

Departement Informatique
Université Caddi Ayad, Faculté de Science Smlalia

Indexation d'un fichier XML

Amri Habiba
Sabri Zineb

13 janvier 2017



Introduction

Problématique

Solution

Outils

Objectifs

Démonstration

Conclusion

References



Le partitionnement de données



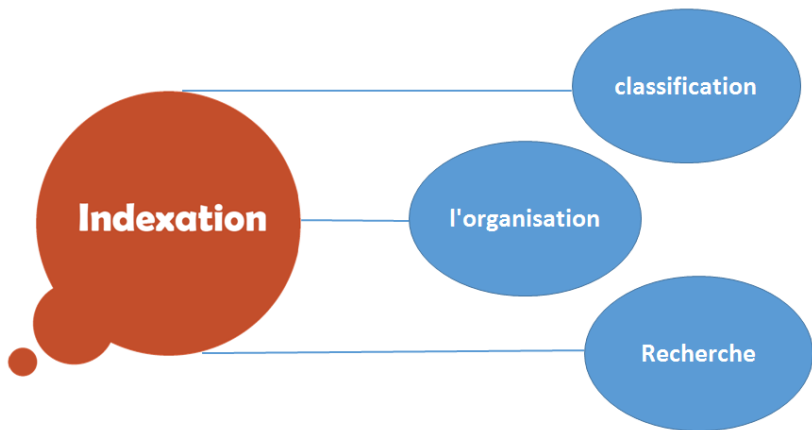
La Quantification



L'extraction



Indexation



A large orange speech bubble with a smaller orange circle at its tail. The word "LUCENE" is written in white, bold, uppercase letters inside the bubble.

LUCENE



Historique

Lucene est d'abord mis en téléchargement par Doug Cutting sur le site SourceForge.net en mars 2000. Son transfert vers Apache Jakarta est annoncé en octobre 2001.

Lucene

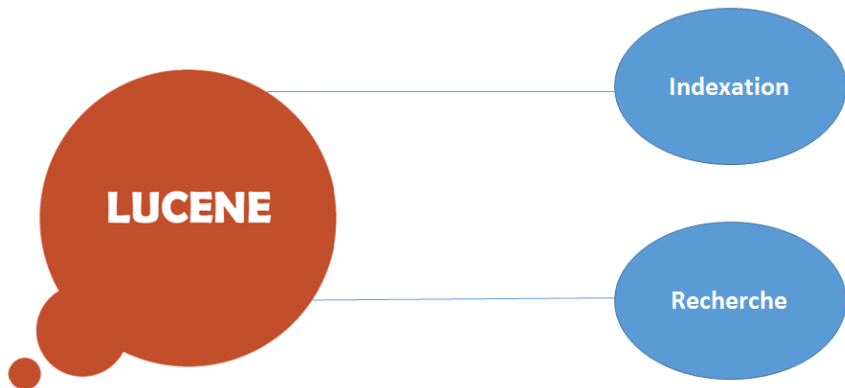
Lucene est une bibliothèque open source écrite en Java qui permet d'indexer et de chercher du texte. Il est utilisé dans certains moteurs de recherche.



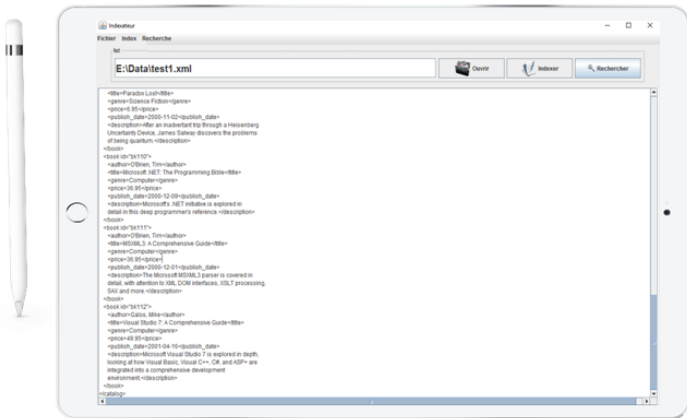
Apache Lucene en bref

- ▶ Langage :Java
- ▶ Interfaces Disponibles : Lucene a été porté dans de multiples langages,dont C, C++, Python, Ruby, Perl, Lisp, PHP,etc
- ▶ Plateformes supportées : Toute plateforme disposant d'une machine virtuelle Oracle Java ou OpenJDK en version 7 et supérieures
- ▶ Licence : Apache 2.0
- ▶ Principaux utilisateurs : Airbus, Apple, Disney, IBM, LinkedIn, Twitter (dans une version modifiée)

Source : JDN



Démonstration





- ▶ Forte communauté, avec des mises à jour fréquentes
- ▶ Faible empreinte mémoire, de nombreux projets open source qui viennent compléter le moteur lui-même
- ▶ Lien vers le Big Data (notamment le projet Blur qui tisse un lien entre Lucene et Hadoop)



- [1] <https://fr.wikipedia.org/wiki/Lucene>
- [2] <http://lucene.apache.org/>
- [3] <http://www.torrefacteurjava.fr/content/>
- [4] <https://wiki.apache.org/lucene-java/>
- [5] <http://www.w3ii.com/fr/lucene/default.html>
- [6] <https://wiki.apache.org/lucene-java/>
- [7] <http://www.w3ii.com/fr/lucene/default.html>
- [8] <http://blog.soat.fr/2010/09/comment-utiliser-lucene-dans-vos-applications/>
- [9] <http://www.journaldunet.com/solutions/saas-logiciel/5-moteurs-de-recherche-open-source/apache-lucene.shtml>
- [10] M. McCandless, E. Hatcher, and O. Gospodnetić.
Livre : *Lucene In Action*

A decorative graphic on the right side of the slide. It features a large, glowing sphere in the center, surrounded by several curved, wavy lines in shades of blue and white. The lines appear to be flowing or radiating from the sphere, creating a sense of motion and energy. The overall style is clean and modern, with a focus on light and color.

Merci pour votre attention ;) ! Questions ?