

BabelNet

March 2021

1 Overview of BabelNet

- BabelNet is a multilingual lexicalized semantic network and ontology. It was automatically created by linking the largest multilingual Web encyclopedia – i.e. Wikipedia to the most popular computational lexicon of the English language – i.e. WordNet [1].
- It contains an API for programmatic access of 5.5 million concepts and a multilingual knowledge-rich Word Sense Disambiguation (WSD). With the aid of this API, we can extract all possible concepts in a piece of text [2].
- BabelNet offers HTTP and Java API as well as the RDF and SPARQL endpoints [3].
- BabelNet offers a better API in Java [4][1] and the founder has devised a BabelNet extract upon java API that extracts the synset with its relations for depth of 3.
- Mostly the BabelNet is operated using a JAVA API compared to other API [3][4][1].
- One can find the BabelNet API guide in the official website [5]
- Multilingual text analytics to get the gist of any text, you can copy paste the content of any language it can translate and give you the meaning in the preferred language [6]

2 Working with BabelNet

- Register in BabelNet [Link]
- Babelcoin: Used as an internal credit system to keep track of the request made against API
 - 1 Babelcoin = 1 Query to BabelNet

– Base account-1000 Babelcoins/day

- For research and non-commercial usage Babel indices can be downloaded upon submitting the request form describing the indented use.
- The API available in open source is HTTP and JAVA API
- The key should be append each time making query using HTTP.

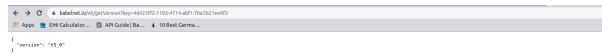


Figure 1: HTTP Version request

- Since it is HTTP, it can be code from Python too [5].

```
jupyter BabelNet_querying Last Checkpoint: 2 minutes ago (autosaved)
File Edit View Insert Cell Kernel Widgets Help Trusted Python 3

In [1]: 1 import urllib3.request
        2 import urllib
        3 import json
        4 import gzip
        5
        6 from io import StringIO, BytesIO

Print ID

In [2]: 1 service_url = 'https://babelnet.io/v6/getSynsetIds'
        2
        3 lemma = 'apple'
        4 lang = 'en'
        5 key = '4d425ff2-1193-4714-abf1-70e2b21ee9f3'
        6
        7 params = {
        8     'lemma': lemma,
        9     'searchLang': lang,
       10     'key': key
       11 }
       12
       13 url = service_url + '?' + urllib.parse.urlencode(params)
       14 print(url)
       15 request = urllib.request.Request(url)
       16 request.add_header('Accept-encoding', 'gzip')
       17 response = urllib.request.urlopen(request)
       18
       19 if response.info().get('Content-Encoding') == 'gzip':
       20     buf = BytesIO(response.read())
       21     f = gzip.GzipFile(fileobj=buf)
       22     data = json.loads(f.read())
       23     for result in data:
       24         print(result['id'])

https://babelnet.io/v6/getSynsetIds?lemma=apple&searchLang=en&key=4d425ff2-1193-4714-abf1-70e2b21ee9f3
bn:00289737n
bn:26558969n
bn:00005054n
bn:00005076n
bn:27110089v
bn:00015676n
bn:00319426n
bn:14792761n
bn:03174949n
bn:00353687n
bn:03740018n
```

Figure 2: Python code for querying using HTTP

- Python code is provided for version-2 which has to be converted to 3 as shown in the above Figure 2.

References

- [1] R. Navigli and S. P. Ponzetto, “Babelnet: The automatic construction, evaluation and application of a wide-coverage multilingual semantic network,” *Artificial Intelligence*, vol. 193, pp. 217–250, 2012.
- [2] A. H. Jadidinejad, “Unsupervised information extraction using babelnet and dbpedia,” in *#MSM*, 2013.
- [3] D. Ustalov and A. Panchenko, “A tool for effective extraction of synsets and semantic relations from babelnet,” pp. 10–13, 04 2017.
- [4] R. Navigli and S. Ponzetto, “Multilingual wsd with just a few lines of code: the babelnet api,” pp. 67–72, 07 2012.
- [5] BabelNet, “Babelnet api guide,” 2022. Accessed on: 2022-02-08. [Online].
- [6] Babelspace, “Extraggo,” 2022. Accessed on: 2022-02-24. [Online].