SAPIS

StilLett API Service

The original `SAPIS` (Fahlborg and Rennes, 2016), was a restful web service based on Java Spring, implementing an API to various component to measure text complexity and automatic text simplification. The serivce has since then been rewritten in Python3, with similar functionality.

The API service can be reached from a server located at: https://sapis.it.liu.se/api, accepting POST requests.

Current components

1. The SurfaceMetrics (Falkenjack et al., 2013) service provides simple readability metrics such as LIX, OVIX, Nominal ratio, Mean sentence length and Mean word length.
2. The LexicalMetrics (Falkenjack et al., 2013) service provides a categorized frequency analysis from word occurrences in the basic Swedish vocabulary SweVoc dictionary.
3. The StructuralMetrics (Falkenjack et al., 2013) provides syntactic and morpho-syntactic features based on part-of-speech tags and dependency tags (MaltParser).
4. The Referential Cohesion, LSA, and Connectives indicies from Coh-Metrix (Graesser et al., 2004). Note that the metric *givenness* (listed among the LSA-indicies) follow the implementation described in TAACO (Crossley et. al., 2016). See http://cohmetrix.memphis.edu/cohmetrixhome/documentation_indices.html for an overview of the indices.
5. StilLett (Rennes and Jönsson, 2015), a rule-based automatic text simplification tool for Swedish, using part-of-speech tags and phrase structure tags (MaltParser 1.2) to identify textual difficulties and execute simplifications as node operations.
6. (Work in progress) A synonym replacement module, that identifies complex words in a text and provide a less complex synonym.

HTTP request

The output response is a JSON object containing all possible variables (Note that some values are not always calculated. For example, metrics that consider multiple paragraphs will not be calculated for texts with a single paragraph). A suggestion is to print the response of the HTTP request to the console of your browser (in javascript by using console.log()), or with the print()-function in Python, in order to easily investigate the response JSON object and the values that is calculated.

Python example

This example require the requests library to be installed. The easiest way to install is by running `pip install requests` or `conda install -c anaconda requests` in your virtual environment.

```python
import requests


def sapis_call(text):
    """
    call the sapis service

    args
        text: a string
    return
        result: dict with results from different SAPIS-modules
    """

    url = "https://sapis.it.liu.se/api"

    r = requests.post(url, data=text.encode('utf-8'))
    result = r.json()[0]

    return result



sapis_call("En liten text att testa.")
```

JavaScript example

This is an HTTP request, using jQuery.ajax, to `SAPIS`

```javascript
var sapis_url = "https://sapis.it.liu.se/api";
var scream_response;
jQuery.ajax({
    headers: {
        'Accept': 'application/json',
        'Content-Type': 'application/json'
    },
    'type': 'POST',
    'url': sapis_url,
    'data': JSON.stringify({
        document : [input text as string]
    }),
    'dataType': 'json',
    'success': function(resp) {
        console.log(resp);
```

```
        scream_response = resp;
    },
    'error': function(xhr, textStatus, errorThrown) {
        alert(xhr.responseText);
    }
});
```

Web interface (Work in progress)

We also provide STePS, a web interface to the text complexity metrics included in `SAPIS`. It is possible to either paste a text or upload a text file for analysis. The results can be downloaded as a .json file, which easily can be read to a dataframe in either Python (with Pandas) or R for statistical analysis. The serivce can be found at https://steps-nlplab.web.app.

Contact

Daniel Holmer, daniel.holmer@liu.se, for technical support.

References

Scott A. Crossley, Kristopher Kyle & Danielle S. McNamara. 2016. The tool for the automatic analysis of text cohesion (TAACO): Automatic assessment of local, global, and text cohesion. Behav Res 48, 1227–1237

Daniel Fahlborg and Evelina Rennes. 2016. Introducing SAPIS-an API service for text analysis and simplification. In The second national Swe-Clarin workshop: Research collaborations for the digital age, Umeå, Sweden.

Johan Falkenjack, Katarina Heimann Mühlenbock, and Arne Jönsson. 2013. Features indicating readability in Swedish text. In Proceedings of the 19th Nordic Conference of Computational Linguistics (NoDaLiDa-2013), Oslo, Norway.

Arthur C. Graesser, Danielle S. McNamara, Max M. Louwerse, and Zhiqiang Cai. 2004. Coh-Metrix: Analysis of text on cohesion and language. Behavior Research Methods, Instruments, & Computers, 36(2):193–202

Evelina Rennes and Arne Jönsson. 2015. A tool for automatic simplification of Swedish texts. In Proceedings of the 20th Nordic Conference of Computational Linguistics (NoDaLiDa-2015), Vilnius, Lithuania.