Carl N. Edwards

Champaign-Urbana, Illinois ● cne2@illinois.edu ● cnedwards.com ● ● ●

Education

University of Illinois Urbana-Champaign — PhD Candidate in Computer Science — Fall 2020 — Present
Focus: Artificial Intelligence, NLP for Molecular and Drug Discovery
Advisor: Professor Heng Ji
RA Experience – NSF Molecule Maker Lab Institute
- Integrating molecular and natural language information.

University of Tennessee-Knoxville — Honors Computer Science BS — May 2020
Honors Math Double Major
STEM GPA: 4.0/4.0 — Overall GPA: 3.99/4.0
Summa Cum Laude — Chancellor's Honors Program
Mathematics Honors Program — Engineering Honors Program

University of Zürich — EuroScholars Research Study Abroad — Fall Semester 2018
Dynamic and Distributed Information Systems Group
Project: Linking Knowledge Graphs and Images Using Embeddings

Experience

University of Illinois Urbana-Champaign – MMLI — PhD Candidate & Research Assistant — Fall 2020-Present
- Developed literature-based systems for drug property improvement (in progress, details withheld).
- Researched integrating molecular and natural language information. Work proposes *Text2Mol* and *MolT5* models enabling novel downstream tasks of cross-modal retrieval of molecules from natural language queries, generation of molecules from textual descriptions, and molecule captioning. Overall research direction is towards language-enabled functional control of molecule design for applications such as drug development, organic photovoltaics, and advanced material design.
- Conducted information extraction, information retrieval, and text mining research to enable the *AlphaSynthesis* platform, which allows AI-assisted synthesis planning for molecular discovery and manufacturing based on large-scale analysis of scientific literature.

Allen Institute for Artificial Intelligence (AI2) – Semantic Scholar — Research Intern — Summer 2022
- Developed an AI model, SynerGPT, which uses a novel pretraining strategy to enable in-context learning for few-shot drug synergy predictions for rare drugs and cell lines. The resulting paper is under review.
- Investigated context optimization for SynerGPT with future applications in creating a standardized assay for patient tumor biopsies that can enable patient-specific drug synergy prediction.
- Developed a methodology for a novel approach to drug structure design for desirable synergy pairs.

Carnegie Mellon University – Auton Lab — Robotics Institute Summer Scholar (NSF REU) — Summer 2019
- Detected organizations in multimodal sex trafficking dataset consisting of over 40 million data points.
- Integrated multiple similarity measures using face detection, word embeddings, and regex-extracted features to detect organizations in tens of millions of escort advertisements scraped from online sources.

University of Zurich – Dynamic and Distributed Information Systems Group — EuroScholar — Fall 2018
- Crafted two linked datasets between knowledge graphs and images.
- Incorporated embedding models based on InceptionNet and TransE into a joint embedding model.

Oak Ridge National Laboratory – CISR — Department of Energy SULI Intern — Summer 2018
- Researched global optimization algorithms for subarrayed phase-only radar beam synthesis.

University of Tennessee – Material Research and Innovation Lab — Undergraduate Research Assistant — Fall 2016-2018

- Performed Brownian dynamics simulations of flowing polymer solutions.
- Created visualizations from resulting data and computed solution physical properties.
- Prepared manuscript for publication in Journal of Molecular Graphics and Modelling.

Oak Ridge National Laboratory — CISR Higher Education Research Experiences Intern Summer 2017

- Developed radar simulations using proprietary technology.
- Designed phased array radar beam width optimization algorithms (GA, particle swarm, simulated annealing).

Nanomechanics, Incorporated. Software Engineering Intern Summer 2016

- Programmed proprietary software to interpret high speed data using C++ .
- Developed programs for Linux, Raspberry Pi, and ALSA libraries.

Preprints

C. Edwards, A. Naik, T. Khot, M. Burke, H. Ji, and T. Hope. "SynerGPT: In-Context Learning for Personalized Drug Synergy Prediction and Drug Design." *bioRxiv preprint. bioRxiv:2023.07.06.547759.* 2023.

Conference Publications †

C. Edwards *, T. Lai*, K. Ros, G. Honke, K. Cho, and H. Ji. "Translation between Molecules and Natural Language." In *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing* (EMNLP2022). 2022.[2]

[2] (Oral presentation, Top 4.8% of submissions)

C. Edwards, CX. Zhai, and H. Ji. "Text2Mol: Cross-modal Molecule Retrieval with Natural Language Queries." In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing* (EMNLP2021). 2021.

C. **Edwards** and H. Ji. "Semi-supervised New Event Type Induction and Description via Contrastive Loss-Enforced Batch Attention." In *Proceedings of The 17th Conference of the European Chapter of the Association for Computational Linguistics* (EACL2023). 2023.

K. Ros*, **C. Edwards***, H. Ji, and CX. Zhai. "Team Skeletor at Touché 2021: Argument Retrieval and Visualization for Controversial Questions." *CEUR Workshop Proceedings*. Vol. 2936. CEUR-WS. 2021.[3]

[3] Ranked 2nd/22 teams for retrieval quality.

X. Du, […], **C. Edwards**, […] and H. Ji. "Resin-11: Schema-guided event prediction for 11 newsworthy scenarios." In *Proceedings of the 2022 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies: System Demonstrations* (pp. 54-63). 2022.

† In my field of computer science, main conference publications at top venues are considered equivalent to journal articles and are often the preferred venue of publication.
* indicates equal contribution

Journal Articles

C. Edwards, M.H. Nafar Sefiddashti, B.J. Edwards, and B. Khomami. "In-plane and out-of-plane rotational motion of individual chain molecules in steady shear flow of polymer melts and solutions." *J. Mol. Graph. Model*., *81*, 184-196. 2018.

C. Edwards, A. Wertz, and A. Dubrawski. "Using Similarity Measures to Detect Organizations in Online Escort Advertisements." *Robotics Institute Summer Scholar' Working Papers Journal*, 7, 43-49. 2019.

Presentations

M.H. Nafar Sefiddashti, **C.N. Edwards**, B.J. Edwards, and B. Khomami. "Out-of-plane rotational motion in shear flow of polymer melts and solutions." The Society of Rheology 89th Annual Meeting. Denver, CO. October 8-12, 2017.

Other

B. Daniel, **C. Edwards**, and A. Anderson. "Phase-Only Beam Broadening of Contiguous Uniform Subarrayed Arrays Utilizing Three Metaheuristic Global Optimization Techniques." *arXiv preprint arXiv:2009.06123.* 2020.

Invited Talks

- First International Conference for the Center of the Transformation of Chemistry (CTC), Ringberg Conference, Max Planck Institute of Colloids and Interfaces, September 2023
- NVIDIA, November 2022, "Translation Between Molecules and Natural Language"

Teaching, Leadership, and Service

- **Program Committee:** ACL-IJCNLP 2021**,** ACL 2022, 2023, NAACL 2022, EACL 2023
- **NSF Molecule Maker Laboratory Institute:** Student and Postdoc Council Educational & Outreach Activities Chair
- **Teaching Assistant Experience**:
  - CS 125 – Introduction to Computer Science (Fall 2020)
  - CS 412 – Introduction to Data Mining (Spring 2021)
- **Undergraduate Mentorship:**
  - **Summer 2023:** Mentoring three undergraduates on projects related to language-enabled protein design, language-molecule association rule mining with language models on scientific literature, and scientific language model factuality evaluation and updating for knowledge base construction.
- **Presentations to High School Students:**
  - **Summer 2019:** Presentation on "Detecting Sex Trafficking Organizations" with AI4All@CMU
  - **April 2021, 2022:** Illinois CS Sail course on "Learning Word Representations"
  - **June 2023:** "AI and NLP for Drug Discovery and My Path There", MMLI Summer Camp
- **Presentations to Middle School Students:**
  - **March 2023:** Presentation on "Intro to AI for Chemistry" for underrepresented middle school girls.
  - **April 2023:** Presentation on "Intro to AI for Chemistry" for underrepresented middle school boys.

Skills

- **Programming**: Proficient in Python, C++, Matlab, and Java; Experience with C, Fortran, LaTeX, HTML, and CSS. Experience with Windows and Linux.
  - Proficient with PyTorch, Lightning, TensorFlow, Keras, pandas, NumPy, SciPy, NLTK, spaCy, RDKit
- **Computational Methods**: Natural language processing (NLP), information retrieval, information extraction, text mining, natural language generation, large language models, representation learning (text, knowledge graph, molecule, image, etc), deep learning, knowledge graphs, multimodal data, computer vision, reinforcement learning, scientific computing, statistical decision theory, cheminformatics, molecule generation, molecule property prediction, quantitative structure-activity relationship (QSAR) prediction.

News Coverage

- MolT5: Chinese News, Social Media (upper bound of 464,449 Twitter followers)
- Undergraduate: Edwards is Top Fellowship/Scholarship Student in Computer Science
  - Math Department: Outstanding Math Honors Students Pursuing Advanced Degrees
  - Goldwater Scholarship Coverage: University-wide, Math Department

Awards

- University of Chicago DSI AI & Science Summer School Acceptance (2023)
- Saburo Muroga Endowed Fellowship, UIUC (2020) – awarded to outstanding computer science graduate students
- **Goldwater Scholarship** (2019) – "the most prestigious undergraduate scholarship given in the natural sciences, engineering and mathematics" in the United States (Wikipedia)

- UTK Outstanding Computer Science Junior (2018) – awarded to a single junior in computer science based on academic merit
- Min H. Kao Scholar ($7,500) (2018, 2019) – roughly six students selected from EE, CS, and CE majors based on academic merit
- Pi Mu Epsilon Math Society Award Membership (2019)
- Schmitt Memorial Scholarship (Math) (2019)
- Thomas & Kathryn Shelton Award (2017, 2018, 2019)
- Volunteer Scholarship (2016-2020)
- UT Provost Scholarship (2016-2020)
- National Merit Scholar (2016)
- National AP Scholar (2016)
- State of Tennessee Governor's School for Computational Physics Attendee (2015)

Relevant Coursework

- Knowledge-Driven Natural Language Generation
- Data Mining Principles
- Natural Language Processing
- Deep Generative and Dynamical Models
- Advanced Information Retrieval
- Text Mining: A New Paradigm
- Transfer Learning
- Biologically-Inspired Computation
- Reinforcement Learning
- Introduction to Pattern Recognition

Activities

- **Illini Dancesport:** 2022-present, Executive Board Member
- **UTK Machine Learning Club:** 2017-2020, Executive Board Member
- **HackUTK**: 2016-2020, UTK cybersecurity organization, VolHacks Hackathon 2016, 2017, 2019
- **Taekwondo**: 2007-2016, martial art, black belt
- **Classical Singing**: 2015-2020
- **FIRST Robotics**: 2011-2016**,** Programming leader and team co-leader