

# TextCritical.net - Bug #2600

## Search fails

02/12/2020 12:28 AM - Luke Murphey

<b>Status:</b>	Closed	<b>Start date:</b>	02/11/2020
<b>Priority:</b>	Normal	<b>Due date:</b>	
<b>Assignee:</b>	Luke Murphey	<b>% Done:</b>	100%
<b>Category:</b>		<b>Estimated time:</b>	0.00 hour
<b>Target version:</b>	3.2		
<b>Description</b>			

### Associated revisions

#### Revision 1084 - 02/13/2020 06:54 PM - luke.murphey

Fixing search issues

Reference #2600

#### Revision 1089 - 02/14/2020 08:10 PM - luke.murphey

Fixing issue where bytes stored in the index causes issues

Reference #2600

### History

#### #1 - 02/12/2020 05:51 PM - Luke Murphey

```
Not Found: /api/search/work:berean-bible section:"Genesis 1" beginning
Performing a search, page=1, page_len=10, include_related_forms=False, search_query="work:berean-bible beginning"
Search query parsed, default_search_field="content", raw_query="(work:<berean-bible> AND content:<beginning>)"

Looking for variations of the search term in order to perform a search, word=beginning
Unable to find work in matched terms, work_slug=b'berean-bible'
Performing a stats search, limit=2000, include_related_forms=False, search_query="work:berean-bible beginning"
Search query parsed, default_search_field="content", raw_query="(work:<berean-bible> AND content:<beginning>)"

Looking for variations of the search term in order to perform a search, word=beginning
Internal Server Error: /api/search/work:berean-bible beginning
Traceback (most recent call last):
  File "/usr/local/lib/python3.7/site-packages/django/core/handlers/exception.py", line 41, in inner
    response = get_response(request)
  File "/usr/local/lib/python3.7/site-packages/django/core/handlers/base.py", line 249, in _legacy_get_response
    response = self._get_response(request)
  File "/usr/local/lib/python3.7/site-packages/django/core/handlers/base.py", line 187, in _get_response
    response = self.process_exception_by_middleware(e, request)
  File "/usr/local/lib/python3.7/site-packages/django/core/handlers/base.py", line 185, in _get_response
    response = wrapped_callback(request, *callback_args, **callback_kwargs)
  File "/usr/local/lib/python3.7/site-packages/django/utils/decorators.py", line 149, in _wrapped_view
    response = view_func(request, *args, **kwargs)
  File "/usr/src/app/reader/views.py", line 857, in api_search
    return render_api_response(request, results_set)
  File "/usr/src/app/reader/views.py", line 636, in render_api_response
    raw_content = json.dumps(content)
  File "/usr/local/lib/python3.7/json/__init__.py", line 231, in dumps
    return _default_encoder.encode(obj)
  File "/usr/local/lib/python3.7/json/encoder.py", line 199, in encode
    chunks = self.iterencode(o, _one_shot=True)
```

```
File "/usr/local/lib/python3.7/json/encoder.py", line 257, in iterencode
    return _iterencode(o, 0)
TypeError: keys must be str, int, float, bool or None, not bytes
```

Qs:

- How am I handling the unicode with the search functionality?
- **[Found]** What field is wrong in views.py?
  - matched\_terms\_verses
  - matched\_terms\_no\_diacritics
- **[None]** What field in search\_stats() is wrong?
  - stats['matches']
  - stats['matched\_works']
  - stats['matched\_terms']
  - stats['results\_count']
- **[Found]** What is the line where the failure is?
- **[No]** Do my tests show that string normalization returns bytes?
- **[No]** Do any of my tests return bytes?
- Search indexing providing bytes?

Obs:

- A search for the following works when diacritics are set to be ignored: work:berean-bible "the beginning"
- The issue likely comes from:
  - VerseSearchResults::\_\_init\_\_()
    - results parameter comes from searcher.search\_page()
      - This comes from Whoosh. It seems like Whoosh is returning bytes
    - I commented out matched\_terms and matched\_terms\_no\_diacritics in VerseSearchResults::\_\_init\_\_(); this didn't seem to change functionality?!
      - It turns out that I'm not even using this functionality
- This looks similar: <https://stackoverflow.com/questions/59006409/haystack-whoosh-convert-string-to-bytes-like-object-cannot-use-a-string-patter>
- The stack is:
  - src/reader/views.py::api\_search
  - src/reader/views.py::search\_verses
  - VerseSearchResults::\_\_init\_\_()
    - searcher.search\_page()

**#2 - 02/13/2020 06:55 PM - Luke Murphey**

- *Status changed from New to Closed*

- *% Done changed from 0 to 100*