JP van Paridon, PhD

+1 619 536 6806 | jvparidon@gmail.com | jvparidon.io | github.com/jvparidon

Profile

I am a computational cognitive scientist and data scientist with a background in language sciences. I use online and in-person experiments and large public datasets to inform statistical and computational models of human behavior and cognition. Currently, I am looking for a data science role where I can apply my skillset to real-world problems.

Experience

Researcher Associate (Postdoctoral), University of Wisconsin-Madison — 2020-Present

- Analyzed behavioral data and public datasets using statistical and ML techniques including **Bayesian hierarchical regression models**, dimensionality reduction, and clustering algorithms
- Used **NLP** models like word embeddings (e.g. word2vec) and transformer models (e.g. **BERT**, ALBERT) to model language use and learning mechanisms
- Designed and deployed **online experiments** using HTML/CSS and JavaScript

Researcher (PhD Candidate), Max Planck Institute for Psycholinguistics — 2015-2019

- Built computational and statistical models in **Python** to model the temporal dynamics of speech
- Developed and taught *Introduction to Python Programming* course for graduate students
- Engaged in international research collaborations as statistical and technical consultant

Education

- PhD, Cognitive Science, 2021 — Radboud University & Max Planck Institute, the Netherlands
- MSc (cum laude), Cognitive Neuroscience, 2015 — Leiden University, the Netherlands
- BSc, Psychology, 2012 — Leiden University, the Netherlands

Open Source Software

- Lead developer and maintainer of lmerMultiMember, an **R package** for modeling multiple membership random effects in **(generalized) linear mixed effects models**
- Developed subs2vec, a **Python package** with **word embeddings** from subtitles in 55 languages

Skills

- Statistical modeling in **frequentist** and **Bayesian** frameworks, including PyMC and a bit of Stan
- **Machine learning,** including **NLP, dimension reduction, feature selection,** and **clustering**
- Data extraction and cleaning using e.g. pandas, dplyr, and **SQL**
- **Data visualization** in ggplot2, matplotlib and seaborn
- **Python** programming, including NumPy, scikit-learn, SpaCy, and Jupyter notebooks
- **R** programming, including lme4, brms, rmarkdown, and various tidyverse packages
- **Version control** and **continuous integration** using git & Github Actions
- Designing and programming behavioral and **online experiments** using **HTML/CSS/JavaScript**
- Presenting technical concepts to non-technical audiences
- Coordinating international and multidisciplinary research collaborations