How to calculate SHA from a string with the Github api in python?

Asked 4 years, 6 months ago   Active 4 years, 6 months ago   Viewed 1k times

▲

1

▼

🔖

2

↺

I want to update a file with the Github API and commit it into a branch. I have troubles creating the commit. The SHA does not match the expected one.

```
{
    'documentation_url':
'https://developer.github.com/enterprise/2.7/v3/repos/contents/',
    'message': 'pom.xml does not match
de42fdd980f9b8067a2af982de46b8d5547e4597'
}
```

I do the following:

```
import hashlib
myfile = "new content of my README"
resulting_file = base64.b64encode(bytes(myfile, "utf-8"))
file_as_str = str(resulting_file.decode('utf-8'))
sha = hashlib.sha1(file_as_str.encode('utf-8')).hexdigest()

url = 'https://someurl.com/someproject/contents/README.md?access_token=' +
access_token
data = '{"message": "bla bla", "content": "'+file_as_str+'", "sha": "'+sha+'",
"branch": "'+branch+'"}'
response = requests.put(url, data=data)
```

I would not like to use a lib to do this to better understand what is happening. probably the SHA is not generated properly, but I cannot identify why. Could someone help?

python   git   github-api

Share Follow

edited Oct 5 '16 at 14:20

kfb
5,272 ● 6 ● 32 ● 46

asked Oct 5 '16 at 12:06

Paul Fournel
8,821 ● 9 ● 31 ● 62

are you sure you need to checksum the base64 encoded version of the file? – Jean-François Fabre ♦ Oct 5 '16 at 12:10

2 Answers

Active | Oldest | Votes

Your privacy

By clicking "Accept all cookies", you agree Stack Exchange can store cookies on your device and disclose information in accordance with our Cookie Policy.

Accept all cookies   Customize settings

Then use the value of `sha` in the `PUT` request to update the file:

```python
with open('myfile', 'rb') as f:
    content = str(base64.b64encode(f.read()), encoding='utf8')
    data = {'message': 'bla bla', 'content': content, 'sha': sha, 'branch': branch}
    r = requests.put(url, json=data)
```

Share Follow

edited Oct 5 '16 at 13:15

answered Oct 5 '16 at 13:10

mhawke
74.3k ● 8 ● 90 ● 120

Thx, for the answer. Now I have 2 different interpretation of the api. None of them seem to work. When I get the SHA as you showed, I still get an error. – Paul Fournel Oct 5 '16 at 14:15

One thing to be aware of when using grabbing the SHA via the file contents API is that the API only works for files up to one megabyte in size, so will fail if the file is larger than that. I expect the `pom.xml` in the original question is smaller than this, but generalising to other files might be problematic. – kfb Oct 5 '16 at 14:19

@PaulFournel: What error do you get? I can assure you that you must use the sha of the file that is being replaced, not the sha of the new content. I didn't show it, but you also need to add authentication to the `put` request - either in the HTTP header, or in the query string. – mhawke Oct 6 '16 at 11:35

@PaulFournel: Also check that you have set the appropriate permissions for the personal access token that you are using for the update request. You need to have at least "Access public repositories " in the "repo" scope enabled. – mhawke Oct 6 '16 at 11:37

I finally got it working. I was using the SHA from master and trying to push to a branch... Could not work. It works now. – Paul Fournel Oct 6 '16 at 11:38

1

GitHub calculates hashes as followes:

```
sha1("blob " + filesize + "\0" + data)
```

so use this code:

```python
with open(filepath, 'rb') as file_for_hash:
    data = file_for_hash.read()
    filesize = len(data)
sha = hashlib.sha1("blob " + filesize + "\0" + data).hexdigest()
```

Share Follow

answered Oct 5 '16 at 12:40

RFV
691 ● 5 ● 18