



MIDS Capstone Project: Request for Proposals

Overview

The Duke University Master in Interdisciplinary Data Science (MIDS) program is requesting proposals for year-long Capstone projects. Capstone projects pair external partners with teams of second year MIDS students to generate data insights and recommendations tailored to the partner's needs. Partners provide a problem directly related to their interests and meet with the student team throughout the year. The student team works on the partner's problem to provide related deliverables that often include both insights/understanding to the problem and relevant code bases.

The Capstone projects allow our partners to explore a domain they may not otherwise have the resources to investigate and to work with (and potentially recruit) our students. For our students, these projects provide an educational project in which the students can investigate real-world problems while interacting with and learning from people with vested interest in the project outcome.

Capstone Projects are one of the most critical components of the MIDS curriculum. Throughout the two semesters of a Capstone, MIDS students participate in a Capstone course dedicated to establishing team expectations, creating work plans, and improving team communication. The MIDS program also provides each team with a faculty mentor to guide the student progress, provide technical expertise, and to help facilitate the connection between the partner and students. MIDS students work collaboratively in Capstones, and each team must achieve specific outcomes of interest for the participating outside party and give a final presentation with an accompanying white paper about the outcome's implications.

This document is a call-for-proposals for Capstone projects during the 2022-2023 academic school year. We encourage proposals that involve previously untested ideas or un-analyzed open-source datasets. Our aim is to provide important proof-of-principle work that may lead to more substantial collaborations in the future. We also welcome proposals that might create tools that facilitate community engagement with data and data-driven questions.

How to Apply

To apply, please prepare a document (three pages maximum) that responds to the following prompts, ideally in this order. If you would like help in developing your proposal, please contact Gregory Herschlag (gjh@math.duke.edu) and Ryan Huang (ryan.huang@duke.edu).

Name of Project: Name of Project: Please use a short name that succinctly describes the nature of the project and is not overly technical. If your project is selected for a MIDS Capstone Project, then this title will be used for the project web page and project listings.

Summary: Please write a project summary, including the basic ideas behind the proposal and the question you are seeking to answer.

Contacts: Please list your name(s), affiliation(s), and email address(es). Please describe the expected benefits from your participation. In successful projects, contacts tend to meet with the students roughly once per week.

Goals: Describe the intended goals and products of the project, in the following manner:

1. Describe entirely reachable goals that you fully expect the students to achieve: these could be answers to a question, explorations of a hypothesis, and things of that nature.
2. Describe a tangible product the students will create in the course of their research, which ideally will be of use both to further your own research and to the students as something they can show off to future employers. This could be, for example, a good piece of well-commented software, or a visualization device, or a detailed curation of previously raw data.
3. Describe a more outrageous goal that you would be quite (pleasantly!) surprised to see the students achieve, along with a plan for them to build a potential roadmap towards that goal. For example, this goal might only be reachable if you had data that you currently do not have, and the students might build a speculative roadmap towards acquiring that data.

Data: Most Capstone Projects involve analysis of datasets. Some of these are publicly available, and some are not. **As it is essential that students be able to analyze the needed data for the project, we are very interested in plans to ensure that this will happen.** Please address this in the following manner:

1. For each dataset that will be analyzed by the student team, please give a high-level description of the dataset (what's in it, how was it collected, and for which purpose, how large is it, etc. . . .).
2. For each dataset, indicate whether you anticipate IRB approval will be needed for student access, and if not, why not. If IRB approval will be needed, indicate whether a protocol already exists and your plan for incorporating the student involvement. If it does not already exist, please describe your plan (including a timeline) for obtaining one.
3. For each dataset, indicate whether it is owned and/or is being provided by an outside party. If so, please describe the intended path towards ensuring that students will be granted the ability to access the dataset (we are often able to assist in crafting Data Use Agreements with outside parties, for example).

Deadline and Contact

Applications and inquires may be submitted on a rolling basis. For full consideration, a completed RFP must be submitted by March 21, 2022. **please note that MIDS Capstone Projects require a full academic year (September – April).**

Please submit your application to Gregory Herschlag (gjh@math.duke.edu) and Ryan Huang (ryan.huang@duke.edu).