

## Ruby master - Bug #17029

### URI.parse considers https://example.com/### invalid when browsers consider it valid

07/13/2020 05:33 PM - Nileshtr (Nilesh Trivedi)

<b>Status:</b>	Closed		
<b>Priority:</b>	Normal		
<b>Assignee:</b>			
<b>Target version:</b>			
<b>ruby -v:</b>	ruby 2.7.1p83 (2020-03-31 revision a0c7c23c9c) [x86_64-darwin19]	<b>Backport:</b>	2.5: UNKNOWN, 2.6: UNKNOWN, 2.7: UNKNOWN

#### Description

I have a form with `<input type="url" required>` and in the backend, I try to extract the domain with `URI.parse(url).host`

A user was able to submit a value like `https://example.com/###` which passed the browser's validation check, but failed by `URI.parse` with this error:

```
3: from /Users/helix/.rbenv/versions/2.7.1/lib/ruby/2.7.0/uri/common.rb:234:in `parse'
2: from /Users/helix/.rbenv/versions/2.7.1/lib/ruby/2.7.0/uri/rfc3986_parser.rb:73:in `parse'
1: from /Users/helix/.rbenv/versions/2.7.1/lib/ruby/2.7.0/uri/rfc3986_parser.rb:67:in `split'
URI::InvalidURIError (bad URI(is not URI?): "https://example.com/###")
```

You can try the browser's behavior at MDN's demo: <https://developer.mozilla.org/en-US/docs/Web/HTML/Element/input/url>

This is what the MDN page says about validation:

The syntax of a URL is fairly intricate. It's defined by WHATWG's URL Living Standard ( <https://url.spec.whatwg.org/> ) and is described for newcomers in our article [What is a URL?](https://developer.mozilla.org/en-US/docs/Learn/Common_questions/What_is_a_URL) ( [https://developer.mozilla.org/en-US/docs/Learn/Common\\_questions/What\\_is\\_a\\_URL](https://developer.mozilla.org/en-US/docs/Learn/Common_questions/What_is_a_URL) )

#### History

##### #1 - 07/13/2020 05:38 PM - jeremyevans0 (Jeremy Evans)

This does seem like a bug to me. It looks like `https://example.com/###` should be a valid URL with a fragment of `##`. However, the `uri` library is maintained in a separate repository. Please submit this as an issue to <https://github.com/ruby/uri/issues/new>.

##### #2 - 07/13/2020 09:01 PM - phluid61 (Matthew Kerwin)

It's not valid according to RFC 3986 (the URI standard) but that is pretty old now. I suspect switching from the IETF URI spec to the WHATWG URL spec would have other consequences, too.

##### #3 - 07/18/2020 03:26 AM - Nileshtr (Nilesh Trivedi)

I filed an issue at the `uri` library's Github repo: <https://github.com/ruby/uri/issues/8>

##### #4 - 07/22/2020 05:46 PM - jeremyevans0 (Jeremy Evans)

- Status changed from Open to Closed