

## w.c.s. - Development #13061

### api de recherche d'utilisateurs: gérer les accents

05 septembre 2016 16:41 - Thomas Noël

<b>Statut:</b> Fermé	<b>Début:</b> 05 septembre 2016
<b>Priorité:</b> Normal	<b>Echéance:</b>
<b>Assigné à:</b>	<b>% réalisé:</b> 0%
<b>Catégorie:</b>	<b>Temps estimé:</b> 0:00 heure
<b>Version cible:</b>	<b>Planning:</b>
<b>Patch proposed:</b> Oui	
<b>Description</b> Actuellement l'API de recherche des utilisateurs impose une égalité de recherche, accents compris (noel et Noël ne sont pas équivalents). Il faudrait rendre le système de recherche plus souple et permettre de retourner les utilisateurs correspondants, accents ou non.  Piste : utilisation de l'extension pg_trgm <a href="https://www.postgresql.org/docs/9.1/static/pgtrgm.html">https://www.postgresql.org/docs/9.1/static/pgtrgm.html</a>	
<b>Demandes liées:</b> Lié à Publik - Development #14159: système de recherche central (utilisateurs... <b>Fermé</b> <b>29 novembre 2016</b>	

#### Révisions associées

##### Révision 1f1be03 - 05 avril 2017 17:03 - Frédéric Péters

general: search for users without looking at accents (#13061)

#### Historique

##### #1 - 05 septembre 2016 19:28 - Benjamin Dauvergne

C'est pas plutôt unaccent<sup>1</sup> qu'on voudrait utiliser ? Les trigrammes ou levensthein c'est plus pour de la recherche approchée.

<sup>1</sup><https://www.postgresql.org/docs/9.1/static/unaccent.html>

##### #2 - 05 septembre 2016 21:37 - Frédéric Péters

Thomas souhaitait également faire de la recherche approchée sur le nom.

##### #3 - 06 septembre 2016 11:49 - Benjamin Dauvergne

```
create extension unaccent;
create extension pg_trgm;
-- unaccent ne peut pas être utilisé directement dans un index parce que pas 'immutable'
CREATE OR REPLACE FUNCTION myunaccent(varchar) RETURNS varchar AS $$ select unaccent('unaccent',$1::text); $$
LANGUAGE 'sql' IMMUTABLE;
create table test (t text);
CREATE INDEX trgm_idx ON test USING gist (lower(myunaccent(t)) gist_trgm_ops);

insert INTO test (t) VALUES ('Thomas Noël'), ('Noël Graët'), ('Noëlle Pernat');

select t, similarity(t, 'noiel thomas') from test order by similarity(t, 'noiel thomas') desc limit 3;
 t          | similarity
-----+-----
 Thomas Noël |    0.5625
 Noël Graët  |  0.0909091
 Noëlle Pernat |    0.08
(3 lignes)
```

À tester avec une plus grosse liste de noms mais c'est déjà pas mal.

##### #5 - 06 septembre 2016 12:06 - Frédéric Péters

Avant de partir sur tout ça, il faut assurer que nos db sont dans les bonnes locales.

Genre :

```
wcs_demarches_alfortville_fr | wcs-au-quotidien | UTF8 | en_US.UTF-8 | en_US.UTF-8 |
```

(c'est connu de Thomas et moi mais je ne sais pas s'il y a un ticket).

**#6 - 29 novembre 2016 19:45 - Thomas Noël**

- Lié à *Development #14159: système de recherche central (utilisateurs, demandes, base de connaissance) ajouté*

**#7 - 09 mars 2017 11:41 - Frédéric Péters**

Proposition de passer outre postgresql et simplement lors de la création des chaines à indexer d'aussi en ajouter des versions sans accent. (et le faire pour la recherche fts sur les formdata aussi).

**#8 - 31 mars 2017 14:58 - Frédéric Péters**

- Fichier *0001-general-search-for-users-without-looking-at-accents-.patch* ajouté

- Statut changé de *Nouveau* à *En cours*

- Patch proposed changé de *Non* à *Oui*

**#9 - 05 avril 2017 16:15 - Thomas Noël**

Ack

**#10 - 05 avril 2017 17:04 - Frédéric Péters**

- Statut changé de *En cours* à *Résolu (à déployer)*

```
commit 1f1e035f7e20ef36c2ac216fa9087f9f15a264
Author: Frédéric Péters <fpeters@entrouvert.com>
Date: Fri Mar 31 14:42:38 2017 +0200
```

```
    general: search for users without looking at accents (#13061)
```

**#11 - 23 décembre 2018 14:52 - Frédéric Péters**

- Statut changé de *Résolu (à déployer)* à *Solution déployée*

**Fichiers**

---

0001-general-search-for-users-without-looking-at-accents-.patch	9,55 ko	31 mars 2017	Frédéric Péters
---	---------	--------------	-----------------