

CRAFT – COMPORTEMENT DE LA RECHERCHE

Emetteurs : Serwan Cravic, Patrick Demellier Date : 19/04/2022

Public : Utilisateurs de l'extranet, des widgets et de l'API

Il y a actuellement deux modes de recherche qui peuvent se combiner dans Craft :

- Recherche par mot clé (le champ de recherche extranet/widget, l'entrée mot clé de la recherche avancée, l'entrée keywords des API)
- Filtre par critère (les filtres et contraintes de la recherche avancée, l'entrée criteria des API, les filtres/légendes des widgets)

Quel que soit le mode utilisé, les résultats de recherche sont régis par un principe de match et de score.

Grands principes

Vu le nombre de champs existant sur une entité et la diversité de leur nature, nous utilisons le mode « most_fields » d'ElasticSearch.

<https://www.elastic.co/guide/en/elasticsearch/reference/7.x/query-dsl-query-string-query.html#query-string-multi-field-parms>

Recherche par mot clé

- Rechercher le mot clé étoile (« * ») revient à rechercher tous les caractères possibles, donc toutes les entités ou éléments recherchés.
- S'il n'y a qu'un seul mot une étoile est ajoutée à la fin du mot.
 - Exemple : la recherche de **pole** cherche le mot clé pole*, soit « pole » suivi éventuellement de n'importe quelle chaîne de caractères
- L'utilisateur peut lui-même mettre des étoiles où il veut.
- L'utilisateur peut utiliser les "" pour effectuer une recherche de la phrase.
 - Exemple : la recherche de **"pole"** cherche strictement le mot « pole », c'est-à-dire précédé et suivi d'un espace, d'une ponctuation ou d'une fin de phrase.
- Les espaces sont interprétés comme des ET inclusifs à l'intérieur d'un même champ (il faut que les mots soient obligatoirement présents dans le même champ).
- Il est possible d'utiliser AND/et/ET, OR/ou/OU explicitement, dans ce contexte l'opérateur est appliqué entre tous les champs (pour l'exemple du ET, il faut que chacun des mots soit présent dans au moins l'un des champs).

Informations fournies par Conjecto, document diffusable par BDI aux utilisateurs concernés. 1/2

- Les majuscules et les accents sont recherchés avec et sans.
 - Exemple : chercher le mot **Santé** équivaut à chercher **Sante, sante, santé, SANTE...**
- Les recherches sont faites en français ou en anglais (avec les analyseurs correspondants, qui en plus de retirer les mots de liaison, utilisent des comparateurs de racine de mot).
- Les () sont interprétées. On peut ainsi définir des blocs de recherche.
 - Exemple : **Santé ET (Brest OU Morlaix)**. Attention : cette recherche n'est cependant pas judicieuse car « Brest » et « Morlaix » ne seront pas recherchés spécifiquement dans la ville, mais dans tous les champs. Passer par une recherche avancée (extranet) ou sur critères (API) est préférable dans ce cas.
- Les autres caractères ("&", "|", "." etc...) ne sont pas interprétés.

Filtre par critères

Les filtres sont appliqués strictement, si le filtre utilise un champ texte et demande la présence de mots, la recherche par mot clé s'applique sur ce champ.

Dans l'API, l'utilisation de parenthèses pour séparer en blocs de conditions les critères recherchés est possible. Voir la documentation correspondante.

Calcul du score

Pour calculer ce score, Elasticsearch va s'appuyer sur plusieurs critères :

- La fréquence du terme recherché dans le document. Plus le terme est fréquent, plus son poids sera élevé.
- La fréquence inverse du terme à travers tous les documents. Plus le terme est fréquent, moins il aura de poids.
- La longueur du champ. Plus le champ est grand, plus le poids sera faible ; inversement, plus le champ est petit, plus le poids sera élevé.
- Contrairement à l'algèbre booléenne, pour la recherche par mot clé, l'ordre a un poids dans le score
 - Exemple : **mer terre** cherche les deux mots dans l'un des champs. Si l'ordre d'apparition de ces mots dans le champ est mer puis terre, il aura plus de poids que si l'ordre est terre puis mer.

<https://www.elastic.co/guide/en/elasticsearch/reference/7.x/query-dsl-query-string-query.html>

<https://www.elastic.co/guide/en/elasticsearch/reference/7.x/query-dsl-query-string-query.html#query-string-top-level-params>