

# STARK: A Tool for Dependency Tree Extraction and Analysis

- Motivation: complement the prevailing hypothesis-based treebank querying with **bottom-up, data-driven linguistic investigation of UD treebanks**
- Python-based [open-source tool](#)
  - takes any CONLL-U treebank as input
  - **produces a list of all trees** meeting the user-defined criteria
  - based on **highly-customizable configuration** settings
  - computes **frequency, association strength** and **keyness**
- Useful for identifying:
  - common / rare treebank phenomena
  - multi-word expressions of various types
  - treebank- or language-specific phenomena