

## w.c.s. - Development #86527

### amélioration de la recherche texte libre de /api/formdefs

05 février 2024 10:29 - Thomas Noël

<b>Statut:</b>	Solution validée	<b>Début:</b>	05 février 2024
<b>Priorité:</b>	Normal	<b>Echéance:</b>	
<b>Assigné à:</b>	Pierre Ducroquet	<b>% réalisé:</b>	0%
<b>Catégorie:</b>		<b>Temps estimé:</b>	0:00 heure
<b>Version cible:</b>		<b>Planning:</b>	Non
<b>Patch proposed:</b>	Non		
<b>Description</b>			
Sur l'ajout d'une recherche sur les formdef <a href="https://git.entrouvert.org/entrouvert/wcs/pulls/707">#81532</a> on fait du "fts" classique dans PostgreSQL (je crois).			
On se retrouve avec des comportements pas toujours heureux, par exemple en cherchant pour des encombrants, « encomb » ne réponds rien alors que « encombr » si.			
Pierre, pourrais-tu analyser cette affaire ? (On peut parfaitement tout revoir au niveau de la table de recherche)			

#### Historique

##### #2 - 05 février 2024 12:40 - Pierre Ducroquet

Et effectivement, il faut revoir ça parce que ce que tu décris est bien une limitation du FTS classique de PostgreSQL (et un peu tout le monde). Sur un projet précédent, nous avons réglé ce problème avec la stratégie suivante (implémentée en triggers et procédures stockées, donc sans impact côté Python):

- extraction des mots depuis les blocs de texte, pour construction d'une unique table listant les mots utilisés
- à l'arrivée d'une requête, récupération des mots "conformes" depuis cette table à l'aide de la recherche par trigrammes de PG
- reconstruction d'une "vraie" requête tsquery depuis les mots conformes

Je vais creuser pour voir les problèmes que cette implémentation avait pu rencontrer, mais de mémoire c'était une révolution sur la qualité des résultats par rapport à du FTS simple ou du trigram simple.

##### #3 - 14 février 2024 19:25 - Frédéric Péters

- *Tracker changé de Bug à Development*
- *Sujet changé de la recherche sur /api/formdefs n'est pas efficace à amélioration de la recherche texte libre de /api/formdefs*

##### #4 - 28 février 2024 11:31 - Robot Gitea

- *Statut changé de Nouveau à En cours*

Pierre Ducroquet (pducroquet) a ouvert une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/wcs/pulls/1201>
- Titre : WIP: wip/86527-better-fts
- Modifications : <https://git.entrouvert.org/entrouvert/wcs/pulls/1201/files>

##### #5 - 04 mars 2024 12:21 - Robot Gitea

- *Statut changé de En cours à Solution proposée*

##### #6 - 04 mars 2024 19:12 - Robot Gitea

- *Statut changé de Solution proposée à En cours*

Frédéric Péters (fpeters) a relu et demandé des modifications sur une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/wcs/pulls/1201>

##### #7 - 05 mars 2024 11:49 - Robot Gitea

- *Statut changé de En cours à Solution proposée*

Pierre Ducroquet (pducroquet) a demandé une relecture de Frédéric Péters (fpeters) sur une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/wcs/pulls/1201>

**#8 - 26 mars 2024 18:04 - Robot Gitea**

- Statut changé de *Solution proposée* à *En cours*

Frédéric Péters (fpeters) a relu et demandé des modifications sur une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/wcs/pulls/1201>

**#9 - 29 mars 2024 14:12 - Robot Gitea**

Frédéric Péters (fpeters) a relu et demandé des modifications sur une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/wcs/pulls/1201>

**#10 - 14 avril 2024 11:31 - Robot Gitea**

Frédéric Péters (fpeters) a relu et demandé des modifications sur une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/wcs/pulls/1201>

**#11 - 17 avril 2024 11:12 - Robot Gitea**

- Statut changé de *En cours* à *Solution proposée*

Pierre Ducroquet (pducroquet) a demandé une relecture de Frédéric Péters (fpeters) sur une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/wcs/pulls/1201>

**#12 - 26 avril 2024 11:12 - Robot Gitea**

- Statut changé de *Solution proposée* à *Solution validée*

Frédéric Péters (fpeters) a approuvé une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/wcs/pulls/1201>