UNIVERSITY OF NOTRE DAME | INTERNATIONAL

SUMMER 2024 iSURE RESEARCH PROJECTS

In the online application, you will be prompted to select one project from this list. The project numbers before each title correspond to the projects in the drop-down menu on the online application questionnaire. When prompted, enter your first-choice project using the appropriate number from the list on the following pages.

Many projects welcome students from across disciplines to apply. Applicants are encouraged to explore the projects that are outside of their majors. For example, our Social Science projects host computer science, statistics, and data science majors. A few Engineering projects provide opportunities for research in machine learning. Most faculty mentors listed the preferred academic backgrounds and skills in the project descriptions. Therefore, the best way to find a project is to use the search function to locate keywords of your ideal research experience, such as *machine learning*, *MATLAB*, *R programming*, etc.

Table of Contents

Note: To ensure fairness of the admissions process, applicants should not contact faculty during the application phase. Applicants who reach out to faculty before admissions decisions are conferred will be disqualified. Please contact globalprograms@nd.edu with any questions.

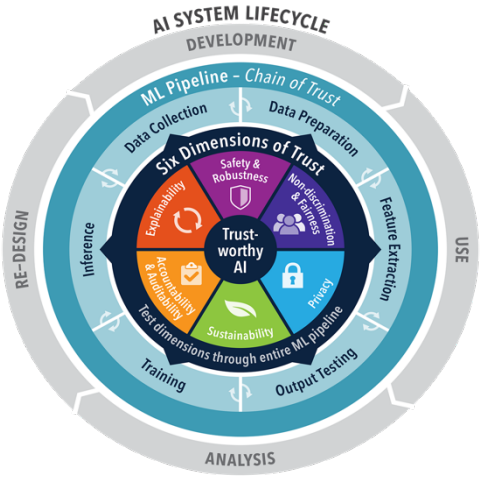
Center for Research Computing

1 - Research Data and FAIR Data APIs

There is more research data than ever being collected and stored. Some of this is "big data", but there is also a lot of small data spread over many repositories. A big problem is how do we find and reuse these datasets? Especially with machine learning, which requires training and evaluation data. Some questions and projects include: How widespread are data APIs on repositories? Can automated tools be developed to evaluate and measure data access? How well do secondary search indexes (e.g. search aggregators) cover various repositories? How suitable are web3 technologies (e.g. IPFS) for enabling identification and sharing? The exact project depends on the participant's interests.

2 - Trusted AI - Cybersecurity and Machine Learning

Identifying the complex causes of potential mission or system failure (or success) and determining effective responses to preventing (or ensuring) such requires leveraging best in class machine learning techniques on rapidly growing, but often poorly structured, data. While the tools available for data science continue to evolve, there remain significant challenges for teams of decision makers trying to wrangle insight from the large and complex data accessible to them. Fortunately, recent advances in natural language processing (NLP) and related machine learning tools such as knowledge graphs (KG) can be harnessed to gain insight. There remain however a significant number of cases where the KG and NLP are not able to easily disambiguate the entities and relations represented in text. This work aims to investigate the contribution of image recognition (computer vision) technology in such cases to provide the most accurate and trusted AI based answers to critical questions.



Note: To ensure fairness of the admissions process, applicants should not contact faculty during the application phase. Applicants who reach out to faculty before admissions decisions are conferred will be disqualified. Please contact globalprograms@nd.edu with any questions.

3 - Curated Pathways to Innovation - Data Analytics and Machine Learning

Research has shown that female adolescents are less likely to display interest and motivation in the pursuit of math and science education, underrepresented minority students score lower averages on math and science tests, and females and URMs are less likely to pursue careers that are heavily math or science dependent. Suggesting, recommending, and measuring the outcomes of diverse students' paths toward computer science through a disparate set of networked programs has not been successful due to the complexity involved. The CPI App incorporates a 'visual map' of where the student is along their 'custom pathway', which is powered by a machine learning algorithm. This project will look at the data generated by the CPI application and ecosystem and develop methods of evaluating the data for efficacy. The project will also develop suggested platform improvements that may improve the CPI platform to better recommend and engage students along their learning pathway.

4 - Dark Web Marketplace Analytics - Machine Learning

The tools and protocols to navigate the modern internet are increasingly complex and selecting the right tools is critical for both internet freedom and privacy. At the same time, those same technical tools and protocols also facilitate elicit behaviors often associated under the term "Dark Web". Students participating in this project will learn about the tools and protocols utilized for both dark web and internet freedom behaviors. As a secondary focus, students will gather dark web marketplace data and perform analytics to understand the types and scale of elicit transactions. The data scraped and analyzed as part of this work will be shared publicly on https://www.dwdata.org/.

5 - Simulation and Data Analytics of Complex Global Social Systems

This team has interest in leveraging modern computational interfaces (such as Jupyter) and machine learning tools (such as Natural Language Processing) to study complex global social systems. He currently has active work in global migration and global news bias research. Additionally, Prof Brenner is leading a new research initiative to develop a high performance yet accessible, multi-scale, human well-being simulation tool on the scale of 7 billion humans. Students participating in research on global social systems will study tradeoffs in social model/data complexity with computational performance limitations.

6 - AI, machine learning, and data science for mosquito surveillance data

Mosquito-transmitted diseases are a growing threat to the United States. This summer saw locally acquired Malaria in the continental US in multiple states – where it had not been seen in decades. As part of regular mosquito control efforts, local governments collect mosquito surveillance of mosquito abatement district data. As part of the VectorByte grant from the US National Science Foundation and the VectorBase project from the National Institutes of Health, Notre Dame has been collating these records from across the United States. In some cases, we have access to live trapping data from networks of automated web-connected smart traps. For this summer project, we will explore using AI, machine learning, and data science techniques on that data. Can we determine minimum surveillance data required to adequately determine true population numbers? Can we correlate data from weather to predict mosquito numbers (and thus disease risk?) Can AI be used to determine true mosquito populations from imperfect or limited field data?

7 - Remote Emerging Disease Intelligence - NETwork (REDI-NET)

The Remote Emerging Disease Intelligence-NETwork (REDI-NET) is envisioned as a long-term, phased initiative to develop a collaborative network among domestic and international partnering institutions to address surveillance needs to effectively DETECT, PREDICT, and CONTAIN potentially emergent zoonosis of human relevance and improve the accuracy and timeliness of the 'data-to-decision' pipeline.

REDI-NET leverages, coordinates and integrates pre-existing and novel, real-time domestic and international xenosurveillance efforts to optimize pathogen discovery and provide a one-stop shop for health decision-makers.
https://redi-net.nd.edu

8 - Forks in Ethereum: Are DLT Democratic? If not, Why?

A distinct innovation of distributed ledger networks (blockchains) is that, unlike other social networks, they are thought to be democratic decision-making mechanisms. This, however, remains an untested claim. The ramifications are quite profound. If the democratic decision making is innate to the decentralized nature of the blockchains, these networks can offer unique solution to garnering the power of "crowd" and efficient consensus building. However, if these networks are subject to power concentration and/or cajoling opinions, then their failings in delivering on the democratic decision making and efficient consensus building deserve a much closer look. Here, we propose two approaches to evaluate the degree of democratic decision making and identifying possible sources of potential failings. We take Ethereum network as the test laboratory. What remain unanswered are: (1) who and how these forks are instigated? (2) how does the discussions lead to these changes? (3) are there common 'influencers' who start and/or guide the discussions which lead to forks?, and if yes, (4) what characterizes these 'influencers'? The answer to these questions—measured by both by the movement of capital as well as GitHub discussions, YouTube videos and more—can help us to understand how do DLTs with significant economic and monetary incentives operate. But more importantly, what makes these networks to behave democratically, if at all.

9 - Systematic review of DID and VC technology, standards and tools

In the evolving landscape of decentralized identity and self-sovereign systems, a structured inventory of various methods, protocols, and libraries is essential for researchers and developers. This project will consist in a systematic review of existing DID (Decentralized Identifiers) and VC (verifiable credentials) specifications, including W3C standards, emerging DID methods, cryptographic techniques and open-source software tools. The goal is to provide a well organized repository describing and categorizing these resources, enabling selection of the

most suitable technologies for their identity and credential needs. This project will contribute to the consolidation of knowledge in the field, fostering growth and adoption of DIDs and VCs in a standardized and interoperable manner.

10 - Open Source Research Software Analysis

At any research institution such as Notre Dame, there are hundreds to thousands of active code repositories in use at any given time. Many of these repositories are closed source, however there are also a large number of open source repositories being created and maintained by individual researchers and groups on campus. It is very hard for institutions to catalog these repositories in order to enhance and recognize the breadth of code being produced by the research community. This project is aimed at 1) creating a registry of open source repositories led or created by Notre Dame researchers, 2) define an attribute/classification structure to tag the domains, languages, licenses and other facets of each repository, and 3) create an analysis report assessing Notre Dame's open source ecosystem. These project tasks will be done in such a manner that the work will be the beginning of a playbook that any institution can also catalog and analyze their open source repositories. The activities in this project are all essential in the journey to starting an OSPO (Open Source Program Office) at Notre Dame that can be a central resource for open source code development, maintenance, and coordination on campus. The report that is generated will be a good candidate for a presentation at US-RSE 2024 for whatever student works on the project.

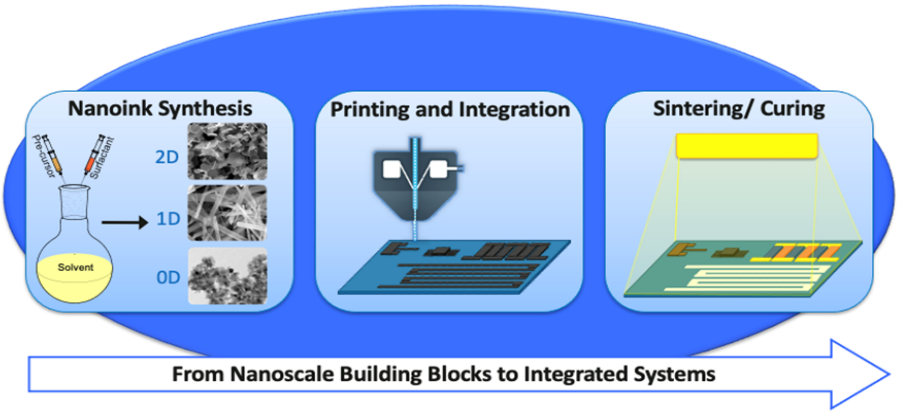
Department of Aerospace and Mechanical Engineering

11 - Scalable nanomanufacturing and hybrid printing of functional devices for energy and health

The mission of this lab is to innovate materials processing methods and transform manufacturing technology in order to improve energy and environment sustainability and individual wellbeing by addressing grand challenges our society is facing. To do so, we have established a transformative nanoscale-to-macroscale engineering approach that synergistically integrates fundamental and applied research programs into a coherent effort.

The overarching goal of this project is to develop and integrate versatile additive manufacturing (AM) and scalable nano-manufacturing (NM) methods in order to transform nanoscale building blocks into macroscale functional systems in a scalable, controllable and affordable manner. Research includes scalable nanoparticle and ink synthesis, and multifunctional device design and printing using a suite of innovative additive manufacturing methods. We aim to harmoniously integrate functional and structural materials into autonomous systems for a range of emerging applications such as clean and sustainable energy, self-powered wireless sensor systems for monitoring of structural health and human health, wearable electronics, and soft robotics.

This research will have a broad impact on materials processing and device engineering across length scales to significantly advance energy, sensing and electronics technology. It will: (1) offer fundamental knowledge on the additive processing of colloidal nanocrystals and their structure and property evolutions across length scales, (2) provide a scalable and low-cost manufacturing process to fabricate efficient and flexible materials for broad applications including thermoelectrics, sensors and others, (3) increase energy efficiency and reduce emission through wide implementation of these low-cost and flexible materials, and (4) advance fast growing technology areas of sensors, energy harvesters, and flexible and wearable electronics.



12 - Scientific Machine Learning for Predictive Modeling Physical Systems

First-principle modeling and simulation of complex systems based on partial differential equations (PDEs) and numerical discretization have been developed for decades and achieved great success. Nonetheless, traditional numerical solvers face significant challenges in many practical scenarios, e.g., inverse problems, uncertainty quantification, design, and optimizations. Moreover, for complex systems, the governing equations might not be fully known due to a lack of complete understanding of the underlying physics, for which a first-principled numerical solver cannot be built. Recent advances in data science. and machine learning, combined with the ever-increasing availability of high-fidelity simulation and measurement data, open up new opportunities for developing data-enabled computational mechanics models. Although the state-of-the-art machine/deep learning techniques hold great promise, there are still many challenges - e.g., requirement of "big data", the challenge in generalizability/extrapolibity, lack of interpretability/explainability, etc. On the other hand, there is often a richness of prior knowledge of the systems, including physical laws and phenomenological principles, which can be leveraged in this regard. Thus, there is an urgent need for fundamentally new and transformative machine learning techniques, closely grounded in physics, to address the aforementioned challenges in computational mechanics problems. This talk will briefly discuss our recent developments of scientific machine learning for computational physics, focusing on the integration of PDE operators into neural architecture to fuse prior knowledge of physics, multi-resolution data, numerical techniques, and neural network through differentiable programming.

Anticipated Student Involvement:
The UG students will work closely with graduate students and postdocs on model development, testing and comparisons.

Preferred discipline(s), expertise, lab skills, ect.

- Preferred disciplines: Mechanical/Aerospace/Civil Engineering, Engineering Mechanics, Applied Mathematics, Computer Engineering/Science, or other related majors
- Preferred lab skills: Numerical Methods, Machine Learning, Python

13 - AI-enabled Computational Modeling of Cardiovascular Fluid Dynamics

Optimization and uncertainty quantification have been playing an increasingly important role in computational hemodynamics. However, existing methods based on principled modeling and classic numerical techniques have faced significant challenges, particularly when it comes to complex 3D patient-specific shapes in the real world. First, it is notoriously difficult to parameterize the input space of arbitrarily complex 3-D geometries. Second, the process often involves massive forward simulations, which are extremely computationally demanding or even infeasible. We propose a novel deep learning solution to address these challenges and enable scalable patient-specific geometric generalization, uncertainty quantification and optimization. Specifically, a statistical generative model for 3-D patient-specific shapes will be constructed based on a handful of available baseline patient geometries. An unsupervised shape correspondence solution is used to enable geometric morphing and a compact geometric design space can then be constructed by the statistical generative shape model. In order to build a fast forward map between geometric input space to the solution space of blood flow, we propose a supervised DL solution, which will facilitate shape optimization and uncertainty quantification analysis in a massively scalable manner.

Anticipated Student Involvement:
The UG students will work closely with graduate students on medical image segmentation and computational fluid dynamics (CFD) simulations.

Preferred discipline(s), expertise, lab skills, etc.

- Preferred disciplines: Mechanical/Aerospace/Civil Engineering, Engineering Mechanics, Applied Mathematics, Computer Engineering/Science, or other related majors
- Preferred lab skills: CFD, numerical methods, Python, machine learning

14 - Differentiable programming for fluid and fluid-structure dynamics

The drive to merge computational physics and machine learning has led to the need for differentiable solvers in fluid dynamics and fluid-structure interactions. This research aims to develop a solver that can easily work with popular machine learning systems, especially for gradient-based optimization and solving complex problems. We aim to build a reliable solver for different flow conditions and designs. Once ready, it should fit well with standard machine learning tools and be tested in real-world conditions. A focus will also be on making sure the solver runs efficiently, can handle large datasets, and works well in parallel computing environments. The final goal is to provide clear guides and training, making this tool helpful for researchers and professionals alike.

Anticipated Student Involvement:
The UG students will work closely with graduate students on model development, testing, and evaluation.

Preferred discipline(s), expertise, lab skills, etc.

- Preferred disciplines: Mechanical/Aerospace/Civil Engineering, Engineering Mechanics, Applied Mathematics, Computer Engineering/Science, or other related majors
- Preferred lab skills: CFD, numerical methods, Python

15 - Soft Growing "Vine" Robots to Move the Human Body

Soft growing "vine" robots are a recently explored class of robots that move by extending from their tip, like natural vine plants. In the most common implementation, the main body of a vine robot is made of a tube of flexible but not stretchable, fluid-tight material (for example, fabric or plastic) that is folded inside of itself; air pressure causes the internal material to turn inside-out and emerge from

the tip. Due to their soft bodies and the fact that they extend from the tip, these robots are well-suited for navigation in confined spaces. They are also very strong in tension, acting like growable steerable ropes, and they can apply high forces using their internal pressure over a large contact area. This project seeks to develop and demonstrate the capabilities necessary to enable these robots to move the human body, either by inflating underneath the body or by wrapping around the body and lifting from above. This work has applications in patient movement (e.g., transfer from a bed to a wheelchair, standing assistance after falling down), as well as other human movement applications.

Anticipated Student Involvement:
An undergraduate student will work closely with an ND graduate student to design and prototype the vine robot designs, as well as to conduct experiments to validate mathematical models of the designs and analyze the resulting data.

Preferred discipline(s), expertise, lab skills, etc.
Mechanical engineering background is preferred. Experience with SolidWorks or other CAD software, Arduino, and MATLAB are helpful.

Department of Biological Sciences

16 - Intellectual Disability in a Culture Dish: Human Neuron Biology Models Kabuki Syndrome

100 billion neurons, each with thousands of synapses, compose the complex framework of the human brain. Brain cells (neurons as well as glia) wire and work together in neural circuits that consist of cells interacting at synapses – specialized cell-cell connections in the brain. Intellectual disability (ID) affects 2–3% of the general population and is often caused by mutations that developmentally impair synapses, neural cells and their circuits. Due to delayed detection time of disease or limited physical access to the developing fetus, potential therapies can only target developed "postnatal" nervous systems. It is unclear to what degree treatments can reverse pathological development. Extensive work has been done in identifying and understanding mutations contributing to or causing ID. However, in vitro systems allowing precise time control of genotypes in human synapses are not available. In this lab we developed a novel cell culture platform to model ID caused by Kabuki Syndrome (KS), to understand the developmental time frame and molecular causes of cognitive impairments. The molecular make-up of affected brain cells and their synapses will reveal optimal timepoints and cellular targets for therapies in the preclinical studies that are outlined here.

Anticipated Student Involvement:
Students will work closely with ND graduate students in the laboratory and learn how to
make induced human neurons out of induced pluripotent stem cells from Kabuki Syndrome patients.

Preferred discipline(s), expertise, lab skills, etc.
Biology, Neuroscience, Molecular Cloning, Cell Culture work. Host professor is a German citizen and can also accommodate German- and English-speaking students.

17 - Identifying novel immuno-oncology targets and developing immunotherapies for metastatic cancers

The goal of the hosting professor's lab is to identify novel immuno-oncology targets and develop immunotherapies to benefit patients with metastatic cancers. They focus on both cancer cell intrinsic and extrinsic mechanisms of tumor immune evasion in a range of cancer types (e.g., prostate, breast, pancreas, kidney). Their research approach integrates genetically engineered mouse models, functional genomics, experimental therapeutics, and cutting-edge experimental and computational methodologies (single cell RNA-seq, spatial transcriptomics, high-throughput drug and CRISPR/cas9 screen, molecular digital pathology, multi-omics integration, etc.). Some recent publications include Wang*, Lu*, et al. Cancer Discovery, 2016 [PMID: 26701088]; Lu et al. Nature, 2017 [PMID: 28321130]; Feng…Lu, PNAS 2018 [PMID: 30232256]; Huang…Lu, Nat Comm, 2020 [PMID: 32358507]; Zhu…Lu, Science Immunology, 2023 [PMID: 36897957]; Zhao…Lu, Cell Metabolism, 2023 [PMID: 37793345].

Located at the Department of Biological Sciences at the University of Notre Dame and affiliated with Indiana University Simon Comprehensive Cancer Center, this lab is seeking highly motivated candidates with interests or expertise in areas such as:

- Cellular, biochemical or metabolic mechanism of cancer progression and metastasis
- Tumor microenvironment and immunology, tumor and immune metabolism, immunotherapy
- Cancer genomics, bioinformatics, single-cell RNA-seq, spatial transcriptomics, CRISPR screen
- Medicinal chemistry, chemical biology, tissue engineering, 3D tumor organoids

The lab research is actively supported by federal agencies and private foundations. A range of available projects include but are not limited to: (1) Mechanistic understanding of immunosuppressive myeloid cells; (2) Heterogeneity of MDSCs, neutrophils and other myeloid cells; (3) Metabolic reprogramming to enhance cancer immunotherapy; (4) Immunological regulation of cancer metastasis; (5) Gene

functions and therapy development for solid tumors; (6) Novel 3D tumor organoids enabling high-throughput screening of immunotherapeutics.

18 - Mitophagy Signaling in ECM-Detached Cancer Cells

Mitophagy and non-apoptotic mechanisms of cell death during ECM-detachment. We have discovered that the activation of RIPK1 can selectively eliminate ECM-detached cells and that cancer cells defective in RIPK1-mediated mitophagy are better able to form tumors in vivo (Hawk et al, Nat Cell Biol. 2018 Mar;20(3):272-284). Currently, we are investigating mechanisms underlying how mitophagy can kill cancer cells, studying links between ECM-detachment-mediated metabolic changes and cell death modalities, and investigating potential vulnerabilities that could be exploited to kill ECM-detached cancer cells. Our discovery of RIPK1-mediated mitophagy as a means to cause death in cancer cells has led us to investigate the importance of this new cell death pathway in additional disease contexts. We have initiated studies to assess mitophagy in Friedreich's Ataxia (FA), a progressive, autosomal-recessive, neuro- and cardio-degenerative disorder. FA is caused by diminished expression of Frataxin (FXN), a mitochondrial matrix protein involved in iron-sulfur (Fe-S) cluster biogenesis. We are conducting pre-clinical studies aimed at understanding how changes in cell metabolism and mitophagy can impact oxidative stress that arises as a consequence of FXN deficiency.

19 - Iron Metabolism Modulation and Ferroptosis Sensitization in ECM-Detached Cells

Iron metabolism during ECM-detachment and ferroptosis markers. We have found that ECM-detached cells are resistant to the induction of ferroptosis, an iron-dependent form of non-apoptotic cell death. Underlying this resistance to ferroptosis is changes in iron metabolism that limit the abundance of redox-active iron necessary for ferroptotic cell death (He et al, iScience. 2023 Jun 16; 26(6):106827). Currently, we are investigating how we could sensitize ECM-detached cells to ferroptosis through modulation of iron metabolism and are studying how changes in iron metabolism might function as markers for cells dying by ferroptosis.

Department of Chemical and Biomolecular Engineering

20 - Reverse engineering embryonic regeneration

Specification of the correct spatial dimensions of tissues within embryos depends on a complex "symphony" of signals that define cellular identities and properties. When subjected to environmental hazards or genetic defects, embryos exhibit a remarkable ability to repair mistakes in specifying their size, either by re-specification of cell identity or the elimination of cells through cell death (apoptosis). Understanding the mechanism of tissue size control is particularly challenging due to the interconnected and complex nature of cell signaling and tissue mechanics. A high degree of nonlinearity in the system occurs due to feedback between cell- and tissue-level processes. Previous studies documenting pattern repair in the embryonic epidermis have principally focused on phenotypic data and tissue-level analysis. Very little is known about the dynamics at the cellular level. The objective of this project is to bridge this knowledge gap by performing live imaging studies at cellular resolution of the embryonic epidermis to track the dynamics of cell membranes during proliferation, apoptosis and cell movement. This project is multi-faceted and includes the design and testing of novel microfluidic organ-on-chips, live-imaging and imaging analysis and cell-based computational models of tissue mechanics. Experiments are compared to computational simulation predictions to identify novel mechanisms of tissue morphogenesis and regeneration.

Prerequisite Skills: Basic biology background and interest in experiments. Image processing/computational work requires previous MATLAB or other programming experience.

21 - Investigating how calcium signaling impacts organ and tumor growth

Cells exhibit impaired calcium (Ca2+) signaling in many diseases including skin diseases, Alzheimer's, and metastatic cancer. The cell's internal "computer" uses calcium ions as messengers to help calculate its response to environmental stimuli. This requires regulation of Ca2+ concentrations in cells to coordinate cellular processes. However, much remains unknown about the functions of time-varying Ca2+ signals in developing or regenerating organs. Our research program seeks to

discover the biochemical and mechanical basis of integrative cell communication mediated by Ca2+ signaling. In this team-based project, students contribute toward a high content, genomic screen identifying and characterizing new calcium-related genes that impact organ growth, regeneration or tumor growth. Students gain experience in functional genomics and quantitative image processing.

Department of Computer Science and Engineering

22 - Hardware/Software Co-Design for Deep Learning Accelerators

Neural networks have benefited our daily life through a range of smartphone applications. However, they have grown exponentially larger (with billions of parameters) and demand higher computing power. On the other hand, the underlying hardware is not scaling at the same pace, leading to an increasing gap between neural network size and hardware capacity. Our group pioneered the research on neural network and hardware co-design on various hardware platforms including ASIC, FPGA, and most recently, quantum computers. Along this research we have published several papers in Nature Communications, Nature Electronics in addition to a few conference papers with best paper nominations. Most recently, our journal article on this topic was designated the 2021 Donald O Pederson Best Paper Award in IEEE. Transactions on Computer-Aided Design. We are looking for talented students to join this expedition.

Good to use this feature

23 - AI for Congenital Heart Disease (CHD) Diagnosis and Treatment

CHD is the most common congenital birth defect, with more than one million babies born with CHD out of the approximately 135 million newborns in the world annually. In developing countries the mortality rate of CHD is currently over 20%, strikingly higher than that of 3%-7% in developed countries. This is because CHD diagnosis, treatment planning and surgery requires complex infrastructures, equipment, and highly skilled doctors, most of which reside in developed countries. For example, it takes even a very experienced radiologist several hours to examine the medical image of a CHD patient, whereas that time is on the order of minutes for common heart diseases. AI can potentially help, but the large heart structural variations in CHD patients and the large number of CHD types render most existing algorithms inaccurate.

Between 2015 and 2019, in collaboration with Guangdong People's Hospital, one of the largest hospitals in China, we developed AI-CHD, the first AI framework for CHD diagnosis, treatment planning and prognosis, which can yield results in a few minutes. Based on a series of innovative machine learning algorithms, the accuracy

of the framework is on par with radiologists most experienced in CHD. The framework enabled the world's first 5G+AI telementoring of cardiac surgery on April 3, 2019. In 2020, the framework was also included in the United Nations Educational, Scientific and Cultural Organization (UNESCO) Digital Anatomy for promotion in developing countries. We are now in the process of further enhancing the framework through the use of unsupervised/transfer learning techniques.

24 - Enhancing AI Accessibility of Pediatric Critical Care

While extensive work has been done in applying AI to healthcare, pediatric critical care is left behind, mostly because of the existence of many rare diseases, and the large individual variations even among the same age group. In collaboration with Boston Children's Hospital, the #1 pediatric hospital and Harvard Medical School, we are developing an AI framework that provides automated quantification of chest wall edema, lung edema, regional lung and cardiomegaly in chest radiographs of pediatric patients, based on a novel temporal contrastive learning algorithm. Upon completion, this framework will for the first time allow physicians to visualize radiographic quantification of edema and heart failure. It will also allow automated identification of hardware presence and position (e.g. endotracheal tube) in real-time, and automatically sends out an alert with abnormal findings that must be acted upon immediately.

25 - Segmentation Problems in Medical Images

Neutrophils are a primary type of immune cells, and their identification is critical in clinical diagnosis of active inflammation. However, in H&E stained histology tissue images, the appearances of neutrophils are highly variable due to large variations in morphology, staining, and locations. Further, the noisy and complex tissue environment causes artifacts resembling neutrophils. Thus, it is challenging to design, in a hand-crafted manner, computerized features that help identify neutrophils effectively. To better characterize neutrophils, we propose to extract features of neutrophils in a machine learning manner, by developing a new deep neural network model for identifying neutrophils. In addition, in clinical practice, neutrophils are identified not only based on their individual appearance, but also on the context formed by multiple related types of cells in their neighborhood. It is not

straightforward for deep learning to capture precisely the complex cell context. Hence, we further propose to combine deep learning with a context model of cell clusters, to extract needed context. Experiments on clinical data will be conducted to evaluate the combination of our deep learning model and context model.

26 - Virtual Reality for the Future of Work

This study will examine how teams can work together using virtual reality devices. After the COVID-19 pandemic, organizations and companies are exploring new alternatives to facilitate remote work and virtual collaborations. The purpose of this study is to develop virtual reality applications and understand they can help teams collaborate better. This work will consist in learning how to operate different head-mounted devices (e.g., Oculus Quest Pro, HTC Focus, HTC Vive Pro), develop applications using Unity, and test those applications in laboratory experiments. The findings and applications of this research contribute to organizations and communities aiming to enhance remote work. Software developers, managers, and designers can use the theoretical and practical implications of this research for new procedures and guidelines for the use of virtual reality for work purposes.

Anticipated Student Involvement:
Students will work closely with the PI and ND graduate students on VR development,
collecting data, and data analysis.

Preferred discipline(s), expertise, lab skills, etc.
Computer Science background, programming skills (e.g., Python, C#), familiarity with
Oculus/HTC Vive, familiarity with Unity.

27 - Knowledge-Enhanced Natural Language Processing

Our research lab has made the earliest attempts in the world to enhance knowledge needed in natural language processing (NLP) technologies. We have offered tutorials in EMNLP 2021, ACL 2022, and WSDM 2023, and organized well-attended workshops in AAAI 2023 and KDD 2023, and an upcoming workshop in ACL 2024.

Knowledge is needed to achieve satisfactory performance in many NLP tasks, such as machine translation, summarization, question answering, and commonsense explanation. External information sources (beyond training data) such as structured knowledge bases and graphs can provide knowledge to improve the precision of target-language wording, factual correctness of summaries, accuracy of open-domain answers, and diversity of explanations. Furthermore, we are interested in retrieval-augmented generation (RAG) in the era of Large Language Models (LLMs). In this project, we will continue exploring the methodology of knowledge-enhanced NLP including LLM with RAG to broaden the impact of NLP technologies in specialized domains such as personalization, mental health, online education, and scientific discovery. We are targeting top data mining, machine learning, and NLP conferences such as KDD, ICLR, ACL, and EMNLP. Our team has twelve Computer Science PhD students.

28 - Large Language Models (LLMs) for information retrieval in scientific documents

Leveraging Large Language Models (LLMs) for advanced information retrieval in scientific literature offers invaluable support to chemists and biologists sifting through extensive research databases. Given the constraints of traditional keyword-based searches and time-consuming manual analyses, our initiative emphasizes the use of LLMs for streamlining scientific data extraction. Interns will engage with models akin to GPT-4, enhancing the system's ability to interpret and answer natural language inquiries accurately. This immersive experience will equip interns with the skills to generate brief, context-rich summaries of intricate chemical and biological studies. Such a method not only amplifies literature review efficiency but also immerses interns in the dynamic crossroads of AI and scientific inquiry, a nexus that's becoming ever more vital in contemporary scientific research.

29 - DL4SciVis

iSURE students familiar with programming in PyTorch (required) and ParaView (preferred) will assist Ph.D. students in designing and implementing deep learning (DL) solutions for scientific visualization (SciVis). We mainly focus on time-varying

multivariate volumetric scalar and vector field data produced by domain scientists from their scientific simulations. They are akin to image and video data but are one-dimensional higher (i.e., 2D images and 3D videos vs. 3D volumes and 4D space-time data) and demand a rendering process for visual display. We aim to augment domain scientists' ability to manage, analyze, and visualize these data sets. Specific project ideas include using neural field representation techniques for view synthesis, style transfer, data reconstruction, and visualization compression.

30 - Machine Learning Systems for Design Assistance

Recent advances in machine learning (ML) models that generate user interface (UI) design have opened up opportunities for more automation in design work. In this project, we will utilize and extend state-of-the-art generative design ML models to support UX designers' work. Specifically, we ask: 1) how can we build effective and efficient human-AI collaboration mechanisms to support UX designers' work? 2) how to utilize and improve current generative ML models to better fit ML into UX designers' realistic workflows?

You will also participate in conducting user studies to evaluate our prototypes. The results of this project will be submitted to premier academic conferences in this interdisciplinary area including CHI, UIST, AAAI, ACL, and NeurIPS.

Skills or Requirements:
Experience with machine learning in Python, OR prototyping by building interactive web applications with React. Understanding of basic human-centered design principles and methodologies is preferred. Interest in or basic knowledge of UX design practices is preferred but not required.

Research Areas in Which This Project Falls:
Human-Computer Interaction; Machine Learning; Human-AI Collaboration; Mixed-Initiative User Interfaces; Creativity-Support Tools.

31 - Human-AI Collaboration in Data Annotation

The quality and the quantity of annotated datasets are crucial to the performance, fairness, robustness, safety, and scalability of data science applications. Machine learning models often learn from and are evaluated against human-annotated ground- truth data. However, data annotation is costly in terms of the money, time, and amount of human labor required. This project seeks to design and develop a human-AI collaboration approach to make data annotation more efficient and more effective. In this project, we will design interactive systems that allow human annotators to work effectively together with ML models in the annotation process. It will leverage an active learning paradigm—when the human annotator processes the outputs from partial- automation models to perform the annotation, the partial-automation models learn from the human behaviors in real-time so that they continue to improve and to adapt to the domain, reducing human efforts in future annotations. The proposed approach will also explore the use of transfer learning, meta learning, and few-shot learning approaches on pre-trained general-purpose models in a human-in-the-loop pipeline.

The results of this project will be submitted to premier academic conferences in this interdisciplinary area including CHI, UIST, ACL, and EMNLP.

Required Skills:
Strong programming skills in Python OR JavaScript

Optional Skills:
Expertise in machine learning; Expertise in natural language processing; Expertise in building interactive systems in React; Expertise in programming language research, especially program synthesis.

Research Areas in Which This Project Falls:
Machine Learning; Human-AI Collaboration; Mixed-Initiative User Interfaces; Human- Centered Data Science.

32 - An Empathy-Based Sandbox for User Privacy and Security Preferences

The privacy research community has been examining how users perceive privacy, their privacy needs, and how they act to protect their privacy. However, due to the abstract and opaque nature of privacy as a concept, users encounter difficulties in articulating their privacy needs, understanding the implications of different possible privacy configurations, and eventually, making effective decisions that align with their privacy goals. Without a proper understanding of their privacy needs and the consequences of their behaviors, users are often forced to make decisions based on their incomplete mental models of privacy. We proposed an empathy-based approach to help users (1) enunciate their real privacy needs and understand the potential implications of their privacy-related choices; and (2) align their decisions and actions with their privacy Goals.

As a member of the project, you will contribute to 1) Building an interactive sandbox to help users build empathy for real-world privacy and cybersecurity incidents. 2) Designing and implementing user-configurable modules to support users' privacy and security decision-making. (3) Longitudinal evaluation of the sandbox and privacy modules.

The results of this project will be submitted to premier academic conferences in this interdisciplinary area including CHI, UIST, S&P, USENIX Security, etc.

Required Skills:
Strong programming skills

Optional Skills:
Expertise in machine learning; Expertise in building interactive systems; Expertise in mobile app development; Expertise or interests in privacy and security.

Research Areas in Which This Project Falls:
Human-AI Interaction; Accessibility; Applied Machine Learning; Computer Vision & Language.

33 - Multimodal exploration of video content for Blind and Visual Impairment (BVI) Populations

Video is one of the most important ways for getting information and entertainment. However, people with blindness or visual impairment cannot understand video content as normal people do. Therefore, in this project, we seek to build an interactive system that can automatically generate accessible captions and descriptions for general types of videos, and enable BVI users to explore the video content through various interaction techniques, making the exploration process more accessible, effective, and immersive.

As a member of the project, you will contribute to 1) Building a machine learning pipeline that automatically generates captions and descriptions for various scenes in the video. 2) Design and implement an interactive system that supports user exploration through mouse, screen reader, gestures, etc.

The results of this project will be submitted to premier academic conferences in this interdisciplinary area including CHI, UIST, CVPR, ASSETS etc.

Required Skills:
Strong programming skills

Optional Skills:
Expertise in machine learning; Expertise in implementing interactive systems; Experience in conducting user studies with BVI populations

Research Areas in Which This Project Falls:
Human-AI Interaction; Usable Privacy; Cybersecurity

34 - Computer Systems, Networks, and Security

The Network Systems Research Lab at the University of Notre Dame works on the intersection of networks, systems, and security with applications to edge computing, Internet of Things, machine learning, and extended reality. The purpose of our projects is to develop state-of-the-art solutions to recent and long-standing

problems in the areas of research of our lab. The results of our research have been published at premier technical conferences and journals.

Anticipated Student Involvement:
Students will work closely with ND graduate students on the design, implementation, and evaluation of these projects.

Preferred discipline(s), expertise, lab skills, etc.
Knowledge of fundamental computer science principles, C/C++, Python

Department of Electrical Engineering

35 - Characterization and Modeling of High Performance Devices

Electronic devices for wireless communication and power applications will be important components of future systems. Students involved in this project will assist with the measurement-based characterization of devices using the extensive suite of electronic measurement instruments in the laboratory. These data will be used to develop behavioral compact models suitable for use in computer-aided design flows. Students will gain first-hand experience with both state of the art measurement techniques and capabilities, as well as design automation.

36 - Developing a general compact model for ferroelectric capacitors

Recent discovery of different ferroelectric materials, such as HfO_2, Sc,AlN, Mg,ZnO, has revived interests in their application as nonvolatile memory. To evaluate their array organization or their application for emerging computing, it is critical to design accurate and computationally efficient compact models. As of now, due to different polarization switching dynamics in different ferroelectric materials, separate models have to be developed for various materials. In this project, we are aiming at developing a unified modeling framework for different materials and calibrating the developed models with experimental data. Using the developed models, we will evaluate the memory array performance.

Anticipated Student Involvement:
Students will work closely with ND graduate students on literature review, device measurement, compact modeling.

Preferred discipline(s), expertise, lab skills, etc.
Semiconductor physics, numerical methods, matlab

37 - Electrothermal modeling of back-end-of-line transistors for monolithic 3D integration

To accommodate the growing demand over computing and storage resources, both high performance logic and memory are adopting monolithic 3D integration by

leveraging the third dimension for scaling. However, as more components are packed into the limited space, there is a growing concern over thermal dissipation more than ever. A critical enabler for monolithic 3D integration is the back-end-of-line (BEOL) transistor that can be stacked on top of Si CMOS. With a typical thin film channel made with BEOL compatible materials, such as amorphous oxides, poly-Si, or 2D materials, understanding the heat generation and dissipation is critical. In this project, the student will conduct electrothermal simulation of BEOL transistors with TCAD tools and understand the impact of logic or memory operation on the temperature of the chip.

Anticipated Student Involvement:
Students will work closely with ND graduate students on literature review, device measurement, modeling.

Preferred discipline(s), expertise, lab skills, etc.
Semiconductor physics, TCAD

38 - Fabrication of HfO_2 ferroelectric capacitor with low coercive field

HfO_2 based ferroelectric memory rises as a prime candidate for nonvolatile memory due to their superior energy efficiency and scalability. However, one challenge with it is the relatively high coercive field (usually around 1.2MV/cm), which hinders its application as a CMOS logic compatible memory and thus requires complex write circuitry. Therefore there is a growing interest in reducing the coercive field of HfO_2 to around 0.5MV/cm. Recently, there have been a few reports showing some promising directions. In this project, we will explore the different experimental approaches to reduce the coercive field and evaluate the performance of demonstrated ferroelectric capacitors.

Anticipated Student Involvement:
Students will work closely with ND graduate students on literature review, device fabrication and measurement.

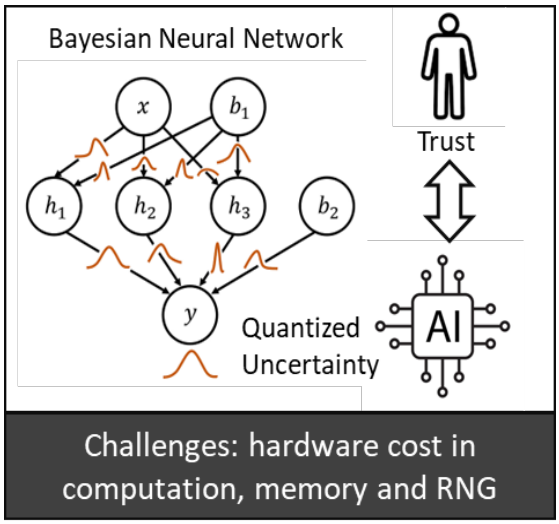
Preferred discipline(s), expertise, lab skills, etc.
Semiconductor physics, nanofabrication

Note: To ensure fairness of the admissions process, applicants should not contact faculty during the application phase. Applicants who reach out to faculty before admissions decisions are conferred will be disqualified. Please contact globalprograms@nd.edu with any questions.

39 - Circuit and System Design for Collaborative Machine Learning

In today's era of rapid technological advancement, machine learning has found its application across a multitude of fields, including but not limited to healthcare, finance, automotive, entertainment, and agriculture. However, a major challenge with many machine learning models, especially deep neural networks, is their "black-box" nature. This opaqueness often poses hurdles in establishing trust between AI systems and human users. One key aspect of this challenge is the inability of these models to effectively communicate the confidence or uncertainty associated with their predictions.

In contrast, statistical machine learning approaches, like Bayesian inference and Bayesian neural networks (BNN), offer a unique advantage. Rather than focusing solely on finding an optimal model with fixed weights, they aim to learn the distribution of model parameters. This intrinsic capability allows Bayesian approaches to quantify and convey the uncertainty related to their learning processes, thereby fostering enhanced trust and collaboration between AI and humans.



While powerful, such statistical machine learning approaches are notably resource-intensive. The requirement for massive random number generation, coupled with extensive computation to assess inference confidence, can be prohibitive. This computational burden makes them challenging to deploy on resource-constrained edge devices, which are often the primary platforms directly interfacing with human users.

This project aims to bridge this gap by offering participants a comprehensive training regimen, focused on the theoretical and practical aspects of analyzing and designing Bayesian Neural Networks on constrained devices.

Anticipated Student Involvement:

1. Algorithm and Simulation: Students will dive deep into the theoretical foundations of Bayesian theory and its application in machine learning with hands-on model building experience from scratch.

2. Hardware Cost Analysis: The training will provide insights into the granular breakdown of BNN computations, covering aspects like random number generation, distribution management, memory access, and more. This will enable students to understand the challenges and bottlenecks associated with deploying BNNs on hardware.

3. Circuit Analysis and Design: A pivotal component of this project is the design and evaluation of innovative circuits aimed at accelerating Bayesian approaches. Students will be guided through the process of circuit analysis, design considerations, and prototyping, ensuring that their solutions are apt for resource-constrained environments.

4. Full-stack Capabilities: To ensure holistic learning, participants will also receive training on python programming, architecture emulation, and various simulation tools. This full-stack approach will empower participants with the versatility needed to tackle real-world challenges.

Preferred Disciplines:

We welcome applications from students with diverse backgrounds, including computer science, computer architecture, machine learning, circuit design, and electronic devices.

40 - Mental stress detection and classification through physiological signal processing

Everyone experiences mental stress from time to time when facing pressure (stressful events) in their daily life. While not all stress is bad, in most cases, stress may alter people's behaviors and may carry physical and mental health risks. It is important to study mental stress and understand how it impacts our bodies. Stress often causes physiological responses that lead to changes in physiological signals,

such as heart rate and electrodermal activities, that can be sensed non-invasively. In this project, we will use physiological signals to detect mental stress. The outcomes of this project may help future human-machine interaction applications to automatically track interaction procedures and adjust interaction content, such as workload, task difficulty levels, and required response time.

Anticipated Student Involvement:
Students will work closely with ND faculty, postdoc fellows, and graduate students on literature review, signal processing and machine learning.

Preferred discipline(s), expertise, lab skills, etc.
MATLAB, digital signal processing, medical signal processing, bio-potential sensing

Note: To ensure fairness of the admissions process, applicants should not contact faculty during the application phase. Applicants who reach out to faculty before admissions decisions are conferred will be disqualified. Please contact globalprograms@nd.edu with any questions.

Social Sciences

41 - Climate Change and Extreme Event Mitigation

This research lab is directed by an applied economist studying the economic aspects of climate change adaptation, disaster risk reduction, as well as other areas related to the sustainable development goals. She has extensive experience in analyzing household-level survey data and integrating GIS maps with vulnerability models and current socioeconomic factors to examine environmental impacts on society. Her recent work includes an NSF-funded project to understand critical infrastructures resilience during hurricanes, an impact evaluation of the US Disaster Response Program for low-income communities, and an assessment of coastal vulnerability regarding climate change risks. This lab is seeking passionate students in the household-level study on climate change issues. Potential research projects include risk perception for future climate change, household risk-averting behaviors to mitigate extreme events, and health impacts caused by the environmental shocks.

Anticipated Student Involvement:
The student will support the household survey analysis and produce statistical reports.

Preferred Qualifications:
Linear regression analysis, knowledge of Stata

Note: To ensure fairness of the admissions process, applicants should not contact faculty during the application phase. Applicants who reach out to faculty before admissions decisions are conferred will be disqualified. Please contact globalprograms@nd.edu with any questions.

Special Projects

42 - Special Projects

Select this project ONLY if you have been instructed to do so by an individual professor or Notre Dame International staff member. These projects are arranged on an individual basis. General iSURE applicants should select one of the projects above for admissions consideration.