

Boris Dev

Stack

- **Language AI:** LangChain, AWS Sagemaker & GroundTruth, annotation pipelines, Instructor, Weaviate, OpenSearch, SNOMED CT Ontology
- **Data Science:** Pandas, Statistics, Numpy, Scikit-learn, PyTorch
- **Data Modeling:** Pydantic
- **Full Stack:** FastAPI, Flask, Django, Angular, HTMX, Bootstrap
- **Backend:** AWS, Docker, Jenkins, Kafka, Grafana, Splunk, ELK, Distributed tracing and OpenTelemetry Instrumentation

Education

PhD in Geography, at SDSU and UCSB, 2015. Data science for social science problems with location referenced data.

Job experience

AI Engineer consultant at Wolf Games, 2023

- Designed and built their story generator using a novel means, motive, opportunity causal DAG

AI Engineer consultant at Intuitive Systems, 2023

- Classified computer vendor receipts fed into AMD's sales analytic pipeline

AI Engineer consultant at SimpleLegal, 2022-2023

SimpleLegal is a legal billing analytics company.

- launched a previously stuck AI feature to flag invoice billing discrepancies
- redesigned and managed the annotation pipeline
- built SageMaker transformer model server and web app endpoint

Lead Analytic Endpoint Engineer at Sight Machine, 2018-2021

Sight Machine is a manufacturing analytics company.

- led backend engineering on biggest public facing analytic feature
- coordinated QA process with sales and engineering
- built company's first distributed tracing
- containerized frontend engineering development environment

Lead Data Engineer at HiQ Labs, 2015-2018

HiQ Labs was a people analytics company.

- led the migration from a monolith to a micro-service paradigm

- built the scraping system
- supported data science team

Developer at Urban Mapping, 2011-2013

Urban Mapping provided geospatial analytics to Tableau.

- CI/CD/QA pipeline
- Observability

Notable success

- For a gaming startup, I designed and built their murder mystery story generation service, starting from a greenfield design of the primitives.
- For a semiconductor analytics startup, I designed a new computer name disambiguation approach for AMD's sale analytics pipeline.
- For the Langchain open-source library, I contributed an experimental [CPAL Chain](#) - first causal LLM chain. [Tweet by LangChain](#).
- For a legal ops company, a new language AI billing feature.
- For a people analytics startup, new microservice paradigm.
- For academics, new location based human inequality metrics in [Assessing Inequality using Geographic Income Distributions](#)
- For several companies, I built observability systems.

Interesting activities

- For my side-project, Nobsmed.com, I am making a website to help people compare their treatment options by summarizing clinical study outcomes along with Reddit personal experience anecdotal comments.
- I climbed Cotopaxi (21,000 ft), survived bodyboarding Mexpipe, worked with students in Medellín, Columbia to make [ClusterPy](#), was a kids snow-board instructor at Vail Resorts, CO, was an assistant manager at Gundy's Grill in Vail Resorts, CO, was a counselor for severely emotionally disturbed children at Seneca Institute, CA.

Papers and code

- [Language AI Evaluation 101: Know your user](#)
- [Langchain PR: Causal Program-aided Language \(CPAL\)](#)
- [Work papers](#)
- [Academic papers](#)