

Passerelle - Development #64313

foederis, connecteur cache des données / recherche

19 avril 2022 11:45 - Benjamin Dauvergne

Statut:	Fermé	Début:	19 avril 2022
Priorité:	Normal	Echéance:	
Assigné à:	Benjamin Dauvergne	% réalisé:	0%
Catégorie:		Temps estimé:	0:00 heure
Version cible:		Planning:	Non
Patch proposed:	Oui		

Description

Le logiciel n'étant pas disponible 100% du temps, il faut mettre en cache la totalité des données nécessaire au fonctionnement du portail avec une mise à jour régulière.

Cf. #62643

Les données :

- le descriptif des postes (les annonces), ws référentiel : https://.../data/annonce?viewIntegrationName=api_public
- différents référentiels :
 - Type d'emploi (permanent, temporaire, saisonnier, stagiaire, apprenti.....), champ annonce: type_emploi_ddr, ws référentiel: https://.../data/type_emploi?viewIntegrationName=api_public
 - Catégorie (A, B, C), champ annonce: categorie_ddr, ws référentiel: https://.../data/categorie1?viewIntegrationName=api_public
 - Filière (administrative, technique, ..), champ annonce: filiere_ddr, ws référentiel: https://.../data/Filiere?viewIntegrationName=api_public
- les fichiers PDF des annonces, ws : http https://.../data/annonce/<champ id de l'annonce>/fields/pdf_ddr?viewIntegrationName=api_public, donnée binaire encodé en base64 dans le champ results.0.pdf_ddr.fileData

Niveau requêtes au niveau des annonces on doit disposer:

- d'une recherche par sous chaîne sur l'intitulé -> __icontains
- d'un filtrage par type d'emploi / catégorie et filière

Les modèles :

- 1 modèle ressource
- 1 modèle document, 4 champs: external_id(text), display_name(text), last_update(datetime), data(json), pdf(file)
 - 1 index JSON sur data

Mettre en cache les référentiels directement dans un seul document (ils sont petits) (external-id = type-emploi, categorie, filiere)

Révisions associées

Révision 2f729a7f - 21 avril 2022 11:37 - Benjamin Dauvergne

misc: force message to be in string in _log (#64313)

It allows translated strings in log calls.

Révision 132185c0 - 21 avril 2022 16:38 - Benjamin Dauvergne

utils/json: move response_schema from rsa13 (#64313)

Révision dea4352c - 21 avril 2022 16:38 - Benjamin Dauvergne

utils/json: add datasource_schema helper (#64313)

Révision 4074f822 - 21 avril 2022 17:11 - Benjamin Dauvergne

templatetags: add support for patternProperties JSON schema (#64313)

Révision 91e7c9af - 21 avril 2022 17:11 - Benjamin Dauvergne

add toulouse_foederis connector (#64313)

Historique

#2 - 19 avril 2022 12:25 - Benjamin Dauvergne

- Description mis à jour

#3 - 19 avril 2022 13:38 - Benjamin Dauvergne

- Assigné à mis à Benjamin Dauvergne

#4 - 19 avril 2022 19:20 - Benjamin Dauvergne

- Fichier 0002-add-toulouse_foederis-connector-64313.patch ajouté
- Fichier 0001-misc-force-message-to-be-in-string-in_log-64313.patch ajouté
- Statut changé de Nouveau à Solution proposée
- Patch proposed changé de Non à Oui

#5 - 19 avril 2022 19:21 - Benjamin Dauvergne

- Assigné à changé de Benjamin Dauvergne à Nicolas Roche

#6 - 20 avril 2022 15:36 - Benjamin Dauvergne

- Fichier 0002-add-toulouse_foederis-connector-64313.patch ajouté
- Fichier 0001-misc-force-message-to-be-in-string-in_log-64313.patch ajouté

#7 - 21 avril 2022 11:23 - Nicolas Roche

- Statut changé de Solution proposée à Solution validée

Ack. Juste quelques suggestions à prendre ou a laisser.

tests :

- lignes non utilisées :

```
HTTP MOCKS = {
  'type-emploi': {
    'path': r'^/.*data/type_emploi$',
    #'query': 'viewIntegrationName=api_publick',
```

- fixture mock_500 non définie
- fixture http_mock surchargée

```
def test_data_annonce_error_500(self, resource, http_mock, caplog):
def test_pdf_error_500(self, resource, http_mock, caplog):
```

models :

- import non utilisé :

```
from django.core.exceptions import ValidationError
```

- log_requests_errors = False
Je vois qu'il est aussi positionné sur smart et que ça n'a pas posé de problème pour déboguer mais je me demande quand même à quoi ça sert ici.
- def announce()
J'aurais mis un 's', mais j'avoue que je m'y perd : update_referentiels met à jour les annonces et les référentiels, et annonce renvoie les annonces et les référentiels.
Sinon, peut-être les renommer update et search ?
- À la lecture de la description du endpoint datasource via l'IHM on imagine qu'il s'utilise avec "ds?name=XXX" alors qu'en fait il faut utiliser le format "ds/XXX".
Idem pour les PDF : "annonce/123/pdf/".

- Pour la recherche textuelle, rien ne précise dans le ticket client sur quel champ la faire, et je me demande s'il ne faudrait pas l'étendre à la description de l'annonce.

#8 - 21 avril 2022 11:29 - Nicolas Roche

Jenkins rouge, je dirais à cause de l'import inutile :

```
passerelle/contrib/toulouse_foederis/models.py:21: [W0611(unused-import), ] Unused ValidationError imported from django.core.exceptions
```

#9 - 21 avril 2022 12:47 - Benjamin Dauvergne

Nicolas Roche a écrit :

Ack. Juste quelques suggestions à prendre ou a laisser.

tests :

- lignes non utilisées :
[...]

C'est utile mais c'est vrai que j'ai oublié de le passer à urlmatch, je corrige ça.

- fixture mock_500 non définie

Défini dans tests/conftest.py mais j'ai préféré mettre un mock_url explicite à la place, je le retire de la signature.

- fixture http_mock surchargée
[...]

En fait il se combine avec le mock_url plus loin.

models :

- import non utilisé :
[...]

Corrigé.

- log_requests_errors = False
Je vois qu'il est aussi positionné sur smart et que ça n'a pas posé de problème pour déboguer mais je me demande quand même à quoi ça sert ici.

C'est pour éviter d'avoir des mails d'erreur HTTP alors que le cron hourly gère déjà les erreurs HTTP. C'est utile quand on préfère gérer les erreurs manuellement.

- def announce()
J'aurais mis un 's', mais j'avoue que je m'y perd : update_referentiels met à jour les annonces et les référentiels, et announce renvoie les annonces et les référentiels.
Sinon, peut-être les renommer update et search ?

On conseille en général de nommer les endpoint REST au singulier, on est pas super cohérent sur ce point.

- À la lecture de la description du endpoint datasource via l'IHM on imagine qu'il s'utilise avec "ds?name=XXX" alors qu'en fait il faut utiliser le format "ds/XXX".
Idem pour les PDF : "announce/123/pdf".

Ok vu, j'ajoute un exemple_pattern pour les deux.

- Pour la recherche textuelle, rien ne précise dans le ticket client sur quel champ la faire, et je me demande s'il ne faudrait pas l'étendre à la description de l'annonce.

Dans le document PUBLIK_LOT1_Diffusion_offres_V1.1.doc (ticket parent) il est indiqué que ça ne concerne que l'intitulé.

Je pousse tout ça ici.

#10 - 21 avril 2022 12:49 - Benjamin Dauvergne

- Assigné à changé de Nicolas Roche à Benjamin Dauvergne

#11 - 21 avril 2022 12:49 - Benjamin Dauvergne

- Fichier 0003-add-toulouse_foederis-connector-64313.patch ajouté
- Fichier 0002-misc-move-response_schema-to-utils.json-64313.patch ajouté
- Fichier 0001-misc-force-message-to-be-in-string-in-_log-64313.patch ajouté
- Statut changé de Solution validée à Solution proposée

J'en ai profité pour ajouter quelques goodies :

- schémas JSON des réponse,
- long_description pour les 3 endpoints (précisant les référentiels dispo et leur nom pour data_sources)
- affichage de la date de renière mise à jour pour les annonces et les sources de donnée sur la page du connecteur

#12 - 21 avril 2022 13:14 - Benjamin Dauvergne

- Fichier 0003-add-toulouse_foederis-connector-64313.patch ajouté
- Fichier 0002-misc-move-response_schema-to-utils.json-64313.patch ajouté
- Fichier 0001-misc-force-message-to-be-in-string-in-_log-64313.patch ajouté

Encore des bouts:

- ajout de la description du champ data_sources au schéma de la réponse du endpoint annonce (ça permet d'afficher la page de recherche en un appel)
- ajout du filtrage par collectivite (les annonces seront vraisemblablement affichées pour chaque collectivité sur des pages différentes)

#13 - 21 avril 2022 15:35 - Nicolas Roche

- Assigné à changé de Benjamin Dauvergne à Nicolas Roche

C'est chouette l'ajout de patternProperties,

<https://json-schema.org/understanding-json-schema/reference/object.html#pattern-properties>

mais je l'aurais bien vu dans un commit à part parce que c'est pas évident de comprendre comment on doit l'utiliser.

Là j'ai l'impression qu'on en fait un usage détourné en ne renseignant pas de regex, et du coup j'ai un affichage approximatif (pattern //):

```
data_sources : object
  pattern // : array object
    id : string
    text : string
```

L'idée étant, si j'ai bien compris, de ne pas alourdir le schéma en lui ajoutant 3 fois le schéma des datasources. Je ne suis pas convaincu que ce soit un bon exemple pour son introduction.

#14 - 21 avril 2022 16:39 - Benjamin Dauvergne

- Fichier 0003-utils-json-add-datasource_schema-helper-64313.patch ajouté
- Fichier 0002-utils-json-move-response_schema-from-rsa13-64313.patch ajouté
- Fichier 0004-templatetags-add-support-for-patternProperties-JSON-.patch ajouté
- Fichier 0001-misc-force-message-to-be-in-string-in-_log-64313.patch ajouté
- Fichier 0005-add-toulouse_foederis-connector-64313.patch ajouté

Allez tout bien séparé.

#15 - 21 avril 2022 17:11 - Benjamin Dauvergne

- Fichier 0003-utils-json-add-datasource_schema-helper-64313.patch ajouté
- Fichier 0002-utils-json-move-response_schema-from-rsa13-64313.patch ajouté
- Fichier 0004-templatetags-add-support-for-patternProperties-JSON-.patch ajouté
- Fichier 0001-misc-force-message-to-be-in-string-in-_log-64313.patch ajouté
- Fichier 0005-add-toulouse_foederis-connector-64313.patch ajouté

Changement au rendu des patternProperties.

#16 - 21 avril 2022 17:15 - Benjamin Dauvergne

Nicolas Roche a écrit :

Là j'ai l'impression qu'on en fait un usage détourné en ne renseignant pas de regex, et du coup j'ai un affichage approximatif (pattern //):

Ce n'est pas fait pour avoir forcément une regexp c'est aussi le seul moyen de spécifier uniquement le schéma des valeurs d'un dictionnaire sans expliciter les clés.

#17 - 21 avril 2022 17:16 - Nicolas Roche

- Statut changé de Solution proposée à Solution validée

Super, merci.

#18 - 21 avril 2022 17:40 - Benjamin Dauvergne

- Assigné à changé de Nicolas Roche à Benjamin Dauvergne

#19 - 21 avril 2022 17:40 - Benjamin Dauvergne

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

```
commit 91e7c9afc9ff26e1ac88a2c70d936e2a40ef35815
Author: Benjamin Dauvergne <bdauvergne@entrouvert.com>
Date: Tue Apr 19 15:20:00 2022 +0200
```

```
add toulouse_foederis connector (#64313)
```

```
commit 4074f8222d702d28b1a6d891fcfd71da60776a4d
Author: Benjamin Dauvergne <bdauvergne@entrouvert.com>
Date: Thu Apr 21 16:35:30 2022 +0200
```

```
templatetags: add support for patternProperties JSON schema (#64313)
```

```
commit dea4352c91a49633c4c8c199c5b04b7b564980ba
Author: Benjamin Dauvergne <bdauvergne@entrouvert.com>
Date: Thu Apr 21 16:37:53 2022 +0200
```

```
utils/json: add datasource_schema helper (#64313)
```

```
commit 132185c03db91133f61349efcfb53d70cc720143
Author: Benjamin Dauvergne <bdauvergne@entrouvert.com>
Date: Thu Apr 21 12:10:37 2022 +0200
```

```
utils/json: move response_schema from rsa13 (#64313)
```

```
commit 2f729a7fb64939b57cdf9e913e8a044ca03c0f36
Author: Benjamin Dauvergne <bdauvergne@entrouvert.com>
Date: Tue Apr 19 15:25:16 2022 +0200
```

```
misc: force message to be in string in _log (#64313)
```

```
It allows translated strings in log calls.
```

#20 - 21 avril 2022 18:17 - Transition automatique

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

#21 - 26 juin 2022 04:42 - Transition automatique

Automatic expiration

Fichiers

0002-add-toulouse_foederis-connector-64313.patch	65,3 ko	19 avril 2022	Benjamin Dauvergne
0001-misc-force-message-to-be-in-string-in-_log-64313.patch	869 octets	19 avril 2022	Benjamin Dauvergne
0002-add-toulouse_foederis-connector-64313.patch	69 ko	20 avril 2022	Benjamin Dauvergne
0001-misc-force-message-to-be-in-string-in-_log-64313.patch	869 octets	20 avril 2022	Benjamin Dauvergne

0003-add-toulouse_foederis-connector-64313.patch	70,7 ko	21 avril 2022	Benjamin Dauvergne
0002-misc-move-response_schema-to-utils.json-64313.patch	2,43 ko	21 avril 2022	Benjamin Dauvergne
0001-misc-force-message-to-be-in-string-in-_log-64313.patch	869 octets	21 avril 2022	Benjamin Dauvergne
0003-add-toulouse_foederis-connector-64313.patch	77,6 ko	21 avril 2022	Benjamin Dauvergne
0002-misc-move-response_schema-to-utils.json-64313.patch	2,43 ko	21 avril 2022	Benjamin Dauvergne
0001-misc-force-message-to-be-in-string-in-_log-64313.patch	869 octets	21 avril 2022	Benjamin Dauvergne
0003-utils-json-add-datasource_schema-helper-64313.patch	1,11 ko	21 avril 2022	Benjamin Dauvergne
0002-utils-json-move-response_schema-from-rsa13-64313.patch	2,1 ko	21 avril 2022	Benjamin Dauvergne
0004-templatetags-add-support-for-patternProperties-JSON-.patch	3,76 ko	21 avril 2022	Benjamin Dauvergne
0001-misc-force-message-to-be-in-string-in-_log-64313.patch	869 octets	21 avril 2022	Benjamin Dauvergne
0005-add-toulouse_foederis-connector-64313.patch	72,6 ko	21 avril 2022	Benjamin Dauvergne
0003-utils-json-add-datasource_schema-helper-64313.patch	1,11 ko	21 avril 2022	Benjamin Dauvergne
0002-utils-json-move-response_schema-from-rsa13-64313.patch	2,1 ko	21 avril 2022	Benjamin Dauvergne
0004-templatetags-add-support-for-patternProperties-JSON-.patch	4,07 ko	21 avril 2022	Benjamin Dauvergne
0001-misc-force-message-to-be-in-string-in-_log-64313.patch	869 octets	21 avril 2022	Benjamin Dauvergne
0005-add-toulouse_foederis-connector-64313.patch	72,6 ko	21 avril 2022	Benjamin Dauvergne