

Développement d'un outil d'analyse et indexation semi-automatique de documents multimédias

Problématique

Le projet MediaTools a été proposé dans le but de développer un outil de partage des documents multimédias, et ce pour faciliter l'indexation et la recherche de documents tel que les fichiers texte, image, vidéo.

Ces documents ont certains points communs entre eux mais en possèdent d'autres qui leur sont propres. Nous devons donc établir un schéma pour définir toutes les spécificités de nos documents, et s'y rapporter lors du développement de la structure.

Il nous faut également développer une interface graphique pour l'utilisateur qui lui servira de guide afin de trouver et de soumettre sa requête. Elle devra être simple et claire.

But du travail de diplôme

Le but de ce projet est de développer un portail web pour l'analyse et l'indexation semi-automatique des documents multimédias afin de permettre leur classification et leur recherche dans un système complet de partage de documents multimédias.

Ce portail devra permettre une description adaptée selon chaque type de documents. Certaines descriptions (telle que le titre ou les mots clés) sont communes aux documents tandis que d'autres (la décomposition d'une vidéo) sont liées aux types de documents.

Un schéma de descriptions des documents multimédias est conçu et représenté via un schéma XML.

Le portail web offrira des interfaces de description et de recherche adaptées à chaque type de documents. Ces interfaces s'occuperont de la création et l'interrogation des instances XML.

Architecture utilisée

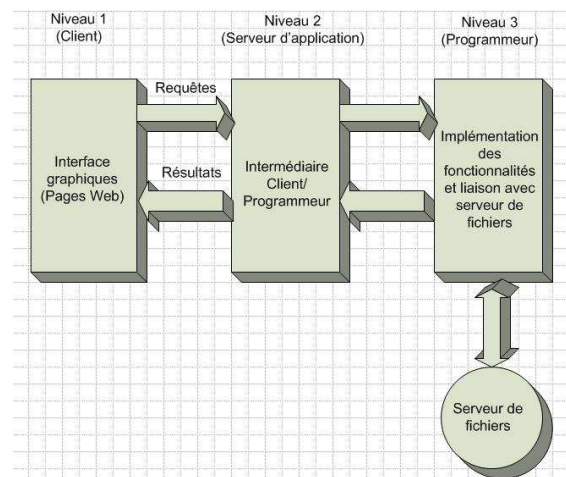
Le modèle utilisé pour cette implémentation est un modèle 3 tiers. Il est basé sur 3 interfaces.

Une première pour l'utilisateur, elle s'occupe de l'aspect graphique et des vues données à ce dernier.

La seconde représente le serveur. Nous utilisons Tomcat pour le déploiement de notre outil. Celui-ci gère la transition entre la vue précédente (utilisateurs) et la suivante (programmeurs).

Et finalement la troisième, concerne la vue du programmeur. Celle-ci s'occupe de toute l'implémentation :

- La recherche (simple et avancée)
- L'indexation : chargement du fichier source, stockage dans le serveur de fichiers, création des instances XML en utilisant un parser Dom qui stocke toutes les informations récupérées dans un fichier XML et ensuite sur le serveur de fichiers.



Auteur: Rajaa MOUJID
Répondant externe: Mme. Fatemi Nastaran
Prof. responsable: Mme. Fatemi Nastaran
Sujet proposé par: HES-SO et HEIG-VD/IIC