

# Projet de Fin d'Etudes – Promo 2022

## Geolocalized Web Usage Mining

nicolas.durand@univ-amu.fr  
mohamed.quafafou@univ-amu.fr

Option : InSI

---

### Description :

Le projet consiste à développer une application web d'analyse de fichiers logs (fichiers journaux), par exemple, de serveurs Web. L'originalité est d'introduire la localisation de différents éléments analysés afin de visualiser les résultats sur une carte.

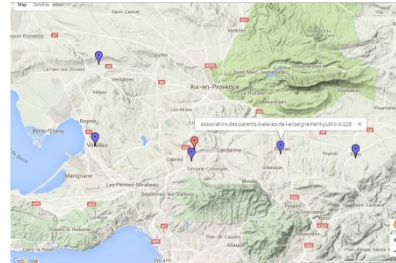


Figure 1. Exemple de visualisation des résultats

**Mots clé :** application web, fouille de données (data mining), fichiers journaux (logs), géolocalisation, visualisation.

---

### Problématique :

L'objectif du projet est de développer une application web d'analyse de logs et de visualisation des résultats.

L'analyse de fichiers journaux (logs) permet, en appliquant des techniques de fouille de données (data mining), de découvrir des motifs intéressants pour comprendre et améliorer les sites et les applications Web. On peut par exemple étudier la structure Web ou les profils d'utilisateurs dans le but d'améliorer l'organisation interne du site ou de détecter les intérêts des utilisateurs. Les retombées sont notamment stratégiques : augmentation des achats pour les boutiques en ligne, amélioration des recherches effectuées par les utilisateurs, ...

Parmi les techniques de fouilles de données utilisées, nous pouvons citer l'extraction de motifs (ensembles fréquents ou séquences fréquentes). Afin de fournir une visualisation des résultats sur une carte, nous introduisons la géolocalisation de différentes informations (adresses IP, adresses de lieux, ...). Un ensemble de filtres ou de contraintes seront utilisés pour faciliter la visualisation (zones géographiques, termes particuliers, ...).

Les langages et les technologies qui seront utilisés sont à discuter.

### Tâches :

Les principales tâches sont :

- 1) définition de l'architecture générale de l'application web,
- 2) prétraitement des logs et extraction des motifs (à l'aide d'outils existants),
- 3) géoréférencement des données,
- 4) visualisation des résultats (motifs) sur une carte,
- 5) mise en place de filtres pour une meilleure visualisation des résultats.

### D'autres informations (utiles) :

- Rapport de stage de M2 Informatique
  - Code source d'un précédent projet
-