

HASK HTTP API Documentation

Piotr Pęzik
pelcra.pl

10 grudnia 2012

1 About this document

This document explains the HTTP access to the HASK collocation dictionary service. Currently, we support basic XML access whereby users specify query parameters in POST and GET requests which are processed by the server to return XML-encoded results from the dictionaries.

This document describes the use of three basic HASK API servlets:

1. EntryTypesXML
2. EntryXML
3. CollocationXML

Authentication is very simple at the moment; the correct login credentials are passed in the user and pass parameters. Demo login details are given below.

user	public
pass	6cc8e93f7a657b484aca56d70f9a7f19

2 Naming conventions

- HTTP_SERVICE_URL is the full HTTP address of the HASK API service, for example:
 - `http://pelcra.pl/hask_pl/`
 - `http://pelcra.pl/hask_en/`

3 EntryXML – retrieving entry information

The following request:

```
HTTP_SERVICE_URL/EntryTypesXML?user=public&pass=6cc8e93f7a657b484aca56d70f9a7f19
```

returns an XML version of the current list of collocations in the dictionary. An example single rule section of the returned XML is given below:

```
<rule n="1">
  <rule_id>1</rule_id>
  <rule_name>N_J</rule_name>
  <node_pos>noun</node_pos>
  <collocate_pos>adjective</collocate_pos>
  <collocate_positions>1</collocate_positions>
  <total_entries>16383</total_entries>
  <total_collocates>267587</total_collocates>
  <description>Nouns immediately preceded by adjectives
    </description>
</rule>
```

4 EntryXML – retrieving entry information

The following request:

```
HTTP_SERVICE_URL/EntryXML?user=public
&pass=6cc8e93f7a657b484aca56d70f9a7f19
&entry_node=finger&rule_id=1&offset=0&limit=20
```

returns an XML list of the first 50 adjectival collocates of the noun finger (rule_id=1).

An example single collocate result in the list is shown below:

```
<col hash="447bfea76114124d3841aa86e9f8588f">
  <collocate>little</collocate>
  <c_forms>
    <c_form>little</c_form>
  </c_forms>
  <pos>N%</pos>
  <a>1277.0</a>
  <ttest>31.31532</ttest>
  <mi3>23.65237</mi3>
  <g2>3218.68433189392</g2>
  <mi>3.0152851248144</mi>
```

```

    <loglog>31.1133486438572</loglog>
    <chisq>8041.36857585891</chisq>
    <jd>0.79455</jd>
    <range>82.0</range>
    <awt>581922.0</awt>
    <fawt>93.41371</fawt>
</col>

```

Collocation measure	Explanation
a	plain frequency of co-occurrence
ttest	association strength measure
mi3	association strength measure
g2	association strength measure
mi	association strength measure
loglog	association strength measure
chisq	association strength measure
jd	dispersion measure (0 to 1)
range	dispersion measure (0 to 100)
awt	dispersion measure (Average Waiting Time in words)
fawt	dispersion measure (frequency scaled by AWT)

5 CollocationXML

The following request:

```

HTTP_SERVICE_URL/CollocationXML?user=public&pass=6cc8e93f7a657b484aca56d70f9a7f19
&col_hash=447bfea76114124d3841aa86e9f8588f&conc_limit=50

```

returns a single XML-encoded collocation record. The value of the col_hash parameter is obtained through the EntryXML service described above. A maximum of 50 concordances are available for a single collocation entry.