

Timothy James Leffel

updated jan01/2019

✉ tjleffel@gmail.com

☎ +1(773)256-6091

<http://lefft.xyz> 

github.com/lefft 

ABOUT

I'm a generative linguist by training, and a data scientist by profession. I work on projects ranging from document classification and topic modeling to the development of automated data ingestion and reporting systems. I approach NLP problems with industry standard tooling and the intuitions of an empirically-focused language scientist. My publications, teaching materials, and (selected) side-projects are all available at <http://lefft.xyz>.
Enthusiast of: Python, R, the NBA, the history of math/logic/computing, math/logic/computing.

WORK EXPERIENCE

- 2016– Data Scientist, Digital Media Collaboratory and ARC. NORC at the University of Chicago.
- >> social media text mining (sentiment analysis, document classification, topic modeling, ...)
 - >> pre-processing; training and evaluation; regression; simulation and resampling; visualization
 - >> development of data ingestion pipelines and automated reporting systems for RCTs
- 2015 Data science intern (p/t). Vantage Sports Inc.
- >> development and implementation of daily fantasy sports lineup optimization algorithm
 - >> evaluation of competing predictive models for daily fantasy performance of NBA players
- 2014–16 Postdoctoral Researcher. Language Processing Lab, Department of Linguistics. University of Chicago
- >> experimental design and implementation (surveys, eye-tracking for visual world studies)
 - >> data analysis and modeling (mixed-effects and logistic regression, RM-ANOVA, others)
 - >> dissemination of research results (conference presentations, proceedings and journal papers)
- 2010–16 Adjunct Instructor and Teaching Assistant in Linguistics and Psychology.
- >> primary instructor or TA for various undergrad courses in linguistics and psychology
 - >> 2016 at the University of Chicago; 2010–2014 at New York University

EDUCATION

- 2014 PhD, Linguistics. New York University
- 2011 MA, Linguistics. New York University
- 2009 BA with honors, *magna cum laude*, Linguistics/Philosophy. The Ohio State University

TECHNICAL SKILLS

core programming languages and computing stack

- >> **Python** – pandas, numpy, scikit-learn, gensim, nltk, keras. Usually use iPython/Jupyter.
- >> **R** – dplyr, ggplot2, lme4, shiny. Happy in base or hadleyverse.
- >> AWS cloud computing (EC2/S3); basic web development; (R)markdown; bash; git(hub); \LaTeX

algorithms, analysis techniques, etc.

- >> NLP: naive Bayes; decision trees; LDA; n -gram language models; word-embeddings; etc.
- >> neural network-based modeling for classification and term representation and clustering tasks
- >> logistic and mixed-effects regression modeling; permutation and randomization testing