

Guus Bouwens

Data Engineer and AI Researcher

Southern California, USA

(253) 409-7155 | guus@bouwens.nl | linkedin.com/in/guus-bouwens | guusbouwens.com

EXPERIENCE

Stealth Startup | Data Engineer Aug 2026 - Present

- Develop multi-agent systems for product data enrichment and ingestion.
- Build lab testing pipelines and workflows that help data analysts review new or flagged items.

Re-Search | Founder June 2026 - Present

- Building a platform that turns searches and saved sources into personalized discoveries and evidence-grounded ideation, preserving decisions and project context across research and development.

Texas Tech University | Statistical Consultant & Graduate Research Assistant Sep 2025 - May 2026

- Advised **30+** graduate students, postdocs and faculty on study design, modeling and interpretation.
- Researched probabilistic AI-generated content detection with Dr. Youngsoo Kim.

McDermott International | LLM Engineering Intern Oct 2024 - Apr 2025

- Improved retrieval accuracy **2x** and speed **5x** with contextual chunking, hybrid search and reranking; deployed the RAG system on Azure with Python, LlamaIndex, ChromaDB and citation-based evaluation.

Invest-NL | AI Engineering Intern Apr 2024 - Sep 2024

- Raised company classification accuracy from **60%** to **95%** using DPO on Llama 3 70B from Hugging Face; automated ingestion, deployment and monitoring with Python, Power BI and Azure CI/CD.

Talpa eCommerce | Data Scientist Mar 2023 - Oct 2023

- Increased activity **10%** with cosine similarity recommendations informed by behavioral analysis; deployed the model with Docker, AWS SageMaker and Jenkins.

EDUCATION

Texas Tech University | M.S. Data Science May 2025 - May 2026

GPA 4.0/4.0; top of class; three scholarships; Honors Beta Gamma Sigma.

Vrije Universiteit Amsterdam | B.S. Econometrics and Data Science Sep 2020 - Mar 2025

Minor in Big Data.

SELECTED PROJECTS

Food Scanner Fine-tuned YOLO11x for multi-ingredient segmentation and deployed a nutrition app integrating Gemini and USDA data to estimate calories and micronutrients from food images.

SwarmGrid a warehouse hazard-detection prototype at a Google DeepMind hackathon, integrating fine-tuned Gemma 3n, agent tool calls, and a live telemetry dashboard.

Palace a browser card game with a single-player campaign and multiplayer for up to four players, implementing server-validated moves and persistent game state.

SELECTED PAPERS

Layer-Wise Category Structure in Large Language Models

Causal State Variables in V-JEPA 2 Latents

Low-Dimensional Document Structure Subspaces in OCR Models

SKILLS AND CERTIFICATIONS

Programming: Python, SQL, R.

AI: PyTorch, TensorFlow, scikit-learn, Transformers, LangGraph, LlamaIndex; RAG, DPO, LoRA, BM25.

Engineering: Azure, AWS SageMaker, Docker, FastAPI, MLflow, CI/CD; PostgreSQL, MongoDB, Supabase, SQLite, ChromaDB, Power BI, Vercel.

Languages: Dutch (native), English (fluent).

Certifications: Microsoft Azure Data Scientist Associate; AWS Fundamentals; AI in Healthcare (Stanford); Agentic AI & RAG (IBM); Deep Learning and NLP (DeepLearning.AI); Machine Learning (UW & ICL).