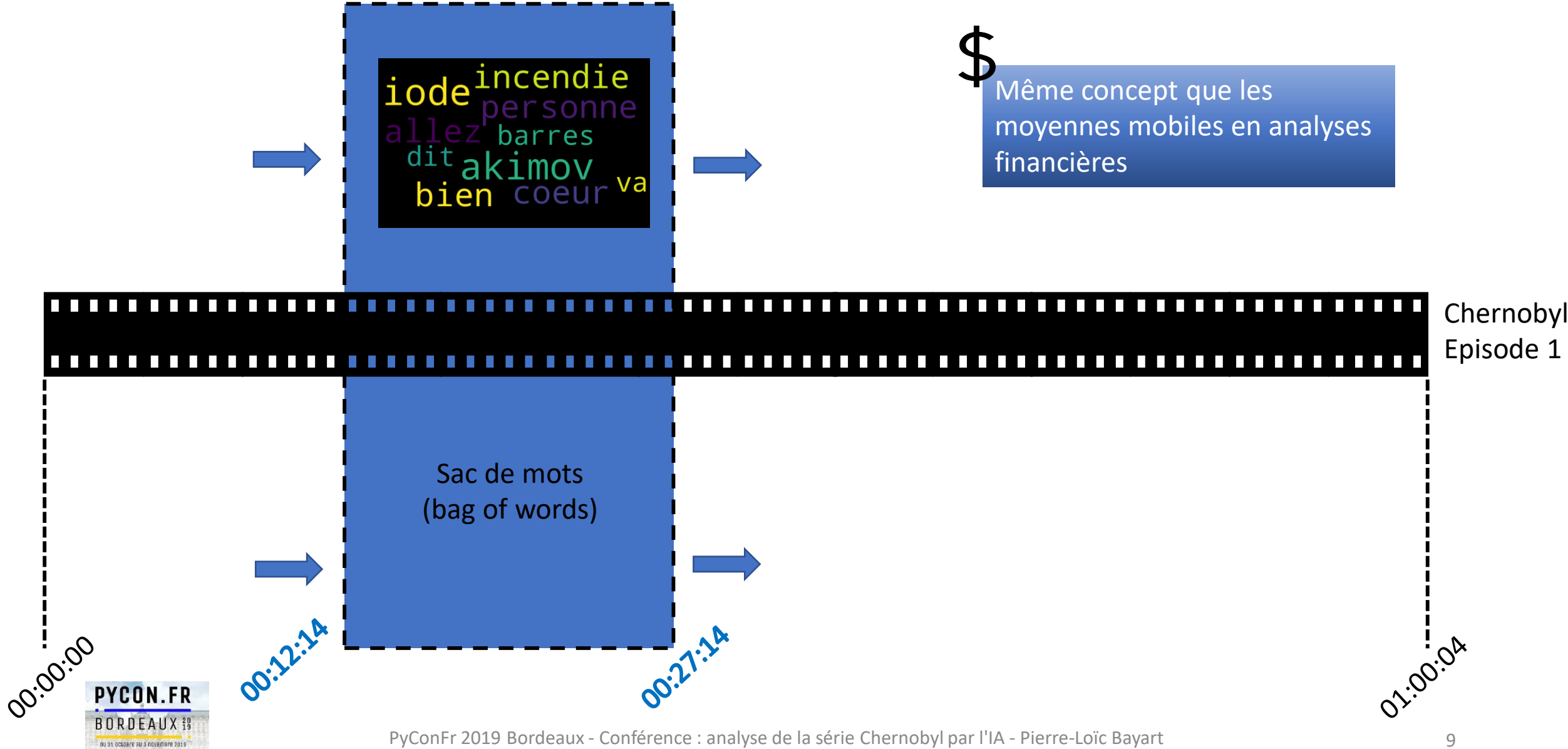
Fichier .avi, .png

Wordcloud, glob, cv2

Architecture du code



Fenêtre d'analyse du vocabulaire pour la génération des fichiers vidéos

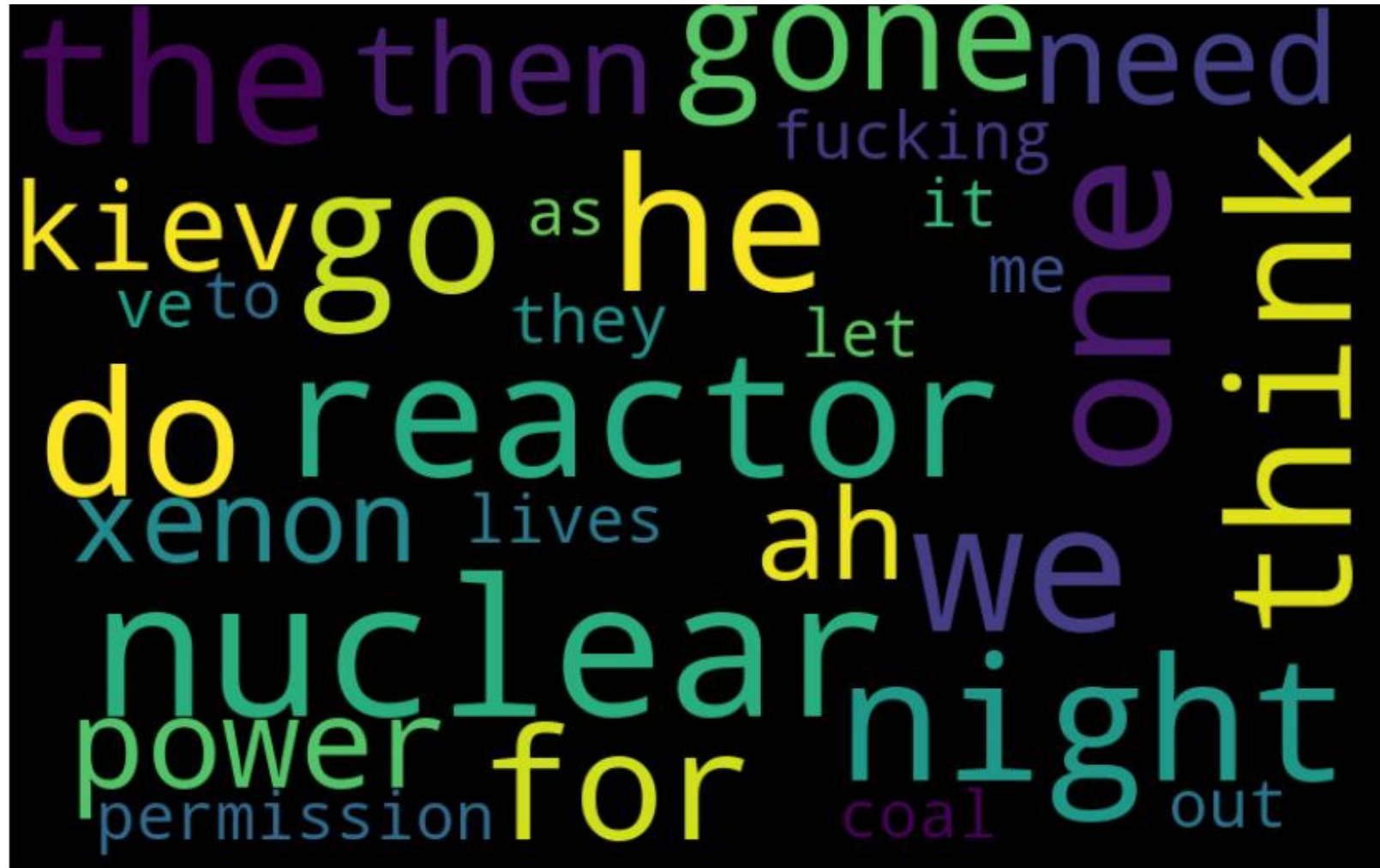


Les résultats

histoires
était
vérité avait
diatlov plus
importe mensonges
coeur explosé

Leonid Toptounov

Ingénieur en charge du contrôle du
réacteur 4 de Tchernobyl



N-grams

Analyse littéraire

Algorithmes plus
évolués de NLP

Analyse d'autres
séries/films

Affinage des
hyperparamètres



Merci pour votre attention

Des questions?

Site support de la conférence : <https://pierre-loic.github.io>

Repository du code de la conférence : https://github.com/Pierre-Loic/Chernobyl_analysis