Redmine - Patch #25861
CSV Importer - handle UndefinedConversionErrors
2017-05-16 11:07 - Jens Krämer

<table>
<thead>
<tr>
<th>Status:</th>
<th>Closed</th>
<th>Start date:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Priority:</td>
<td>Normal</td>
<td>Due date:</td>
</tr>
<tr>
<td>Assignee:</td>
<td>Jean-Philippe Lang</td>
<td>% Done:</td>
</tr>
<tr>
<td>Category:</td>
<td>Importers</td>
<td>Estimated time:</td>
</tr>
<tr>
<td>Target version:</td>
<td>3.2.7</td>
<td></td>
</tr>
</tbody>
</table>

**Description**

The CSV import already handles a couple of Exceptions that may happen when the user-selected encoding does not match the real encoding of the CSV (or it contains otherwise illegal byte sequences). We recently encountered `Encoding::UndefinedConversionError`, which is not yet caught by that error handling. The attached patches add a test case and a fix for that.

**Associated revisions**

Revision 16572 - 2017-05-27 10:40 - Jean-Philippe Lang

Rescue any `EncodingError` exceptions (#25861).

Revision 16617 - 2017-06-06 22:56 - Jean-Philippe Lang

Merged r16572 (#25861).

Revision 16618 - 2017-06-06 22:56 - Jean-Philippe Lang

Merged r16572 (#25861).

**History**

#1 - 2017-05-16 11:50 - Go MAEDA
- Target version set to 3.2.7

We can reproduce the problem with the test in attachment:0001-adds-failing-test-case-for-import-with-wrong-encoding.patch. Setting target version to 3.2.7.

By the way, the encoding of the file `invalid-Shift_JIS.csv` is valid CP932. Shift_JIS and CP932 are almost identical, but CP932 includes more characters.

```bash
$ iconv -f cp932 -t utf8 test/fixtures/files/invalid-Shift_JIS.csv
```

#2 - 2017-05-27 09:51 - Jean-Philippe Lang
- Status changed from New to Needs feedback

Go MAEDA wrote:
We can reproduce the problem with the test in attachment:0001-adds-failing-test-case-for-import-with-wrong-encoding.patch. Setting target version to 3.2.7.

The test passes for me with ruby 2.0 (I would get a Encoding::InvalidByteSequenceError). Maybe we could simply rescue EncodingError which is the base class for encoding errors?

#3 - 2017-05-27 09:58 - Go MAEDA

Jean-Philippe Lang wrote:

> Maybe we could simply rescue EncodingError which is the base class for encoding errors?

I agree, caching EncodingError is better. We can simplify the code and catch all errors caused by iconv.

#4 - 2017-05-27 10:41 - Jean-Philippe Lang
- Status changed from Needs feedback to Resolved
- Assignee set to Jean-Philippe Lang

#5 - 2017-06-06 22:56 - Jean-Philippe Lang
- Status changed from Resolved to Closed

Files

<table>
<thead>
<tr>
<th>Patch Name</th>
<th>Size</th>
<th>Date</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>0001-adds-failing-test-case-for-import-with-wrong-encoding.patch</td>
<td>1.69 KB</td>
<td>2017-05-16</td>
<td>Jens Krämer</td>
</tr>
<tr>
<td>0002-adds-rescue-from-Encoding-UndefinedConversionError.patch</td>
<td>1.05 KB</td>
<td>2017-05-16</td>
<td>Jens Krämer</td>
</tr>
</tbody>
</table>