

Appendix A: Installation

In this appendix installation is introduced.

Documentations is available at <https://github.com/Mailis/EstNer/blob/master/cloud/readme.md> .

The software is available at <https://github.com/Mailis/EstNer/tree/master/cloud>.

Requirements

Master server and virtual machines in the cloud of worker instances both use

- Ubuntu operating system
- Apache2 server,
- php 5.x,
- python 2.7 and
- python 3.4.

Installation in master server

This installation, for master, is tested on Ubuntu 14.10 and Ubuntu 15.04

```
sudo apt-get update
sudo apt-get install Apache2
sudo apt-get install php5
sudo apt-get install python-pip
sudo pip install --upgrade google-api-python-client
sudo apt-get install python3-pip
sudo pip3 install rdflib
-----
sudo apt-get install python-pip
sudo pip install --upgrade google-api-python-client

--- download master folder files from
https://github.com/Mailis/EstNer/tree/master/cloud/master
into folder /var/www/html ---

cd /var/www/html
mv index.html apache2.html
sudo mkdir upload_logfile
sudo mkdir logfiles
sudo mkdir css
sudo mkdir js
sudo mkdir datasets
sudo mkdir SPARQLendpoint
sudo mkdir errors
sudo mkdir updates
sudo mkdir statistics
sudo mkdir statistics/processed_logfiles
sudo mkdir statistics/monthly_updates
```

Installation in worker server

This installation, for worker, is tested on Ubuntu 14.10 and Ubuntu 15.04

```
sudo apt-get update
sudo apt-get install Apache2
sudo apt-get install php5

sudo apt-get install python3-pip
sudo pip3 install numpy>=1.8.2
sudo python3 -m pip install nltk==3.0.0
sudo pip3 install python-crfsuite>=0.8.1
sudo pip3 install jsonpath-rw>=1.3.0
sudo pip3 install six>=1.7.3
sudo pip3 install tempdir>=0.6
sudo pip3 install xmltodict>=0.9.0
sudo pip3 install beautifulsoup4>=4.3.2
sudo pip3 install pandas>=0.14.1
sudo pip3 install scikit-learn>=0.15.1
sudo pip3 install xlrd>=0.9.2
sudo pip3 install xlswriter>=0.5.7
sudo pip3 install pytz>=2014.4
sudo pip3 install python-dateutil>=2.2
sudo pip3 install nose>=1.3.3
sudo pip3 install pyparsing>=2.0.2
sudo pip3 install matplotlib>=1.2.1
sudo pip3 install xlwt-future>=0.7.5
sudo apt-get install libfreetype6-dev libxft-dev
sudo apt-get install g++ swig python3-setuptools libfreetype6-dev
python3-pip liblapack-dev libblas-dev python3-dev gfortran default-jre
sudo python3 -m pip install estnltk==1.1
python3 -m nltk.downloader punkt
```

find /path/to/nltk_data

```
find / -xdev 2>/dev/null -name "nltk_data"
```

copy this folder to /var/www

```
cp -a /path/to/nltk_data /var/www
```

```
# e.g. cp -a /home/marfa_majakas/nltk_data /var/www
```

remove it from old location

```
rm -rf /path/to/nltk_data
```

```
#e.g. rm -rf /home/marfa_majakas/nltk_data
```

```
sudo apt-get install python3-lxml
sudo apt-get install python-httpplib2
sudo pip3 install PyPDF2
sudo pip3 install rdflib
sudo apt-get install python-pip
sudo pip install --upgrade google-api-python-client
--- download worker folder files from
https://github.com/Mailis/EstNer/tree/master/cloud/worker
into folder /var/www/html ---
```