

Installationsanleitung

Neo4j Installation:

- Java installieren
- Neo4j installieren
 - o Apt install neo3j
- Als Dienst aktivieren, sodass Neo4j beim Systemstart gestartet wird
 - o Systemctl enable neo4j
- Datei /etc/neo4j/neo4j.conf editieren und folgenden Inhalt hinzufügen
 - o dbms.connectors.default_listen_address=0.0.0.0
 - o dbms.connector.bolt.listen_address=0.0.0.0:7687
 - o dbms.connector.http.listen_address=0.0.0.0:7474
- Dienst neustarten
 - o Service neo4j restart
- Per Shell an Neo4j anmelden
 - o Cypher-shell -u neo4j -p neo4j
 - o Neues Passwort setzen: CALL dbmc.changePassword(„neues Passwort“);
 - o Shell wieder verlassen: „:exit“

CoreNLP installieren und als Service einrichten:

- In dem Ordner CoreNLP/stanford-corenlp-full-2018-10-05 ist ein Textdokument hinterlegt, das beschreibt wo CoreNLP geladen werden kann. Der Ordnerinhalt mit jar-Dateien etc. auf einen Server in den Ordner /opt/corenlp kopieren.
- Neuen Benutzer anlegen
 - o Useradd nlp
- Berechtigungen und Besitzer des Ordners auf den neuen Benutzer ändern
 - o Chown -R nlp:nlp /opt/corenlp
 - o Chmod 755 -R /opt/corenlp
- Den Inhalt von CoreNLP/start_corenlp.txt in die Datei /etc/init.d/corenlp kopieren und speichern
- Symlink erzeugen
 - o ln -s /etc/init.d/corenlp /etc/rc2.d/S75corenlp
- Server neustarten. CoreNLP wird automatisch beim Systemstart mitgestartet und läuft auf dem Port 9000. Soll der Port verändert werden, muss dieser Vorgang in der Datei /etc/init.d/corenlp durchgeführt werden. Der Dienst kann mit „sudo service corenlp start“ oder „... stop“ manuell gestartet oder beendet werden.

Modul für Drupal einrichten

- Aus dem Ordner Modul Drupal Umgebung hochladen (in den modules Ordner), entsprechende Dateirechte vergeben.
- In dem Ordner nlp_search/nlp_python/ClausIE/clausie_lib ist ein Textdokument, in dem steht woher die Ressourcen für ClausIE bezogen werden können.
- Ggf. Java auf dem Webserver installieren, falls noch nicht vorhanden
- Mindestes Python 3.5 sollte verwendet werden
- Für Python3 pip installieren
 - o Apt install python3-pip
- Mit Pip folgende Abhängigkeiten installieren

- Python3 -m pip install -U spacy
 - python3 -m spacy download en
 - python3 -m spacy download en_core_web_lg
- Python3 -m pip install neo4j
- Python3 -m pip install nltk
- Python3 -m pip install annoy
- Python3 -m pip install flask
- Python3 -m pip install unicode
- Auf dem Webserver in den Ordner des Moduls wechseln und dort in den Ordner nlp_python
- Hier folgenden Code ausführen
 - python3 nltkDownload.py
- Modul in Drupal aktivieren. Zur Einstellungsseite gehen und konfigurieren
- Mit dem Button „Index All“ die Nodes für die initiale Indexierung exportieren.
- Die Datei aufrufen, um den initialen Index zu starten
 - Python3 nlpStart.py
 - Alle Nodes sollten nun iteriert werden und die noch zu verarbeitenden Nodes angezeigt werden. Dieser Prozess kann je nach Inhalt in der Drupal-Umgebung etwas Zeit in Anspruch nehmen
- Cronjob hinzufügen
 - Crontab -e
 - 24 21 * * * python3
/var/www/html/women_d8/web/modules/custom/nlp_search/nlp_python/nlpStart.py
 - Ruft jeden Tag um 21:24 Uhr den Job auf, Pfad anpassen
 - Der Cronjob sollte für einen normalen Benutzer, nicht root eingerichtet werden.

Flask Anwendung (Schnittstelle zwischen den Python-Klassen und dem Drupal-Modul) als Service einrichten:

- Unicorn installieren
 - Python3 -m pip install unicorn
- In den Ordner /var/www/html/women_d8/web/modules/custom/nlp_search/nlp_python wechseln
- Befehl starten: `unicorn --bind 0.0.0.0:5000 wsgi:app`
- Schauen, ob es keine Fehlermeldung gibt
- Datei erstellen: `nano /etc/systemd/system/myproject.service`
- Folgenden Inhalt einfügen und speichern


```
[Unit]
Description=Gunicorn instance to serve nlpInterfaceFlask
After=network.target

[Service]
User=corin
Group=www-data
WorkingDirectory=/var/www/html/women_d8/web/modules/custom/nlp_search/nlp_python
ExecStart=/usr/local/bin/gunicorn --workers 3 --bind 0.0.0.0:5000 nlpInterfaceFlask:app

[Install]
```

WantedBy=multi-user.target

- Dienst starten
 - `sudo systemctl start nlpInterfaceFlask`
- Dienst aktivieren, sodass dieser beim Systemstart gestartet wird.
 - `sudo systemctl enable nlpInterfaceFlask`
- Status des Dienstes überprüfen
 - `sudo systemctl status nlpInterfaceFlask`