

Moteur de recherche

Module CM3133 - MRM

1

Présentation

Benoît Terradillos
Développeur indépendant
benoit.terradillos@heig-vd.ch

2

Objectifs

Comprendre les enjeux de l'indexation des données.
Différencier les métadonnées et les données multimédia.
Comprendre l'importance de l'identification unique et universelle des données multimédia (GUID/UUID).

3

Objectifs

Expérimenter l'installation et la configuration d'un moteur de recherche multimédia.
Intégrer un moteur de recherche multimédia dans un projet web.

4

Contenu du cours

Préparation par groupes de délivrables médias de différents types (audio, texte, vidéo, image).

Association de métadonnées à ces délivrable, par l'intermédiaire de fichier XML descriptifs.

Contenu du cours

Présentation du projet Solr, moteur de recherche flexible et puissant.

Apprentissage de la syntaxe des requêtes Lucene, utilisée par Solr.

Contenu du cours

Installation et configuration de Solr

Type des champs et fonctionnement de l'indexation.

Contenu du cours

Préparation d'un schéma de données commun à tous les délivrables médias.

Transformation des fichiers XML (via XSLT) vers le format d'importation commun.

Importation des métadonnées dans le moteur de recherche.

Contenu du cours

Interface du moteur de recherche Solr avec un projet web.

Exemples en Javascript (via JSON), php, etc.

Exemples de fonctions avancées de Solr.

9

Évaluations

19 décembre - en groupe & individuel : création du modèle et des délivrables média

9 janvier - en groupe & individuel : création de la XSLT pour l'indexation des données multimédia

20 janvier - individuel : requêtes Lucene

30 janvier - en groupe : interface de recherche

10
