DEREK C. ANDERSEN

Computational Linguist

✉ derek@derekandersen.net   ⌨ dechrissen   👤 dechrissen.github.io   in derekcandersen

EXPERIENCE

NLU Developer
Cerence Inc.

📅 2021 – Present   📍 Remote

- Improved the accuracy of AI-based language models used in automobile voice assistant technology. Performed error analysis and bug fixing to correct patterned problems in source data.
- Led data-cleaning effort which resulted in a ∼9% increase in accuracy on customer test sets.
- Developed context-free grammars in the Java Speech Grammar Format (JSGF) for artificial input data synthesis to be used in machine learning (seq2seq) model training.
- Facilitated the weekly publication of customer deliverables in the form of voice assistant models which are deployed in automobiles from manufacturers such as Audi, BMW, Mercedes, and Toyota. Ensured on-time delivery as the leader of this project, addressing any road-blocks which prevented publication.
- Collaborated on development with various internal teams (automatic speech recognition team, language teams, and software teams) in order to build a cohesive language model for our customers.
- Developed Python scripts for data processing and task automation.
- Reviewed colleagues' development work and offered constructive criticism and suggestions before merging it into our resources.
- Worked daily in a Linux environment utilizing Bash, Git, regular expressions, and several other tools and techniques on the command line.

Linguist
Defined AI

📅 2020   📍 Remote

- Annotated prosody in English utterance audio data using the ToBI annotation guidelines to create a dataset for human speech model training.
- Performed spectrogram analysis using Praat and annotated consonants and vowels using the International Phonetic Alphabet.
- Collaborated with a team of linguists to meet customer deadlines efficiently.

EDUCATION

M.A. in Computational Linguistics
Stony Brook University

📅 Sept. 2019 – January 2021   Cumulative GPA: 3.93

B.A. in Linguistics, *Summa Cum Laude*
Stony Brook University

📅 Sept. 2016 – Dec. 2018   Cumulative GPA: 3.98

Coursework in Computer Science
Suffolk County College

📅 Sept. 2013 – May 2016   Cumulative GPA: 4.0

SKILLS

Python   Linux   Git   GitHub/GitLab   Bash   Regex   Scripting   JSGF   NLP   OCR   Annotation   Data processing   JavaScript   Node   HTML/CSS   Markdown   LaTeX   SQLite   Jira   Confluence   Documentation

PROJECTS

News Headline Toolkit
Personal Project

📅 2024

- Toolkit for creating and interfacing with a database of news headlines.
- Utilizes Tesseract OCR backend for image-to-text conversion.
- Includes language modeling, text synthesis, data analysis, and data visualization tools.

Blog Engine
Personal Project

📅 2021

- Backend engine for my personal blog, built with Node.
- Utilizes Node packages Showdown.js and Mustache.js to auto-generate HTML blog posts via Markdown conversion.

Recommendation Engine
Master's Thesis Project

📅 2020

- Content-based filtering recommendation engine for academic papers in linguistics from the LingBuzz database.
- Utilizes similarity matrix based on TF-IDF vectorization values to score and measure similarity between papers.

INTERESTS

- Language technology
- Web development
- Writing & typography
- Running