

Übung 10: Stanford Parser

1. Finden und aktivieren Sie den Stanford Parser in unseren Ressourcen! Welches ist die neueste Version, die bei uns installiert ist?
2. Starten Sie den Stanford Parser ohne Argumente.
3. Wenn Sie mit einem Browser die Datei `index.html` im Verzeichnis des Parsers öffnen, können Sie die Dokumentation aufrufen. Machen Sie sich mit Hilfe der Dokumentation der `LexicalizedParser`-Klasse mit der Bedienung des Parsers vertraut.
4. Parsen Sie nun die Datei `Darth_Vader.txt` mit dem Stanford Parser. Verwenden Sie dazu das vortrainierte Modell `englishPCFG.ser.gz` im `edu/stanford/nlp/models/lexparser/-` Verzeichnis des Stanford Parsers. Probieren Sie verschiedene Ausgabeformate aus. Erstellen Sie eine Datei `Darth_Vader.trees` im Ausgabeformat `penn`.
5. Konvertieren Sie `Darth_Vader.trees` ins `conll`-Format.

Zusatzaufgabe: Parsen Sie die `.conll`-Datei mit dem MaltParser und vergleichen Sie die Ausgabe des MaltParsers mit der Ausgabe des Stanford Parsers.