

Solr - Implémenter un moteur de recherche scalable

3 j (21 heures)

Ref : SOLR

Public

Architectes et développeurs

Pré-requis

Avoir une expérience de développement informatique

Moyens pédagogiques

Formation réalisée en présentiel ou à distance selon la formule retenue
Exposés, cas pratiques, synthèse, assistance post-formation pendant trois mois
Un poste par stagiaire, vidéoprojecteur, support de cours fourni à chaque stagiaire

Modalités de suivi et d'évaluation

Feuille de présence émargée par demi-journée par les stagiaires et le formateur
Exercices de mise en pratique ou quiz de connaissances tout au long de la formation permettant de mesurer la progression des stagiaires
Questionnaire d'évaluation de la satisfaction en fin de stage
Auto-évaluation des acquis de la formation par les stagiaires
Attestation de fin de formation

Objectifs

- Démarrer un projet de recherche d'information
- Modéliser les unités d'indexation
- Développer les services de recherche
- Analyser les performances de votre moteur de recherche
- Déployer le moteur de recherche suivant différentes topologies

Programme détaillé

FONDAMENTAUX DE LA RECHERCHE D'INFORMATION (RI)

- Définitions
- RI Web vs RI Entreprise
- Concepts de base
- Structure et construction de l'index

Solr - Implémenter un moteur de recherche scalable

Modèle booléenne de recherche d'information

Recherche ordonnée

Modèle vectoriel de recherche d'information

INDEXATION DU WEB - UN ETAT DE L'ART

Historique de l'innovation des principaux moteurs de recherche

Organisation des documents du Web

Construction du dictionnaire des termes

Stockage de l'index

Répondre à une requête de l'utilisateur

Mise à l'échelle du moteur de recherche

Le cas Google Search Engine

SOLUTIONS OPEN SOURCE LUCENE / SOLR

Présentation de la librairie Apache Lucene

Présentation du serveur Apache Solr

INDEXATION AVEC SOLR

Structure du répertoire d'installation de Solr

Comprendre le concept Solr Core

Les fichiers de configuration

Schéma des documents et analyse des textes

Les modes de communication avec Solr

Le framework Data Import Handler (DIH) de Solr

Indexer les fichiers avec Solr Cell

RECHERCHE AVEC SOLR

Les paramètres de recherche

La syntaxe des requêtes

Parseur de requête Lucene vs parseur de requête Dismax

Recherche géospatiale

Influencer la pertinence des résultats

Recherche par facettes pour une meilleure expérience de l'utilisateur

Les composants Highlight et MoreLikeThis

Aller au-delà de la recherche par mots-clés

MISE A L'ECHELLE DE SOLR

Evaluer les performances de Solr avec SolrMeter

Optimiser une instance unique de Solr

Passer à plusieurs serveurs avec Solr Replication et/ou Solr Cloud

