

Kurt Thorn, Ph.D.

San Francisco, CA · 415-298-3495 · kurt.thorn@gmail.com

Summary

Proven builder of AI systems for real-world life sciences decision-making, with deep experience converting biological and chemical workflows into structured data, model objectives, and computational environments. As second employee and CTO of an ML-enabled drug discovery company, owned AI/ML research strategy, scientific data infrastructure, and agentic workflow orchestration.

Strengths

Scientific AI strategy · ML-guided drug discovery · LLM/agentic workflow orchestration · Scientific data standards · Cheminformatics · Biological assay design · Model evaluation · Ranking / virtual screening · AWS-scale ML infrastructure · Cross-functional scientific leadership

Technical

Python · AWS · SQL · Pandas · scikit-learn · XGBoost · RDKit · Smina/Gnina · Docker · Athena · S3

Experience

ArrePath, Inc.

Chief Technology Officer 1/2024 – Present

VP, Technology and Data Science 12/2021 – 12/2023

Second employee and CTO of a seed-stage ML-enabled antibiotic discovery company. Own AI/ML research strategy, scientific computing, data infrastructure, and AI-assisted discovery workflows across target/MOA prediction, hit identification, and lead optimization. Built systems that translate biological assays, cheminformatics, medicinal chemistry hypotheses, and model predictions into repeatable computational workflows used for real experimental decisions. ~50% of time hands-on in code.

Key accomplishments:

- Built an agentic AI system to allow orchestrating complex workflows utilizing our 30+ ML models and cheminformatics tools in natural language, enabling non-technical users to independently carry out sophisticated analyses.
- Deployed AI/ML-guided design and scoring workflows where AI-generated compounds matched or exceeded the potency of medicinal chemist designs. Routine use of ML models doubled the fraction of synthesized compounds meeting potency goals.
- Ingested large public compound and assay datasets and defined standardized data interfaces connecting molecular structures, assay results, model predictions, provenance, and agentic workflow execution.
- Developed first-in-class, patentable ML models for key drug parameters including oral bioavailability, cytotoxicity, and antibacterial activity, enabling differentiated IP around our AI/ML stack.
- Oral bioavailability model successfully predicted the first orally bioavailable molecule in our lead program, increasing bioavailability from 10% to 80%+.
- Developed ML-guided HTS approach that increased antibacterial screening hit rates by 3x while avoiding cytotoxic molecules and prioritizing chemically novel ones.
- Set company-level AI/ML platform goals and presented quarterly progress, strategy, and technical milestones to the board.
- Architected scalable data infrastructure using AWS S3 data lake containing >100M commercially available compounds, with AWS Athena integration for efficient querying of both public data and data from our LIMS ingested through na ETL pipeline.

- Developed containerized execution framework using AWS Batch for complex ML pipelines, orchestrating 40+ interdependent jobs for chemistry model inference.
- Established a rapid model-development workflow capable of delivering a highly predictive ML model for a new property in under a week.
- Advanced the company's licensed image-based phenotyping technology from manual agarose pad imaging to automated 384-well plate screening with 100x oil immersion imaging, enabling hundreds of compounds to be tested with walk-away automation.
- Developed and validated a supervised ML pipeline for mechanism-of-action (target/MOA) prediction from imaging data (70% accuracy) and novelty detection (90% accuracy), using rigorous leave-one-compound-out and leave-one-mechanism-out cross-validation.
- Built and led a fully-remote technical organization integrating wet lab scientists, ML engineers, and platform developers.
- Directly supervised both technical and laboratory personnel, successfully bridging the gap between computational and experimental workflows.
- Recruited and managed talent from diverse backgrounds, including strategic hire of chemistry ML leader from petroleum industry and identification of high-potential ML talent from fintech.

Zymergen, Inc.

Senior Director, Data Science	1/2021 – 10/2021
Director, Data Science	1/2019 – 1/2021
Associate Director	7/2018 – 1/2019
Senior Scientist	3/2017 – 7/2018

Zymergen Data Science supported a diverse range of activities including: data reduction and decision support for HTS data, algorithmic recommendations for strain and protein designs, machine learning for material property prediction, and technoeconomic analysis of potential new products.

- Led a 30-person Data Science / Modeling organization supporting HTS analytics, ML-guided strain and protein design, product-property prediction, technoeconomic modeling, and executive decision support.
- Reorganized fragmented modeling and data-science functions into a unified team with shared standards, reusable tools, and a single interface to biological project teams.
- Built standardized operational metrics and automated reporting pipelines for high-throughput biological testing platforms, with monthly reporting to the CEO.

Department of Biochemistry and Biophysics, UCSF

Associate Adjunct Professor	2013 – 2017
Assistant Adjunct Professor	2008 – 2013
Director, Nikon Imaging Center	2006 – 2017

- Principal investigator on a \$1.3M / 5 year R01 grant.
- Managed a lab developing a novel assay platform incorporating microfluidics, chemistry, microscopy, and image analysis for multiplexed biochemical assays. This platform measured 48 protein-peptide binding interactions simultaneously, with the potential to scale to >1000.
- Inaugural director of the major microscopy core at UCSF Mission Bay, consisting of eight microscopes serving 300 researchers/year and resulting in 40 papers per year. Fully responsible for budget (\$300,000/year), operations, instrument acquisition, and industry partnerships.

Bauer Center for Genomics Research, Harvard University

Research Fellow	2001 – 2006
-----------------	-------------

Education PhD in Biophysics : University of California, San Francisco

BA in Chemistry : Princeton University

Publications 35 total; h-index 31. Full list at: <https://scholar.google.com/citations?hl=en&user=IX23KMMAAAAJ>
Inventor on 5 patents.