Queries

You can query the corpus for very different pieces of information such as tokens in the SMS, part of speech annotations, demographic information like the age of the informant, etc.

Please keep in mind that all the fields in the corpus are text fields, i.e. you cannot query for age:21-23 but have to start three independent queries, one per age.

The following options for querying the corpus are described in more detail in the sub-sections of this document:

- Simple queries: These are basically queries for words e.g. est or ich etc.
- RegEx queries: These are used for more complex patterns such as alternatives (man and men), for patterns with different endings (Man and Manchester) etc.
- Queries for meta data
- Combined queries: These are used whenever you want information from different layers, e.g. the word man written by only females.

From:
https://sms.linguistik.uzh.ch/ -

Permanent link:
https://sms.linguistik.uzh.ch/02_browsing/04_queries?rev=1641473937

Last update: **2022/06/27 09:21**