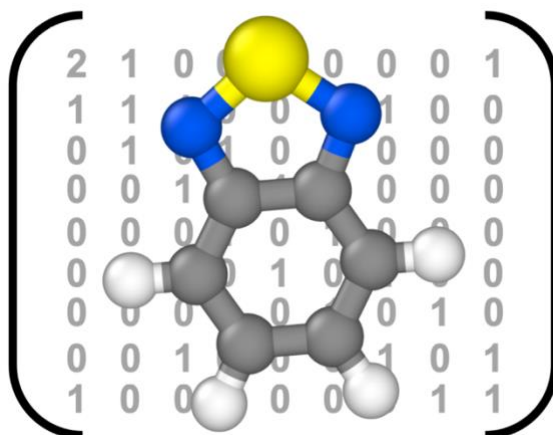


Thomas Remec and Susan Morisato Undergraduate Scholarship in Chemical Data Science

This scholarship will support two undergraduate students to perform research in Summer 2023 at the interface of data science and chemistry. Awards will be \$3,500.

Eligibility:

- Applicants must be current UIUC undergraduates (freshman, sophomore, or junior) in good academic standing.



Application:

All applications must include a resume, unofficial transcript, and **one page** statement of purpose. In the statement of purpose, please include:

- A description of your future career goals and what you hope to gain from the summer research experience.
- Any relevant background, academic experiences, skills, or previous research experiences.
- Your ranked first and second choice research projects (see back side) and your reasons for selecting them.

The resume, transcript, and statement of purpose must be emailed to jacksonn@illinois.edu as a single .pdf file on or before February 1.

Timeline:

- November 1 – Applications open.
- February 1 – Deadline for applications.
- April 1 – Awards announced.

Applications will be reviewed by a committee of Chemistry Department faculty members. Please direct any application questions to Prof. Jackson (jacksonn@illinois.edu).

Available Research Projects (Alphabetical)

(1) Prof. Mikael Backlund (Chemistry)

Title: *Machine learning of quantum measurement schemes to saturate fundamental information bounds in single-molecule microscopy*

Description: Single-molecule fluorescence microscopy is a powerful tool in the study of biophysics, materials, and molecular systems. In these experiments, photons are a precious resource, and the experimenter must think critically about how to best divide photons in time and space for the desired statistical inference task. In this project we will use machine learning algorithms to reveal metrological schemes that attain the quantum limits of single molecule microscopy.

(2) Prof. Scott Denmark (Chemistry)

Title: *Graph-based feature reduction methods for conformer-dependent grid-based descriptors*

Description: Machine learning is a powerful tool for optimizing asymmetric catalytic systems. Conformer-dependent grid-based descriptors will be calculated for catalysts spanning various classes of reactivity for which high quality experimental data exists. These descriptors will be processed with graph-based feature reduction methods to produce information-rich chemical feature spaces of use for QSPR models of catalyst-dependent asymmetric induction.

(3) Prof. Martin Gruebele (Chemistry)

Title: *Physics-constrained machine learning for in-cell protein dynamics*

Description: All-atom computation of in-cell dynamics of proteins and RNAs generates large datasets (TB) that require data mining and AI techniques to help extract information. We are particularly interested in physical model-constrained machine learning to extract information about protein-protein interactions and solvation dynamics from these data sets.

(4) Prof. Nick Jackson (Chemistry)

Title: *Extracting simplified quantum mechanical models using unsupervised machine learning*

Description: Quantum chemical calculations bear exorbitant computational costs for even moderately sized molecular systems. We will apply combine deep kernel learning with systematic, physics-based dimensionality reduction algorithms to develop computationally efficient coarse-grained quantum mechanical models of molecular systems.

(5) Prof. Diwakar Shukla (Chemical and Biomolecular Engineering)

Title: *Elucidating the kinetics and thermodynamics of cyclopamine binding to the smoothed receptor through multiagent reinforcement learning.*

Description: Smoothed (SMO) is a key transmembrane protein for hedgehog signaling pathway, which is a pathway critical for embryogenesis and cancer. Therefore, it is important to know the role of cyclopamine binding to SMO. We will employ Molecular Dynamics (MD) simulation along with a Reinforcement Learning algorithm, a machine learning method that concerns how intelligent agents take actions in a specific environment to maximize cumulative reward, to understand SMO binding.

(6) Prof. Antonia Statt (Materials Science & Engineering)

Title: *Inverse Design of Sequence Defined Macromolecules*

Description: The overarching goal of this project is to predict which sequence (e.g. monomers, amino-acids) on a macromolecule (e.g. copolymer, protein, peptide) will self-assemble into which large-scale disordered aggregate. We will use a simple model system to perform MD simulations of a peptide with a specific sequence in dilute conditions. The simulation results are the basis to perform dimensionally reduction using UMAP to find appropriate order parameters and to enable the inverse design.