



ITgate

Training

Your Gateway to Excellence

# Formation Moteur de recherche Lucene-SolR, prise en main

## Objectifs de la formation Lucene Solr

SolR est un moteur de recherche Open Source largement déployé en entreprise, dans de très grands comptes. Basé sur le composant d'indexation Lucene, il permet de mettre en place des solutions d'index de recherche performantes, incluant des fonctionnalités avancées comme la navigation par facettes, l'autocomplétion, la recherche approchante ou la gestion des synonymes.

Ces index forment la base de portails de recherche documentaires, et peuvent remplacer les bases de données dans des solutions NoSQL. SolR peut s'intégrer avec de nombreux CMS comme Drupal ou Typo3.

Cette formation Lucene Solr couvre à la fois les aspects installation et paramétrage de SolR, et les aspects d'intégration et la syntaxe des requêtes à SolR.

Comment installer SolR ?

Quels sont ses principaux éléments de configuration ?

Quels types de champs sont disponibles dans un index, quand et comment les utiliser ?

Comment ajouter une autocomplétion ?

Une recherche par facettes ?

Comment paramétrer SolR pour indexer une base de données SQL ?

Quels sont les opérateurs de recherche disponibles dans SolR ?

Comment paramétrer le moteur pour répondre à des recherches spécifiques ?

Comment demander un affichage de facettes ?

Comment ajouter du surlignement dans les résultats de recherche ?

Quelles bibliothèques sont disponibles dans les différents langages pour interagir avec SolR ?

## À qui s'adresse cette formation ?

**Public :**

---

**Capital Social:** 50000 DT **MF:** 1425253/M/A/M/000 **RC:** B91211472015  
**Tél. / Fax.:** +216 73362 100 **Email:** contact@itgate-training.com **Web:** www.itgate-training.com  
**Adresse :** 12 Rue Abdelkadeur Daghri - Hammam Sousse 4011 – Tunisie



ITgate

Training

Your Gateway to Excellence

Ce cours Lucene Solr s'adresse aux développeurs ou chefs de projet impliqués dans un projet de développement de portail de recherche documentaire.

### **Prérequis :**

Connaissance de XML obligatoire. Une connaissance de Java est un plus mais n'est pas obligatoire. Notions de développement et/ou d'administration d'applications.

## **Contenu du cours Lucene Solr**

### **Introduction à SolR**

Présentation : Lucene, une librairie d'index, SolR, une encapsulation serveur de Lucene  
Comment installer SolR ? quels sont ses principaux fichiers de configuration ? Une première indexation à l'aide des données d'exemple Utilisation des interfaces d'administration et de navigation : SolR admin, Solr browse, Solr Explorer

### **L'index de recherche**

Qu'est-ce qu'un schéma d'index ? comment le spécifier ? Le fichier schema.xml : types de champs et options L'analyse de texte dans un champ : tokenizer, filtres, stemming, stopwords, synonymes, analyse phonétique

### **Indexer des données dans SolR**

Le format XML de mise à jour d'un document SolR : ajouter, modifier, supprimer des entrées  
Utiliser SolR Cell pour indexer des documents plein-texte (pdf, word)

### **Rechercher dans SolR depuis une application tierce**

Qu'est-ce qu'un Request Handler ? Quel est le format d'un résultat de recherche SolR ? La syntaxe et les différents opérateurs de requête Comment associer des poids aux différents champs de l'index ? Comment améliorer la pertinence des résultats de recherche ?

### **Les facettes**

Comment paramétrer et interroger des facettes sur une valeur numérique ? une date ? Comment paramétrer un interroger des facettes sur une valeur texte ?

## **Indexer une base relationnelle**

Le DataImport Handler pour indexer une base relationnelle Faire une indexation incrémentale

## **Vue sur quelques autres composants de recherche**

La surbrillance des résultats de recherche L'autocomplétion "More Like This" : proposer des résultats similaires

## **Travaux Pratiques**

Toutes les étapes de la formation sont illustrées en manipulant et configurant un serveur SolR version 4, sur des données d'exemple.