Guus Bouwens

(253) 409-7155 • guus@bouwens.nl • linkedin.com/in/guus-bouwens • guusbouwens.com
Open to: US (OPT/EAD), EU (Dutch Citizen)

EDUCATION

Texas Tech University, Rawls College of Business — **Lubbock, TX, USA**
Master of Science in Data Science — May 2025 - May 2026
- *Three academic scholarships;* ***Top of Class*** *(GPA: 4.0/4.0); Honors: Beta Gamma Sigma*

Vrije Universiteit Amsterdam, School of Business and Economics — **Amsterdam, NL**
Bachelor of Science in Econometrics and Data Science, Minor in Big Data — Sep 2020 – Mar 2025

PROFESSIONAL EXPERIENCE

Re-Search & Episteme — **San Diego, CA, USA**
Founder — Jan 2026 - Current
- Engineered LLM-orchestrated multi-agent systems for autonomous research discovery, building a secure local-first RAG workspace (Episteme) and a parallel-retrieval novelty engine (Re-Search) that evaluates and benchmarks hypotheses against millions of papers in scientific literature corpora.

Texas Tech University, Graduate School — **Lubbock, TX, USA**
Statistical Consultant — Sep 2025 – Apr 2026
- Advise **20+** graduate students, postdocs, and faculty on all stages of the quantitative research lifecycle, from experimental design and data cleaning to model selection and interpretation of results.

Graduate Research Assistant — Nov 2025 - Apr 2026
- Research probabilistic methods for AI-generated content detection with Dr. Youngsoo Kim.

McDermott International, Ltd — **The Hague, NL**
AI Engineering Intern — Oct 2024 - Apr 2025
- Boosted retrieval accuracy **2x** and speed **5x** by architecting an advanced RAG system with contextual chunking, hybrid retrieval, and reranking, deployed on **Azure** using **Python, LlamaIndex, ChromaDB** and **CI/CD** pipelines.
- Architected a security-focused evaluation framework with in-text citations to guarantee response factuality, while collaborating with PhD researchers and project owners in an Agile workflow.

Invest-NL — **Amsterdam, NL**
AI Engineering Intern — Apr 2024 – Sep 2024
- **Fine-tuned** a Llama-3-70B model using **DPO** for Deep Tech investment teams, boosting accuracy from **60**% to **95**% for company classification while cutting API costs by optimizing token output.
- Automated the MLOps lifecycle from data ingestion to deployment using **Python, Power BI,** and **Azure CI/CD**, enabling continuous monitoring and dashboard updating.

Talpa eCommerce — **Amsterdam, NL**
Data Science Intern — Mar 2023 - Oct 2023
- Improved recommendation engine performance, increasing activity by **10%,** by adding cosine similarity measures, informed by customer insights uncovered via **Fourier analysis** of multi-year behavioral data.
- Engineered a scalable deployment by containerizing the model with **Docker** and deploying on **AWS SageMaker** with **Jenkins** for CI/CD, within a **Jira**-managed Agile workflow.

PROJECTS

Agentic Workflow *(AI Agents)*: Built a multi-agent orchestrated system using **LangGraph** and **Tavily** for agentic search, incorporating **SQLite**-backed state persistence and human-in-the-loop oversight via **Gradio**.

Food Scanner App *(YOLO11x, GPU)*: Converted FoodSeg103 to polygons and fine-tuned a model for accurate multi-ingredient food segmentation, now deployed as a live-service app also incorporating Gemini prompt engineering.

Google DeepMind x InstaLILY AI Hackathon *(Gemma-3n, Finetuning)*: Built SwarmGrid, a multi-agent warehouse orchestration prototype for dynamic task allocation, routing, and operational coordination.

SELECTED PAPERS

Identifying Compassion in LLMs *(ISR, under review)*: Aligned internal representations across open-source LLMs and showed semantic resolution predicts compassionate behavior, revealing negative priming risks.

Layer-Wise Cognitive Specialization in LLMs *(ICML, accepted)*: Mapped cognitive capabilities across transformer layers using probing and causal interventions, finding shared hierarchies and architecture-specific dynamics.

Late Semantic Repair in I-JEPA *(TPAMI, under review)*: Identified a late semantic repair mechanism in I-JEPA via causal activation patching, with a bottleneck at encoder layer 29 for object-level reconstruction.

CERTIFICATES

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

SKILLS

Programming: Python, SQL, R.
Generative AI & LLMs: RAG, Agentic Workflows, DPO, LoRA, Mechanistic Interpretability, Hybrid Search (BM25).
ML/AI Frameworks: PyTorch, LangGraph, LlamaIndex, Hugging Face Transformers, TensorFlow, Scikit-learn.
MLOps & Cloud: Azure, AWS SageMaker, Docker, MLflow, CI/CD (GitHub Actions, Jenkins), FastAPI.
Databases & Tools: Vector DBs (ChromaDB), PostgreSQL, MongoDB, SQLite, Power BI.
Languages: Dutch (native), English (fluent).