

Using the API

The Swissdox@LiRI API allows you to submit a query, to check the status of submitted queries and to download the retrieved data.

Endpoint	Description
/query	Endpoint for submitting a query. Required parameters are name and query . Parameter test can be used to check if the query is correct. Endpoint will return id of a submitted query.
/status	Returns list of all submitted queries.
/status/<query_id>	Returns status of query with specific id query_id .
/download/<filename>	Download of the retrieved dataset.

For using the API, you first need to create an API key in the Swissdox@LiRI web application (page **Projects**) for each project separately. To pass the API key to the server you need to specify the **X-API-Key** and **X-API-Secret** headers.

Submitting a query

The query is defined using a YAML format. Optional arguments are query name, comment, expiration date and a flag to specify whether the query should be run or not (for syntax checking). Below you find a simple example in Python for submitting a query using the API.

```
import requests

headers = {
    "X-API-Key": "<your-api-key>",
    "X-API-Secret": "<your-api-secret>"
}
API_BASE_URL = "https://swissdox.linguistik.uzh.ch/api"
API_URL_QUERY = f"{API_BASE_URL}/query"

yaml_example = """
query:
  sources:
    - ZWA
    - ZWAS
  dates:
    - from: 2022-12-01
      to: 2022-12-31
  languages:
    - de
    - fr
  content:
    AND:
      - OR:
          - COVID
```

```
        - Corona
        - NOT: China
        - NOT: chin*
result:
  format: TSV
  maxResults: 100
  columns:
    - id
    - pubtime
    - medium_code
    - medium_name
    - rubric
    - regional
    - doctype
    - doctype_description
    - language
    - char_count
    - dateline
    - head
    - subhead
    - content_id
    - content
  version: 1.2
"""

data = {
  "query": yaml_example,
  "test": "1",
  "name": "Query name 1",
  "comment": "Query comment",
  "expirationDate": "2023-02-28"
}

r = requests.post(
  API_URL_QUERY,
  headers=headers,
  data=data
)
print(r.json())
```

Checking the status of submitted queries

It is possible to check the status of all submitted queries, as well as status of a certain query with a specific id. The following example shows how to list all submitted queries with their respective statuses:

```
import requests

headers = {
```

```
"X-API-Key": "<your-api-key>",
"X-API-Secret": "<your-api-secret>"
}
API_BASE_URL = "https://swissdox.linguistik.uzh.ch/api"
API_URL_STATUS = f"{API_BASE_URL}/status"

r = requests.get(
    API_URL_STATUS,
    headers=headers
)
print(r.json())
```

Download of the retrieved dataset

When you were checking status of your query (like shown in the example above), you also got a download URL in a response, for those queries which are completed. By using this URL in an API request you are able to download your dataset:

```
import requests

headers = {
    "X-API-Key": "<your-api-key>",
    "X-API-Secret": "<your-api-secret>"
}
API_BASE_URL = "https://swissdox.linguistik.uzh.ch/api"
API_URL_DOWNLOAD = f"{API_BASE_URL}/download/6399ae50-0d80-4304-9d2a-
d92fa3dc753c__2022_02_28T10_26_03.tsv.xz"

r = requests.get(
    API_URL_DOWNLOAD,
    headers=headers
)
if r.status_code == 200:
    print("Size of file: %.2f KB" % (len(r.content)/1024))
    fp = open("./dataset.tsv.xz", "wb")
    fp.write(r.content)
    fp.close()
else:
    print(r.text)
```

From:
<https://liri.linguistik.uzh.ch/wiki/> - **LiRI Wiki**

Permanent link:
<https://liri.linguistik.uzh.ch/wiki/langtech/swissdox/api?rev=1679500708>

Last update: **2023/03/22 15:58**

