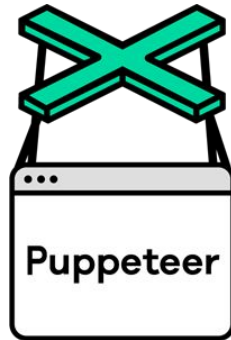


# Gestion automatique de Cookies

Maxence Neus

Tuteurs :  
Jeremie Dequidt - Naif Mehanna - Walter Rudametkin

# Contexte



*Inria*

### Quantcast

We value your privacy

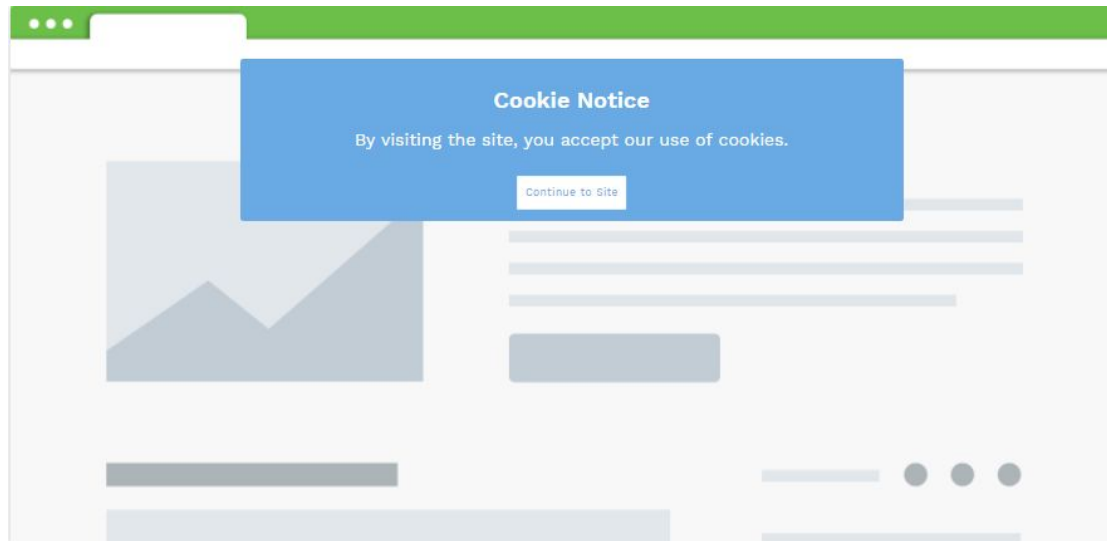
We and our [partners](#) store and/or access information on a device, such as cookies and process personal data, such as unique identifiers and standard information sent by a device for personalised ads and content, ad and content measurement, and audience insights, as well as to develop and improve products. With your permission we and our partners may use precise geolocation data and identification through device scanning. You may click to consent to our and our partners' processing as described above. Alternatively you may access more detailed information and change your preferences before consenting or to refuse consenting.

Please note that some processing of your personal data may not require your consent, but you have a right to object to such processing. Your preferences will apply to this website only. You can change your preferences at any time by returning to this site or visit our [privacy policy](#).

[MORE OPTIONS](#)[AGREE](#)

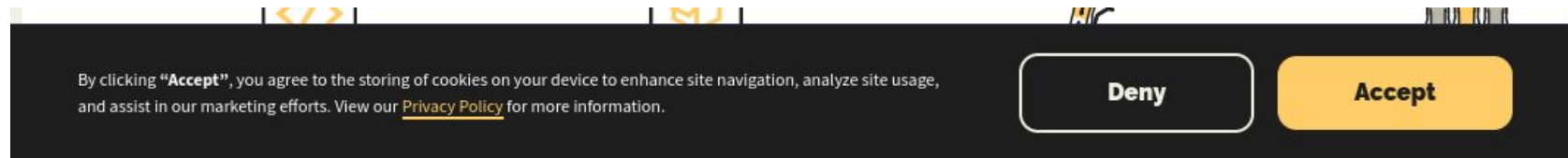


# Les bannières de cookies





# Les bannières de cookies





# Les bannières de cookies

## Quantcast

### We value your privacy

We and our partners store or access information on devices, such as cookies and process personal data, such as unique identifiers and standard information sent by a device for the purposes described below. You may click to consent to our and our partners' processing for such purposes. Alternatively, you may click to refuse to consent, or access more detailed information and change your preferences before consenting. Your preferences will apply to this website only. Please note that some processing of your personal data may not require your consent, but you have a right to object to such processing. You can change your preferences at any time by returning to this site or visit our [privacy policy](#).

[REJECT ALL](#)[ACCEPT ALL](#)

Store and/or access information on a device	OFF >
Select basic ads	OFF >
Create a personalised ads profile	OFF >
Select personalised ads	OFF >

[PARTNERS](#)[LEGITIMATE INTEREST](#)[SAVE & EXIT](#)



## Vue d'ensemble

**Détection**

**Parsing**

**Interaction**

# Parsing

## Quantcast

### We value your privacy

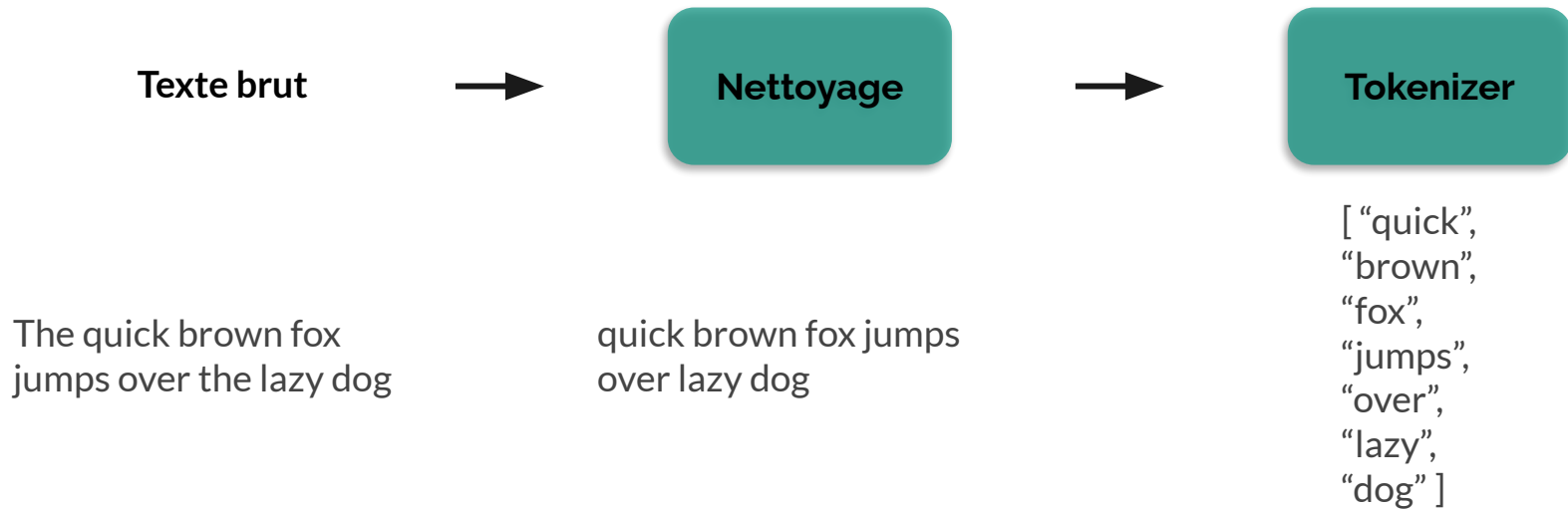
We and our partners store or access information on devices, such as cookies and process personal data, such as unique identifiers and standard information sent by a device for the purposes described below. You may click to consent to our and our partners' processing for such purposes. Alternatively, you may click to refuse to consent, or access more detailed information and change your preferences before consenting. Your preferences will apply to this website only. Please note that some processing of your personal data may not require your consent, but you have a right to object to such processing. You can change your preferences at any time by returning to this site or visit our [privacy policy](#).

[REJECT ALL](#)[ACCEPT ALL](#)

Store and/or access information on a device	OFF >
Select basic ads	OFF >
Create a personalised ads profile	OFF >
Select personalised ads	OFF >

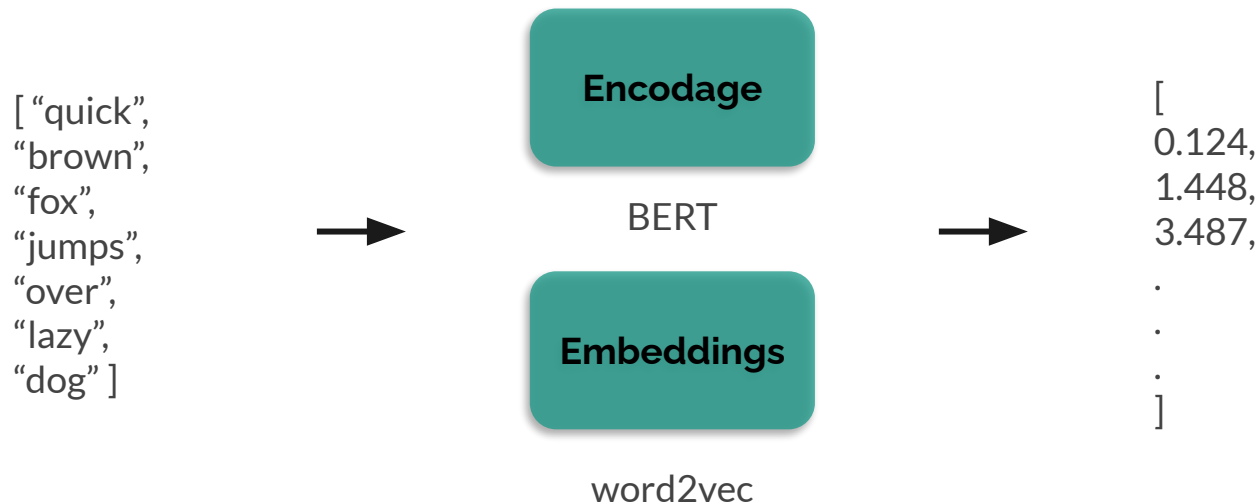
[PARTNERS](#)[LEGITIMATE INTEREST](#)[SAVE & EXIT](#)

# Natural Language Processing (NLP)





# Natural Language Processing (NLP)



# Collecte de données

Tranco  
Top 1M sites les  
plus visités

1,google.com  
2,gtld-servers.net  
3,youtube.com  
4,facebook.com  
5,microsoft.com  
6,akamaiedge.net  
7,netflix.com  
8,twitter.com  
...





## Détection d'une bannière sur la page

- AdBlockers avec 'Easylist cookies'
  - Difficile d'identifier si un élément a été bloqué
  - Si il y a blocage de la requête HTML la bannière n'est jamais affichée
  - Syntaxe des listes de filtres difficile à traduire en JS efficace



## Détection d'une bannière sur la page

- Extension 'I Don't Care About Cookies'
  - Trois catégories de traitement :

**Blocage de  
la requête**

**Injection de  
JS**

**Injection de  
CSS**



## Détection d'une bannière sur la page

- Extension 'I Don't Care About Cookies'
  - Trois catégories de traitement :

**Blocage de  
la requête**

**Injection de  
JS**

**Injection de  
CSS**



# Pre-Processing

- Supprimer les 'faux positifs'
- Extraire les morceaux de texte à classifier
- Classifier à la main les éléments du dataset



# Classes de cookies

Catégories proposées par Consent-O-matic:

- Préférences et fonctionnalité
- Performance et analytics
- Stockage et accès à l'information
- Sélection, transmission et reporting de contenu
- Sélection, transmission et reporting de pubs
- Autres



# **Merci de votre attention**

Place aux questions