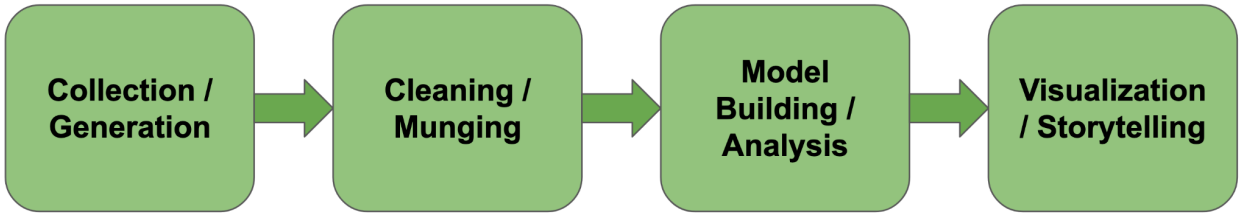
Data Science Capstone

Department of Computational Mathematics, Science, and Engineering
Michigan State University
Spring 2023 (January 9 - May 5)

The Department of Computational Mathematics, Science, and Engineering (CMSE) at Michigan State University (MSU) is seeking sponsors[1] to support projects in the new Data Science Capstone Course. The broad goals of this capstone course are: (1) to provide practical, team-based, project development experiences to MSU's senior undergraduate data science students, and (2) to help our partners make connections with data science students that may support the sponsors' technical and/or recruiting needs.

Overview of Sponsored Projects and Expectations

Sponsors will have the opportunity to work with the course instructor to develop a data science project appropriate for a team of 4-5 seniors who have completed the majority of their Data Science degree at MSU. These students are familiar with multiple programming languages (e.g., Python, R, MATLAB, Bash, C++) and have experience in a range of data science skills and topics, including: databases, data manipulation, statistical analysis, optimization methods, computational modeling, large scale computing, scientific visualization, and machine learning. Each member of the team will work approximately 10-15 hours a week during the 10-12 week project period (totalling 400-900 person-hours per project). Unless the sponsor makes other arrangements in advance, students will use university computational resources to complete the project, and results will be shared publicly (final reports, presentations, etc.).



Ideal projects include each element of the data science pipeline, from data collection to presentation, as outlined above. However, projects may focus more heavily in some areas than others, depending on the sponsors' goals. Examples of the types of projects that may be appropriate include:

- ***Sponsors have an existing dataset and some key questions:*** students will write software to read in the data, generate models (e.g., using Machine Learning), and show results - all with the goal of trying to answer the sponsor's questions.
- ***Sponsors have an existing dataset, but no specific questions:*** students will build an interface to parse the data, generate models, visualize the dataset, and explore the data - all with the goal of identifying categories or types of questions for which the data could provide answers.
- ***Sponsors have questions, but no data:*** students will generate, query, or mine desired data from existing resources (such as the internet), clean up the data, visualize data, and apply

[1] Note: In addition to project descriptions, project sponsorship typically includes a $5,000 donation to the CMSE department; this donation may be waived for projects sponsored by units at MSU or in the state government.

models - all with the goal of trying to identify appropriate datasets and answer the sponsor's questions.

In an effort to build relationships and try to obtain the best project outcomes, sponsors are encouraged to engage with students and instructors regularly. At a minimum, sponsors are asked to meet monthly with student teams and provide weekly feedback by email. Additional informational, networking, and/or project-related activities will be available on a weekly basis for sponsors interested in more engagement. For highly engaged sponsors, the time commitment is expected to average 1.5 hours a week (~20 hours total).

Overview of Timeline and Next Steps
Interested sponsors are encouraged to contact Dr. Dirk Colbry (colbrydi@msu.edu) as far in advance as possible. An initial meeting will be scheduled to share more information and determine whether this opportunity is a good fit for the sponsor's interests and needs.

In general, sponsors who choose to move forward would work with the instructor to complete the following **tasks during Fall 2022**:

- Develop an appropriate project description
- Prepare any necessary agreements for the project (e.g., expectations about intellectual property, non-disclosure agreements, software/hardware resources, data/information security restrictions)
- Identify an individual within the sponsoring organization to act as the primary point of contact and "client" for the student project team
- (optional) Provide a logo to include on the CMSE 495 course website
- (optional) Approve a press release and article about the sponsorship that will be posted on the CMSE department website, and may be picked up and distributed by other news outlets

Key **tasks during Spring 2023** include:
- Provide regular feedback on project deliverables, and guidance and mentoring to the student team during the semester (mid January 2023 through early May 2023)
- Provide feedback on the final project report (including a video and support materials), which will be provided to the sponsor at the end of the project
- (optional) Share information about the sponsor's organization at a seminar for appropriate MSU audiences (e.g., undergraduates, graduate students, postdocs, faculty, or staff in CMSE, the College of Engineering, and/or the College of Natural Science)
- (optional) Participate in networking activities with students from the course (beyond the sponsor's project team, as desired)

The tentative project timeline for Spring 2023 is outlined below:

Week	**Description**
January 16-20	Teams assigned to each project
January 23-27	First Sponsor/Student/Faculty networking event
February 6-10	Project proposal presentations
March 20-24	Mid-Semester Progress Report and Presentation
April 24-28	Final presentations

Potential sponsors should reach out to Dr. Dirk Colbry (colbrydi@msu.edu) as soon as possible.