Kristel Uiboaed

### Praktikumi jaoks on vaja paigaldada järgmised programmid:

* [R](https://cran.r-project.org/)
* [RStudio](https://www.rstudio.com/products/rstudio/download/)
* [Python, 3.6 verisoon](https://www.python.org/downloads/)
* [Anaconda](https://www.continuum.io/downloads)
* [Notepad++](https://notepad-plus-plus.org/download/v7.3.1.html)

### Lemmatiseerimiseks vajalikud lisasammud (teeme praktikumis koos)

Windowsis on selline terminal nagu PowerShell ning vastavad käsud tuleks trükkida sellesse aknasse. Alustuseks paigaldame Anaconda keskkonda ka Python 3.5 versiooni, millel töötab [Estnltk](https://estnltk.github.io/estnltk/1.4.1/index.html), vaikimisi tuleb Anacondaga kaasa Pythoni versioon3.6. Paigaldame eesti keele töötlemiseks loodud Pythoni teegi [Estnltk](https://estnltk.github.io/estnltk/1.4.1/index.html).

conda install python=3.5

Kui on juba olemas vanem Anaconda versioon, piisab [Estnltk](https://estnltk.github.io/estnltk/1.4.1/index.html) paigaldamiseks järgneva käsu trükkimisest.

conda install -c estnltk -c conda-forge estnltk

Kuna Estnltk veel Pythoni uusimas, 3.6 versioonis, mis tuleb Anacondaga vaikimisi, ei tööta, siis teegi kasutamiseks on vaja teha mõned lisasammud. Tuleb luua Pythoni 3.5 keskkond, mis tuleb terminali iga uue sessiooni käivitamisel uuesti aktiveerida.

# estnltk\_env on keskkonna nimi, mille võib ise valida vabalt, praegu jätkame sellega  
conda create --name estnltk\_env python==3.5  
  
# aktiveeri keskkond, see käsk tuleb sisestada iga uue sessiooni alustamisel keskkonna aktiveerimiseks terminaliaknasse uuesti  
activate estnltk\_env  
  
# paigaldame estnltk  
conda install -c estnltk -c conda-forge estnltk  
  
# installi järel võib vajalik olla veel järgmine samm  
python -m nltk.downloader punkt  
  
# keskkonna deaktiveerimine  
deactivate estnltk\_env

Käivitame lemmatiseeriva skripti. **NB!** Kõik töödeldavad failid ja skript peavad olema samas kaustas. Sisendfailid peavad oleam utf-8 kodeeringus (või eemalda vastav parameeter skriptist). Käsurealt tuleb liikuda samuti sellesse töökausta (cd) ning seejärel sisestada käsurealt järgnev käsk.

py -3 lemmatizeFiles.py

Lisaks vaata kindlasti veel [Estnltk dokumentatsiooni](https://estnltk.github.io/estnltk/1.4.1/index.html).

Alternatiivne viis lemmatiseerimiseks [Keeleliinis](https://keeleliin.keeleressursid.ee/#/public/definition/32).

### Stilomeetrilise analüüsi jaoks vajalikud lisasammud (teeme praktikumis koos)

Stilomeetrilise analüüsi jaoks kasutame programmi R ja selle graafilist keskkonda RStudio. Alustuseks on vajalik paketti stylo installeerimine ja aktiveerimine keskkonnas käsuga library.

install.packages("stylo")  
library("stylo")  
  
# Paketi stylo graafilise kasutajaliidese käivitamine  
stylo(gui=T)