

# THOMAS TO

Oakland, CA, 510-387-5408, [thomas.to.bcheme@gmail.com](mailto:thomas.to.bcheme@gmail.com), [Linkedin](#), [AI/ML Portfolio](#)

## EXPERIENCE

---

**Founder | Open Source Access** | Oakland, CA | Oct 2023 – Present

- Multi-agentic orchestration (**langGraph**) with tool-use (**MCP**, **langChain**) to plan-and-execute tasks initiated with **retrieval augmented generation (RAG)**, optimized with **fine-tuning** for domain specific use case action items.
- **Github** Monolith for **CI/CD RAG** frontend on **AWS** (vercel), and **CT/CD Python ML models (scikit, pytorch, tensorflow)** on **Huggingface (FastAPI, Docker)**, with **vector database** for end-to-end cross platform ecosystem.
- **Deep learning (NLP)** of Kaggle & Huggingface datasets for keyword grouping to support new graduates to industry.

**Research Engineer | Stanford (Dunn Lab)** | Stanford, CA | Dec 2025 – Present

- **Deep Learning** and **Generative Physics** for de novo binder design.

**Founding Fullstack Engineer | Canvanta Life Sciences** | Emeryville, CA | Oct 2023 – Present

- Architected a revenue optimization system using **Snowflake** to integrate a predictive machine learning model with a **RAG AI agent** to reduce stakeholder decision-making from hours to minutes.
- Augmented natural knowledge by **enriching RAG fine-tuning** with Atlassian (Confluence) to enhance **GenAI** (data-to-text, text-to-image) context improving learning rate up to 80% (Wright's Law: Stanford-B model).
- Deployed in-house fullstack **DevOps SaaS** on **GCP** via **CI/CD**, converting serial to concurrent workflows reducing daily calculation time by 87% (-40 min), minimized calculation risk, and forecasted production within 3 of actual.
- Architected **ETL/ELT** pipelines with **Python** and **Google Apps Script** to digitize 5+ years of unstructured data.
- Reverse engineered data pipelines using **Tableau**, **Snowflake**, **dbt**, **fivetran**, **GCP**, **SAP**, and **SQL** for gap analysis to deploy stop-gap data pipelines and scope scalable & sustainable long-term cloud architecture process improvements.

**Founder | Proprietary FinTech** | Oakland, CA | Jun 2022 – Present

- Optimized Python Algorithmic & Agentic trading frequency to Coinbase's allowable public **REST API** limit.

**Laboratory Technician | Canvanta Life Sciences** | Emeryville, CA | Jan 2023 – Oct 2023

- Leveraged hands-on bioprocessing expertise to architect data systems & data models to physical reality.
- Modeled empirical data using **python (numpy, pandas)** to scale validated datasets over 300k datapoints.

**Software Engineer | Genentech** | South San Francisco, CA | Jun 2022 – Dec 2022

- Engineered scalable full-stack applications using backend **Python REST API**, frontend **Vue.js**, and **Node.js** using **git** & **GitHub** for version control to streamline workflows for medical writing teams.
- Designed knowledge management platforms using **HTML**, **CSS**, **JavaScript** to incorporate **UI/UX** feedback.

**Process Engineer | Genentech** | Vacaville, CA | Jun 2021 – Jun 2022

- Developed automated data consolidation algorithms using **Python**, **R**, and **SQL**, reducing time by over 99%.
- Deployed data-driven process monitoring tools on Google Cloud Platform (**GCP**) to track KPIs.

**Research Engineer | UC Davis (Nandi, McDonald, Wan, Siegel Labs)** | Davis, CA | Sep 2019 – Jun 2021

- Optimized model using **Python (numpy, pandas)** and **numerical methods** to reduce costs by \$63.2 Million.
- Quantified organoid growth using **Python (OpenCV, ImageJ, Fiji)** for image processing analysis.
- [Published](#) novel variants using **pyRosetta**, **pyMol**, and **Benchling** with wet-lab techniques to biomanufacture.

## TECHNICAL SKILLS

---

**AI & Machine Learning:** LLMs, RAG, NLP, Computer Vision, LangGraph, LangChain, PyTorch, TensorFlow, Scikit-learn, Hugging Face, AI SDK, n8n, Opal, Claude, cursor, codeWiki.

**API & Web Frameworks:** FastAPI, Django (DRF), Flask, GraphQL, OpenAPI (Swagger), OAuth2, Postman.

**Middleware (CTMS/CRM/ERP):** SNPGlue (SAP), MuleSoft (salesforce), Open API (CRIO), dbt, Fivetran, Zapier.

**Databases & Backend:** Snowflake, ETL/ELT Data Pipelines, Node.js, Django REST, PostgreSQL.

**Cloud & IoT:** GCP, AWS, Docker, Kubernetes, CI/CD, Git, GitHub, Atlassian, SAP, Raspberry Pi, AntMiner.

**Languages:** Python, SQL, R, Java, TypeScript, JavaScript (JSX), MATLAB, HTML/CSS, Google Apps Script.

**Frontend & Visualization:** Next.js, React.js, Vue.js, Tableau, Fiji, ImageJ, Python (matplotlib, seaborn, streamlit).

## EDUCATION

---

**Bachelor of Science, Biochemical Engineering** University of California, Davis (Graduated)

Davis, CA

Advanced Machine Learning with Agentic AI (on going)

AWS Cloud Technical Essentials (completed)