Austin Clyde

Data Science and Learning Division
Argonne National Laboratory
9700 South Cass Avenue, Bldg. 240
Lemont, IL 60439
Phone: (630) 252-3738
aclyde@anl.gov

3506 N Reta Ave, FL 1
Chicago, IL 60657
Phone: (585) 406-9146
aclyde11@gmail.com
austinclyde.com

RESEARCH INTERESTS

AI for science, high-performance computing, large language models, science and technology studies (STS).

EMPLOYMENT

Assistant Computational Scientist, Argonne National Laboratory 2022 to present
Research Assistant, Argonne National Laboratory 2019 to 2022
Desk Developer, XR Trading 2018 to 2019
Investment Associate Intern, Bridgewater Associates 2018 to 2018

EDUCATION

University of Chicago

Ph.D. Computer Science 2022
Thesis: *Artificial Intelligence and High-Performance Computing for Accelerating Structure-Based Drug Discovery*
Advisor: Prof. Rick Stevens

M.S. Computer Science 2019

B.A. Mathematics, with general honors 2019

FELLOWSHIPS

Visiting Research Fellow, Science, Technology, and Society, Harvard Kennedy School 2021 to 2022

AWARDS AND RECOGNITION

- **Pozen Family Center for Human Rights Graduate Lectureship** '22
  The Pozen Center awards up to three lectureships per academic year to advanced doctoral students, each of whom teaches one undergraduate Human Rights course of their design.
- **Special recognition in Department of Energy Secretary's Honor Award for COVID-19 Research '21**
  Annual award recognizing DOE employees and contractors for their service and contributions to the Department's mission and the benefit of the Nation.
- **ACM Gordon Bell Special Prize for High Performance Computing-Based COVID-19 Research**
  - **'22 Awardee**, *enSLMs: Genome-scale language models reveal SARS-CoV-2*
  - **'21 Finalist**, *#COVIDisAirborne: AI-Enabled Multiscale Computational Microscopy of Delta SARS-CoV-2 in a Respiratory Aerosol* and *Intelligent Resolution: Integrating Cryo-EM with AI-driven Multi-resolution Simulations to Observe the SARS-CoV-2 Replication-Transcription Machinery in Action*
  - **'20 Awardee**, *AI-driven multiscale simulations illuminate mechanisms of SARS-CoV-2 spike dynamics*
  Gordon Bell Prize, referred to as the 'Nobel Prize' of supercomputing, is awarded each year at Supercomputing to recognize outstanding achievement in high-performance computing.
- **Impact Argonne Award** for Discovery '20 and for Innovation '20.
  Internal Argonne Award for the first sighting of new knowledge which appears to lead the way to finding more knowledge (development of drug screening pipeline), and for the initial development of a new concept (reinforcement learning for drug discovery).

GRANTS

Active Grants *(PI)*

9/22 – 5/24 *Title:* "Finding Ligands Targeting the Conserved RNA Binding Site of SARS-CoV-2 NSP13"
Principal Investigator: **Austin Clyde**
Funder: CACHE (Critical Assessment of Computational Hit-finding Experiments) Challenge 2
Total Direct Costs: $31,500

12/21 – 10/23 *Title:* "Predict Hits for the WDR Domain of LRRK2"
Principal Investigator: **Austin Clyde**
Funder: CACHE (Critical Assessment of Computational Hit-finding Experiments) Challenge 1
Total Direct Costs: $31,500

Active Grants *(Non-PI, key personnel, and milestone author)*

1/22 – 12/25 *Title:* "IMPROVE: Innovative Methodologies and New Data for Predictive Oncology Model Evaluation"
Principal Investigators: Rick Stevens
Funder: National Cancer Institute
Total Annual Direct Costs: $2,400,00

4/22 – 9/23 *Title:* "Autonomous Discovery for Science"
Principal Investigators: Thomas Brettin
Funder: Argonne National Laboratory Directed Research and Development
Total Annual Direct Costs: $1,700,000

6/20 – 6/23 *Title:* "AI-driven Data Integration, Inference and Multi-scale Modeling Approaches to Low-dose Radiation Effects Understanding (RadBio-AI-3)"
Principal Investigators: Rick Stevens, Anuj Kapadia, Francis Alexander
Funder: Office of Science, Department of Energy
Total Annual Direct Costs: $5,000,000, Argonne: $2,000,000

PUBLICATIONS

Under review

2. **AI and Synthetic Biology: Towards Collaborative Regulation**
   Nicholas Nolan and **Austin Clyde.**
   MIT Science Policy Review on AI

1. **The Black Box Civic Epistemology: AI Senselessness and Democratic Participation**

In Preparation

6. **Certainty Through Uncertainty: Aleatoric Uncertainty in Computational Cancer Cell Line Drug Response Prediction**
5. **Forward-Forward Algorithm for Online Docking**
4. **The Hermeneutic Philosophy of Jasanoffian Co-Production**
   Austin Clyde and Aishani Aatresh.
3. **Reproducibility is not Usability: A Case Study in Cancer Drug Response Models**
2. **An Interpretable and Human-Centric Solution to the Chemical Enumeration Problem**
1. **Human Rights and AI for Science**

Peer-reviewed Journals

13. **AI-accelerated Protein-Ligand Docking for SARS-CoV-2 is 100-fold Faster With No Significant Change in Detection**
    Austin Clyde et al.
    Scientific Reports 2023
12. **Deep learning methods for drug response prediction in cancer: predominant and emerging trends**
    Alexander Partin, Thomas Brettin, Yitan Zhu, Oleksandr Narykov, **Austin Clyde**, et al.
    Frontiers in Medicine (Precision Medicine)
11. **AI-Accelerated Design of Targeted Covalent Inhibitors for SARS-CoV-2**
    Rajendra Joshi et al.
    Journal of Chemical Informatics 2022

10. **SARS-CoV-2 infects the human kidney and drives fibrosis in kidney organoids**
Jitske Jansen, et al. (incl. COVID Moonshot Consortium)
Cell stem cell '22.

9. **Scaffold-Induced Molecular Graph (SIMG): Effective Graph Sampling Methods for High-Throughput Computational Drug Discovery**
Austin Clyde, et al.
BMC Bioinformatics.

8. **Intelligent Resolution: Integrating Cryo-EM with AI-driven Multi-resolution Simulations to Observe the SARS-CoV-2 Replication-Transcription Machinery in Action**
Anda Trifan, Defne Gorgun, Zongyi Li, Alexander Brace, Maxim Zvyagin, Heng Ma, **Austin Clyde**, et al.
The International Journal of High Performance Computing 2022; Finalist for Gordon Bell Special Prize for HPC-Based COVID-19 Research '21.

7. **#COVIDisAirborne: AI-Enabled Multiscale Computational Microscopy of Delta SARS-CoV-2 in a Respiratory Aerosol**
Abigail Dommer, Lorenzo Casalino, Fiona Kearns, Mia Rosenfeld, Nicholas Wauer, Surl-Hee Ahn, John Russo, Sofia Oliveira, Clare Morris, Anthony Bogetti, Anda Trifan, Alexander Brace, Terra Sztain, **Austin Clyde**, et al.
The International Journal of High Performance Computing (to appear); Finalist for Gordon Bell Special Prize for HPC-Based COVID-19 Research '21.

6. **High Throughput Virtual Screening and Validation of a SARS-CoV-2 Main Protease Non-Covalent Inhibitor**
Austin Clyde, et al.
Journal of Chemical Informatics 2022 (front cover).

5. **Structural, electronic and electrostatic determinants for inhibitor binding to subsites S1 and S2 in SARS-CoV-2 main protease**
Daniel Kneller, Hui Li, Stephanie Galanie, Gwyndalyn Phillips, Audrey Labbe, Kevin Weiss, Qiu Zhang, Mark Arnould, **Austin Clyde**, et al.
Journal of Medicinal Chemistry 2021 (appeared as supplemental cover).

4. **A cross-study analysis of drug response predictions in cancer cell lines**
Fangfang Xia, Jonathan Allen, Prasanna Balaprakash, Thomas Brettin, Cristina Garcia-Cardona, **Austin Clyde**, et al.
Briefings in Bioinformatics 2021.

3. **Pandemic Drugs at Pandemic Speed: Accelerating COVID-19 Drug Discovery with Hybrid Machine Learning-and Physics-based Simulations on High Performance Computers**
Agastya Bhati, Shunzhou Wan, Dario Alfè, **Austin Clyde**, et al.
Interface Focus 2021.

2. **AI-driven multiscale simulations illuminate mechanisms of SARS-CoV-2 spike dynamics.**
Lorenzo Casalino, Abigail C Dommer, Zied Gaieb, Emilia P Barros, Terra Sztain, Surl-Hee Ahn, Anda Trifan, Alexander Brace, Anthony T Bogetti, **Austin Clyde**, et al.
The International Journal of High-Performance Computing Applications, Gordon Bell Special Prize for HPC-Based COVID-19 Research '20.

1. **Learning Curves for Drug Response Prediction in Cancer Cell Lines**
Alexander Partin, Thomas Brettin, Yvonne A Evrard, Yitan Zhu, Hyunseung Yoo, Fangfang Xia, Songhao Jiang, **Austin Clyde**, et al.
BMC Bioinformatics 2021.

Conference Proceedings

9. **enSLMs: Genome-scale language models reveal SARS-CoV-2**

Maxim Zvyagim, Alexander Brace, Kyle Hippe, Yuntian Deng, Bin Zhang, Cindy Orozco Bohorquez, **Austin Clyde**, et al.
Supercomputing 2022 (*Gordan-Bell Prize '22)*

8. **Spatial Graph Attention and Curiosity-driven Policy for Antiviral Drug Discovery**
Yulun Wu, Nicholas Choma, Andrew Chen, Mikaela Cashman, Érica T Prates, Manesh Shah, Verónica G Melesse Vergara, **Austin Clyde**, et al.
The 10th International Conference on Learning Representations (ICLR) 2022.

7. **Autodock on Enamine Hit Locator Library and Drugbank**
Hubertus van Dam, Martin Purschke, Dean Hidas, Oleg Tchoubar, Maksim Rakitin, Xiaohui Qu, **Austin Clyde**, et al.
IEEE Healthcare Summit Data Hackathon 2021.

6. **IMPECCABLE: Integrated Modeling PipelinE for COVID Cure by Assessing Better LEads**
Aymen Al Saadi, **Austin Clyde**, et al.
50th International Conference on Parallel Processing (ICPP 21).

5. **Stream-AI-MD: Streaming AI-driven Adaptive Molecular Simulations for Heterogeneous Computing Platforms**
Alexander Brace, Michael Salim, Vishal Subbiah, Heng Ma, Murali Emani, Anda Trifan, **Austin Clyde**, et al.
Platform for Advanced Scientific Computing (PASC '21).

4. **Scalable HPC and AI Infrastructure for COVID-19 Therapeutics" in the Platform for Advanced Scientific Computing**
Hyungro Lee, Andre Merzky, Li Tan, Mikhail Titov, Matteo Turilli, Dario Alfe, Agastya Bhati, Alex Brace, **Austin Clyde**, et al.
Platform for Advanced Scientific Computing (PASC '21).

3. **Targeting SARS-CoV-2 with AI-and HPC-enabled lead generation: a first data release**
Yadu Babuji, Ben Blaiszik, Tom Brettin, Kyle Chard, Ryan Chard, Austin Clyde, et al.
Platform for Advanced Scientific Computing (PASC '21).

2. **Benchmarking Machine Learning Workloads in Structural Bioinformatics Applications**
Heng Ma, **Austin Clyde**, et al.
First International Workshop on Benchmarking Machine Learning Workloads on Emerging Hardware '20.

1. **Virtual screening with deep learning using cancer cell line dose-response data**
Austin Clyde, et al.
American Association for Cancer Research '20.

Workshops

6. **Deep Surrogate Docking: Accelerating Automated Drug Discovery with Graph Neural Networks**
Ryien Hosseini, Filippo Simini, **Austin Clyde**, et al.
NeurIPS Workshop on AI for Science: Progress and Promises '22.

5. **Scaffold embeddings: Learning the structure spanned by chemical fragments, scaffolds and compounds**
Austin Clyde, et al.
NeurIPS Workshop on Learning Meaningful Representation of Life '21.

4. **Created in our Likeness: Is Open-Source AI at Odds with Animal Rights?**
Austin Clyde
MANCEPT Workshop on Politics, Animals, and Technology '21.

3. **Scaffold-Induced Molecular Graph (SIMG): Effective Graph Sampling Methods for High-Throughput Computational Drug Discovery**
Austin Clyde, et al.
Sixth Computational Approaches to Cancer Workshop at Supercomputing '20.

2. **Integrating High-Performance Simulations and Learning toward Improved Cancer Therapy**
   Austin Clyde, et al.
   Fifth Computational Approaches to Cancer Workshop at Supercomputing '19.
1. **Combining molecular simulation and machine learning to INSPIRE improved cancer therapy**
   David Wright, Adrian Devitt-Lee, **Austin Clyde**, et al.
   CompBioMed Conference '19.

Book Chapters

3. **Accelerating COVID-19 Drug Discovery with High-Performance Computing**
   Austin Clyde, to appear in *High Performance Computing (HPC) for Drug Discovery and Biomedicine* (ed. Heifetz).
2. **Large Language Models for Science**
   Austin Clyde, Arvind Ramanathan, Rick Stevens
   Forthcoming in *AI for Science.*
1. **Ultrahigh Throughput Protein–Ligand Docking with Deep Learning**
   Austin Clyde
   Artificial Intelligence in Drug Design, Methods in Molecular Biology. Springer, New York, NY, 2022.

Inventions and Patents

1. **Systems and Methods for Reinforcement Learning Molecular Modeling**
   Austin Clyde, Arvind Ramanathan, Rick Stevens
   Patent App. 17244711, 2022

Pre-prints

3. **Open Science Discovery of Oral Non-Covalent SARS-CoV-2 Main Protease Inhibitor Therapeutics**
   COVID Moonshot Consortium.
2. **Regression enrichment surfaces: a simple analysis technique for virtual drug screening models.**
   Austin Clyde, et al.
1. **A Systematic Approach to Featurization for Cancer Drug Sensitivity Predictions with Deep Learning**
   Austin Clyde, et al.

Public Opinion Pieces

4. **Are Social Media Platforms a Legitimate Component of Democracy?**
   Tech Policy Press '22.
3. **Algorithmic Systems Designed to Reduce Polarization Could Hurt Democracy, Not Help It**
   Tech Policy Press '22.
2. **AI for science and global citizens**
   Patterns, Cell Press '22.
1. **Human-in-the-Loop Systems Are No Panacea for AI Accountability**
   Tech Policy Press '21.

Public Comments

1. **Equitable Data Engagement and Accountability (87 FR 54269)**
   For the White House Office of Science and Technology Policy, '22.

Short Stories

1. **MyMuse**
   Future Humans: An Anthology (read at Harvard Kennedy School's STS Program's 20th Anniversary)

Presentations and Invited Talks

24. **A Possible Solution to the Chemical Enumeration Problem**
    ATOM Consortium 2023
23. **AI for Science: Drug Design**

ALCF Intro to AI-Driven Science Training, Chicago, 2022.

22. **Safeguarding Information Integrity**
Harvard Data Science Institute Trust in Science Workshop, Boston, 2022.
21. **Realigning Deep Learning with Virtual Ligand Screening**
AI-Driven Drug Discovery Summit, Boston, 2022.
20. **AI and Biology: A 100 Year History & AI-Driven Drug Discovery as an Emerging Technology**
Non-Proliferation Emerging Technology Seminar Series, Depts. of Commence and Energy, 2022.
19. **The Boundaries of AI Ethics in American Undergraduate Computer Science Education**
Science and Democracy Network, Harvard Kennedy School, Boston, 2022.
18. **Sharing What's Possible, Not Just What's Done**
Democratizing HPC for Science, Platform for Advanced Scientific Computing, Basel, 2022.
17. **End-to-End Computational Drug Design for COVID-19: From Screening to Series and Back Again**
CompBioMed2 Workshop, Rome, 2022.
16. **Automating Biological Discoveries at Scale**
ISC High Performance, Hamburg, 2022
15. **The Demos and the Expert: On Techniques of Self-Rule**
Night of Ideas, French Embassy and Harvard University, 2022.
14. **Throwing Away the Mirror: Why Algorithmic Disgorgement is an Epistemic Rights Violation**
A Right to Truth? Information, Communication, and Democracy in the 21st Century, HKS, 2022.
13. **How Does AI Change the Ethics of Synthetic Biology?**
Convergence of SynBio and Data Science: *Prospects, Promise and Perils*, Northwestern University, 2022.
12. **AI and Democracy**
University of Chicago Department of Computer Science Student Seminar 2021.
11. **The Human-in-the-Loop Learning Fallacy**
Harvard Kennedy School STS Fellows 2021.
10. **Virtual Screening with Deep Learning: Understanding Speed and Accuracy in Context**
Frontline Genomics Webinar 2021.
9. **AI for Drug Discovery**
ATOM Consortium, June 2021.
8. **Panelist**
Intellectual Property Management: A Seminar Series, Argonne National Laboratory 2021.
7. **Ensemble Models and Consensus Scoring for Computational Docking**
National Virtual Biotechnology Laboratory Seminar 2021.
6. **Integrating Simulations and Learning Towards Improved Cancer Therapy**
CompBioMed Webinar 2020.
5. **Tiered BFE Estimation Workflow**
CompBioMed Conference 2020.
4. **Accelerating Virtual Docking Screens with Deep Learning**
Janssen Pharmaceuticals Seminar 2020.
3. **Virtual Drug Screening with Deep Learning**
NIH.AI Workshop on Applications of Machine Learning for NGS and Drug Data 2019.
2. **Computational Drug Discovery at Scale**
ComBioMed Kick-off meeting 2019.
1. **Molecule Generation, Search, and Optimization**
ATOM Technical Meeting 2019.

TEACHING

University of Chicago

6. **Is the Public Square Online? Human Rights and Democracy on Social Media** (HMRT)
Instructor, Spring 2023.
5. **AI, Algorithms, and Human Rights** (HMRT 23450, CMSC 10450, MAAD 13450)

Instructor, Fall 2022
Awarded Pozen Family Center for Human Rights Graduate Lectureship
Cross-listed in Computer Science and Media Arts and Design.

4. **Introduction to Computer Science II** (C/C++) (CMSC 15200)
   Instructor, Summer 2021
   Department of Computer Science.
3. **Reinforcement Learning for Drug Discovery Practicum** (MPCS 57010)
   Instructor, Spring 2020
   Master's Program in Computer Science.
2. **Computational Thinking** (CAAP)
   Instructor, Summer 2022, 2021 & 2020
   University of Chicago Academic Achievement Program.
1. **Machine Learning in Biology and Medicine** (CMSC 35440)
   Graduate Teaching Assistant, Fall 2019
   Department of Computer Science.

TRAINING, MENTORING, AND ADVISING

1. Max Zvyagin, post-bac at Argonne
   Taught in MPCS course, mentored, and hired at ANL, advised research on LLMs and genomics.
2. Ryien Hosseini, MS student, University of Michigan,
   Advised internship project on high-performance computing and graph-neural networks.
3. Alex Brace, Ph.D. student, University of Chicago
   Mentored teaching skills for Python, advised research on LLMs and genomics.
4. Bharat Kale, Ph.D. student, Northern Illinois University
   Advised doctoral research on reinforcement learning for drug discovery. Mentored through publication process, research project design, and grant submission.
5. Xuefeng Liu, Ph.D. Student, University of Chicago
   Advised doctoral research on reinforcement learning for drug discovery. Mentored through publication process, research project design, and grant submission.
6. Ege Halac, high-school student, University of Chicago Laboratory School,
   Mentored research project on virtual ligand drug discovery for transient receptor potential cation channel subfamily V member 1 (TRPV1).

SERVICE AND OUTREACH

Outreach

- **Advisor**, SEED AI, 2022-.
- **Co-host** of Science Technology and Society Podcast for New Books Network, 2022-.
- **Member**, Association for Computing Machinery (ACM).
- **Panelist**, Socioeconomic Diversity Summit, University of Chicago, 2018.
- **Volunteer**, Hour of Code, 2018.

Service

- **Organizer**, AI and human rights workshop series between University of Chicago's Pozen Family Center for Human Rights and Harvard Kennedy School's Program on Science, Technology, and Society.
- **Organizer**, panel, AI Across America, 2022.
  First-generation low-income students presented to members of the congressional AI Caucus and technology company executives.
- **Organizer**, minisymposium titled "Democratizing AI" at Platform for Advanced Scientific Computing (PASC '22).
- **Committee member**, AI for Science and Security Initiative, U.S. National Laboratory System
- **Organizer**, Harvard STS Writing Group for graduate students, fellows, and visitors.

Chairing and Reviewing

- **Area chair** (chemistry and small molecules), NeurIPS workshop, Learning Meaningful Representations of Life at NeurIPS (2022).
- **Reviewer**, Cell Patterns (2021, 2022).
- **Reviewer,** Journal of Biological Engineering (2022).
- **Reviewer**, NeurIPS workshop on Machine Learning and the Physical Sciences (2021, 2022).
- **Reviewer**, NeurIPS dataset and benchmark track (2022).
- **Reviewer**, ICML (2022).

REFERENCES

4. **Professor Rick Stevens**
   Professor of computer science at the University of Chicago and associate laboratory director for Computing, Environment and Life Sciences at Argonne National Laboratory
   Email: stevens@anl.gov

3. **Dr. Arvind Ramanathan**
   Computational Science Group Leader, Data Science and Learning Division at Argonne National Laboratory and a senior scientist at the University of Chicago Consortium for Advanced Science and Engineering
   Email: ramanathana@anl.gov

2. **Professor Eric Jonas**
   Assistant Professor of computer science at the University of Chicago
   Email: ericj@uchicago.edu

1. **Professor Shantenu Jha**
   Professor of Computer Engineering at Rutgers University and the Chair of the Department (Center) for Data Driven Discovery at Brookhaven National Laboratory
   Email: shantenu.jha@rutgers.edu