School of Integrative Science and Technology
GROUP SUMMER SCHOLARS RESEARCH PROGRAM

2026 Research Streams *(Tentative & Subject to Change*)*

4 week session Research Streams, July 6–July 30, 2026:

Evolution and Developmental Biology (4-week session)
Dr. Matthew Niepielko
During animal development, a handful of highly conserved cell signaling pathways regulate when and where genes are expressed within a developing tissue. This process, called tissue patterning, is an essential step to ensure that animals' organs develop in the correct location and at the right time. A growing body of evidence suggests that changes in gene regulation during tissue patterning drive the evolution of new animal structures and morphologies. In our lab, we use Drosophila, the common fruit fly, as a model system to investigate the underlying mechanisms that tightly regulate gene expression during development. Additionally, we explore how these mechanisms evolve and give rise to new morphologies using a combination of computational and molecular techniques.

Molecular & Cellular Biology: Exploring Cancer Malignancy (4-week session)
Dr. Salvatore Coniglio
Participants in this research stream will learn the basics in molecular biology and host discussions surrounding how cells acquire and integrate information from the extracellular environment with a focus on cancer biology. Activities that focus on signal transduction, gene expression and protein expression in cancers will be the primary focus. Using a hands-on approach with image analysis software, students will learn how scientists quantify changes in gene and protein expressions in tumors.

Developmental Biology and Alcohol Studies (4-week session)
Dr. Maria Agapito
GSSRP summer students will analyze the effect of chronic alcohol exposure on the developing embryo. We will monitor the overall health of the animals using established assays (Body length, brood size, egg hatching & laying) and develop new assays specific to neurons using an existing neuronal actin reporter.
C. elegans is a good model system for studying human disease. C. elegans have over 100 genes that are associated with human disease genes, making it a potential model system for investigating the conserved cellular roles of these genes. The genome of these animals has been fully sequenced and the nervous system wiring is fully mapped. In the last decade, C. elegans have been used to study alcohol addiction and alcohol dependency.
Studies have shown that chronic drug use can lead to compulsive drug-seeking habits and drug

addiction (Everitt & Robbins 2005). Continuous alcohol intoxication results in increased alcohol consumption, tolerance and sometimes alcohol dependency in mammals (Koob 1998; Roberts et al. 2000; Koob 2003; Rimondini et al. 2003, Rimondini et al. 2007). However, there is limited information about the mechanism of action of alcohol during neuronal development.

Acoustics and Digital Signal Processing (4 week session)
Dr. Edward Farnum

Physically, all sound is produced by periodic vibrations in air pressure. At first glance, the frequency of vibration determines the pitch of a particular musical note. In fact, most sounds are a much more complicated combination of sine waves, each with a different frequency and amplitude. Mathematically, these can be described by something called a Fourier Series. Using a short-time Fourier Transform, we can analyze audio signals to better understand the production of sound waves and also modify those signals to create new sounds. We will focus on the motion induced from nearby vibrating bodies – this is called sympathetic resonance in coupled oscillators. Students in this group will explore how the sounds of neighboring guitar strings can be encouraged to interact and generate a drone sound. This will aid in the design and construction of new and strange musical instruments, which we will build. This project can move in many directions, including theoretical, experimental, or computational. This research stream is appropriate for students with interests in applied math, physics, computational science, engineering, and especially music. Any familiarity with any of the above fields, as well as interests in math, physics, computation, or music, will be beneficial.

Human-centered AI (4-week session)
Dr. Iyadunni Adenuga

Artificial Intelligence (A.I.) systems such as Large Language Models (LLMs) like ChatGPT, are ubiquitous in today's environment, but there is still hesitancy about their use in consequential areas like education. General beliefs are that students, especially first-time learners, do not learn or gain the required skills when they utilize these systems. This may be because even though LLMs are created to generate and retrieve information, they are not primed to aid learning. Despite these limitations, LLMs are powerful entities, which, due to the access to vast amounts of data, provide intelligent responses to user queries on a simple chat-based interface. It is imperative that these benefits are leveraged for the learning process. This would require that LLMs are enhanced to prioritize human-centered values such as agency in its responses.

Exploring AI Language Models for Health-Related Insights (4-week session)
Dr. Mahlihe Aliasgari

In this project, students will investigate how large language models (LLMs) and basic machine learning techniques can be applied to extract meaningful insights from text in health-related contexts. The focus will be on understanding how AI systems analyze language, recognize patterns, and support research questions using real-world examples.

Students will gain hands-on experience with simplified machine learning concepts, including text classification, pattern detection, and data exploration. They will also learn how LLMs interpret

information, summarize content, and generate helpful outputs that can support scientific thinking and health-related studies.

By working through guided activities and small projects, students will discover how AI and machine learning can contribute to improving quality of life and inspiring future innovations. The project emphasizes ethical awareness, creativity, teamwork, and practical AI skills, making it an accessible and exciting introduction to the world of intelligent technologies.

Identification of Phytochemicals as Potential Drugs to Block Zika Virus Replication Targeting NS5 RNA-Dependent RNA Polymerase (4-week session)

Dr. Supratik Kar

Zika virus (ZV) continues to affect communities around the world, and safe, effective treatments are still limited. This project focuses on one of ZV's most important proteins, called NS5 RNA-Dependent RNA Polymerase, which the virus needs to copy its genetic material and spread. If this protein can be blocked, the virus cannot multiply. To search for possible blockers, the study will focus on natural plant compounds (phytochemicals) already known to have some antiviral activity. Their chemical structures and reported effects will be used to build a computer model that predicts which types of molecules are likely to work best against ZV. After that, computer screening will be used to sort through many plant-based compounds and identify the most promising ones. The strongest candidates will be examined more closely using additional computer techniques that show how well each compound fits into the NS5 protein and whether it could stay bound long enough to block the virus. Safety-related features, such as how the body might absorb or remove each compound, will also be checked. By combining all these computer-based steps: prediction, screening, binding analysis, and safety checks, the goal of the project is to highlight a small number of plant compounds that look most promising as starting points for future laboratory testing and eventual treatment development for ZV.

Introduction to Machine Learning and AI Applications (4-week session)

Dr. Zimeng Lyu

The research program introduces students to artificial intelligence and machine learning through a combination of foundational skill development and applied project work. The program is designed to provide hands-on experience with modern programming tools and machine learning techniques while exploring real-world applications across multiple domains.

All participants will receive training in essential technical skills:

Python Programming for Machine Learning: Introduction to Python fundamentals, data structures, and common machine learning libraries (NumPy, Pandas, Scikit-learn)

Version Control with GitHub Classroom: Training in collaborative software development practices using Git and GitHub for project management and code sharing

Visualization and Analytics Development for Data Insights (4-week session)
Dr. Ching-Yu (Austin) Huang

In this program, students will develop essential skills in data visualization, analytics, and web development, all of which are highly valued in today's data-driven world. They will master SQL to manage and query real-world datasets, practicing data manipulation, aggregation, and filtering to understand how data is structured and analyzed in modern systems.

Students will build dynamic web applications using PHP, MySQL, JavaScript, and HTML. They will create visualizations, including line charts, bar charts, scatter plots, and pie charts, using graphing libraries to develop interactive dashboards. Additionally, students will design web pages that connect to databases, collecting user inputs through forms, dropdowns, checkboxes, and radio buttons, allowing users to filter and interact with data while updating visualizations in real-time. Students will also gain hands-on experience with cookies for login authentication, equipping them with the skills needed to build functional, data-driven applications.

The program emphasizes the importance of interdisciplinary research, demonstrating how data visualization and analytics are critical tools across fields like healthcare, business, and social sciences. Students will apply their coding and analytical skills to formulate research questions, extract insights, and solve real-world problems with complex datasets. Along the way, they will learn the art of data storytelling, effectively presenting their findings through visualizations and clear narratives. By honing these skills, students will be able to communicate insights compellingly and make data-driven decisions that can influence research and industry outcomes.

By the end of the program, students will complete a research poster that showcases their ability to integrate SQL, PHP, MySQL, JavaScript, data visualization techniques, and insights from their selected datasets. Additionally, students will be equipped to independently explore advanced topics in data visualization, data mining, and web development, empowering them to continue developing these high-demand skills and apply them to more complex challenges in the future. A laptop is required for this program, as students will need to install necessary software, access datasets, and actively engage in hands-on exercises.

Predicting Student Performance Using Machine Learning Based on Python (4-week session)
Dr. Yousef Nejatbakhsh

This research project utilizes machine learning techniques to predict student performance by analyzing key factors, including study hours, attendance, parental involvement, and socioeconomic status. Python will be utilized for data analysis, visualization, and implementing machine learning. The project begins with data preprocessing, including cleaning, handling missing values, and normalizing features, followed by Exploratory Data Analysis (EDA) to identify trends and correlations, such as the relationship between attendance and academic success. Machine learning models, including linear regression, Decision Trees, and Random Forests, will be trained and evaluated using metrics such as accuracy and mean squared error (MSE), with hyperparameter tuning applied to optimize performance. Results will include a feature of importance analysis and visualizations that highlight key insights, culminating in a comprehensive

report and presentation. This 4-week program introduces students to practical data science and machine learning techniques, equipping them with hands-on skills while addressing real-world educational challenges. By leveraging predictive models, the project aims to support data-driven decision-making for educators and policymakers, improve academic outcomes, and inspire students to explore technology-driven problem-solving.

2 week session Research Streams, July 20–July 30, 2026:

Topographic and Hydrodynamic measurements (2-week session)
Dr. Jun Cheng
This two-week summer program with Dr. Jun Cheng will immerse students in the study of river and beach environments through hands-on, field-based learning. The program will take place primarily on campus, where students will gain practical experience using modern surveying and mapping technologies to monitor environmental change, including RTK GPS, total stations, LiDAR sensors, and photogrammetry for building high-resolution 3D models. The program will also introduce students to acquiring and processing satellite imagery to detect and quantify landscape change, with exposure to newly emerging AI-based methods for data analysis.

A half-day field excursion to Sandy Hook Beach may be arranged to provide a real-world application of the concepts and skills learned on campus, allowing students to directly connect theory, technology, and coastal processes. And a research poster will be completed by the end of this program.

Investigating Fibonacci numbers and applications to data security (2-week session)
Dr. Eilidh McKemmie
The Fibonacci sequence can be found all over nature, art and computer science. It can be used to crack some cryptographic hash functions which are used to protect password security and verify the authenticity of data. The Fibonacci sequence is connected to many other interesting sequences like the Pell numbers and the Tribonacci numbers. In this project, you will explore the connections between these interesting sequences. You will use the Python programming language to test your ideas, come up with new mathematical facts, and prove you are right. You will compare different methods for cracking hash functions using your mathematical discoveries.

Mycobiome of dead wood (2-week session)
Dr. Maria Shumskaya
Forests rely on biodiversity of decomposers to control storage of carbon and mineral nutrients. When dead wood is decomposed, carbon and nutrients are released. Dead wood-inhabiting fungi are saproxylic organisms responsible for this decomposition, making dead wood available for other organisms to use. In this research stream we will investigate biodiversity of saproxylic fungi during colonization of wooden pins (MycoPins) that were decaying over time in a swamp in Finland. We will test the hypotheses of type of wood and amount of time affecting the fungal colonization in pins that were placed in a swamp in Finland in 2022, and then collected in certain

time frames throughout the year. We will use DNA metabarcoding method which consists of DNA isolation and PCR to amplify fungal DNA from decomposed wood, and Next Generation Sequencing technique paired with statistical analysis to analyze the communities of fungi over the time. We will also grow our own dead wood fungi in a lab! This stream provides excellent experience in modern molecular biology techniques applied to research of microbial communities. Students interested in molecular biology and integrative biology are encouraged to join this stream.

Population genomics (2-week session)

Dr. Brenna Levine

Population genomics is the study of DNA variation across many individuals to understand how evolutionary forces like migration, natural selection, genetic drift, and population history shape biological diversity. By analyzing genomes from different populations, scientists can reconstruct where populations came from, how they spread, and how they adapt to new environments—questions that are central to evolution, conservation, and human health. In this stream, students will:

- Analyze real genomic datasets to investigate genetic diversity and population structure
- Learn how scientists use genomes to track evolutionary history, adaptation, and biological invasions
- Gain hands-on experience with computational tools used in modern genetics research
- Develop transferable skills in data science, critical thinking, and scientific communication