Plan Overview

A Data Management Plan created using DMPTool

DMP ID: https://doi.org/10.48321/D1N64F

Title: DMSP for "RTG: Transdisciplinary Training in Mathematical and Computational Biology at NJIT: From Data to Theory and Back"

Creator: Cristo Yanez leon - **ORCID:** 0000-0002-0930-0179

Affiliation: New Jersey Institute of Technology

Principal Investigator: Casey Diekman, Horacio G. Rotstein, Kristen E. Severi , Enkeleida Lushi , Ines Figueiras

Contributor: Simon J. Garnier , Victoria J Livingstone, Amitabha K. Bose , Jorge P. Golowasch , James N. MacLaurin , Victor V. Matveev , Cristo Leon, Yuan-Nan Young , Farzan Nadim , Ines Figueiras

Funder: National Science Foundation (nsf.gov)

Funding opportunity number: NSF 23-579

Grant: https://www.nsf.gov/pubs/2023/nsf23579/nsf23579.html

Template: NSF-DMS: Mathematical Sciences

Project abstract:

Overview: The goal of this project is to create a modern and innovative cross-disciplinary training program at NJIT built on the idea that mathematics is a fundamental tool for understanding biology and at the same time, biology is a source of original mathematical problems. The former, more classical side of this mutual relationship focuses on applying existing mathematical tools to address biological issues. The other side, the one we aim to achieve, focuses on creating “new mathematics” to describe biological data and explain the mechanisms and biological processes underlying the generation of these data. To achieve this goal, we propose to provide a type of training where cross-disciplinary research is not simply seen as the interaction of two different activities but as the concerted activity of experimental design, data collection, and theory/modeling creation where mathematics plays a fundamental role in all of these aspects in a given biological context. To this end, we will establish two vertically-integrated research groups in computational neuroscience and collective behavior, composed of mathematical sciences trainees from the community college level up through the postdoctoral level that will be co-mentored by faculty members in mathematical sciences and biological sciences. These working groups will address different biological scales but share a cohesive focus on data-

driven mathematical modeling, dynamical systems theory, stochastic processes, and scientific computation. Associated with these working groups, we will take the following actions: (1) create a new Mathematical Biology Option within the Applied Math Track of our PhD program featuring extensive new curriculum development and research rotations in both mathematical groups and biological labs and; (2) host an annual Summer Research and Professional Development Academy that will provide undergraduates with authentic interdisciplinary research experiences, training in scientific writing, and exposure to the industry while also affording graduate students a mentored teaching opportunity; and (3) organize annual 3-day Hackathon-Style Collaborative Workshops in Computational Neuroscience & Collective Behavior designed to foster novel collaborations between mathematicians and biologists and provide networking opportunities and enhanced career prospects for RTG graduate students and postdocs.

Intellectual Merit: The working groups will advance the understanding of how system components (cells in a neuronal network or organisms in a shared environment) interact to generate functional behavior at higher levels of the organization. Knowledge will be created on (i) how to identify interaction rules from data, (ii) how to model them dynamically, stochastically and statistically, (iii) how to estimate model parameter values from data, (iv) how to simulate the models, compare the simulations to data and modify the models accordingly, (v) how to analyze the data and, more generally, how to compatibilize both the models with the data and the different modeling approaches (dynamic, stochastic, statistical) among themselves.

Broader Impacts: This project will increase the number of minority students pursuing undergraduate and graduate degrees in mathematical sciences through partnerships with the GS-LSAMP and NNJ-B2B programs at 7 universities and 5 community colleges in New Jersey that are dedicated to increasing the retention, graduation, and success of students from racial and ethnic groups that are historically underrepresented in STEM fields. This project will increase the number of female students pursuing undergraduate and graduate degrees in mathematical sciences through partnerships with the Murray Center for Women in Technology and the Albert Dorman Honors College at NJIT. This project will equip students and postdoctoral associates with the transdisciplinary skills to tackle some of society’s most complex challenges.

Start date: 05-01-2024

End date: 04-30-2029

Last modified: 01-19-2024

Copyright information:

DMSP for "RTG: Transdisciplinary Training in Mathematical and Computational Biology at NJIT: From Data to Theory and Back"

Primarily, the proposed work will generate model data. Mathematical models produced for the purpose of mathematical analysis and numerical simulations will be described fully and included in refereed publications. Codes will be made available to other researchers by depositing them on GitHub and ModelDB (a curated database of published models in computational neuroscience). The dissemination of the model code will comply with the New Jersey Institute of Technology policies governing intellectual property, copyright, and the dissemination of research products.
The project will also generate experimental data, which will be managed as described below.
The sub-award organization, Essex County College, will not generate data on their campus. Essex County College students, as well as RTG participants from other institutions, will be granted access to NJIT data and computer systems in accordance with university policies on security.

Golowasch/Nadim Labs: Primary data samples from electrophysiological experiments are the main "product" and will be shared freely upon request. This will be explicitly indicated in the publications derived from this work.
Severi Lab: This proposal will generate calcium imaging and physiology data, behavior tracking data, with and without optogenetic stimulation, and computational results, as well as routine documentation of laboratory procedures. We will document the production of data and laboratory procedures in a laboratory notebook that each lab member backs up on a web-based Google document (NJIT provides unlimited space on Google Drive).
Garnier Lab: Raw data will consist of videos (saved as .MP4 files) and still images (saved as .JPG files) of field and laboratory experiments. This data will require significant storage space and, therefore, possibly be compressed before storage, provided that it does not impair the extraction of observations and measurements from them. Raw data will also consist of manual observations and measurements that will be recorded in paper notebooks and electronic databases.

All data, metadata, and source code collected during the proposed project will be made publicly available as per NSF guidelines within 3 years of collection via published manuscripts, publicly available final reports to NSF, and/or from online data sharing systems, including GitHub, Google Drive, and the Open Science Framework (OSF).

We will reasonably facilitate the access by other researchers to software and data providing "physical samples" or methodologies derived from this research in response to their request and as reasonably quickly as possible. Reasonable exceptions may include work or applications that may need to be protected for intellectual and copyright reasons, in which case we will proceed in line with our Offices of Research and in accordance with the Bayh-Dole Act. All samples (if long-term preservation applies) and data will be kept for at least 3 years following the programs' completion. University Environmental Health and Safety officers will appropriately dispose of the physical samples generated.
Creative commons CC-ND-BY-SA will be used.

Golowasch/Nadim Labs: The results of experimental manipulations and analysis will be preserved permanently using NJIT-wide backup systems and locally in a data backup system located in each lab. The results will be carefully organized and indexed (both digitally and in appropriate laboratory notebooks) and preserved to allow for their reproducibility and verification. The experimental samples will be archived in standard formats

derived from the acquisition software. Notebooks with summary descriptions of experiments and results and their location will be routinely kept and updated.
Severi Lab: Short-term storage will take place on imaging and tracking computers. All lab computers back up daily to a local Network Associated Storage unit, followed by a deep back-up to NJIT servers for long-term storage. Paper laboratory notebooks are kept as a hard copy backup for procedures and observations. We will organize and archive the data in the OSF system.

Garnier Lab: All data will be archived at 3 locations: (1) in the lab, (2) outside the lab on NJIT's campus, and (3) off campus. Data will be stored in the lab on specialized RAID 6 storage arrays. On NJIT's campus, data will be stored on the Andrew File System servers that are backed up daily. Off-campus data will be archived in GitHub repositories, including GitHub's Large File Storage system, when necessary.

All data, metadata, and source code collected during the proposed project will be made publicly available as per NSF guidelines within 3 years of collection via published manuscripts, publicly available final reports to NSF, and/or from online data sharing systems, including GitHub, Google Drive, and the Open Science Framework (OSF), and STEM for Sucess institutional repository.

Planned Research Outputs

Text - "RTG: First Article (preliminary date)"

Articles for academic dissemination on journals

Audiovisual - "RTG Videos"

Videos showcasing the research projects

Planned research output details

Title	Type	Anticipated release date	Initial access level	Intended repository(ies)	Anticipated file size	License	Metadata standard(s)	May contain sensitive data?	May contain PII?
RTG: First Article (preliminary date)	Text	2025-06-29	Open	ModelDB	5 MB	Creative Commons Attribution 4.0 International	Dublin Core	No	No
RTG Videos	Audiovisual	2025-06-29	Open	Open Science Framework STEM for Success	600 MB	Creative Commons Attribution Non Commercial No Derivatives 4.0 International	Dublin Core	No	No