Calebasse - Development #4842

Mise à jour de la base de données des lieux de socialisation

23 mai 2014 12:38 - Jérôme Schneider

Statut:	Fermé	**Début:**	23 mai 2014
Priorité:	Normal	**Echéance:**	
Assigné à:	Jérôme Schneider	**% réalisé:**	90%
Catégorie:		**Temps estimé:**	0:00 heure
Version cible:	1.2 Lieux de socialisation		
Patch proposed:	Oui	**Planning:**	

Description

Supprimer les entrées actuelles et peupler la base avec les fichiers tableurs disponibles sur le site du rectorat

Historique

#1 - 30 juin 2014 12:32 - Mikaël Ates (de retour le 29 avril)

- Patch proposed mis à Non

"""J'ai trouvé les adresses sur le site http://www.ia42.ac-lyon.fr/

Sommaire
Etablissements et circonscriptions
Premier degré:
Télécharger le fichier écoles publiques et privées
Second degré:
Télécharger le fichier collèges et lycées privés et publics

Voilà je regarde si je peux trouver une liste des crèches et jardin d'enfants"""

Annuaire officiel : http://www.education.gouv.fr/pid24301/annuaire-accueil-recherche.html

#2 - 17 juillet 2014 18:05 - Jérôme Schneider

- Statut changé de Nouveau à En cours

#3 - 17 juillet 2014 18:36 - Jérôme Schneider

- Statut changé de En cours à Information nécessaire

Je vois plusieurs points qui demandent réflexion sur ce ticket :

- la suppression des entrées actuelles : ça aura un impact fort sur l'existant. Tous les onglets lieux de socialisation vont se retrouver vide. Ca ne me semble pas être la bonne solution. On peut essayer de faire une migration intelligente qui va essayer de faire un matching entre les anciennes entrées et les nouvelles mais je pense qu'on aura beaucoup d'erreurs. La troisième idées serait de supprimer que les lieux de socialisation qui ne sont pas utilisés et de garder les autre en laissant le soin aux secrétaires de nettoyer la base.
- dans le fichier ANNUAIRE_ECOLES_PUBLIQUES.xls il manque le nom d'une école sur deux. Je propose qu'on injecte pas ces écoles.

#4 - 18 juillet 2014 18:36 - Jérôme Schneider

- Fichier import_school_from_xls_files.py ajouté

- % réalisé changé de 0 à 40

Voici un premier script qui ne fait que le mapping entre les xls et calebase. Je le mets dans le ticket pour valider que je pars bien dans le bonne direction. Les xls sont loin d'être parfait mais l'extraction n'est pas catastrophique.

#5 - 18 juillet 2014 18:51 - Frédéric Péters

Sur le fond je vais laisser Mikaël, sur la forme moi je trouve qu'il y a beaucoup de scripts qui se trouvent ainsi faits sur le côté, et que ça pourrait très bien aller en commande de management (et comme dans docbow ça pourrait être appelable depuis le script dans /etc/init.d/, comme ça on est sûr d'avoir un seul endroit qui s'occupe de l'environnement, et donc les settings les mêmes).

Sinon dans un survol rapide alors je le note, même si ça devrait se voir en testant pour de vrai : ici lycee/college devraient être des chaînes :

```
        elif school_type.lower() in [u"lp", u"lycée", u"sep", u"lycée professionnel"]:
            value['school_type'] = lycee
```

```
        elif school_type.lower() == u"collège":
            value['school_type'] = college
```

Le isinstance multiplé sur str et unicode, Thomas m'a récemment fait remarqué qu'il était avantageusement remplacé par un seul test sur le type basestring.

#6 - 01 août 2014 18:57 - Jérôme Schneider

- Fichier 0001-ressources-new-command-reinit_schools_from_xls-which.patch ajouté

- Patch proposed changé de Non à Oui

J'ai suivi les remarques de Fred en utilisant basestring et ajoutant une commande dans ressources. De plus cette version vide la base des lieux de socialisations et injecte ceux des xls.

#7 - 01 août 2014 18:57 - Jérôme Schneider

- % réalisé changé de 40 à 70

#8 - 02 août 2014 00:09 - Benjamin Dauvergne

Je n'ai pas regardé le schéma mais il me semble que si tu supprimes toutes les school tu perds toutes les affectations existantes. Si c'est le cas il faudrait plutôt essayer d'utiliser la clé unique fournie par l'educnat, le code RNE ou code école, pour pouvoir faire des mises à jour. Ça donnerait un truc comme ça

```
    def save_schools(self, schools):
        for school in schools:
            mschool, created = School.objects.get_or_create(code_rne=school['code_rne'], defaults=school)
            if not created:
                mschool.__dict__.update(schools)
                mschool.save()
            mschool.services = self.services
```

Les types d'établissements devraient plutôt être récupérés par chaîne que par id (ça rejoint la remarque de fred) vu que l'objet ne sert qu'à stocker un nom, ou alors il faut transformer school_type en champ varchar.

```
    creche = SchoolType.objects.get(id=1)
    maternelle = SchoolType.objects.get(id=2)
    primaire = SchoolType.objects.get(id=3)
    college = SchoolType.objects.get(id=4)
    lycee = SchoolType.objects.get(id=5)
    inconnu = SchoolType.objects.get(id=8)
    services = Service.objects.all()
```

Il faudrait ça:

```
    creche = SchoolType.objects.get_or_create(name=u'crèche')
    maternelle = SchoolType.objects.get(name=u'maternelle')
    etc...
```

Aussi je ne sais pas si il y a une ligne d'entête dans le XLS mais si il y en a une je ferai un assert pour vérifier que les titres ne changent pas; sûr ça va péter lorsqu'ils corrigent une typo ou en font une mais si le format change vraiment on le saura.

Autre petit conseil j'aime bien mettre le décorateur commit_on_success sur la méthode handle() ou la méthode qui effectue des modifications sur la base dans mes commandes, ça me rassure de savoir que quand ça foire il n'y a pas d'effet sur la base et ça permet d'implémenter facilement un mode -dry ou -fake qui ne fait rien, en lançant simplement une exception.

#9 - 11 août 2014 11:22 - Frédéric Péters

> Je n'ai pas regardé le schéma mais il me semble que si tu supprimes toutes les school tu perds toutes les affectations existantes. [...]

Jérôme, la nouvelle commande a été poussée; le problème évoqué par Benjamin n'existe pas ?

#10 - 13 août 2014 11:26 - Jérôme Schneider

Benjamin Dauvergne a écrit :

> Je n'ai pas regardé le schéma mais il me semble que si tu supprimes toutes les school tu perds toutes les affectations existantes. Si c'est le cas il faudrait plutôt essayer d'utiliser la clé unique fournie par l'educnat, le code RNE ou code école, pour pouvoir faire des mises à jour. Ça donnerait un truc comme ça

Le code RNE n'est pas présent dans les fichiers xls et ce script n'est pas prévu pour faire une mise à jour mais injecter une fois les lieux de socialisation.

> Les types d'établissements devraient plutôt être récupérés par chaîne que par id (ça rejoint la remarque de fred) vu que l'objet ne sert qu'à stocker un nom, ou alors il faut transformer school_type en champ varchar.
>
> [...]

J'ai modifié avant de pousser dans le master.

> Aussi je ne sais pas si il y a une ligne d'entête dans le XLS mais si il y en a une je ferai un assert pour vérifier que les titres ne changent pas; sûr ça va péter lorsqu'ils corrigent une typo ou en font une mais si le format change vraiment on le saura.

Vu que c'est du one shot je ne pense pas que ca soit utile.

> Autre petit conseil j'aime bien mettre le décorateur commit_on_success sur la méthode handle() ou la méthode qui effectue des modifications sur la base dans mes commandes, ça me rassure de savoir que quand ça foire il n'y a pas d'effet sur la base et ça permet d'implémenter facilement un mode -dry ou -fake qui ne fait rien, en lançant simplement une exception.

On repart de 0 pour cette commande donc je ne pense pas que soit utile. Au contraire ça permet de voir directement ou ça foire.

#11 - 19 août 2014 18:01 - Jérôme Schneider

- Statut changé de Information nécessaire à Résolu (à déployer)

- % réalisé changé de 70 à 90

#12 - 28 août 2014 16:13 - Mikaël Ates (de retour le 29 avril)

- Statut changé de Résolu (à déployer) à Fermé

Ok

Fichiers

import_school_from_xls_files.py	4,26 ko	18 juillet 2014	Jérôme Schneider
0001-ressources-new-command-reinit_schools_from_xls-which.patch	6,67 ko	01 août 2014	Jérôme Schneider