

w.c.s. - Development #60777

API stats, avoir des statistiques par formulaire, puis par champ à choix fermé

18 janvier 2022 15:07 - Benjamin Dauvergne

Statut:	Fermé	Début:	18 janvier 2022
Priorité:	Normal	Echéance:	
Assigné à:	Valentin Deniaud	% réalisé:	0%
Catégorie:		Temps estimé:	0:00 heure
Version cible:		Planning:	Non
Patch proposed:	Oui		
Description			
Pour équivalence avec les pages de statistiques actuelles. Il faut gérer pour chaque champ :			
<ul style="list-style-type: none">• filtre par une valeur• faire un GROUP BY sur ce champ			
Les champs actuellement gérés sont les case à cocher, liste à choix et liste à choix fermé, i.e. les sous-classes de Field ayant une méthode stats().			
PS: prendre en compte que pour les source de donnée avec un filtrage dynamique (qui dépend d'un autre champ via un critère, vue personnalisé ou source JSON paramétrable) il faut récupérer l'ensemble des valeurs (Cf. problème #60571 sur l'ancienne interface).			
Demandes liées:			
Lié à w.c.s. - Bug #60751: stats : les statistiques des champs listes à choix...		Rejeté	18 janvier 2022

Révisions associées

Révision 3117afcb - 16 février 2022 10:06 - Valentin Deniaud

statistics: filter forms count by fields (#60777)

Révision 913fabb5 - 16 février 2022 10:06 - Valentin Deniaud

statistics: group forms count by field (#60777)

Historique

#1 - 18 janvier 2022 15:07 - Benjamin Dauvergne

- Lié à Bug #60751: stats : les statistiques des champs listes à choix vers des sources de donnée dynamique sont généralement vides ajouté

#3 - 18 janvier 2022 15:09 - Benjamin Dauvergne

- Description mis à jour

#6 - 19 janvier 2022 15:24 - Benjamin Dauvergne

- Description mis à jour

#7 - 25 janvier 2022 18:19 - Valentin Deniaud

- Assigné à mis à Valentin Deniaud

#8 - 25 janvier 2022 18:19 - Valentin Deniaud

- Fichier 0001-wip.patch ajouté

- Statut changé de Nouveau à Solution proposée

- Patch proposed changé de Non à Oui

Je commence par les filtres sur une valeur, voilà ce que ça pourrait donner, avec [#61083](#) en face côté combo pour la gestion des sous-filtres.

Je me suis très fortement inspiré des méthodes FormPage.stats et FormPage.get_filter_sidebar.

Pour info je fais le choix de ne supporter que les champs qui ont un identifiant alors que ce n'est pas obligé. Je me demande ce qu'il se passe quand plusieurs champs ont le même mais je n'ai pas encore essayé.

Questions :

- L'approche ici, OK ? (il manque de la peinture des tests et la gestion du group by)
- Est-ce qu'il faut garder les `is_using_postgresql()`, si oui est-ce que c'est juste pour se protéger ou il faut aussi écrire du code pour gérer correctement le cas contraire ?
- Pour le filtre par statut, est-ce qu'on veut exposer tous les statuts ou juste les statuts simplifiés ? (les status simplifiés ça permettrait d'avoir le filtre au niveau global plutôt qu'en sous-filtre)

#9 - 25 janvier 2022 18:21 - Frédéric Péters

Pour info je fais le choix de ne supporter que les champs qui ont un identifiant alors que ce n'est pas obligé.

Dans un début d'idée de plan posé nulle part j'étais pour être explicite, avec a priori une nouvelle option "indicateurs statistiques" dans "display_locations" (qui ne serait pas coché par défaut ça me semble très bien).

(ça aiderait aussi bijoe aujourd'hui, [#60951](#)).

Je me demande ce qu'il se passe quand plusieurs champs ont le même mais je n'ai pas encore essayé.

Sans doute ça se passe mal.

Dans la suite du plan, j'étais pour ajouter un JSONField "indicateurs statistiques" dans les tables des demandes, en recopiant dedans les infos; c'était pour dans la suite reprendre cette colonne JSONField dans la nouvelle table `wcs_all_forms`, pour avoir des indicateurs couvrant plusieurs démarches à la fois.

Est-ce qu'il faut garder les `is_using_postgresql()`, si oui est-ce que c'est juste pour se protéger ou il faut aussi écrire du code pour gérer correctement le cas contraire ?

On peut avoir les fonctionnalités uniquement quand `postgresql` est là, faut juste lever une erreur propre dans le cas contraire.

Pour le filtre par statut, est-ce qu'on veut exposer tous les statuts ou juste les statuts simplifiés ? (les status simplifiés ça permettrait d'avoir le filtre au niveau global plutôt qu'en sous-filtre)

Oui on peut commencer comme ça. (et plus loin il y aura une notion de "zone" de workflow qui sera utile globalement etc. mais c'est une autre histoire).

#10 - 25 janvier 2022 20:19 - Benjamin Dauvergne

Frédéric Péters a écrit :

Je me demande ce qu'il se passe quand plusieurs champs ont le même mais je n'ai pas encore essayé.

Sans doute ça se passe mal.

Mes deux centimes: il y a une tension permanente entre les identifiants choisis par un humain et donc modifiables, victime de duplications, etc.. mais bien pratiques et un identifiant automatique aléatoire style UUID, aucun n'est satisfaisant en permanence. On a pu avoir ce souci avec les rôles ou les usagers lors d'un SSO ou d'un lien avec un annuaire aussi et ce qui marche bien c'est d'avoir les deux et de jongler (pour les utilisateurs on aura l'email/id de fédération/uuid technique d'annuaire qui jouent le même genre de rôle).

Ici si on avait un uuid par champ (un truc caché pas forcément ce qu'on utilise pour générer le nom des colonnes de table, comme pour les données de traitement qui n'a pas forcément été une bonne idée) on pourrait stocker la paire `uuid/varname` coté combo lors de la configuration, comme référence, et conserver le lien avec les champs du formulaire au gré des renommages, écrasement de formulaire, etc.. ce n'est pas parfait, si on écrase un formulaire avec un formulaire dont le champ a changé à la fois de `varname` et d'`uuid` ça ne marchera pas; mais dans la grande majorité des cas ça marche très bien. Il faut prévoir de rafraîchir les `varname` en moissonnant `w.c.s.` régulièrement pour que ça marche bien.

À mon avis chaque fois qu'on a un élément de structure avec une identité, on devrait lui assigner un uuid, qu'on y voit un intérêt ou pas dans l'immédiat, ça finit toujours par servir. Il n'est pas important de s'en servir réellement comme clé primaire dans le code, mais on doit le conserver comme référence supplémentaire chaque qu'on conserve une référence.

Coté bijoe ça aurait bien simplifié des choses aussi.

#11 - 25 janvier 2022 21:08 - Frédéric Péters

Ici si on avait un uuid par champ

(ne pas se lancer dans ça dans ce ticket)

#12 - 31 janvier 2022 15:05 - Valentin Deniaud

- Statut changé de Solution proposée à En cours

Frédéric Péters a écrit :

Pour info je fais le choix de ne supporter que les champs qui ont un identifiant alors que ce n'est pas obligé.

Dans un début d'idée de plan posé nulle part j'étais pour être explicite, avec a priori une nouvelle option "indicateurs statistiques" dans "display_locations" (qui ne serait pas coché par défaut ça me semble très bien).

Ouep très bien je vais faire ça.

Je me demande ce qu'il se passe quand plusieurs champs ont le même mais je n'ai pas encore essayé.

Sans doute ça se passe mal.

Dans la suite du plan, j'étais pour ajouter un JSONField "indicateurs statistiques" dans les tables des demandes, en recopiant dedans les infos; c'était pour dans la suite reprendre cette colonne JSONField dans la nouvelle table wcs_all_forms, pour avoir des indicateurs couvrant plusieurs démarches à la fois.

Là je décroche, je vais probablement faire simple ici, volontairement mal gérer le cas identifiants dupliqués et ouvrir un autre ticket pour faire mieux.

#13 - 03 février 2022 18:02 - Valentin Deniaud

- Fichier 0002-statistics-group-forms-count-by-field-60777.patch ajouté

- Fichier 0001-statistics-filter-forms-count-by-fields-60777.patch ajouté

- Statut changé de En cours à Solution proposée

0001 pour ajouter les filtres sur un champ, 0002 pour ajouter le regroupement par un champ.

Si le champ fait partie d'un bloc, le filtrage est supporté mais pas le regroupement : j'ai flairé les difficultés à permettre un GROUP BY sur un json field alors je n'ai pas essayé, ça se fera dans un autre ticket.

Principal truc qui reste à mettre au clair, le coup des identifiants de champ, varname ou id ?

Mon patch utilise les varnames. Au début c'était le critère discriminant pour que le champ apparaisse dans les filtres, maintenant qu'il y a un display_locations explicite cette raison ne tient plus. Je ne vois plus vraiment d'avantage à imposer ça, j'ai l'impression que les ids sont stables à l'import/export donc pas de problème non plus de ce côté là. Resterait seulement le bonus d'avoir des paramètres d'URL lisibles.

Et de l'autre côté il y a l'inconvénient de pouvoir avoir des varnames dupliqués ce qui ne serait pas le cas avec des ids.

OK pour que je modifie le patch pour utiliser les ids ? (c'est pas mal de code à bouger, pour ça que je demande)

Frédéric Péters a écrit :

Pour le filtre par statut, est-ce qu'on veut exposer tous les statuts ou juste les statuts simplifiés ? (les status simplifiés ça permettrait d'avoir le filtre au niveau global plutôt qu'en sous-filtre)

Oui on peut commencer comme ça.

Finalement non je garde le statut comme sous filtre avec Ouvert/Fermé + les autres status du workflow (en fait je ne sais pas faire autrement).

#14 - 03 février 2022 20:45 - Frédéric Péters

OK pour que je modifie le patch pour utiliser les ids ? (c'est pas mal de code à bouger, pour ça que je demande)

Non on a bien plus d'avenir à utiliser les varname. (particulièrement des perspectives de stats agrégeant plusieurs démarches).

#15 - 15 février 2022 14:23 - Frédéric Péters

```
+ options = [  
    ('validation', _('Validation Page')),  
    ('summary', _('Summary Page')),
```

```

    ('listings', _('Management Listings')),
]
+     if self.allow_statistics and self.varname:
+         options.append(('statistics', _('Statistics')))

```

Je serais pour toujours afficher statistics mais pour lever une erreur quand c'est coché et qu'il n'y a pas de varname (l'idée étant qu'ainsi la personne qui sait peut faire les deux en une seule édition).

#16 - 15 février 2022 17:27 - Valentin Deniaud

- Fichier 0001-statistics-filter-forms-count-by-fields-60777.patch ajouté
- Fichier 0002-statistics-group-forms-count-by-field-60777.patch ajouté

Yep, revoici.

#17 - 15 février 2022 19:02 - Frédéric Péters

- Statut changé de Solution proposée à Solution validée

#18 - 16 février 2022 10:06 - Valentin Deniaud

- Statut changé de Solution validée à Résolu (à déployer)

```

commit 2c94311d880bb5804ce4fb06848c86d545b77082
Author: Valentin Deniaud <vdeniaud@entrouvert.com>
Date: Tue Feb 1 16:39:02 2022 +0100

```

```

statistics: group forms count by field (#60777)

```

```

commit f9a59663c6dc9e948b7b55a97f9f6d5b283b4e4f
Author: Valentin Deniaud <vdeniaud@entrouvert.com>
Date: Mon Jan 24 18:14:05 2022 +0100

```

```

statistics: filter forms count by fields (#60777)

```

#19 - 17 février 2022 21:16 - Transition automatique

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

#20 - 24 avril 2022 04:42 - Transition automatique

Automatic expiration

Fichiers

0001-wip.patch	8,04 ko	25 janvier 2022	Valentin Deniaud
0002-statistics-group-forms-count-by-field-60777.patch	18,2 ko	03 février 2022	Valentin Deniaud
0001-statistics-filter-forms-count-by-fields-60777.patch	25 ko	03 février 2022	Valentin Deniaud
0002-statistics-group-forms-count-by-field-60777.patch	18,5 ko	15 février 2022	Valentin Deniaud
0001-statistics-filter-forms-count-by-fields-60777.patch	26,5 ko	15 février 2022	Valentin Deniaud