

Amos Teo

Data Scientist | Language Specialist | Linguistics PhD

Email: amosbteo@gmail.com

Website: www.amosbteo.com

EXECUTIVE SUMMARY

Experienced data scientist with over 9 years of experience in academic and professional settings. Skilled in leveraging data to understand customer behavior and drive business decisions, with expertise in statistical modeling, data visualization, and machine learning. Adept at collaborating with interdisciplinary teams. Proven track record of conducting original research and communicating complex concepts to diverse audiences.

Core competencies: Data analysis, Data cleaning and preprocessing, Statistical modeling, Programming (SQL, Python, R), Data visualization.

PROFESSIONAL EXPERIENCE

Senior Language Data Researcher, Alexa Home Productivity, Amazon	2023-2024
Language Data Researcher, Alexa Home Productivity, Amazon	2019-2023
- Data scientist for Alexa alarms, kitchen, calendar, notes & weather products	
Graduate Employee, Department of Linguistics, University of Oregon	2013-2019
- Lecturer, Teaching assistant, Research assistant	
Researcher, Laboratoire Dynamique du langage (DDL), Lyon	2016
Researcher, Laboratoire Langues et civilisations à tradition orale (LACITO), Paris	2015

KEY PROJECTS

Alexa Home Productivity speech recognition improvement	2023
<ul style="list-style-type: none">Applied clustering and classification techniques to target potential speech recognition errors in live data.Identified and curated test sets for Alexa Home Productivity products across 9 languages.Drove down average test set word error rate by 300 bps (17% to 14%).	
Alexa Arabic language launch	2021
<ul style="list-style-type: none">Identified bugs and coverage gaps in ML models through a combination of quantitative and qualitative analysis.Drove improvements in goal success rate by up to 2,500 bps (60% to 85%) for key Home Productivity features.Compiled and submitted executive-level reports that led to a successful Arabic launch decision.	
Linguistic clustering analysis for quantifying differences between languages	2017
<ul style="list-style-type: none">Developed a corpus-based method for measuring cross-linguistic distances.Used cluster analysis and dimensionality reduction techniques.Presented results at the 12th Association of Linguistic Typology meeting (Dec 2017).	
Alexa Data Services annotation workflow launch	2022
<ul style="list-style-type: none">Identified inefficiencies in existing annotation workflow using a combination of dashboard monitoring and ad-hoc analysis of data samples.Collaborated with Alexa Data Services team to add additional fields to create a new workflow.New workflow resulted in an average data yield increase of 3,400 bps (49% to 83%).	
Alexa annotation budget management automation	2022
<ul style="list-style-type: none">Managed weekly budget and data targeting strategies for 5 annotation workflows.Validated data targeting strategies by creating Tableau dashboards to display aggregated output labels, in combination with random sampling to inspect data quality.Wrote Python scripts to compile annotation budget and usage data from multiple spreadsheets and calculate weekly excess and shortfalls; scripts reduced manual work by 5-8 hours per week.	

EDUCATION

PhD in Linguistics, University of Oregon	2019
MA in Linguistics, University of Melbourne	2010
BA (Linguistics & Russian majors) with Honours, University of Melbourne	2007
Diploma of Modern Languages in French, University of Melbourne	2005

SKILLS

Technical skills and software:

- Database query & Statistical analysis: SQL, Python (e.g. numpy, pandas, scikit-learn); R (tidyr, dplyr, lme4)
- Data visualization: Python (matplotlib); Tableau dashboard building; R (ggplot2, ggmap)

Natural language competencies:

- **English** (N. American, Australia, Singapore): native
- **French**: excellent reading/written/spoken/listening
- **Mandarin**: excellent reading/written/spoken/listening
- **Russian**: good reading/written; basic spoken

ENGAGEMENT & KNOWLEDGE SHARING

- Alexa Home Productivity Diversity Equity and Inclusion: committee member & advocate (2022-2023)
- Glamazon: Vancouver Pride march co-captain (2022)
- Talk: "How to construct a language" at UO open house days for high school students (4 May 2018, 3 May 2019)
- UO Linguistics blog posts: [Linguistics in pop culture: Arrival \(film\)](#), [Linguistics in the news – Election season](#).
- Selected personal blog posts: [Issues with Ice Age Linguistics](#), ["Futureless languages?"](#)
- Other interests: birdwatching, gardening, urban design, stage management, language learning.