

w.c.s. - Bug #78824

sérialisation pétée lxml, revenir à ElementTree (XMLSyntaxError: PCDATA invalid Char value 2, line 29, column 59 (<string>, line 29))

21 juin 2023 16:56 - Sentry lo

Statut:	Fermé	Début:	21 juin 2023
Priorité:	Haut	Echéance:	
Assigné à:	Frédéric Péters	% réalisé:	0%
Catégorie:		Temps estimé:	0:00 heure
Version cible:		Planning:	Non
Patch proposed:	Non		

Description

<https://sentry.entrouvert.org/entrouvert/gplexpert/issues/112395/>

```
XMLSyntaxError: PCDATA invalid Char value 2, line 29, column 59 (<string>, line 29)
(17 additional frame(s) were not displayed)
...
File "src/lxml/parser.pxi", line 1777, in lxml.etree._parseDoc
File "src/lxml/parser.pxi", line 1082, in lxml.etree._BaseParser._parseUnicodeDoc
File "src/lxml/parser.pxi", line 615, in lxml.etree._ParserContext._handleParseResultDoc
File "src/lxml/parser.pxi", line 725, in lxml.etree._handleParseResult
File "src/lxml/parser.pxi", line 654, in lxml.etree._raiseParseError
```

Révisions associées

Révision 28fdbd69 - 21 juin 2023 17:15 - Frédéric Péters

general: revert switch to lxml, go back to ElementTree (#78824)

Révision 06b11f29 - 21 juin 2023 17:38 - Frédéric Péters

snapshots: store full serialization if latest snapshot is broken (#78824)

Historique

#1 - 21 juin 2023 16:57 - Frédéric Péters

Déjà vu [#78549](#) avec là analyse comme quoi c'était l'indentation et bien non ça foire aussi sans indentation :

```
In [21]: ET.fromstring(ET.tostring(workflow.export_to_xml(include_id=True)).decode('utf-8'))
Traceback (most recent call last):
```

```
File "/usr/lib/python3/dist-packages/IPython/core/interactiveshell.py", line 3427, in run_code
exec(code_obj, self.user_global_ns, self.user_ns)
```

```
File "<ipython-input-21-2358d7f3a5c7>", line 1, in <module>
ET.fromstring(ET.tostring(workflow.export_to_xml(include_id=True)).decode('utf-8'))
```

```
File "src/lxml/etree.pyx", line 3237, in lxml.etree.fromstring
```

```
File "src/lxml/parser.pxi", line 1896, in lxml.etree._parseMemoryDocument
```

```
File "src/lxml/parser.pxi", line 1777, in lxml.etree._parseDoc
```

```
File "src/lxml/parser.pxi", line 1082, in lxml.etree._BaseParser._parseUnicodeDoc
```

```
File "src/lxml/parser.pxi", line 615, in lxml.etree._ParserContext._handleParseResultDoc
```

```
File "src/lxml/parser.pxi", line 725, in lxml.etree._handleParseResult
```

```
File "src/lxml/parser.pxi", line 654, in lxml.etree._raiseParseError
```

```
File "<string>", line 29
```

```
XMLSyntaxError: PCDATA invalid Char value 2, line 29, column 59
```

#2 - 21 juin 2023 17:02 - Frédéric Péters

Sans doute opportun ici d'annuler [#78281](#) et revenir au parseur python (xml.etree.ElementTree) et quand on sera passé à bookworm et une nouvelle version de libxml2 réévaluer ça.

#4 - 21 juin 2023 17:12 - Frédéric Péters

- Assigné à mis à Frédéric Péters

#5 - 21 juin 2023 17:17 - Robot Gitea

- Statut changé de Nouveau à Solution proposée

Frédéric Péters (fpeters) a ouvert une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/wcs/pulls/416>
- Titre : revenir au parseur python (xml.etree.ElementTree) ([#78824](#))
- Modifications : <https://git.entrouvert.org/entrouvert/wcs/pulls/416/files>

#6 - 21 juin 2023 17:17 - Robot Gitea

- Statut changé de Solution proposée à En cours

Frédéric Péters (fpeters) a commencé à travailler sur une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/wcs/pulls/416>
- Titre : WIP: revenir au parseur python (xml.etree.ElementTree) ([#78824](#))
- Modifications : <https://git.entrouvert.org/entrouvert/wcs/pulls/416/files>

#7 - 21 juin 2023 17:18 - Frédéric Péters

- Projet changé de Suivi des traces à w.c.s.

- Sujet changé de sérialisation pétée lxml (XMLSyntaxError: PCDATA invalid Char value 2, line 29, column 59 (<string>, line 29)) à sérialisation pétée lxml, revenir à ElementTree (XMLSyntaxError: PCDATA invalid Char value 2, line 29, column 59 (<string>, line 29))

#8 - 21 juin 2023 17:38 - Robot Gitea

- Statut changé de En cours à Solution proposée

#9 - 21 juin 2023 21:05 - Frédéric Péters

- Priorité changé de Normal à Haut

Ce ticket est important avant la mise à jour, au risque d'avoir du XML pété de partout et mille galères.

#10 - 21 juin 2023 21:41 - Robot Gitea

- Statut changé de Solution proposée à Solution validée

Benjamin Dauvergne (bdauvergne) a approuvé une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/wcs/pulls/416>

#11 - 22 juin 2023 07:10 - Robot Gitea

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

Frédéric Péters (fpeters) a mergé une pull request sur Gitea concernant cette demande :

- URL : <https://git.entrouvert.org/entrouvert/wcs/pulls/416>
- Titre : revenir au parseur python (xml.etree.ElementTree) ([#78824](#))
- Modifications : <https://git.entrouvert.org/entrouvert/wcs/pulls/416/files>

#12 - 22 juin 2023 08:14 - Transition automatique

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

#13 - 27 août 2023 04:42 - Transition automatique

Automatic expiration