Austin Clyde

Chicago, IL USA
aclyde@uchicago.edu
austinclyde.com
585-406-9146

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

EMPLOYMENT
Assistant Computational Scientist, Argonne National Laboratory, Data Science and Learning Division, 2022-.
Research Assistant, Argonne National Laboratory, Data Science and Learning Division, 2019-2021.
Desk Developer, Cryptocurrencies, XR Trading, 2018 to 2019.

EDUCATION
PhD in Computer Science (est. December 2022) – University of Chicago (advisor: Rick Stevens).
MS in Computer Science, June 2019 – University of Chicago.
BA in Mathematics, June 2019 – University of Chicago.

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

AWARDS AND RECOGNITION
- Pozen Family Graduate Lectureship in Human Rights '22.
- Recognition in the Secretary's Honor Awards for Department of Energy National Virtual Biotechnology Laboratory for COVID-19 effort '21.
- Finalist, ACM Gordon Bell Special Prize for High Performance Computing-Based COVID-19 Research '21.
- ACM Gordon Bell Special Prize for High Performance Computing-Based COVID-19 Research '20.
- Impact Argonne Award for Discovery '20.
- Impact Argonne Award for Innovation '20.

PUBLICATIONS
In Preparation
1. **Protein-Ligand Docking Surrogate Models: A SARS-CoV-2 Benchmark for Deep Learning Accelerated Virtual Screening**
   Austin Clyde, Xuefeng Liu, Thomas Brettin, Alexander Partin, Hyunseung Yoo, Yadu Babuji, Ben Blaiszik, Andre Merzky, Matteo Turilli, Shantenu Jha, Arvind Ramanathan, Rick Stevens.

2. **If Algorithms are not Black Boxes, then What is the Ideal AI Citizen?**
   Austin Clyde.

3. **Uncertainty Calibrated Cancer Drug-Response Prediction**
   Austin Clyde, Rajeev Jain, Fangfang Xia, Tom Brettin, Rick Stevens.

4. **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).

Peer-reviewed Journals
1. **Scaffold-Induced Molecular Graph (SIMG): Effective Graph Sampling Methods for High-Throughput Computational Drug Discovery**
   Austin Clyde, Ashka Shah, Max Zvyagin, Arvind Ramanathan, Rick Stevens
   BMC Bioinformatics.

2. **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**, David A Clark, Michael Salim, David Hardy, Tom Burnley, Lei Huang, John McCalpin, Murali Emani, Hyunseung

Yoo, Junqi Yin, Aristeidis Tsaris, Vishal Subbiah, Jessica Liu, Noah Trebesch, Geoffrey Wells, Venkatesh Mysore, Tom Gibbs, James Phillips, S Chakra Chennubhotla, Ian Foster, Rick Stevens, Anima Anandkumar, Venkatram Vishwanath, John E Stone, Emad Tajkhorshid, Sarah A Harris, Arvind Ramanathan
The International Journal of High Performance Computing (to appear); Finalist for Gordon Bell Special Prize for HPC-Based COVID-19 Research '21.

3. **#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**, Heng Ma, Chakra Chennubhotla, Hyungro Lee, Matteo Turilli, Syma Khalid, Teresa Tamayo-Mendoza, Matthew Welborn, Anders Christensen, Daniel GA Smith, Zhuoran Qiao, Sai Krishna Sirumalla, Michael O'Connor, Frederick Manby, Anima Anandkumar, David Hardy, James Phillips, Abraham Stern, Josh Romero, David Clark, Mitchell Dorrell, Tom Maiden, Lei Huang, John McCalpin, Christopher Woods, Alan Gray, Matt Williams, Bryan Barker, Harinda Rajapaksha, Richard Pitts, Tom Gibbs, John Stone, Daniel Zuckerman, Adrian Mulholland, Thomas Miller III, Shantenu Jha, Arvind Ramanathan, Lillian Chong, Rommie Amaro
The International Journal of High Performance Computing (to appear); Finalist for Gordon Bell Special Prize for HPC-Based COVID-19 Research '21.

4. **High Throughput Virtual Screening and Validation of a SARS-CoV-2 Main Protease Non-Covalent Inhibitor**
Austin Clyde, Stephanie Galanie, Daniel W Kneller, Heng Ma, Yadu Babuji, Ben Blaiszik, Alexander Brace, Thomas Brettin, Kyle Chard, Ryan Chard, Leighton Coates, Ian Foster, Darin Hauner, Vilmos Kertesz, Neeraj Kumar, Hyungro Lee, Zhuozhao Li, Andre Merzky, Jurgen G Schmidt, Li Tan, Mikhail Titov, Anda Trifan, Matteo Turilli, Hubertus Van Dam, Srinivas Chennubhotla, Shantenu Jha, Andrey Kovalevsky, Arvind Ramanathan, Marti Head, Rick Stevens
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**, Heng Ma, Arvind Ramanathan, Marti Head, Coates, John M. Louis, Peter Bonnesen, Andrey Kovalevsky
Journal of Medicinal Chemistry 2021 (appeared as supplemental cover).

6. **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**, Judith Cohn, James Doroshow, Xiaotian Duan, Veronika Dubinkina, Yvonne Evrard, Ya Ju Fan, Jason Gans, Stewart He, Pinyi Lu, Sergei Maslov, Alexander Partin, Maulik Shukla, Eric Stahlberg, Justin M Wozniak, Hyunseung Yoo, George Zaki, Yitan Zhu, Rick Stevens
Briefings in Bioinformatics 2021.

7. **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**, Mathis Bode, Li Tan, Mikhail Titov, Andre Merzky, Matteo Turilli, Shantenu Jha, Roger Highfield, Walter Rocchia, Nicola Scafuri, Sauro Succi, Dieter Kranzlmüller, Gerald Mathias, David Wifling, Yann Donon, Alberto Di Meglio, Sofia Vallecorsa, Heng Ma, Anda Trifan, Arvind Ramanathan, Tom Brettin, Alexander Partin, Fangfang Xia, Xiaotan Duan, Rick Stevens, Peter Coveney
Interface Focus 2021.

8. **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**, Heng Ma, Hyungro Lee, Matteo Turilli, Syma Khalid, Lillian T Chong, Carlos Simmerling, David J Hardy, Julio DC Maia, James C Phillips, Thorsten

Kurth, Abraham C Stern, Lei Huang, John D McCalpin, Mahidhar Tatineni, Tom Gibbs, John E Stone, Shantenu Jha, Arvind Ramanathan, Rommie E Amaro
The International Journal of High-Performance Computing Applications, Gordon Bell Special Prize for HPC-Based COVID-19 Research '20.

9. **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**, Maulik Shukla, Michael Fonstein, James H Doroshow, Rick Stevens
BMC Bioinformatics 2021.

Conference Proceedings

1. **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**, Thomas S Brettin, Wibe A de Jong, Neeraj Kumar, Martha S Head, Rick L Stevens, Peter Nugent, Daniel A Jacobson, James B Brown
The 10th International Conference on Learning Representations (ICLR) 2022.

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

2. **IMPECCABLE: Integrated Modeling PipelinE for COVID Cure by Assessing Better LEads**
Aymen Al Saadi, **Austin Clyde**, Dario Alfe, Yadu Babuji, Agastya Bhati, Ben Blaiszik, Thomas Brettin, Kyle Chard, Ryan Chard, Peter Coveney, Anda Trifan, Alex Brace, Ian Foster, Tom Gibbs, Shantenu Jha, Kristopher Keipert, Thorsten Kurth, Dieter Kranzlmüller, Hyungro Lee, Zhuozhao Li, Heng Ma, Andre Merzky, Gerald Mathias, Alexander Partin, Junqi Yin, Arvind Ramanathan, Ashka Shah, Abraham Stern, Rick Stevens, Li Tan, Mikhail Titov, Aristeidis Tsaris, Matteo Turilli, Huub Van Dam, Shunzhou Wan, David Wifling
50th International Conference on Parallel Processing (ICPP 21).

3. **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**, Corey Adams, Thomas Uram, Hyunseung Yoo, Andrew Hock, Jessica Liu, Venkatram Vishwanath, Arvind Ramanathan
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**, Peter Coveney, Heng Ma, Arvind Ramanathan, Rick Stevens, Anda Trifan, Hubertus Van Dam, Shunzhou Wan, Sean Wilkinson, Shantenu Jha
Platform for Advanced Scientific Computing (PASC '21).

5. **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**, Ian Foster, Zhi Hong, Shantenu Jha, Zhuozhao Li, Xuefeng Liu, Arvind Ramanathan, Yi Ren, Nicholaus Saint, Marcus Schwarting, Rick Stevens, Hubertus van Dam, Rick Wagner
Platform for Advanced Scientific Computing (PASC '21).

6. **Benchmarking Machine Learning Workloads in Structural Bioinformatics Applications**
Heng Ma, **Austin Clyde**, Anda Trifan, Venkatram Vishwanath, Arvind Ramanathan, Debsindhu Bhowmik, Shantenu Jha
First International Workshop on Benchmarking Machine Learning Workloads on Emerging Hardware '20.

7. **Virtual screening with deep learning using cancer cell line dose-response data**
Austin Clyde, Arvind Ramanathan, Rick Stevens

American Association for Cancer Research '20.

Workshops
1. **Scaffold embeddings: Learning the structure spanned by chemical fragments, scaffolds and compounds**
   Austin Clyde, Bharat Kale, Maoyuan Sun, Michael Papka, Arvind Ramanathan, Rick Stevens
   NeurIPS Worb kshop on Learning Meaningful Representation of Life '21.

2. **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, Ashka Shah, Max Zvyagin, Arvind Ramanathan, Rick Stevens
   Sixth Computational Approaches to Cancer Workshop at Supercomputing '20.

4. **Integrating High-Performance Simulations and Learning toward Improved Cancer Therapy**
   Austin Clyde, David Wright, Shantenu Jha
   Fifth Computational Approaches to Cancer Workshop at Supercomputing '19.

5. **Combining molecular simulation and machine learning to INSPIRE improved cancer therapy**
   David Wright, Adrian Devitt-Lee, **A Clyde**, Kiruba Palani, Fangang Xia, Matteo Turilli, John Karanicolas, Shantenu Jha, Rick Stevens, John Chodera, Peter Coveney
   CompBioMed Conference '19.

Book Chapters.
1. **Large Language Models for Science**
   Austin Clyde, Arvind Ramanthan, Rick Stevens
   Forthcoming in *AI for Science.*

2. **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. **RLMM: Reinforcement Learning for Molecular Mechanics**
   Austin Clyde, Arvind Ramanathan, Rick Stevens
   ANL-IN-20-123, patent filed.

Pre-prints
1. **Open Science Discovery of Oral Non-Covalent SARS-CoV-2 Main Protease Inhibitor Therapeutics**
   COVID Moonshot Consortium.

2. **COVID Moonshot: Open Science Discovery of SARS-CoV-2 Main Protease Inhibitors by Combining Crowdsourcing, High-Throughput Experiments, Computational Simulations, and Machine Learning**
   COVID Moonshot Consortium.

3. **Regression enrichment surfaces: a simple analysis technique for virtual drug screening models.**
   Austin Clyde, Xiaotian Duan, Rick Stevens.

4. **A Systematic Approach to Featurization for Cancer Drug Sensitivity Predictions with Deep Learning**
   Austin Clyde, Tom Brettin, Alexander Partin, Maulik Shaulik, Hyunseung Yoo, Yvonne Evrard, Yitan Zhu, Fangfang Xia, Rick Stevens.

Public Opinion Pieces
1. **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.

3. **Human in The Loop and Democracy**
   Tech Policy Press '21.

Short Stories
   1. **MyMuse**
      Future Humans: An Anthology

TEACHING

1. **AI, Algorithms, and Human Rights** (HMRT 23450, CMSC 10450, MAAD 13450)
   Instructor, Fall 2022
   Awarded Pozen Family Center for Human Rights Graduate Lectureship, University of Chicago
   Cross-listed in Computer Science and Media Arts and Design.

2. **Introduction to Computer Science II** (C/C++) (CMSC 15200)
   Instructor, Summer 2021
   Department of Computer Science, University of Chicago.

3. **Reinforcement Learning for Drug Discovery Practicum** (MPCS 57010)
   Instructor, Spring 2020
   University of Chicago Master's Program in Computer Science.

4. **Computational Thinking** (CAAP)
   Instructor, Summer 2022, 2021 & 2020
   University of Chicago Academic Achievement Program.

5. **Machine Learning in Biology and Medicine** (CMSC 35440)
   Graduate Teaching Assistant, Fall 2019
   University of Chicago.

SERVICE AND OUTREACH

- Organizer of 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.
- Panel organizer for AI Across America with first-generation low-income students along with members of the congressional AI Caucus and technology company executives.
- Co-host of Science Technology and Society Podcast for New Books Network
- Organizer for minisymposium titled "Democratizing AI" at Platform for Advanced Scientific Computing (PASC '22).
- Reviewer for Cell Patterns (2021, 2022), NeurIPS workshop on Machine Learning and the Physical Sciences (2021, 2022), NeurIPS dataset and benchmark track (2022), ICML (2022)
- Volunteer, Hour of Code, December 2018.
- Panelist, Socioeconomic Diversity Summit, University of Chicago, 2018.
- Association for Computing Machinery (ACM) Member.

GRANTS

- CACHE Challenge (PI: 10,000$), 2021

TRAINING, METORING, AND ADVISING

1. Max Zvyagin, post-bac student at ANL, deep learning for discovery.
2. Ryien Hosseini, master's student, deep learning for discovery.
3. Bharat Kale, Ph.D. student, chemical informatics, and visualization.

AREAS OF EXPERTISE

Artificial intelligence and machine learning, scientific high-performance computing, scientific workflows, AI hardware accelerators, C/C++/Python, computational thinking, algorithms, and data science.

AI ethics, science and technology studies (STS), and human rights.

PRESENTATIONS AND TALKS

1. **AI for Science: Drug Design**
   ALCF Intro to AI-Driven Science Training, Chicago, 2022.

2. **Safeguarding Information Integrity**
   Harvard Data Science Institute Trust in Science Workshop, Boston, 2022.

3. **Realigning Deep Learning with Virtual Ligand Screening**
   AI-Driven Drug Discovery Summit, Boston, 2022.

4. **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.

5. **The Boundaries of AI Ethics in American Undergraduate Computer Science Education**
   Science and Democracy Network, Harvard Kennedy School, Boston, 2022.

6. **Sharing What's Possible, Not Just What's Done**
   Democratizing HPC for Science, Platform for Advanced Scientific Computing, Basel, 2022.

7. **End-to-End Computational Drug Design for COVID-19: From Screening to Series and Back Again**
   CompBioMed2 Workshop, Rome, 2022

8. **Automating Biological Discoveries at Scale**
   ISC High Performance, Hamburg, 2022

9. **The Demos and the Expert: On Techniques of Self-Rule**
   Night of Ideas, French Embassy and Harvard University, 2022.

10. **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.

11. **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.

13. **The Human-in-the-Loop Learning Fallacy**
    Harvard Kennedy School STS Fellows 2021.

14. **Virtual Screening with Deep Learning: Understanding Speed and Accuracy in Context" at Emerging AI/ML Technologies for Drug Discovery**
    Frontline Genomics Webinar 2021.

15. **AI for Drug Discovery**
    ATOM Consortium, June 2021.

16. **Panelist**
    Intellectual Property Management: A Seminar Series, Argonne National Laboratory 2021.

17. **Ensemble Models and Consensus Scoring for Computational Docking**
    National Virtual Biotechnology Laboratory Seminar 2021.

18. **Integrating Simulations and Learning Towards Improved Cancer Therapy**
    CompBioMed Webinar 2020.

19. **Tiered BFE Estimation Workflow**
    CompBioMed Conference 2020.

20. **Accelerating Virtual Docking Screens with Deep Learning at**

Janssen Pharmaceuticals Seminar 2020.

21. **Virtual Drug Screening with Deep Learning**
NIH.AI Workshop on Applications of Machine Learning for NGS and Drug Data 2019.

22. **Computational Drug Discovery at Scale**
ComBioMed Kick-off meeting 2019.

23. **Molecule Generation, Search, and Optimization**
ATOM technical Meeting 2019