Thomas Wyndham Bush, M.S.

Gualtieri, IT • thomasbush52@gmail.com

Github, Personal Website, Linkedin, Work Website

Engineer Fellow – Machine Learning *@Sabatini Lab, Kempner Institute (US)* *(Jan 2026 – Present)*

Objective: Develop scalable distributed ML pipelines and models for large-scale biological data processing and computational neuroscience applications.

Experience

Summer Research Fellow – Bioinformatics *@Sabatini Lab, Harvard Medical School (US)* *(Jul 2025 – Aug 2025)*

Objective: Investigate how genetic mutations in neuronal cells contribute to epilepsy, and whether computational screening can support early identification of pathogenic variants and their functional effects.

- Fine-tuned a deep learning model inspired by AlphaFold to predict the structural and functional consequences of genetic mutations in neural proteins, with the goal of classifying mutations as benign, loss-of-function, or gain-of-function.

Freelancer Consultant – AI Development for Systematic Reviews *@EUDA European Union Drugs Agency* (PT) *(Jan 2025 – Jan 2026)*

Objective: Develop a cloud-based AI system tailored for systematic reviews in policy and healthcare domains.

- Developed a production-ready Retrieval-Augmented Generation (RAG) system using Microsoft Azure, integrating document upload, preprocessing, chunking, embedding, and query answering.
- Implemented a full-stack pipeline with FastAPI, Azure Cognitive Search, Azure OpenAI, and SQLite; managed database logic, deployment, and front-end integration.
- Used the DSPy framework to modularize and optimize the RAG architecture, enabling explainable answer generation and chain-of-thought tracing.
- Integrated citation-based grounding and model transparency features to improve factual reliability and user trust in model outputs.
- Worked closely with Microsoft Azure engineers, IT team, and non-technical stakeholders to align system behavior with domain-specific review workflows.

Research Intern – Computational Neuroscience *@Iurilli Lab , Italian Institute of Technology (IT)* *(April 2024 – Jan 2026)*

Objective: Developed a deep learning-based pipeline to extract and analyze 3D postural features of mice for behavioral modeling during spontaneous and hunting tasks.

- Built a multi-camera behavioral analysis pipeline: 2D pose estimation using CNN-based keypoint detectors (SLEAP, ResNet50), 3D triangulation, temporal filtering, and optimization using reprojection and anatomical constraints.
- Applied autoencoder-based denoising techniques and keyframe selection methods to improve 2D tracking quality and downstream 3D reconstruction accuracy.
- Investigated behavioral syllables from 3D pose sequences using dimensionality reduction (PCA, t-SNE) and unsupervised clustering to support interpretable modeling of action motifs.

AI Intern – NLP for Systematic Reviews *@EUDA (PT)* *(August 2024)*

Objective: Designed an early prototype of a Retrieval-Augmented Generation (RAG) system to assist researchers during systematic reviews.

Technical Skills

Programming Languages: Python, R, Bash
Machine Learning & Deep Learning: PyTorch, Scikit-learn, Transformers (Hugging Face), Einops, Einsum, DSPy
Scientific Computing & Data Analysis: NumPy, Pandas, Xarray, SciPy, OpenCV, Napari, Matplotlib, Seaborn
Mathematical & Statistical Methods: Linear Algebra, Probability Theory, Tensor Calculus, Generalized Linear Models, Regression (linear, logistic), Cross-Validation, Bootstrapping, Hypothesis Testing.
Developer Tools & Infrastructure: Git, Docker, Conda, SQLite, SLURM, VIM

Education

Tilburg University Tilburg, Netherlands
M.Sc. in Data Science *(2023 – 2025)*
Relevant Coursework: Machine Learning, Deep Learning, Data Mining, Advanced Data Processing, Image Analysis

University of Trento Trento, Italy
M.Sc. in Cognitive Science (Computational Modelling of Language and Cognition) *2023 – Present (expected 2025)*
Relevant Coursework: Machine Learning for Natural Language Processing (advanced), Computational Linguistics, Computational Modeling of Perception

Tilburg University Tilburg, Netherlands
Pre-M.Sc. in Data Science *2022 – 2023*
Grade: 8.5 / 10
Relevant Coursework: Tensor Calculus, Linear Algebra, Statistics

University of Bologna Bologna, IT
BA *2019 – 2022*
Grade: 107/110
Relevant Coursework: Philosophy of Mind, Principles of AI, Cognitive PSy

---

Awards and Public Speaking

Armenise Harvard Summer Fellowship: *Harvard Medical School (2025)*, Selected for a competitive 2-month research fellowship in computational neuroscience at the Sabatini Lab.
Talk "AI Systems for Systematic Reviews" (45 min.) *European Union Drugs Agency Sector Meeting (Oct. 2024)*, Presented architecture and use case of a could-based RAG pipeline for evidence sysnthesis.

Language Proficiency

English: C2
Italian: Native