

TP6 : Recherche

Il est conseillé d'installer Kibana pour profiter de sa Dev Console et donc effectuer les requêtes plus facilement

6.1 Requêtes DSL

Effectuer les requêtes DSL sur la base documentaire :

- Document PDF triés par date
- Documents dont le champ *content* répond à « administration »
- Documents dont le champ *content* ou *title* répond à « administration »
- Documents dont le champ *content* ou *title* répond à « administration » et dont la date de création est comprise dans un intervalle
- Documents PDF dont le champ *content* ou *title* répond à « administration » et dont la date de création est comprise dans un intervalle
- Documents dont le champ *content* répond à « Administration » ou à « Oracle »
- Documents dont le champ *content* répond à « Administration » et éventuellement à « Oracle »

6.2 Contrôle du score

Effectuer des recherches avancées avec le paramètre *explain* :

- En utilisant le boosting, récupérer les Documents dont le champ *content* ou *title* répond à « administration ». Les documents ayant « administration » dans le champ *title* apparaissant en premier
- Même requête en utilisant des clauses *should*, comparer les scores
- Utiliser un autre mode de calcul avec *dis_max*

6.3 Matching partiel

- Préparer un nouvel index qui utilise une indexation multiple pour le champ *attachment.title* :
 - Type *keyword*
 - Type *text* avec standard analyseur
 - Type *text* avec un analyseur utilisant *edge_ngram*
- Effectuer des requêtes de matching partiel en utilisant un des 3 champs du mapping. Comparer les résultats et les temps de réponse

6.4 Phrases

Effectuer des requêtes avec des phrases :

- Récupérer les documents contenant la phrase « framework java », autoriser une distance de 5 mots
- Récupérer les documents dont le titre commence par « administration j »

6.5 Fuzzy, Langage naturel

- Effectuer des recherches fuzzy avec des fautes de typo
- Préparer un nouvel index avec un filtre phonétique, effectuer des recherches avec des fautes d'orthographe

6.5 Surbrillance

Effectuer les requêtes précédentes en ajoutant la surbrillance sur le champ content (limiter la taille du fragment)