



**Université Chouaïb Doukkali**  
**Faculté des Sciences, Département d'informatique**  
**El JADIDA**

**Master BIBDA**  
**Information Retrieval**

---

**Thème**

---

**CONSTRUCTION D'UN SYSTEME DE RECHERCHE**  
**D'INFORMATION (ENGIN DE RECHERCHE)**  
**SIMPLIFIE**

---

**RÉALISÉ PAR : DOUAIOUI Hamza**

**Année universitaire : 2020/2021**

## Introduction

Un système de recherche d'information (RI) est un système qui permet de retrouver les documents pertinents à une requête d'utilisateur, à partir d'une base de documents volumineuse. Ce travail a pour objectif la construction d'un système de recherche d'information (engin de recherche) simplifié. Ce système doit indexer les fichiers textes que l'utilisateur lui donne, construire des structures d'index et les utiliser pour retrouver les textes contenant les mots d'une requête de l'utilisateur, en utilisant une interface graphique pour permettre à l'utilisateur d'interagir avec le système.

## Système de recherche d'information

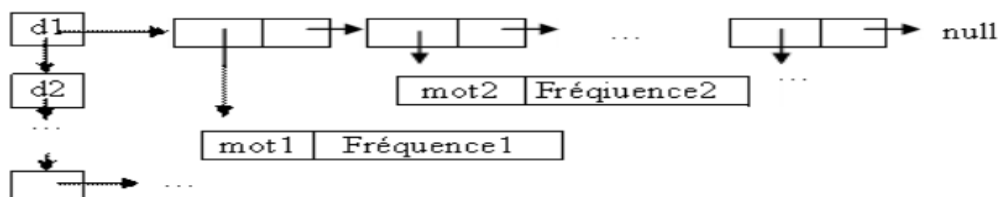
### 1. Langage de programmation

#### Java

Java est complètement orienté objet, il est puissant, également connu comme le "roi de la programmation". Il présente un certain nombre d'avantages pour la réalisation de ce travail à savoir la déclaration des classes, l'interface graphique et la gestion des événements...

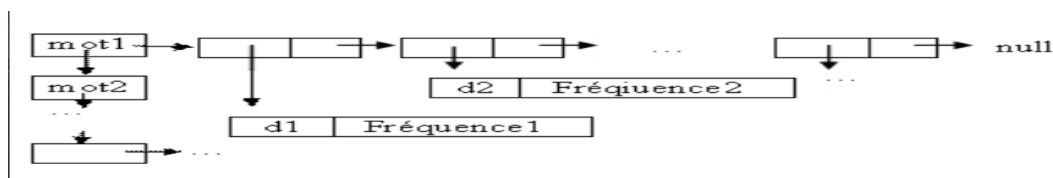
### 2. Indexation

Index de document : Pour représenter cette schéma, on a créé 3 classes, la première classe **Mot.java** qui contient le mot et leur fréquence, la deuxième classes **Nœud.java** qui contient Mot et un attribut Nœud 'next' pour pointer sur le Nœud suivant, et la troisième classe **Liste.java** qui contient tous les nœuds.



La Classe **Index.java** a pour objectif, l'indexation du document donné et la récupération d'une TreeMap qui contient tous les documents et les listes qui contiennent les mots et leurs fréquences.

#### Fichier inversé :



Pour représenter cette schéma on fait la même chose que l'index de document, on a créé 3 classes, la première classe **Document.java** qui contient le nom du document et la fréquence

du mot, la deuxième classes **NœudDoc.java** qui contient Document et un attribut NœudDoc next pour pointer sur le NœudDoc suivant, et la troisième classe **ListeDoc.java** qui contient tous les nœudDoc.

Pour créer le fichier inverse, on a créé une classe **Inverse.java** qui prend la TreeMap indexé et génère une TreeMap qui contient tous les mots et les listes qui contiennent les documents et les fréquences.

### La recherche d'une requête

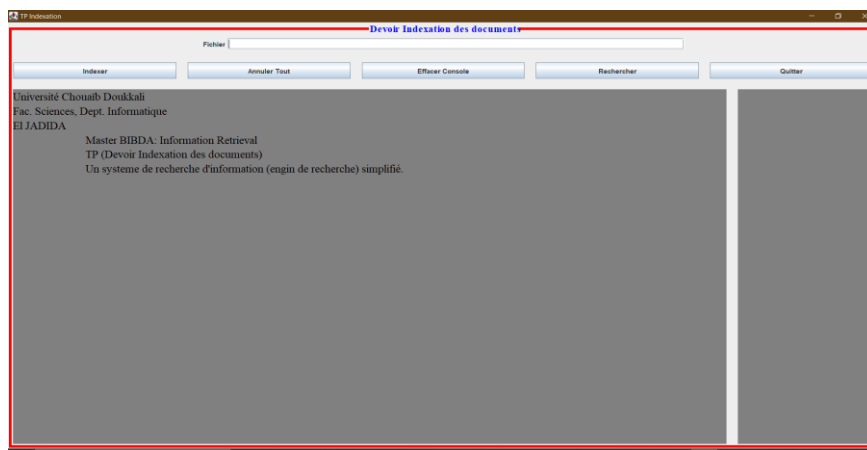
La classe **Action.java** contient la fonction de recherche qui prend comme arguments le fichier inversé et la requête et renvoie la liste des documents résultats.

### L'interface graphique :

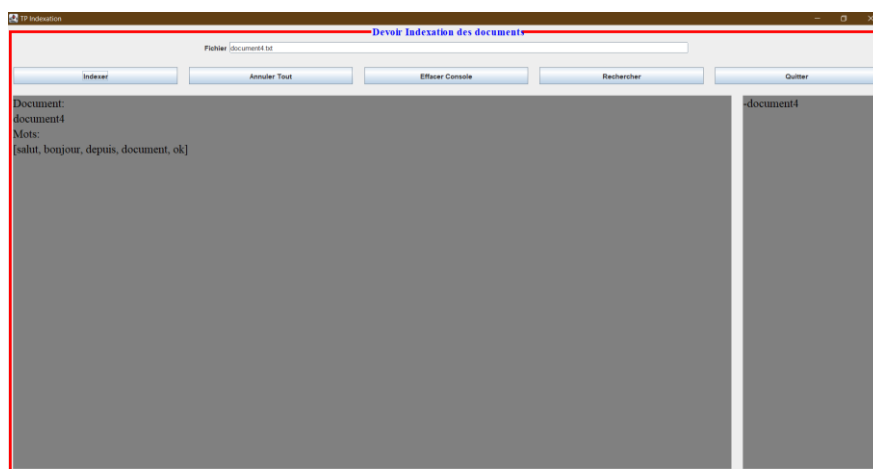
La classes **Interface.java** représente l'interface graphique qui permet à l'utilisateur d'interagir avec le système.

## 3. Illustrations des résultats

### La fenêtre principale.



### Le résultat du bouton indexer.



Ce screen représente la schéma demandé de 'index de document'

```
*****from index *****
Document3 : (bin,1)--->(bonjourno,1)--->(habe,1)--->(ich,2)--->(keine,1)--->(lolo,1)--->(mamamia,1)--->(probleme,1)--->>null
document1 : (bonjour,1)--->(douaioui,1)--->(hamza,1)--->(je,1)--->(suis,1)--->>null
document2 : (bibda,1)--->(douaioui,1)--->(hamza,1)--->(je,1)--->(master,1)--->(salut,1)--->(suis,1)--->>null
document4 : (bonjour,1)--->(depuis,1)--->(document,1)--->(ok,1)--->(salut,1)--->>null
```

Ce screen représente la schéma demandé de 'index de document'

```
*****From inverse *****
bibda : (document2,1)----->>null
bin : (Document3,1)----->>null
bonjour : (document1,1)----->(document4,1)----->>null
bonjourno : (Document3,1)----->>null
depuis : (document4,1)----->>null
document : (document4,1)----->>null
douaioui : (document1,1)----->(document2,1)----->>null
habe : (Document3,1)----->>null
hamza : (document1,1)----->(document2,1)----->>null
ich : (Document3,2)----->>null
je : (document1,1)----->(document2,1)----->>null
keine : (Document3,1)----->>null
lolo : (Document3,1)----->>null
mamamia : (Document3,1)----->>null
master : (document2,1)----->>null
ok : (document4,1)----->>null
probleme : (Document3,1)----->>null
salut : (document2,1)----->(document4,1)----->>null
suis : (document1,1)----->(document2,1)----->>null
```

Le résultat de la requête 'hamza douaioui'

