

```
In [ ]: import requests
        from bs4 import BeautifulSoup
        import re
        import unicodedata
```

code for writing xml from Elsevier papers through paper's doi

```
In [ ]: import requests
        dois = ['10.1016/j.apsusc.2016.12.052'] # make list of all dois available
        j = 1
        for i, doi in enumerate(dois): # here queries is the list of dois for writing xml
            xml_url = 'https://api.elsevier.com/content/article/doi/' + doi + '?view=FULL'

            headers = {
                'X-ELS-APIKEY': '410695f2c877b1beb1a44d0b4c363ceb', # api keys from elsevier site
                'Accept': 'text/xml'
            }

            r = requests.get(xml_url, stream = True, headers=headers, timeout=30)
            print(r.status_code)
            if r.status_code == 200: # to check either request has been succeeded or not
                try:
                    writefile = open("G:\\publisherless_elsvier_xml\\xml_" + str(j) + '.xml', 'wb')
                    j = j+1

                    for chunk in r.iter_content(2048): # writing of content of paper to xml file
                        writefile.write(chunk)
                except:
                    continue
```

code for writing html from Nature's papers through paper's doi

```
In [ ]: dois = ['10.1557/mrc.2020.34', '10.1557/mrc.2018.208']
        j = 0
        for i, doi in enumerate(dois):

            download_url = 'https://doi.org/' + doi
            headers = {
                'Accept': 'text/html',
                "User-Agent": "Mozilla/5.0 (Windows NT 10.0; Win64; x64)"
            }
            r = requests.get(download_url, stream=True, headers=headers, timeout=60)
            print(i, r.status_code)

            if r.status_code == 200:
                writefile = open("G:\\New folder\\springer_papers\\springer_" + str(j) + '.html', 'wb')
                j = j+1
                try:
                    for chunk in r.iter_content(2048):
                        writefile.write(chunk)
                except:
                    continue
```