# WS1 Data Mining Clinical Notes

Feb 24, 2018

Co-facilitator

Nao Ichiahra

[nai807@mail.harvard.edu](mailto:nai807@mail.harvard.edu)

# Tools for learning RegEx

RegEx works in the same way with Japanese text as with English.

There are handy tools free of charge that help learn RegEx.

Google “free regex app.” I use the following tools, which both are good.

Expresso: <http://www.ultrapico.com/expresso.htm>

.Net Regular Expression Designer: <https://www.workshell.co.uk/products/netregex>

# Tools for NLP with Japanese

Because of grammatical difference, NLP with Japanese requires different set of tools than that with English and other Latin languages.

I recommend using the below open source tool on R:

MeCab: <http://taku910.github.io/mecab/>

# Web apps to easily create a word cloud

There are easy-to-use we apps that let you create “word clouds” with any given English text. This also works for other languages in which words are separated with spaces. Try the following and have fun.

Word Cloud Generator: <https://www.jasondavies.com/wordcloud/>

Wordclouds.com: <https://www.wordclouds.com/>