

Formation Elasticsearch, Indexation de contenu

Présentation

Devenu extrêmement populaire depuis sa sortie en 2010, Elasticsearch est un moteur de recherche open source puissant permettant d'indexer et de rechercher des données.

Cette formation certifiante vous permettra d'apprendre à développer avec Elasticsearch.

Durée : 14,00 heures (2 jours)

Tarif INTRA : Nous consulter

Objectifs de la formation

- Découvrir les enjeux et les concepts clés d'Elasticsearch dans la recherche en plein texte
- Monter des indices et indexer des documents
- Maîtriser les requêtes DSL et les agrégations pour la recherche en plein texte
- Appliquer les scripts et les boosts pour personnaliser les résultats de recherche
- Utiliser les fonctionnalités avancées pour répondre aux besoins courants : auto-complétion, géolocalisation
- Bien choisir la relation des documents dans la production
- Savoir administrer et configurer les clusters
- Coupler Java à Elasticsearch via l'API client

Prérequis

- Connaissances solides en administration système Unix/Linux.
- Connaissances basiques en SGBD, JSON et REST.

Public

- Ce cours est destiné aux futurs développeurs d'applications Big Data intégrant Elasticsearch comme moteur de recherche.

Programme de la formation



Concepts de base

- Concept du moteur de recherche
- Introduction à Elasticsearch
- Concepts basics d'Elasticsearch
- Installation et configuration

Indexation de documents

- Vue générale de l'API REST
- CRUD API : Index, Get, Exist, Delete, Update, Multi Get, Bulk
- Le versionage des documents

Elasticsearch vs Lucene : les valeurs ajoutées

- Index inversé
- Trace de transaction
- Segment

Analyse de texte

- Le besoin d'une analyse de texte
- Anatomie d'un Analyzer
- Analyze API
- Configuration d'un Analyzer
- Les Tokenizers (Parseurs en Token)
- Filtre de Tokens
- Filtre de Synonyme
- ICU Tokenizers & FiltersSQL avec Spark

Mapping

- Les fondamentaux du Mapping
- Mapping API
- Les différents attributs et types de champs

Recherche de documents

- Recherche de documents

Spark – SQL découverte

- Les dataframes
- Chargement de données avec un schéma
- Agrégation
- Jointure
- SQL avec Spark

Spark – streaming, recherche avancée

- Agrégation : Scope
- Agrégation métrique
- Agrégation bucket
- Pertinence
- Modèle d'espace vectoriel
- TF/IDF
- Booster la requête
- Scoring personnalisé
- Highlighting
- Suggestion

Relation des documents

- Problématique
- Dénormalisation
- Nested Object :
 - Mapping
 - Requête, Filtre, Agrégations
- Parent / Child :
 - Mapping
 - Requête, Filtre, Agrégations

Fonctionnalités avancées

- Percolator
- Geolocation

Vers la production

- Configuration d'un cluster

- Scalabilité et volumétrie de données
- Bonnes pratiques matérielles
- Supervision

Java et Elasticsearch

- Requête DSL
- Requête full-text
- Requête aux termes
- Requête composée
- BitSet, Filtre et Lucene
- Pagination
- Ordonnancement
- Types de recherche
- Warmer Api

Organisation

Formateur

Les formateurs de Docaposte Institute sont des experts de leur domaine, disposant d'une expérience terrain qu'ils enrichissent continuellement. Leurs connaissances techniques et pédagogiques sont rigoureusement validées en amont par nos référents internes.

Moyens pédagogiques et techniques

- Alternance d'apports théoriques et de mises en pratique contextualisées.
- Progression pédagogique structurée, de la compréhension à la mise en œuvre opérationnelle.
- Pédagogie immersive favorisant une mise en application immédiate des compétences.
- Formation interactive et participative s'appuyant sur des ateliers pratiques, études de cas et échanges collectifs (+ 60 % du temps de formation.)
- Séquences pédagogiques rythmées et évaluations formatives régulières pour renforcer l'ancrage des apprentissages.
- Supports de cours, ressources documentaires et références mis à disposition par le formateur

Dispositif de suivi de l'exécution et de l'évaluation des résultats de la formation

En amont de la formation *(évaluation en ligne dans l'extranet apprenant)*

- Recueil des besoins et des attentes des participants.
- Auto-positionnement permettant d'évaluer le niveau initial des apprenants.

Tout au long de la formation

- Évaluations continues réalisées par le formateur travers des questions orales, exercices, QCM, études de cas et mises en situation pratiques.

À l'issue de la formation *(évaluations en ligne dans l'extranet apprenant)*

- Auto-positionnement afin de mesurer l'acquisition des compétences visées.
- Validation de l'acquisition des compétences par le formateur
- Questionnaire de satisfaction « à chaud » à l'issue de la formation
- Questionnaire de satisfaction "à froid", 45 jours après la formation afin d'évaluer l'ancrage des acquis et leur mise en application en situation professionnelle.

NB : dans le cadre d'une Action collective, chaque stagiaire bénéficiaire sera contacté par un prestataire choisi par l'Opco Atlas afin d'évaluer « à chaud » la qualité de la formation suivie.

Accessibilité

Nos formations peuvent être adaptées à certaines conditions de handicap. Nous contacter pour toute information et demande spécifique.