

Frazier N. Baker

AI Researcher & Ph.D. Candidate

frazierbaker.com • github.com/frazierbaker • linkedin.com/in/frazierbaker

Grove City, Ohio
frazier@frazierbaker.com
+1 740 586 7553

Professional Profile

AI researcher specializing in **agentic AI for scientific discovery**. Builds intelligent systems for complex science and engineering applications, including retrosynthesis, drug design, and end-to-end drug-discovery workflows. Experienced leading interdisciplinary research programs and deploying research software. Ph.D. Candidate at The Ohio State University with expected graduation in Summer 2027.

Education

The Ohio State University

Ph.D. Candidate, Computer Science and Engineering

Advisor: Xia Ning, Ph.D. • Passed candidacy July 2026 • Part-time 2021–2023; full-time Graduate Research Assistant since 2023

- Dissertation in progress: *AI Agents for Retrosynthesis*.
- Coursework includes machine learning, artificial intelligence, reinforcement learning, and foundations of speech and language processing.

University of Cincinnati

M.S., Computer Science — Accelerated Engineering Degree BS/MS Program (ACCEND)

Advisor: Alexey Porollo, Ph.D. • GPA: 3.97/4.00

Thesis: Mining and Visualization of Amino Acid Coevolution Data

University of Cincinnati

B.S., Computer Science — University Honors Student; Dean's List

GPA: 3.96/4.00

Ohio University Zanesville

High School Post-Secondary Program; GPA: 4.00/4.00

Tri-Valley High School

Honors High School Diploma; GPA: 4.00/4.00

August 2021–Present

Columbus, Ohio

December 2019

Cincinnati, Ohio

May 2019

Cincinnati, Ohio

August 2012–August 2014

Zanesville, Ohio

May 2014

Dresden, Ohio

Research & Professional Experience

The Ohio State University

Graduate Research Assistant, Department of Biomedical Informatics

August 2023–Present

Columbus, Ohio

- Developed **agentic AI** systems for drug discovery applications, including retrosynthesis, drug design, and end-to-end drug-discovery workflows.
- Developed and demonstrated a framework for designing, constructing, and testing **multi-agent systems** for multi-objective retrosynthesis planning.
- Designed, implemented, and tested an **LLM Agent-as-a-Judge** framework for constrained retrosynthesis planning under practical safety constraints, the first AI framework for retrosynthesis in that setting; this work received a **Best Paper Award** at AIAS 2025.
- Architected and trained a multi-agent reinforcement-learning system for synthon completion in retrosynthesis; published in *Journal of Chemical Information and Modeling*.
- Collaborated with **interdisciplinary teams** of medicinal chemists, computational chemists, biologists, computer scientists, and pharmaceutical-industry partners; mentored junior students and researchers.
- Organized journal-club-style lab meetings and lab retreats, coordinating student and faculty schedules.

Georgia Tech Research Institute, Advanced Concepts Laboratory

Research Scientist II

July 2022–October 2023

Dayton, Ohio

Georgia Tech Research Institute, Advanced Concepts Laboratory

Research Scientist I

June 2019–July 2022

Dayton, Ohio

- Architected, designed, built, and tested AI and ML solutions across diverse application domains, including vehicle autonomy,

radar, mis/dis-information, and chemical and biological threat detection.

- Fine-tuned **language models** to identify and annotate types of extremism within information-warfare contexts according to a literature-based ontology.
- Developed and maintained research applications using Python, C++, ReactJS, Docker, and GitLab CI/CD DevOps pipelines.
- Designed and implemented low- and no-code systems for studying neural-network architectures and transfer learning on radar data.
- Captured and executed more than **\$1M** in externally funded research for autonomy **modeling and simulation** and an AI contest.
- Led hybrid and remote teams of up to **12 researchers**; supervised 10 undergraduate interns and mentored full-time new hires and junior researchers.
- Contributed to proposals generating more than **\$6M** in external research funding and more than \$75K in internal research funding.
- Initiated and led an organization-wide grassroots AI-focused virtual journal club, leveraging connections across laboratories and divisions.
- Led and participated in **Agile** development teams using issue boards, daily stand-ups, sprint meetings, and point estimation.
- Presented to high-profile government sponsors, including AFRL, NRL, DTRA, Big Safari, and AFLCMC.

Cincinnati Children's Hospital Medical Center

2015–2016; 2017–2019

Student Researcher, Center for Autoimmune Genomics and Etiology

Cincinnati, Ohio

- Designed and implemented a well-cited visualization platform for amino-acid coevolution data.
- Built **deep-learning models** for protein metal-binding-site prediction and analyzed datasets containing over 695K records.
- Developed **proteomics** and **toxicogenomics** tools using JavaScript, Perl, and R; first author on **three peer-reviewed journal articles**.
- Named Employee of the Month in February 2016.

Kinetic Vision

2016–2017

Software Development Co-op, Web Development Team

Evendale, Ohio

- Developed full-stack web applications for corporate customers using Ruby on Rails, JavaScript, SCSS, ActiveRecord migrations, MVC design, and AWS cloud deployment.
- Collaborated in Agile development teams using issue boards, daily stand-ups, biweekly sprint meetings, and time/difficulty estimation.
- Designed backend database schemas and migrations; partnered with graphic designers to implement user-friendly frontend designs.

Ohio University Zanesville

January–August 2014

Peer Tutor, Student Success Center; Laboratory Assistant, Physics Laboratory

Zanesville, Ohio

- Tutored undergraduate mathematics from pre-algebra through univariate calculus.
- Built a Python visualization pipeline to download, extract, and analyze nuclear-physics data for more than 3,000 atomic nuclei; presented the work at the 2014 APS Ohio Section meeting.

Publications

Journal Articles

- J1. **F. N. Baker**, Z. Chen, D. Adu-Ampratwum, and X. Ning. “RLSynC: Offline-Online Reinforcement Learning for Synthon Completion.” *Journal of Chemical Information and Modeling*, 64(17), 6723–6735, 2024.
- J2. D. Concoran, N. Maltbie, S. Sudalairaj, **F. N. Baker**, J. Hirschfeld, and A. Porollo. “CoeViz 2: Protein Graphs Derived From Amino Acid Covariance.” *Frontiers in Bioinformatics*, 2021, 19.
- J3. **F. N. Baker** and A. Porollo. “CoeViz: A web-based integrative platform for interactive visualization of large similarity and distance matrices.” *Data*, 3(1), 4, 2018.
- J4. **F. N. Baker**, M. T. Cushion, and A. Porollo. “A quantitative model to estimate drug resistance in pathogens.” *Journal of Fungi*, 2(4), 30, 2016.
- J5. **F. N. Baker** and A. Porollo. “CoeViz: A web-based tool for coevolution analysis of protein residues.” *BMC Bioinformatics*, 17, 119, 2016.

Peer-Reviewed Conference Proceedings

- P1. **F. N. Baker**, D. Adu-Ampratwum, R. Averly, B. Yu, H. Sun, and X. Ning. “LARC: Towards Human-level Constrained

Retrosynthesis Planning through an Agentic Framework.” *Tianqiao and Chrissy Chen Institute Symposium on AI for Accelerated Science*, 2025. **Best Paper Award**.

- P2. R. Averly*, **F. N. Baker***, and X. Ning. “LIDDIA: Language-based Intelligent Drug Discovery Agent.” *EMNLP*, 2025. **Equal contribution*.
- P3. Z. Chen et al., including **F. N. Baker**. “ScienceAgentBench: Toward Rigorous Assessment of Language Agents for Data-Driven Scientific Discovery.” *International Conference on Learning Representations (ICLR)*, 2025.
- P4. B. Yu, **F. N. Baker***, Z. Chen*, G. Herb, B. Gou, D. Adu-Ampratwum, X. Ning, and H. Sun. “Tooling or Not Tooling? The Impact of Tools on Language Agents for Chemistry Problem Solving.” *Findings of NAACL*, 2025.
- P5. B. Yu, **F. N. Baker***, Z. Chen*, X. Ning, and H. Sun. “LlaSMol: Advancing Large Language Models for Chemistry with a Large-Scale, Comprehensive, High-Quality Instruction Tuning Dataset.” *Conference on Language Modeling (COLM)*, 2024.

Preprints

- A1. **F. N. Baker**, T. Nguyen, R. Averly, B. Yu, D. Adu-Ampratwum, H. Sun, and X. Ning. “MMORF: A Multi-agent Framework for Designing Multi-objective Retrosynthesis Planning Systems.” *arXiv:2604.05075*, 2026.
- A2. R. Gao, J. Peng, Z. Chen, **F. N. Baker**, D. C. Kombo, J. L. Kane, A. A. Scholte, Y. Li, M. J. LaMarche, L. I. Iconaru, H.-P. Biemann, M. Hong, and X. Ning. “Generating Developable 3D Molecules via Pocket-Conditioned Diffusion and Property-Aware Optimization.” *arXiv:2607.12349*, 2026.

Theses and Dissertations

- T1. **F. N. Baker**. *AI Agents for Retrosynthesis*. Ph.D. dissertation in progress, The Ohio State University.
- T2. **F. N. Baker**. *Mining and Visualization of Amino Acid Coevolution Data*. Master’s thesis, University of Cincinnati, 2019.

Inventions

- I1. “Systems and Methods for Pocket-Based Drug Design.” Provisional patent application submitted to the USPTO; assignee: The Ohio State University, 2025.
- I2. “Systems and Methods for Molecule Design.” Patent application submitted to the USPTO; assignee: The Ohio State University, 2025.

Research Funding

Multi-platform Autonomy Simulation Suite (MASS) <i>ARCNET/AFRL/DoW R&D contract; Project Director / Co-PI, GTRI/GTARC</i>	2020–2021 \$946K
Swarm and Search AI Hackathon (SSAI) <i>Wright Brothers Institute/AFRL/DoW R&D contract; Project Director / PI</i>	2021–2022 ~\$150K
Title not approved for public release <i>GTRI-ACL/DoW internal funding; Project Director / PI</i>	2021–2022 ~\$50K
Karat: A Visual Framework for Constructing Neural Network Architectures <i>GTRI-HIVES/DoW internal early-career funding; Project Director / PI</i>	2020–2021 \$25K

Presentations, Abstracts, & Posters

- C1. **F. N. Baker** et al. “LARC: Towards Human-level Constrained Retrosynthesis Planning through an Agentic Framework.” ACS Spring, Generative and Agentic Systems in Chemistry Session, 2026.
- C2. **F. N. Baker** et al. “LARC: Towards Human-level Constrained Retrosynthesis Planning through an Agentic Framework.” Ohio State CSE Graduate Student Visit Day, 2025.
- C3. Z. Chen, **F. N. Baker**, et al. “Methodological Advances in Generative AI for Drug Discovery: An AI-Centric Perspective.” AI Driven Drug Discovery Summit, 2024.
- C4. **F. N. Baker**, Z. Chen, and X. Ning. “RLSynC: Offline-Online Reinforcement Learning for Synthon Completion.” Ohio State CSE Graduate Student Visit Day, 2024.
- C5. **F. N. Baker**, Z. Chen, and X. Ning. “RLSynC: Offline-Online Reinforcement Learning for Synthon Completion.” Molecular Machine Learning Conference, 2023.

- C6. **F. N. Baker**. “Title not approved for public release.” GTRI IRAD Extravaganza, 2022.
- C7. M. Farrell, **F. Baker**, R. Hutchins, V. Srikanth, and B. Quick. “Population-based Fever Surveillance for Early Warning of Infectious Disease Outbreaks.” Chemical and Biological Defense Science and Technology Conference, 2022.
- C8. R. Hutchins, **F. Baker**, C. Thomas, and M. Lavelle. “Edge Computing for Global Genomics-based Threat Detection, Assessment, and Decision Support.” Chemical and Biological Defense Science and Technology Conference, 2022.
- C9. **F. N. Baker**. “Karat: A Visual Framework for Constructing Neural Network Architectures.” GTRI IRAD Extravaganza, 2021.
- C10. **F. N. Baker**. “Karat: A Visual Framework for Constructing Neural Network Architectures.” AI for Systems Engineering and Systems Engineering for AI Workshop, 2020.
- C11. **F. N. Baker**, N. Maltbie, J. Hirschfeld, and A. Porollo. “Metalloproteome Landscape from the Amino Acid Covariance Perspective.” 26th Conference on Intelligent Systems for Molecular Biology, 2018.
- C12. **F. N. Baker** and A. Porollo. “A Covariance-based Model to Estimate Antimicrobial Drug Resistance.” The Protein Society 30th Annual Symposium, 2016.
- C13. **F. N. Baker** and A. Porollo. “CoeViz: A Web-based Tool for Coevolution Analysis of Protein Residues.” 24th Conference on Intelligent Systems for Molecular Biology, 2016.
- C14. **F. N. Baker** and A. Porollo. “Accurate Prediction of Metal Binding Sites in Proteins.” 24th Conference on Intelligent Systems for Molecular Biology, 2016.
- C15. **F. N. Baker** and A. Porollo. “Prediction of Metal Binding Sites in Proteins Using Coevolution Data.” Great Lakes Bioinformatics Conference, 2016.
- C16. **F. N. Baker** and A. Porollo. “Comparative Analysis of Gene Expression Profiles Perturbed by Lupus-inducing Drugs.” Great Lakes Bioinformatics Conference, 2015.
- C17. G. Popa and **F. Baker**. “Systematics of Nuclear Level Densities.” International Workshop on Nuclear Theory, 2014, pp. 187–192.
- C18. **F. Baker**, B. Hina, L. Schneider, and G. Popa. “Automated Evaluation of Experimental Data for Nuclear Level Densities.” APS Ohio-Region Section Annual Spring Meeting, 2014.

Professional Service

Peer Review

- 2027: AAAI
- 2026: *Bioinformatics*; AAAI; ACM Multimedia (Main Conference and Dataset Track); COLM.
- 2025: COLM; NeurIPS; AAAI.
- 2024: BIBM; The Web Conference; WSDM; SIGIR.
- 2023: RecSys; ICDM; SIGIR.
- 2022: MICCAI; RecSys.

Conferences listed by date of meeting, not date of review. Journals listed in *italics*.

Panels and Committees

- Curriculum Committee, Ohio State Department of Computer Science and Engineering, Autumn 2024–Spring 2026.
- Ph.D. Student Panel, Ohio State CSE Graduate Student Visit Day, 2024.
- AI Expert Panelist, Swarm and Search Artificial Intelligence Competition, 2022.

Teaching

(Guest Lecturer) Quantitative Methods Capstone Seminar on Machine Learning
 Guest Lecturer / Panelist, The Ohio State University

February 26, 2025
 Columbus, Ohio

Awards, Honors, and Nominations

Best Paper Award, Chen Institute Symposium on AI for Accelerated Science	2025
Travel Grant, Chen Institute Symposium on AI for Accelerated Science	2025
College of Medicine Next Generation Innovator of the Year Nominee, Ohio State ERIK	2025
Mantai/Mae Scholarship, University of Cincinnati EECS	2018
Eta Kappa Nu Inductee, University of Cincinnati Chapter	2017
Employee of the Month (February), CAGE, Cincinnati Children's Hospital Medical Center	2016
Cincinnati University Scholar, University of Cincinnati	2014–2019
J. Donald Kincaid and Julianne Kincaid Scholarship	2014–2018
T. R. Murphy and Dorothy K. Murphy Scholarship	2014–2018
National Society of Collegiate Scholars Inductee	2014

Methods & Skills

Research areas	Agentic AI; generative AI; reinforcement learning; AI for science; natural language processing; bioinformatics; cheminformatics
Programming	Python; JavaScript; LaTeX; C++; C; R; MATLAB; Julia; Ruby on Rails; Perl; SQL; HTML; CSS; Bash
Frameworks	vLLM; Hugging Face; OpenAI API; Anthropic API; PyTorch; ReactJS
Engineering	AI agents; machine learning; deep learning; data visualization; Docker; cloud computing; GitLab CI/CD; Agile development; DevOps pipelines

Professional Memberships

Association for Computing Machinery (ACM) • American Chemical Society (ACS) • Ohio Academy of Science (OAS)
Institute of Electrical and Electronics Engineers (IEEE)

Literary Publications

L1. **F. Baker**. “Paper” [poem]. *Zane Traces Literary Journal*, 2014.

Frazier Baker

Last updated: August 2026