

Big Data - Moteurs de recherche

## Solr - Implémenter un moteur de recherche scalable

3 jours (21h00) | ★★★★★ 4,6/5 | SOLR-IMPL | Évaluation qualitative de fin de stage | Formation délivrée en présentiel ou distanciel <sup>(1)</sup>

Formations Informatique > Big Data > Big Data - Moteurs de recherche



### À l'issue de ce stage vous serez capable de :

- 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.

### Niveau requis

Avoir une expérience de développement informatique.

### Public concerné

Architectes et développeurs.

### Cette formation :

- Est animée par un consultant-formateur dont les compétences techniques, professionnelles et pédagogiques ont été validées par des diplômes et/ou testées et approuvées par l'éditeur et/ou par M2i Formation
- Bénéficie d'un suivi de son exécution par une feuille de présence émarginée par demi-journée par les stagiaires et le formateur.

#### (1) Modalité et moyens pédagogique :

Formation délivrée en présentiel ou distanciel \* (e-learning, classe virtuelle, présentiel à distance). Le formateur alterne entre méthodes \*\* démonstrative, interrogative et active (via des travaux pratiques et/ou des mises en situation). La validation des acquis peut se faire via des études de cas, des quiz et/ou une certification.

Les moyens pédagogiques mis en oeuvre (variables suivant les formations) sont : ordinateurs Mac ou PC (sauf pour les cours de l'offre Management), connexion internet fibre, tableau blanc ou paperboard, vidéoprojecteur ou écran tactile interactif (pour le distanciel). Environnements de formation installés sur les postes de travail ou en ligne. Supports de cours et exercices.

\* Nous consulter pour la faisabilité en distanciel. \*\* Ratio variable selon le cours suivi.

# Programme

## Fondamentaux de la Recherche d'Information (RI)

- Définitions
- RI Web vs RI Entreprise
- concepts de base
- Structure et construction de l'index
- Modèle booléenne de recherche d'information
- Recherche ordonnée
- Modèle vectoriel de recherche d'information

## Indexation du Web - Un état 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 à l'échelle 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

## Modalités d'évaluation des acquis

L'évaluation des acquis se fait :

- En cours de formation, par des études de cas ou des travaux pratiques
- Et, en fin de formation, par un questionnaire d'auto-évaluation ou une certification (M2i ou éditeur)