Boris Dev

Stack

- **Data modeling and Knowledge graphs:** Pydantic, SNOMED CT ontology, causal discovery, causal graph based story generation
- **Language AI:** LangChain, AWS Sagemaker & GroundTruth, annotation pipeline, Sklearn, Pandas, Numpy, Instructor, Weaviate vector search DB, OpenSearch
- **Engineering:** Python backend (FastAPI/Flask/Django), micro-services, distributed systems, observability, analytic endpoints

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

- murder mystery story compositional grammar
- murder mystery story events service

AI Engineer consultant at Intuitive Systems, 2023

- Normalize vendor receipts fed into AMD's sales analytic pipeline
- Human-in-the-loop product design

AI Engineer consultant at SimpleLegal, 2022-2023

SimpleLegal is a legal billing analytics company.

- launched a stuck AI feature
- shifted focus from training example quantity to quality
- led offshore annotators, data scientist, and the subject-matter expert
- built SageMaker inference server

Lead Analytic Endpoint Engineer at Sight Machine, 2018-2021

Sight Machine is a manufacturing analytics company.

- analytic endpoints
- started new checklist process to lock down successful sales demos
- built distributed tracing

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

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, I am deriving normal forms for clinical studies and murder mysteries.
- For my side-project, Nobsmed.com, I am building a new causal discovery approach to RAG to help ordinary people evaluate their medical treatment options.
- I climbed Cotopaxi (21,000 ft), survived bodyboarding Mexpipe, worked with students in Medellín, Columbia to make ClusterPy, was a kids snowboard 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