

Projet Recherche Documentaire

Enseignant : Bernard Levrat

Groupe : Dorian Coffinet, Thibault Gauthier, Yassine Badih

- 
- I) Organisation du projet
 - II) Application
 - III) Tests et Améliorations
 - IV) Démonstration

Organisation du projet

Outils de développement

- Unix/Ubuntu
- Java 7
- Git/GitHub

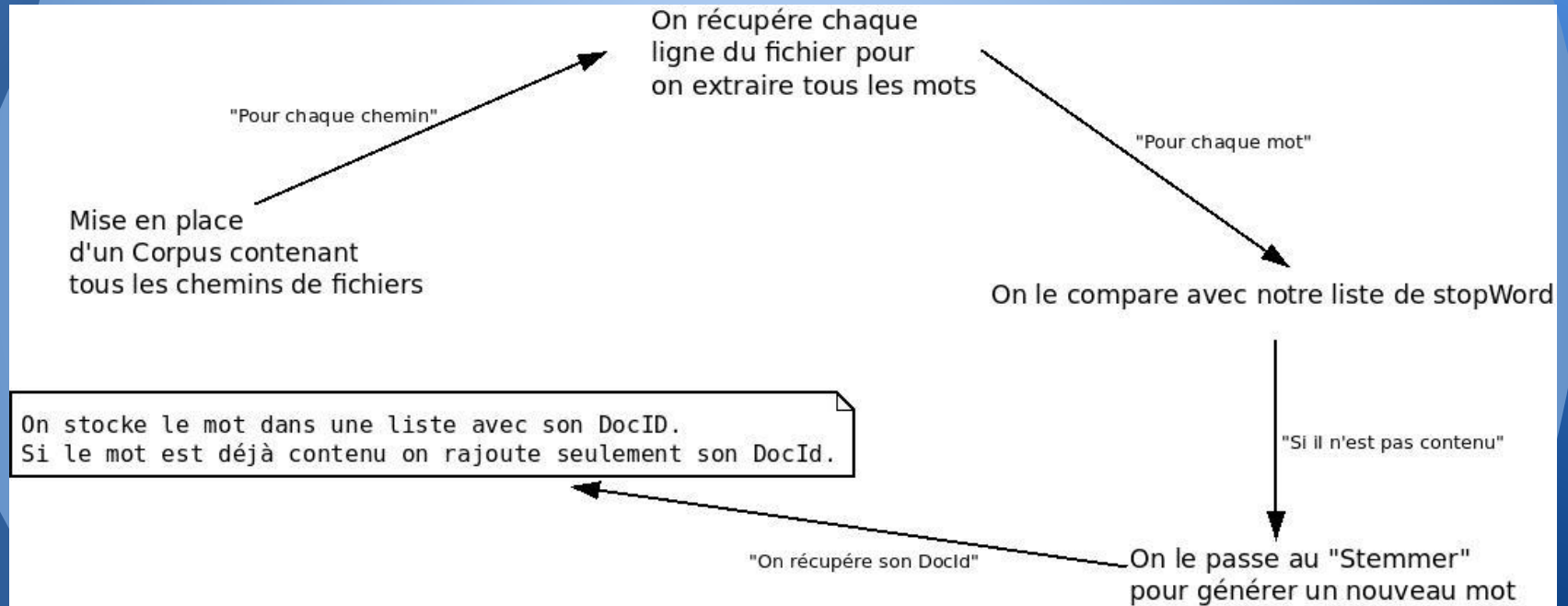


Découpage en phases

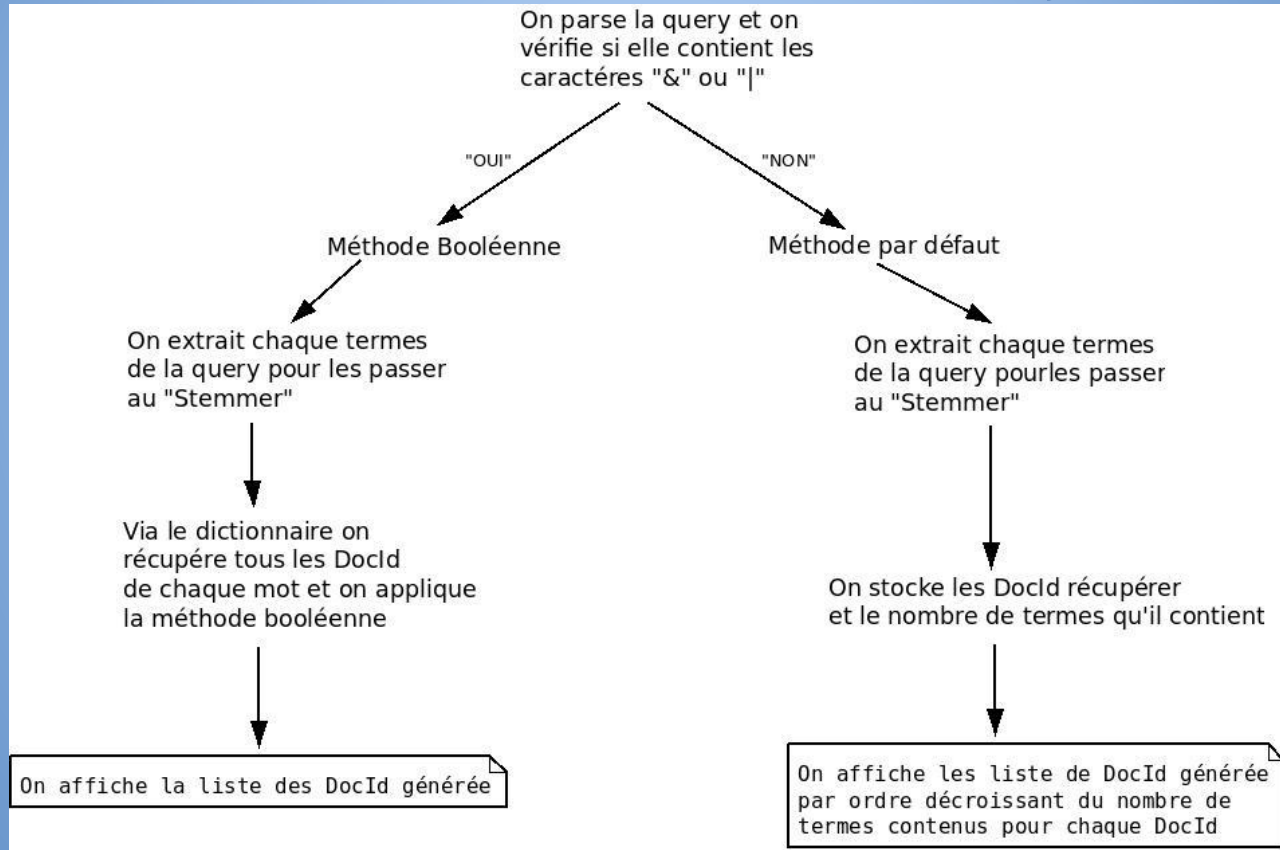
1. Lecture des fichiers
2. “Stemming” pour un fichier donné
3. Gestion du dictionnaire
4. Gestion de la “Query” utilisateur
5. Interface graphique
6. JavaDoc

Application

Génération du dictionnaire



Traitement d'une Query



Tests et Améliorations

Tests

Corpus de 364 documents = ~ 250Mo

Dictionnaire 128383 entrées = ~15Mo

Temps de création du dictionnaire = ~20
sec

Temps de réponse d'une Query =
~instantané

Améliorations

Prendre en compte la négation

Prendre en compte des Bi-words ou expressions

Passage au modèle vectoriel

Démonstration