

TextCritical.net - Bug #578

Unicode decode error prevents work from being opened

03/15/2013 12:00 AM - Luke Murphey

Status:	Closed	Start date:	03/14/2013
Priority:	Urgent	Due date:	
Assignee:	Luke Murphey	% Done:	100%
Category:		Estimated time:	0.00 hour
Target version:	0.7		
Description			
To view the error, go to: http://textcritical.com/work/the-extant-works-of-aretaeus-the-cappadocian			
<pre>2013-03-09 07:07:08,557 CST [ERROR] django.request: Internal Server Error: /work/the-extant-works-of-aretaeus-the-cappadocian Traceback (most recent call last): File "/usr/local/lib/python2.7/site-packages/django/core/handlers/base.py", line 111, in get_response response = callback(request, *callback_args, **callback_kwargs) File "/usr/local/lib/python2.7/site-packages/django/utils/decorators.py", line 91, in _wrapped_view response = view_func(request, *args, **kwargs) File "/var/webapps/textcritical.com/reader/views.py", line 321, in read_work context_instance=RequestContext(request)) File "/usr/local/lib/python2.7/site-packages/django/shortcuts/__init__.py", line 20, in render_to_response return HttpResponse(loader.render_to_string(*args, **kwargs), **httpresponse_kwargs) File "/usr/local/lib/python2.7/site-packages/django/template/loader.py", line 176, in render_to_string return t.render(context_instance) File "/usr/local/lib/python2.7/site-packages/django/template/base.py", line 140, in render return self._render(context) File "/usr/local/lib/python2.7/site-packages/django/template/base.py", line 134, in _render return self.nodelist.render(context) File "/usr/local/lib/python2.7/site-packages/django/template/base.py", line 823, in render bit = self.render_node(node, context) File "/usr/local/lib/python2.7/site-packages/django/template/debug.py", line 74, in render_node return node.render(context) File "/usr/local/lib/python2.7/site-packages/django/template/loader_tags.py", line 123, in render return compiled_parent._render(context) File "/usr/local/lib/python2.7/site-packages/django/template/base.py", line 134, in _render return self.nodelist.render(context) File "/usr/local/lib/python2.7/site-packages/django/template/base.py", line 823, in render bit = self.render_node(node, context) File "/usr/local/lib/python2.7/site-packages/django/template/debug.py", line 74, in render_node return node.render(context) File "/usr/local/lib/python2.7/site-packages/django/template/loader_tags.py", line 123, in render return compiled_parent._render(context) File "/usr/local/lib/python2.7/site-packages/django/template/base.py", line 134, in _render return self.nodelist.render(context) File "/usr/local/lib/python2.7/site-packages/django/template/base.py", line 823, in render bit = self.render_node(node, context) File "/usr/local/lib/python2.7/site-packages/django/template/debug.py", line 74, in render_node return node.render(context) File "/usr/local/lib/python2.7/site-packages/django/template/loader_tags.py", line 62, in render result = block.nodelist.render(context) File "/usr/local/lib/python2.7/site-packages/django/template/base.py", line 823, in render bit = self.render_node(node, context) File "/usr/local/lib/python2.7/site-packages/django/template/debug.py", line 74, in render_node return node.render(context) File "/usr/local/lib/python2.7/site-packages/django/template/loader_tags.py", line 62, in render</pre>			

```

    result = block.nodelist.render(context)
File "/usr/local/lib/python2.7/site-packages/django/template/base.py", line 823, in render
    bit = self.render_node(node, context)
File "/usr/local/lib/python2.7/site-packages/django/template/debug.py", line 74, in render_node
    return node.render(context)
File "/usr/local/lib/python2.7/site-packages/django/template/defaulttags.py", line 476, in rende
r
    output = self.nodelist.render(context)
File "/usr/local/lib/python2.7/site-packages/django/template/base.py", line 823, in render
    bit = self.render_node(node, context)
File "/usr/local/lib/python2.7/site-packages/django/template/debug.py", line 74, in render_node
    return node.render(context)
File "/usr/local/lib/python2.7/site-packages/django/template/defaulttags.py", line 185, in rende
r
    nodelist.append(node.render(context))
File "/usr/local/lib/python2.7/site-packages/django/template/defaulttags.py", line 281, in rende
r
    return nodelist.render(context)
File "/usr/local/lib/python2.7/site-packages/django/template/base.py", line 823, in render
    bit = self.render_node(node, context)
File "/usr/local/lib/python2.7/site-packages/django/template/debug.py", line 74, in render_node
    return node.render(context)
File "/usr/local/lib/python2.7/site-packages/django/template/debug.py", line 84, in render
    output = self.filter_expression.resolve(context)
File "/usr/local/lib/python2.7/site-packages/django/template/base.py", line 599, in resolve
    new_obj = func(obj, *arg_vals)
File "/var/webapps/textcritical.com/reader/templatetags/reader_extras.py", line 87, in perseus_x
ml_to_html5
    converted_doc = convert_xml_to_html5( value, language=language, text_transformation_fx=text_tr
ansformation_fx, node_transformation_fx=transform_perseus_node )
File "/var/webapps/textcritical.com/reader/shortcuts/__init__.py", line 159, in convert_xml_to_h
tml5
    converter.feed(xml_str)
File "/usr/local/lib/python2.7/HTMLParser.py", line 114, in feed
    self.goahead(0)
File "/usr/local/lib/python2.7/HTMLParser.py", line 152, in goahead
    if i < j: self.handle_data(rawdata[i:j])
File "/var/webapps/textcritical.com/reader/shortcuts/__init__.py", line 101, in handle_data
    transformed_text = self.text_transformation_fx( text=data, parent_node=self.current_node, dst_
doc=self.dst_doc )
File "/var/webapps/textcritical.com/reader/templatetags/reader_extras.py", line 85, in <lambda>
    text_transformation_fx = lambda text, parent_node, dst_doc: transform_perseus_text(text, paren
t_node, dst_doc, language)
File "/var/webapps/textcritical.com/reader/templatetags/reader_extras.py", line 179, in transfor
m_perseus_text
    txt_node = dst_doc.createTextNode( transform_text(s, language).decode( "utf-8" ) )
File "/usr/local/lib/python2.7/encodings/utf_8.py", line 16, in decode
    return codecs.utf_8_decode(input, errors, True)
UnicodeEncodeError: 'ascii' codec can't encode character u'\xe6' in position 2: ordinal not in ran
ge(128)

```

Associated revisions

Revision 437 - 03/24/2013 04:02 AM - Luke Murphey

Fixed issue where some works were not displayed properly because the text we being incorrectly encoded and decoded. Closes #578.

Revision 437 - 03/24/2013 04:02 AM - Luke Murphey

Fixed issue where some works were not displayed properly because the text we being incorrectly encoded and decoded. Closes #578.

Revision 427 - 03/24/2013 04:02 AM - Luke Murphey

Fixed issue where some works were not displayed properly because the text we being incorrectly encoded and decoded. Closes #578.

History

#1 - 03/24/2013 03:08 AM - Luke Murphey

- *Status changed from New to In Progress*

#2 - 03/24/2013 03:35 AM - Luke Murphey

This work contains words that are not ASCII: prācipu and quæ.

I recall dealing with this same issue during the import process. I believe it was resolved by loading the text as UTF-8.

#3 - 03/24/2013 04:15 AM - Luke Murphey

- *Status changed from In Progress to Closed*

- *% Done changed from 0 to 100*

Applied in changeset [r437](#).