LO17 – Indexation et recherche d'information Performance du système

Phases du projet

Performance

Joints positifs

Indexation:

- Analyse/Extraction
- Normalisation
- Fichiers inverses

Recherche:

- Création de la correction syntaxique
- Création de la grammaire

IHM:

- Exécution des requêtes SQL
- Création de l'IHM/Usage de Tomcat

<u>Indexation:</u>

- Perl
- XML

Recherche:

- Antlr
- Java
- SQL

IHM:

- Bootstrap Twitter

echnologies

Améliorations

Points négatifs

- Tomcat
- HTML
- Javascript

- Question
 d'énumération/de listage
- Jointure d'un nombre infini de tables
- Possibilité de retourner des pages, des articles ou les deux
- Filtre : contenu, auteur, date, thème
- Interface web sexy!

- Amélioration du système de correction (phonétique, etc.)
- Gestion des questions booléennes
- Possibilité de retourner des dates, auteurs, etc.
- Laisser le choix de la correction syntaxique à l'utilisateur

- Apprentissage du Perl
- Connaissance de base en indexation, algorithme de correction et grammaire
- Génération des systèmes complets permettant de voir toutes les phases du projet

- Problème avec Tomcat
- Très difficile de travailler depuis chez soi

LO17 – Indexation et recherche d'information Présentation du système

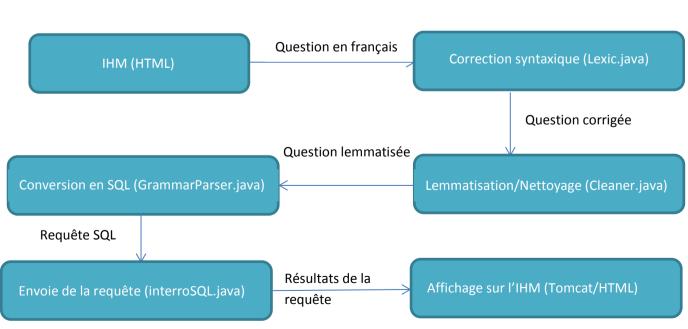
Corpus LCI
Analyse/extraction des informations importantes

Corpus extraits
Normalisation du corpus
Corpus normalisé

Corpus Stoplist
Lemmatisation

Fichier inverses

Création de la base de donnée



Aucun post-traitement n'est nécessaire

Recherche d'informations