

Principe du projet : Efrei / Markethings

Réaliser un code qui doit être capable de comprendre un prompt (donc une phrase en français) sur des demandes avec des informations légales et de produire une liste de champs permettant l'édition d'une commande mysql avec des AND et des OR.

En résumé, remplacer les checkbox par une analyse des phrases.

Commencer par la fonction d'identification de la présence des critères ci-dessous avec des pourcentages de probabilité pour chacune d'elles. Et si présent transcoder la demande dans la grammaire mysql précisée ci-dessous.

Localisation

Code_postal = '92500'
Departement = '23'
Region = 'Ile-de-France'
Donc presence d'un besoin de localisation

Activité

Libelle_secteur = 'Administration publique'
Activite_entreprise = '2512Z'
Donc detection d'une demande en lien avec l'activité ou pas

Tailles d'entreprise

Tranche effectif ex plus de 10 salaires, entre 10 et 50 salariés mais aussi acronymes TPE, pme, eti, grand groupe

Critères financiers

CA_le_plus_recent = '100000'
Resultat_net_le_plus_recent = '10000'
Rentabilite_la_plus_recente = '5.5'

Critères juridiques (moins prioritaires)

Categorie_juridique = 'Société commerciale'
Idem presence ou pas
Siege_entreprise = 'oui' par défaut
Date_creation_entreprise = '2022-20-02'
Capital = '15000'
Date_de_changement_dirigeant = '2020-10-22'
Nombre_etablissements = '3'

Exemple de commande mysql:

```
select count(*) from table where Siege_entreprise = 'oui' AND
(Categorie_juridique = 'Société commerciale') AND ( Activite_entreprise =
'2511Z' OR Activite_entreprise = '2512Z' OR Activite_entreprise = '2561Z' OR
Activite_entreprise = '2812Z' OR Activite_entreprise = '2893Z' ) AND (
(Region = 'Ile-de-France') ) AND ( Tranche_effectif_entreprise = '50 a 99
salaries' )
```

Le travail doit intégrer une liste la plus grande possible de prompt avec un critère de qualité sur le résultat. Une analyse humaine des résultats pour indiquer les cas de vrai et faux positifs et négatifs est impérative.

Markethings fournira la liste des requêtes utilisateurs déjà collectées sur le SaaS et le moteur en page d'accueil

Critères de votre réponse

1. Solution envisagée avec ou sans recours à des API externes, ou algorithmes open source ?
2. Quel niveau de contrainte sur le coût de calcul ? (ex : recours à des GPU dans le cloud)
3. Décrire la MVP (minimum value product) envisagée : quelle fonctionnalités
4. Tests envisagés (une liste des tests passés est impérative)
5. Temps de reponse maximum envisagé...
6. Indiquer si le travail intègre (ou pas) la création d'un bot qui permet de valider progressivement le besoin de l'utilisateur
7. Indiquer votre organisation projet : rapporteur ? Reporting ? Kpi projet ? Couverture des tests ?
8. Modalités d'aide à la création, debug, tests du code avec les ChatGPT, Claude, cursor,...
9. Indiquez si parmi vous certains ont déjà travaillé sur des projets proches de la demande (LLM / bot...)

L'appréciation finale intégrera **avec un fort coefficient** la livraison d'un code fonctionnel déployable dans une fonction beta. En clair, faire simple au départ et améliorer ensuite car le temps est contraint.

Exemples de vues d'écran des parcours actuels

- Module étude de marché : <https://www.markethings.io/formulaire-comptage/>
- SaaS markethings

Choix de l'activité

1. SÉLECTIONNEZ LES ENTREPRISES PAR LEUR DOMAINÉ D'ACTIVITÉ

RECHERCHER PAR SECTEUR OU CODE NAF

OU

RECHERCHER PAR MOT CLEF (SITES WEB ET OBJETS SOCIAUX)

1. Saisir le code ou rentrer un mot-clé

2. Cocher les suggestions proposées

Cliquer dans l'espace sélection pour supprimer un choix

Secteur ou code NAF

- > 1 Culture et production animale, chasse et services annexes
- > 2 Sylviculture et exploitation forestière
- > 3 Pêche et aquaculture
 - 0311Z Pêche en mer
 - 0312Z Pêche en eau douce
 - 0321Z Aquaculture en mer
 - 0322Z Aquaculture en eau douce
- > 5 Extraction de houille et de lignite
- > 6 Extraction d'hydrocarbures

Votre sélection de Secteur / Code NAF

Important : l'absence de sélection de l'un des sous-menus (ex. Localisation, structures juridiques, données financières...) entraîne la sélection par défaut de l'ensemble du sous-menu. La modification des curseurs "Données financières" et "Autres filtres" inclut la sélection des bornes inférieures et supérieures.

LOCALISATION

- > France
- > Île-de-France
 - Essonne
 - Hauts-de-Seine
 - Paris
 - Seine-et-Marne
 - Seine-Saint-Denis
 - Val-de-Marne
 - Val-d'Oise
 - Yvelines
- > Auvergne-Rhône-Alpes
- > Bourgogne-Franche-Comté
- > Bretagne
- > Centre-Val de Loire
- > Grand Est
- > Hauts-de-France

DONNÉES FINANCIÈRES

Chiffre d'affaires

Pourcentage CA Export

Excédent Brut d'Exploitation

Résultat net

Rentabilité nette

Capital

Date changement capital

Sélectionner uniquement les entreprises avec un résultat net négatif ⓘ

AUTRES FILTRES

Nombre d'établissements (hors siège social)

Nombre d'années d'existence

Siège entreprise uniquement

Intégrer les sociétés en procédure collective

Uniquement les sociétés en procédure collective

Lots ultérieurs...

Evolution du service avec un agent capable d'apporter une réponse contextualisée à la demande de recherche, là où la recherche sémantique répond exclusivement au mot-clé recherché. Ex. je cherche des cabinets de conseil => pouvoir intégrer les agences conseil ou les sociétés de conseil.

Possibilité : réutiliser les popups des lots précédents pour interagir et orienter la réponse
Ex : je cherche des cabinets de conseil => ouverture d'une fenêtre proposant des options à valider