

Passerelle - Development #43201

Ne pas copier les entêtes de réponse tel-quel dans les logs au niveau DEBUG (était: UnicodeDecodeError: 'utf8' codec can't decode byte 0xe9 in position 14: invalid continuation byte)

22 mai 2020 09:10 - Sentry lo

| | | | |
|------------------------|--------------------|----------------------|-------------|
| Statut: | Fermé | Début: | 22 mai 2020 |
| Priorité: | Normal | Echéance: | |
| Assigné à: | Benjamin Dauvergne | % réalisé: | 0% |
| Catégorie: | | Temps estimé: | 0:00 heure |
| Version cible: | | Planning: | Non |
| Patch proposed: | Oui | | |

Description

<https://sentry.entrouvert.org/entrouvert/publik/issues/2372/>

```
UnicodeDecodeError: 'utf8' codec can't decode byte 0xe9 in position 14: invalid continuation byte
(27 additional frame(s) were not displayed)
```

```
...
```

```
File "psycopg2/_json.py", line 96, in getquoted
    s = self.dumps(self.adapted)
File "django/contrib/postgres/fields/jsonb.py", line 25, in dumps
    return json.dumps(obj, **options)
File "__init__.py", line 244, in dumps
    return _default_encoder.encode(obj)
File "json/encoder.py", line 207, in encode
    chunks = self.iterencode(o, _one_shot=True)
File "json/encoder.py", line 270, in iterencode
    return _iterencode(o, 0)
```

Error occurred while processing request

Révisions associées

Révision b8b864a8 - 25 mai 2020 11:38 - Benjamin Dauvergne

misc: log HTTP response headers safely (#43201)

Historique

#1 - 22 mai 2020 09:11 - Lauréline Guérin

- *Projet changé de Suivi des traces à Passerelle*

A priori c'est lié au passage à django jsonfield

#2 - 22 mai 2020 10:41 - Benjamin Dauvergne

J'ai l'impression que ça n'a rien à voir avec jsonb, c'est dans un simple appel à json.dumps() et le contenu contient visiblement une chaîne bytes qui n'est pas de l'UTF-8; je dirai que c'est le header 'Erreur Webservice' (le nom ne me paraît même pas légal).

```
Erreur Webservice : 'Aucun champ clé renseigné : impossible de charger l'objet (gf.tiers.BsTiersDonneesGenerales). URL: hdnCodeTrt=AFFICHE&CodTrt=Tiers&CodDom=GF&_hdnCodOrg=CG06&_hdnCodBud=01&_hdnCodExe=2020&comboOrganisme=|CG06|',
```

ça n'a rien à faire dans un header sans encodage et on les copie sans les décoder ou les vérifier :

```
# passerelle/utils/__init__.py
189     if logger.level == 10: # DEBUG
190         extra['response_headers'] = dict(response.headers.items())
```

durant une petite période le niveau de log DEBUG était activé, ça explique pourquoi on ne voit plus l'erreur :

https://passerelle-departement06.test.entrouvert.org/manage/astregs/associations/logs/?log_level=&q=20%2F05%2F2020+19%3A52#logs

#3 - 22 mai 2020 10:42 - Benjamin Dauvergne

- *Sujet changé de UnicodeDecodeError: 'utf8' codec can't decode byte 0xe9 in position 14: invalid continuation byte à Ne pas copier les entêtes de réponse tel-quel dans les logs au niveau DEBUG (était: UnicodeDecodeError: 'utf8' codec can't decode byte 0xe9 in position 14: invalid continuation byte)*

#4 - 22 mai 2020 10:47 - Serghei Mihai (congrés, retour 15/05)

Benjamin Dauvergne a écrit :

durant une petite période le niveau de log DEBUG était activé, ça explique pourquoi on ne voit plus l'erreur :

https://passerelle-departement06.test.entrouvert.org/manage/astregs/associations/logs/?log_level=&q=20%2F05%2F2020+19%3A52#logs

Je plaide coupable, j'avais passé le connecteur en mode debug pour voir les réponses XML à mes requetes.

#5 - 22 mai 2020 12:13 - Benjamin Dauvergne

Serghei Mihai a écrit :

Je plaide coupable, j'avais passé le connecteur en mode debug pour voir les réponses XML à mes requetes.

T'as rien fait de mal, il y a un bug c'est tout.

#6 - 22 mai 2020 13:44 - Benjamin Dauvergne

- *Assigné à mis à Benjamin Dauvergne*

#7 - 22 mai 2020 14:11 - Benjamin Dauvergne

- *Fichier 0001-misc-log-HTTP-response-headers-safely-43201.patch ajouté*

- *Tracker changé de Bug à Development*

- *Statut changé de Nouveau à Solution proposée*

- *Patch proposed changé de Non à Oui*

#8 - 24 mai 2020 18:17 - Nicolas Roche

- *Statut changé de Solution proposée à Solution validée*

... c'est dans un simple appel à json.dumps() et le contenu contient visiblement une chaîne bytes qui n'est pas de l'UTF-8;

Je reproduis bien l'erreur :

```
$ python2
>>> import json
>>> json.dumps({'coucou': b'\xe9'})
Traceback (most recent call last):
  File "/usr/lib/python2.7/json/__init__.py", line 244, in dumps
    return _default_encoder.encode(obj)
  File "/usr/lib/python2.7/json/encoder.py", line 207, in encode
    chunks = self.iterencode(o, _one_shot=True)
  File "/usr/lib/python2.7/json/encoder.py", line 270, in iterencode
    return _iterencode(o, 0)
UnicodeDecodeError: 'utf8' codec can't decode byte 0xe9 in position 0: unexpected end of data
```

Je plaide coupable, j'avais passé le connecteur en mode debug

et donc ça va dans le sens de l'analyse faite ici.

#9 - 25 mai 2020 11:38 - Benjamin Dauvergne

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

commit b8b864a863042e8dd2d3f7f75d36e055e3ec8ddf
Author: Benjamin Dauvergne <bdauvergne@entrouvert.com>

Date: Fri May 22 14:10:54 2020 +0200

misc: log HTTP response headers safely (#43201)

#10 - 25 mai 2020 20:16 - Frédéric Péters

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

Fichiers

| | | | |
|--|---------|-------------|--------------------|
| 0001-misc-log-HTTP-response-headers-safely-43201.patch | 4,09 ko | 22 mai 2020 | Benjamin Dauvergne |
|--|---------|-------------|--------------------|