

732A94 Advanced R Programming

Computer lab 5: API R projects

Krzysztof Bartoszek, Bayu Brahmantio, Marc Braun, Arash Haratian
(designed by Leif Jonsson, Måns Magnusson and Héctor Rodríguez Déniz)
(updates and converted to L^AT_EX by Arash Haratian))

14 September 2023

Seminar date: **28 September 10:15** (S41)
Lab deadline: **30 September 23:59**

Below are some API that can be used as a part of the assignment to implement an R package to connect to an API. In connection to each API there is a short description and the difficulty of implementation is indicated using one, two or three stars with one star (simpler) to three stars (difficult). Some projects are about implementing XML parsers for more complex XML files of interest.

In general there exist many different web API:s and you can choose any API you may find interesting. Most web API:s in Sweden can be found in [API-katalogen](http://apikatalogen.se/api) (<http://apikatalogen.se/api>). If you choose any of these other than the ones listed below you need to confirm your choice with the teacher, the only restriction is that the API should not already have an R package.

* **Valmyndigheten XML** After each election the Swedish election agency (Valmyndigheten) release a detailed file with all election results in an XML format. Implement a package to get the election data (that is published as excel files) out of these files direct.

Historical information can be found [here](https://historik.val.se/val/val2014/statistik/index.html) (<https://historik.val.se/val/val2014/statistik/index.html>).
Current data can be found [here](https://www.val.se/valresultat/riksdag-region-och-kommun/2022/radata-och-statistik.html) (<https://www.val.se/valresultat/riksdag-region-och-kommun/2022/radata-och-statistik.html>).

* **OpenStreetMaps via Stamen Tiles** is a free alternative to the Google Maps API. Stamen provides three tiles that uses OpenStreetMap data, and that can be freely used. More information can be found [here](http://maps.stamen.com) (<http://maps.stamen.com>). The ggmap library can provide access to Stamen maps, although you can still download maps from Google if you have an API key. More information [here](http://github.com/dkahle/ggmap) (<http://github.com/dkahle/ggmap>).

* **Kolada API** is an API to the Kolada database that contains key indicators regarding the Swedish municipalities. More information can be found [here](https://www.kolada.se/om-oss/api) (<https://www.kolada.se/om-oss/api>).

** **Riksdagen** or the Swedish parliament have an API to get documents, debates and information about members of parliament. Information about the API can be found [here](http://data.riksdagen.se/) (<http://data.riksdagen.se/>).

The data is mainly textual data so the R package should convert the data to the datastructures used in the `tm` R package.

** **Språkbanken XML** is a web resource with corpora in XML format. As with the other textual resources the purpose with this corpus is to read any XML file into R as a `tm` type corpus object with all textual metadata included. The Swedish corpora can be found [here](http://spraakbanken.gu.se/swe/resurser/corpus) (<http://spraakbanken.gu.se/swe/resurser/corpus>) while in multiple languages (including Swedish) [here](https://spraakbanken.gu.se/resurser/corpus) (<https://spraakbanken.gu.se/resurser/corpus>).

*** **Thenmaps** Historical map in geo.JSON format. Need to handle spatial data, see spatial.ly for more information on geographical data. More information can be found [here](http://www.thenmap.net/) <http://www.thenmap.net/>.