Boris Dev, AI Data Products Engineer

Tool Stack

- **Product:** Python FastAPI/Flask/Django, SQLAlchemy, Postgres, AngularJS, HTMX, Boostrap, NextJS (learning)
- **Data Science:** Pandas, Sklearn, Numpy, Plotly, PySpark, DataBricks, Tableau
- **Langauge AI:** LangSmith, Langchain, InstructorAI, Embedding DBs (Weaviate, ElasticSearch), SageMaker, PyTorch, GroundTruth, Pydantic Agents (learning), Sentence Transformers (fine-tuning)
- **System:** Jenkins, Kafka, Rancher, Docker, Helm charts, kubectl, Splunk, ELK, Grafana, OpenTelemetry, RabbitMQ

Education

PhD in Quantitative Human Geography, at SDSU and UCSB, 2015. Data science for location referenced social science problems. Dissertation: Assessing Inequality using Geographic Income Distributions

Experience

Nobsmed, 2024, Founder

Built https://nobsmed.com/ for users to search for people's health hacking experiences found in clinical studies and Reddit comments.

- Integrated an LLM into the ETL data pipeline for enrichment parsing, relevance scoring, and classification.
- Developed an LLM based topic modeling tool, bertopic-easy.
- Customer discovery and user interviews.
- Experimented with fine-tuning opensource embeddings model and ML classification.
- Prompt engineering using InstructorAI.
- Embedding DBs for search (Weaviate, ElasticSearch).

AI Engineer consultant at Wolf Games, 2023

- Built an automated DAG of LLM prompts for story generation
- Worked closely with the CTO on requirements
- Built the story generation Flask API service

AI Engineer consultant at Intuitive Systems, 2023

- NED (Name Entity Disambiguation) of vendor receipts fed into AMD's sales analytic pipeline. Used LangSmith for evaluation.

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
- was first to provide metric documentation at the meetings
- trained 8 ML classification models to flag suspicious invoice line items
- redesigned and managed 5 workers on the annotation pipeline of the labels
- built the inference server (SageMaker, PyTorch) and its product integration

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.

- built and owned the scraping system and observability (Splunk)
- led the migration from a monolith to a micro-service paradigm
- supported the data science team's migration to DataBricks PySpark and microservices

Developer at Urban Mapping, 2011-2013

Urban Mapping provided geospatial analytics to Tableau.

- Formulated first map rendering and query performance metrics integrated into our CI/CD pipeline

Impactful projects

- persuaded team that our culprit was data quality, not data quantity, to get a stuck AI feature launched.
- proposed and built observability systems to reduce firefighting.
- proposed and built containerized FE dev env for faster onboarding.
- migrated a data science monolith to a microservice paradigm

Papers and code

- LLM Topic Modeling: bertopic-easy.
- Language AI Evaluation 101: Know your user
- Langchain PR: Causal Program-aided Language (CPAL). See Harrison Chase's Tweet.

- Work papers
- Academic papers