Mississippi State University

Bachelor of Science in Data Science

Capstone Project Guidelines

Drafted By:

University Data Science Committee

Capstone and Partnerships Subcommittee

Dr. Mahdi Ghafoori
Dr. Mimmo Parisi
Dr. Julie Shedd
Dr. Merrill Warkentin (Chair)

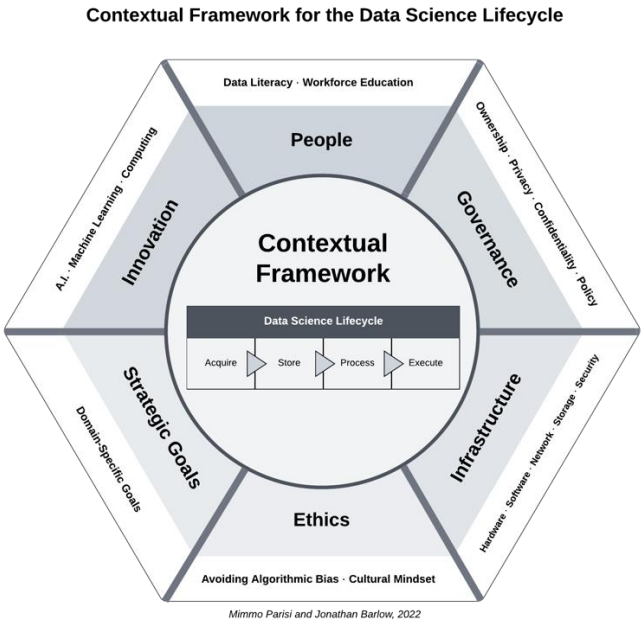
Last Updated, November 2024

https://dsci.msstate.edu/

INTRODUCTION

These guidelines describe the general expectations for the Mississippi State University (MSU) Bachelor of Science in Data Science Capstone Projects and outline the process, a framework for timelines, and types of deliverables. **Each concentration will determine how to implement these guidelines (as well as to what extent) to meet the specific needs of their core curriculum and student outcomes.** This document provides a plan to guide students in the capstone work.

Capstone projects are a unique and essential experiential learning component of the Bachelor of Science in Data Science at MSU. The goal is to enable students to apply the knowledge, skills, and competencies gained through their coursework to real-world situations. The University Data Science Committee (UDSC) Capstone and Partnerships Subcommittee developed the capstone process in response to prospective employer requests that MSU's graduates possess real-world skills in data science and its applications. The MSU Bachelor of Science in Data Science degree provides three categories of skills: 1) traditional liberal arts skills in communications, analytical thinking, and imagination; 2) skills in navigating diversity and teamwork while gaining technical fluencies that enable success in an increasingly digitally connected and operated world; and 3) the foundations of data science theory and practice. The capstone project provides students an opportunity to work with faculty to design, conduct, and present a data science solution in the student's area of concentration.



Students pursuing the Bachelor of Science in Data Science degree must successfully complete a capstone project (6 credit hours), which is specified in their concentration. The two capstone courses typically cannot be completed in the same semester. **While some concentrations use their own courses (where the scope and enforcement of this document shall not apply to their students)**, many concentrations use the two DSCI capstone courses:

- DSCI 4553 (or equivalent) – Data Science Capstone 1 (3 credit hours)
- DSCI 4663 (or equivalent) – Data Science Capstone 2 (3 credit hours)

Students who pursue capstone projects should collaborate with MSU faculty members. Students are responsible for initiating the search for and identifying a faculty advisor for their capstone projects at least one semester before enrolling in their first capstone course. The prospective faculty advisor should ordinarily be from the college that offers the student's concentration. The faculty advisor should help the student identify a problem that can be addressed and solved within the expectations of a two-semester capstone project. The capstone should use the data lifecycle within the overall contextual framework used in the MSU Bachelor of Science in Data Science (see diagram above).

The capstone project ordinarily requires:

- An environment in which the students can learn to operate as independent, professional individuals;
- A spirit of trust and collaboration by all involved parties;
- Healthy and professional communication and rapport among all involved parties; and
- The ability for students to choose courses of action, make mistakes, and learn from those experiences.

CAPSTONE PROJECT SUGGESTED TIMELINE OVERVIEW

Below are the general tasks that generally should be accomplished in each semester of a student's capstone project. **Note that if a specific concentration uses its own courses instead of DSCI 4553/4663, the scope and enforcement outlined in this document shall not apply to their students.** Students are expected to develop a timeline to complete the tasks outlined below. The instructor can request additional milestones for drafts or other materials in addition to the tasks listed here.

Semester of Enrollment in First Capstone Course	
Task 1	Student submits *Design and Implementation Plan* to the instructor.
Task 2	Faculty advisor provides feedback on *Design and Implementation Plan* to the student.
Task 3	*Design and Implementation Plan* reviewed by the faculty.
Task 4	Student presents the *Design and Implementation Plan* in a group setting (1 hour).
Task 5	Student submits final version of *Design and Implementation Plan* to the instructor.
Semester of Enrollment in Second Capstone Course	
Task 6	Student submits the outline of *Technical Documentation* to the faculty advisor.
Task 7	Student completes the draft of *Technical Documentation* and *Project Repository* (if applicable) to the faculty advisor.
Task 8	Student completes the capstone project final presentation (1 hour).
Task 9	The student submits the finalized *Technical Documentation* (PDF version) and *Project Repository* (if applicable) to the instructor.

GENERAL INFORMATION

A. Capstone Project Suggested Timeline

MSU Bachelor of Science in Data Science students complete their capstone project over two semesters. *Design and Implementation Plan* faculty reviews are held at the end of the first capstone semester. A student's Capstone Project Final Presentation is held at the end of the second capstone semester, with the *Technical Documentation* and *Project Repository* (if applicable) also due at the end of the second capstone semester. The timeline overview above provides the suggested tasks and deliverables for each semester. The student and the faculty advisor may define their internal timelines for establishing the student's progress toward the completion of the project.

B. Academic Credit Hours and Grading

Students generally register for their two capstone courses during their final two semesters. The first capstone course should be completed before the second. **Students must achieve a grade of C or better in both capstone courses to meet the requirements for the MSU Bachelor of Science in Data Science degree, which only apply to DSCI 4553/4663.** If a specific concentration uses its own courses instead of DSCI 4553/4663, this requirement shall not apply to their students.

C. Student Time Commitment

Students should expect to devote a **considerable number of hours per week** to their capstone project, although more time can be expected for some tasks. Work on capstone projects should be evenly allocated over the two semesters to avoid excess workload at the end of the project.

D. Deliverables for Whole Project (Capstone 1 and 2)

The primary deliverables for the Capstone Project are typically the following:

- *Design and Implementation Plan* (document);
- *Design and Implementation Plan* Presentation (1 hour);
- *Technical Documentation*;
- Capstone Project Final Presentation (1 hour);
- *Project Repository*, which includes as appropriate:
    - Completed analysis or model, including data description;
    - Smart system design, prototype, or logic design as applicable;
    - Documentation and workflows;
    - Interface and/or data visualization;
    - Software source code, if applicable;
    - Note: *Project Repository* is not needed when the student works with non-Data Science students in their project;
- Presentation of Capstone Project at an MSU Undergraduate Research Symposium or similar conference.

E. Publishing

Students desiring to publish their work must discuss this intent with their advisors. Faculty advisors are experts in peer-reviewed publications, and students should take advantage of their knowledge and experience. Publishing peer-reviewed literature requires interfacing with a larger scholarly community, which should be done in a way that reflects well on the student, their advisor(s), the original sources of

data, and MSU. Note that it often takes a prolonged period (potentially years) to achieve publication of a paper. Data collection for the purposes of publishing research results may initiate the need for ethical review through MSU's Institutional Review Board if humans or animals are used in the research process (see **Section H** below).

F. Authorship/Ownership

Students are encouraged, but not required, to present their capstone project findings in formal conferences outside the university. Participation in such conferences gives students valuable experience and increases the visibility of MSU's Bachelor of Science in Data Science program and its students. Students may also wish to publish results in a peer-reviewed journal. Students may also include their faculty advisor or others who contributed substantially to their capstone project research as co-authors. Students and faculty advisor(s) may collectively develop criteria for authorship of these supplemental materials. The faculty advisor may choose not to be included as a co-author on a publication. After the project has ended, if a student adapts the deliverables for presentation at a meeting or submission to a journal, the Bachelor of Science in Data Science program should be appropriately acknowledged.

G. Data Distribution

Depending upon the nature and context of the student's capstone project, datasets obtained or derived during the work may be covered by licenses, copyrights, or confidentiality agreements. Students should not make such data available to third parties or generally available online without authorization from their faculty advisor and the original source of data and/or its respective owner(s). The use of this data should also follow the guidelines and policies established by MSU or the data contributor/owner. If a student is working with confidential data, a Non-Disclosure Agreement (NDA) will likely be necessary through formal agreement between the original source of data and MSU. NDAs are developed, approved, and signed only by MSU's Office of Technology Management. Under no circumstances can a student sign an NDA or comparable agreement with an original source of data.

H. Use of Human Subjects

Faculty and students who engage in research involving human subjects must obtain prior approval from the MSU Office of Research Compliance and Security (ORCS) – Institutional Research Board (IRB). "Human Subject" means a living individual about whom an investigator (whether professional or student) conducting research obtains (i) data through intervention or interaction with the individual or (ii) identifiable private information. This means that ORCS approval MUST be obtained in advance if a survey is conducted. Approval is required no matter with whom the student will interact – even friends or family! Students and faculty advisors must undergo IRB training before seeking approval for such research. (Note: Similar requirements may be required for research involving the use of animals.)

Students who use human subjects must review the ORCS site: https://www.orc.msstate.edu/compliance/human-subjects. They must understand and abide by the policies and procedures. The process for obtaining ORCS approval for the proposed research is not immediate and cannot be addressed at the last minute. There are serious consequences if a student is not in compliance. Human subjects cannot be interviewed, surveyed, or contacted in any way without prior approval from ORCS.

COMPOSITION OF THE CAPSTONE PROJECTS

The following are the requirements for the DSCI courses. Other concentration capstones may have different requirements.

A. Faculty Advisors

Each student working on a Capstone Project must identify an MSU faculty member as their project advisor, ordinarily in the student's selected concentration. If a student is enrolled in more than one concentration, the student may need to fulfill the Bachelor of Science in Data Science program's requirements by carrying out **a separate six-hour capstone project for each concentration**. The UDSC Capstone and Partnerships Subcommittee will review situations involving students pursuing multiple concentrations on a case-by-case basis.

The faculty advisor is generally expected to monitor progress and provide technical assistance, domain expertise, project feedback, project evaluations, and deliverable quality evaluations. **The student is responsible for project leadership, management, and the quality of their final products.**

Faculty advisors do not serve as project managers; their role is similar to that of a consultant. The advisor meets with the student regularly in the designated capstone course. The advisors may offer reactive advice, respond to activities, and provide guidance when asked. The advisors may also give proactive advice regarding deficiencies and timelines if appropriate. It is crucial that students understand the role of the advisors and the limited, though important, role they serve in directing a student's capstone project. Each faculty advisor will likely have a unique approach. Students should expect variability in engagement, expectations, and feedback from one advisor to the next. During the first capstone semester, each student should clarify the expected level of interaction with their advisors. The student should include this information in their *Design and Implementation Plan*.

B. External Advisors/Partnerships

Interacting and networking with the professional community may be valuable to the capstone project process. We encourage students to seek expertise from other professionals to assist them as informal external advisors/partners, such as individuals from government agencies, industry, non-governmental organizations, universities, or private citizens who may be interested in the project, its data, or its deliverables. An external advisor/partner is someone who has knowledge about the project topic and can provide unbiased feedback.

If an external advisor/partner is desired, the student will be responsible for identifying external advisors/partners and maintaining professional contact with the advisors/partners for the duration of the project. The external advisors/partners may be invited to meet with the faculty advisor(s) and the student. The student may invite external advisors/partners to attend the Capstone Project Final Presentation.

C. Data Science Academic Coordinator

The Data Science Academic Coordinator is an MSU staff member who assists students, faculty advisors, and the Capstone and Partnerships Subcommittee in facilitating the Capstone Project process. A recommended disagreement resolution process between the student and the faculty advisor is described in the next section.

PROJECT MANAGEMENT

The following are the requirements for the DSCI courses. Other concentration capstones may have different requirements.

A. Capstone Project Meetings

The student and the faculty advisor are expected to meet as often as necessary. It is the responsibility of the student, not the advisor, to schedule the meetings and make necessary arrangements. As a matter of courtesy, a student should provide advance notification of meeting absences.

B. Disagreement Resolution

If there is a disagreement on how to pursue the project between the student and the faculty advisor during the capstone project, and both parties are unable to resolve the disagreement after internal discussion, the student and the faculty advisor may seek assistance from the Capstone and Partnerships Subcommittee or the Data Science Program Executive Director. Further resolution through the "Student Advocate/Ombudsman" is available at https://www.students.msstate.edu/resources/student-advocate-ombudsman/ for students.

PROJECT DELIVERABLES

The following are the requirements for the DSCI courses. Other concentration capstones may have different requirements.

Refer to the suggested timeline overview for a summary of typical capstone project tasks and deliverables. Students must pass capstone courses **with a C or better** to fulfill the requirements of the Bachelor of Science in Data Science degree. Students must actively interact with their faculty advisors throughout the project to receive a passing grade.

The following sections provide detailed suggestions for how to structure the project tasks and deliverables. **Please note that these are suggestions only and should not be construed as universal requirements.**

A. First Capstone Course (DSCI 4553 or Equivalent)

Bachelor of Science in Data Science students begin capstone projects ideally in the penultimate semester of their degree and must register for DSCI 4553 or the equivalent course specified by their concentration. This section proposes a potential structure that faculty advisors and students can use to organize their work together in the first semester.

The University Data Science Committee recommends that the first semester of the capstone project generally involves the completion of the following components:

1. Scoping of the Project

As part of their first capstone course, each student should develop a feasible *Design and Implementation Plan*. Students should immediately begin investigating the objective presented in the plan. **A student should schedule regular meetings with their faculty advisor, ideally starting during the first week of their first capstone semester.** Students and faculty advisors should aim to determine early on what is feasible to complete within an academic year and to discuss expectations for project work.

The first capstone semester should focus on the design of the data science solution, ensuring that the proposed design addresses the following:

- Data science solutions that contain:
    - A clear challenge, problem, or opportunity to address;
    - A clear connection to theory and best practices within the field of concentration;
    - Robust, reproducible workflows;
    - Reliable, sustainable products that are easy to use and maintain;
    - Quality assurance and error-checking; and
    - A plan describing how the products could feasibly be incorporated into a smart system.
- Interface development, e.g., data visualization;
- Testing approach;
- Documentation;
- Data, model, workflows, and documentation that are archivable and retrievable; and
- Project management strategy for design and implementation.

A critical element of a successful capstone project is clear, frequent, and open communication between the student and the faculty advisor. External advisors/partnerships may be involved with the Capstone Projects to a greater or lesser degree, depending on their availability, expertise, and desired level of engagement. Central to deciding what the project will and will not include is a scoping meeting that is suggested for weeks 1–2 with the student and faculty advisor(s). The meeting between the student and faculty advisor aims to develop a shared understanding of the project objectives. If any objectives in the original proposal are not feasible, the faculty advisor and students should discuss this to develop an understanding of how the student plans to proceed. **The meeting(s) with the faculty advisor should result in an explicit agreement about what is possible and not possible, given the available time, resources, data, external advisor/partnership support, and faculty and student capabilities.** Following these meetings, students should document the project objectives and produce a short list of deliverables.

Students may schedule additional meetings with the faculty advisor(s) as needed and appropriate during the scoping phase.

2. Design and Implementation Plan

The *Design and Implementation Plan* should be a concrete and realistic statement of what the student will do to solve the data science problem identified by the student and approved by the faculty advisor. The construction and articulation of this plan are significant parts of the work essential to solve the problem. **While the page limit of the *Design and Implementation Plan* is subject to the faculty advisor, the recommendation is ten (10) single-spaced pages of written text.** The recommended page limit does not include the title page, executive summary, graphics/figures, or references cited. All graphics and figures should be included as appendices and referenced in the body of the document. A draft of the *Design and Implementation Plan* should be submitted to the faculty advisor by the due date that is developed internally between the student and the faculty advisor). The final *Design and Implementation Plan* should be submitted to the faculty advisor and the Data Science Academic Coordinator by the last day of the first capstone semester.

The *Design and Implementation Plan* typically includes the following components:

a. Title Page
b. Executive Summary (should not exceed one page)
c. Problem Statement and Objectives
d. Theory

e. Solution Design:
    a. Approach and Methods
    b. Data Management Plan (DMP)
    c. Proposed Software or Tools
    d. Proposed Smart System Implementation or Prototype
f. Products and Deliverables
g. Documentation Plan
h. Testing Plan
i. Timeline and Milestones
j. References

Description of *Design and Implementation Plan* components:

a. Title Page (see formatting guidelines in Appendix I)
The title page typically includes the following information:
i. Title (no more than ten words recommended), including the name of concentration;
ii. Name of the student;
iii. Name(s) of faculty advisor(s);
iv. Capstone course number and title – Design and Implementation Plan;
v. Mississippi State University;
vi. Date (month and year of completion, e.g., December 2023);

b. Executive Summary (should not exceed one page)
The executive summary should include background information, objectives, proposed approach, and anticipated results to give readers a high-level overview. The executive summary should be text only; do not include graphs, tables, or photos.

c. Problem Statement and Objectives

d. The problem statement and objectives typically provide a brief description of the problem to be addressed through data science in the enrolled concentration(s). It should include a short list of objectives based on the project proposal and the feedback received during meetings between the faculty advisor and the student. Typically, students are expected to address only one objective for a Capstone Project. **Theory**
This section usually provides a brief overview of the dominant theories within the field of inquiry that impact the identified problem and describes how these theories relate to the likely success of the proposed solution design.

e. Solution Design

i. Approach and Methods
This section often begins with an overall strategy and the specific tasks necessary to accomplish the objectives. It can describe in detail how data is organized and managed to support analyses that will inform the creation of a smart system to address the stated objectives. It should use schematics, conceptual models, or workflow diagrams where possible and appropriate. It should also include information on software tools to be used. It may explain the following: how the data will be prepared for analysis; the nature of the results and how results will be communicated to users; a detailed plan for the visualization of results and a plan for ensuring the reliability of the data analysis; how testing will be incorporated; how sustainability and maintenance are being considered in the design, tool selection, and development of the product; how documentation of workflows will look. Overall, the section should summarize how software or data analysis will be used to meet the overall project objectives by gaining insight through analysis and using this insight to design a smart system.

ii. Data Management Plan (DMP)

Data and metadata emerging from the project will typically be archived for public use according to a license agreed upon by all parties and the originator of the data. This expectation does not apply to data restricted by a Non-Disclosure Agreement (NDA). The student's *Design and Implementation Plan* should typically include a Data Management Plan for public and restricted data. The DMP usually describes how research data will be managed during the project and, if appropriate, made available to others after completion. There are several major topics to discuss in the DMP:

1. *Description of the research data*. What data are needed? How will the data be securely gathered/transferred from the data owner to the student and then securely stored? Describe the data the student will use, including the file types, data set size, number of expected files or sets, content, and source of the data (creator and method of collection).
2. *Data standards.* Are there any standard formats in the specific research field for managing or disseminating the data sets that have been identified (e.g., CSV, PNG, JPG, PY, BAG)?
3. *Metadata.* Metadata is the documentation that helps make data sets reusable. Think about what details (metadata) someone would need to be able to understand and use these files. For example, perhaps a *readme.txt* file is necessary to explain variables, the structure of the files, etc.
4. *Data sharing and access.* The data may have significant value for other researchers beyond this project, and sharing this data is often part of the student's responsibility as a member of the scientific community. Specify the extent to which data can be reused, including any licensing and access limitations. List any proprietary software that might be needed to read the files. If there is data that are not appropriate for sharing due to confidentiality, NDA, or disclosure risk, describe that here.
5. *Intellectual property (IP) and re-use.* If data were collected from outside the university (e.g., external partnerships, government agencies, non-government organizations), does the student have the right to redistribute it? If so, are there any restrictions on redistribution? If students create their data files, will they assign a Creative Commons license to their data? For advice on how to move forward, contact the instructor. The Libraries' Scholarly Communication Team can provide additional insight; email sct@library.msstate.edu.
6. *Data archiving and preservation.* Throughout the project, the student may produce a large number of files. At the end of the project, students typically submit data used in the project (except data protected by an NDA) and associated metadata. Not all data need to be saved, such as interim results derived from an original data set. If another researcher wanted to replicate the student's work or re-use the student's data, what data and documentation would be required for them to do so? Where will the data and metadata be stored after the project is completed? Is there a subject-specific and/or open-access repository appropriate for the data? Necessary files should be archived in Scholars Junction (but can also be submitted to subject-specific repositories). Contact the Scholarly Communication Team at sct@library.msstate.edu for help evaluating subject-specific repositories, preservation options, and documentation best practices.

iii. Proposed Software or Tools

Describe any open-source or proprietary software expected to be used in the project. This may include modeling, spatial analysis, relational database, repository, or any other software, computer language, or library/module that will be used. Describe how this software will be accessed and how it will be used to meet the objectives of the capstone project.

iv. Proposed Smart System or other Data Science Artifact Implementation or Prototype
Describe how the project's analysis will be incorporated into the design for a smart system or into a working prototype (alternatively, a logical data flow model and data structure) of a smart system. Describe specifically how the system will help automate processes that include decision-making or taking actions to solve the problem.

f. Products and Deliverables
There are two types of capstone project deliverables: (1) products and/or (2) academic deliverables. This section should typically list the products that will be created and the academic deliverables, which can include the *Design and Implementation Plan*, project data/metadata, design and implementation plan faculty review presentation, *Technical Documentation*, *Project Repository*, and capstone project final presentation. This section may include detailed descriptions of the expected interface, demonstration, or data visualization.

g. Documentation Plan
A detailed documentation plan should usually be included. The documentation approach should include a description of the documentation format (e.g., GitHub, Wiki, User Manual) and an outline of key topics to be included in the documentation.

h. Testing Plan
A plan for testing should usually be included. Strategies for testing should include functionality to be tested, real or synthetic data to be used for testing, and specific tests to be completed. Depending on the project, testing may also involve other human testers (e.g., to test the usability of an interactive tool).

i. Timeline and Milestones
Milestones will help students complete their capstone projects by identifying key tasks and products they will create as the project proceeds. In this section, students and faculty should make mutually agreed-upon timelines by which they intend to complete the types of tasks outlined in the *Approach and Methods* section, which may include data acquisition, analyses, evaluation of results, product development, and testing and documentation. The timeline may specify dates by which the student intends to complete drafts and final versions of each of its products and deliverables. These timelines may be reviewed and adjusted as needed after the first semester.

j. References
Include a bibliography of references used to support the student's research. Information derived from other authors must be appropriately cited. The reference citations should be presented in a standard and consistent format (e.g., APA, IEEE, Chicago) suggested by the project faculty advisor and can be managed using citation software (e.g., Zotero Reference Manager, Mendeley Reference Manager, EndNote online).

The final *Design and Implementation Plan* is due to the faculty advisor by the end of the first capstone semester.

3. Design and Implementation Plan Faculty Review Presentation

In the first capstone semester, the student will typically present their project *Design and Implementation Plan* for faculty review. By this point, students should typically have revised their draft of their *Design and Implementation Plan* based on feedback from their faculty advisor. The faculty review is an opportunity for the students to present their approach to solving the problem and expected products and documentation. The faculty review can serve as a checkpoint in the capstone project process so that students receive verbal feedback from their faculty advisors with a primary emphasis on theory, methodologies, technical approaches, and proposed products and documentation, such as data visualization or demonstration. At the time of the review, the project will typically still be underway, and students should integrate feedback from faculty advisors to the extent possible. Students are expected to prepare a 30-minute presentation (including a Q&A session). This presentation should include the project's significance, data, methods, and proposed

products and documentation. Students can also include a mockup of an expected interface or data visualization component. Students also may include any outstanding questions for which they are seeking guidance. Following the presentation, the faculty advisor can provide additional comments or suggestions to that can help improve the presentation.

Students should expect questions and constructive comments from their faculty advisor that may result in additional revisions to their *Design and Implementation Plan*. It is the student's responsibility during the faculty review to explain their work to the faculty advisor(s). The faculty advisor will usually provide oral feedback to the student summarizing the strengths of the project and/or recommendations for improvement. Students can invite their external advisors/partners to attend this presentation, if any.

B. Second Capstone Course (DSCI 4663 or Equivalent)

Bachelor of Science in Data Science students complete the second capstone course ideally in the final semester of their degree and must register for DSCI 4663 or the equivalent course specified by their concentration. In the second capstone semester, students typically focus on the implementation of the *Design and Implementation Plan* and execution of product development and delivery, including *Technical Documentation*. By the end of the second capstone semester, students typically should deliver final products and academic deliverables, such as *Technical Documentation* and the *Project Repository*.

The University Data Science Committee recommends that the second semester of the capstone project generally involves the completion of the following components:

1. Outline for Technical Documentation

In the second capstone semester, each student is recommended to submit an outline for the *Technical Documentation* to the faculty advisor. The outline can include the following:

a. Problem Statement
b. Theoretical Context
c. Revised Project Objectives
d. Summary of Solution Design
e. Products and Deliverables
f. Summary of Testing
g. User Documentation
h. Archive Access
i. References

For each section of the outline, students can include as much information as completed to date and note any remaining obstacles to its completion. As an appendix to the outline, students can describe the work to be completed during the remainder of the second capstone semester and a mutually agreed timeline for remaining tasks and deliverables.

2. Capstone Project Final Presentation

Capstone project final presentations celebrate the completion of solution development and offer Bachelor of Science in Data Science students the opportunity to share their work with faculty, peers, potential employers, community members, family, and friends. The final presentation is an opportunity for the students to present their products and the significance of the products in the context of the field and the problem or challenge addressed. Final presentations contribute to the reputation and prestige of the MSU

Bachelor of Science in Data Science program and, hence, the value of the students' degrees. Participants are usually advised to dress in business attire.

Students should present their final projects in the second semester of the capstone project (the timeline should be developed internally between the student and the faculty advisor). At the time of presentation, the project should be nearing completion. Students usually prepare a 1-hour presentation (including a Q&A session) with a product-oriented focus; students can include a demonstration of their product interface or data visualization. Following the presentation, students will typically engage in approximately 10 minutes of Q&A and discussion with the audience.

The audience at the final presentations can be different from the audience at the *Design and Implementation Plan* faculty review. The final presentations should typically be understandable to a diverse audience (e.g., employers, experts, non-experts, clients, regulators, family, friends, etc.). Students usually prepare a presentation for an audience more interested in their concentration area-related data science products and less interested in, for example, analytical methods or data management. Before the final presentation, it is the student's responsibility to submit the project-related information (including an abstract, along with the project title, student's name, faculty advisor(s), and acknowledgments) to the instructor.

The entire Bachelor of Science in Data Science program community will usually be invited to attend the capstone project final presentations. The Data Science Academic Coordinator can make all arrangements, including setting the presentation schedule and facilitating presentations and Q&A. Students can invite their external advisors, if any, to attend this presentation.

3. Technical Documentation

Technical Documentation is an academic deliverable that usually provides a functional discussion of the project's objectives, solution design, products, testing, and user documentation. A complete draft of the *Technical Documentation* is typically submitted by the end of the second capstone semester.

The *Technical Documentation* should reflect the student's ability to articulate in writing the 1) problem and objectives, 2) underlying field-specific theory, 3) summarized solution design, 4) description of products, 5) testing approach and results, and 6) user manual or instructions. The documentation usually demonstrates that the student can develop data science solutions that are robust, reproducible, reliable, and quality-assured. If project data is restricted by an NDA, the student should include sufficient information in the user manual section so that potential users are aware of the data's restricted nature and can contact the issuer of the NDA for access.

The goal of the *Technical Documentation* is typically not to simply repeat information from the *Design and Implementation Plan* but rather to communicate critical information so that a user may understand product development and reproduce results.

The *Technical Documentation* should typically include the following:

a. Title Page (sample provided in Appendix I)
b. Signature Page (ample provided in Appendix II)
c. Abstract (typically 200 words or less)
d. Executive Summary (typically four pages or less)
e. Table of Contents
f. Problem Statement
g. Theoretical Context
h. Specific Objectives

i. Summary of Solution Design
j. Products and Deliverables
k. Summary of Testing
l. User Documentation
m. Archive Access
n. References

A complete draft of the *Technical Documentation* report will typically be due to the faculty advisor in the second semester of the capstone project. The revised final documentation will typically be due at the end of the second semester of the capstone project.

The *Technical Documentation* should acknowledge any individuals or organizations that have supported the project in any significant way, professionally or financially. However, students must first obtain permission to include such acknowledgment; supporters have the right not to be publicly associated with the final documentation and *Project Repository*.

Students should choose a method of citation and use it consistently (e.g., APA, IEEE, Chicago).

Students should usually expect multiple rounds of revision before their *Technical Documentation* is finalized. The report should typically be presented in a professional format. It is the student's responsibility to make sure that the style and level of detail are consistent throughout the report.

Students and faculty advisors are encouraged to discuss and agree upon a realistic internal timeline and consistent expectations for the review process. Some advisors will expect drafts earlier and may require more extended periods for review and comment than others. In general, students should expect that their advisors may require at least 1–2 weeks to review the documentation thoroughly. More than one iteration will likely be necessary before faculty advisors find the *Technical Documentation* acceptable. The project timeline should ideally take into account the time needed for multiple reviews and revisions of the documentation and *Project Repository*. The faculty advisor will often be the only person who reads and provides feedback on the complete draft of the *Technical Documentation*. Some external advisors also may have the interest, time, and expertise to review and provide feedback on some or all of the draft documentation and/or *Project Repository*.

The page limit of *Technical Documentation* is subject to the faculty advisor. The recommendation is that it should be at most 25 pages and be free of typographical, formatting, and other errors. All documentation should typically be formatted in compliance with MSU ETD Standards and Templates at https://scholarsjunction.msstate.edu/td-standards/.

Students should typically provide an electronic copy of their approved *Technical Documentation*, including the completed signature page, to the Data Science Academic Coordinator and their faculty advisors. The *Technical Documentation* will usually be posted to the MSU Bachelor of Science in Data Science program website unless an NDA restricts public sharing. The faculty advisor may also request hard copies of the documentation.

4. Project Repository

Students may be requested to submit a *Project Repository* of their work. A *Project Repository* is not a required component when the student works with non-Data Science students in their project or the concentration has its own requirement. This repository should typically be accessible to the faculty advisor. This may include multiple source code files, data, other resources the project uses, and the user documentation section from the *Technical Documentation*.

A working draft of the *Project Repository* will often be required to be submitted during the second capstone semester, although students are not usually expected to submit a complete repository at this point. The completed repository is typically due at the end of the second capstone semester if applicable.

EVALUATIONS

The following are the requirements for the DSCI courses. Other concentration capstones may have different requirements.

A. Faculty Evaluation of Students

Each student will receive a separate grade for each semester of the capstone project, whether they are enrolled in the DSCI 4553/4663 sequence or equivalent courses specified by the concentration.

A student's performance will be evaluated on the ability of the student to demonstrate the ability to develop data science solutions that are robust, reproducible, reliable, and quality-assured. Specific outcomes that faculty advisors may assess in assigning project grades include:

1. Problem-oriented data science design and implementation design.
2. Project management.
    a. Students may be evaluated on their attendance at all project meetings, their punctuality for such meetings, and their preparation and engagement with their faculty advisors. Students may be evaluated on their ability to manage and deliver intermediate and final products on schedule, both to meet MSU Bachelor of Science in Data Science suggested timelines as well as internal timelines that are mutually agreed upon by both sides.
3. Data processing/analysis/manipulation.
    a. Students may be evaluated on their demonstrated knowledge of coursework and data science techniques to address the problems in the selected concentration areas.
4. Data science solutions.
    a. Students may be evaluated on their ability to produce data science solutions that contain robust, reproducible workflows; reliable and easy-to-use products; quality assurance and error-checking; and a clear connection to theory within the field of inquiry.
5. Interface development (e.g., data visualization).
6. Documentation.
    a. Students may be evaluated on the degree to which their project documentation and demonstrations are well organized, scholarly, and well communicated.
7. Archive/retrievability.
    a. Students may be evaluated on how accessible and retrievable all data, models, workflows, and documentation are to various audiences.
8. Participation.
    a. Students may be evaluated on their ability and willingness to participate and actively contribute to meetings and training sessions.
9. Resourcefulness.
    a. Throughout the project, students may be expected to demonstrate initiative in finding information and identifying tools necessary to achieve the scope of their project.
10. Professional relationships.
    a. Students may be expected to demonstrate the highest level of professionalism and respect in their dealings with their faculty advisors and other stakeholders, if any.

B. Student Assessment of Capstone Project Experience

At the end of the project, all capstone project students will be asked to provide some insights about their experiences with the Capstone Project. The Data Science Academic Coordinator will compile all comments and provide them in a summarized format to faculty only after the final grade for the second capstone course is submitted. This is only for internal use.

C. Bulldog Experience Reflection

Completion and submission of a Bulldog Experience Culminating Reflection which addresses the five Bulldog Experience Reflection prompts. Students will be provided with the reflection prompts at the start of the Capstone experience. Reflections should be submitted through an assignment in the Capstone Canvas course. At the end of the Capstone Project, the faculty would be responsible for sending the Office of Bulldog Experience all of the reflections from the students enrolled in the course.

COMPUTER RESOURCES

The following section describes the computer resources typically available for each capstone project and recommended management practices. Most of these resources do not require any special privileges.

A. Microsoft OneDrive/Office 365 Document Collaboration

If a student needs to collaboratively author electronic documents (e.g., with the faculty advisor) as a deliverable or in support of the capstone project, there is a recommended solution that supports this:

i. MSU Microsoft OneDrive with Office 365 integration

The suggested option allows files to be shared among multiple collaborators and for those collaborating to edit the document simultaneously. MSU OneDrive has the benefit of Office 365 integration, allowing a more fully featured set of authoring tools. Each faculty/staff/student should have 5 TB of free could space in OneDrive.

B. References

Each student must maintain a file of bibliographic references that can be incorporated into project deliverables, papers, etc. This document must be submitted as a plain text file to Scholars Junction with the rest of the capstone project. MSU currently provides institutional access to the Mendeley Reference Manager and Zotero Reference Manager bibliographic software. Some students may prefer to use an online citation manager (e.g., EndNote online); basic accounts are often free.

Appendix I

Sample Title Page for *Design and Implementation Plan* and *Technical Documentation*

MISSISSIPPI STATE UNIVERSITY

PROJECT TITLE

CONCENTRATION

A CAPSTONE PROJECT SUBMITTED IN PARTIAL SATISFACTION OF THE REQUIREMENT FOR THE DEGREE OF BACHELOR OF SCIENCE IN DATA SCIENCE

BY

STUDENT NAME

COMMITTEE IN CHARGE:

ADVISOR NAME

ADVISOR NAME (IF MORE THAN ONE)

MONTH AND YEAR OF FILING

Appendix II

Sample Signature Page for *Technical Documentation*

PROJECT TITLE

As the developer of this capstone project documentation, I archive this documentation on the Mississippi State University Bachelor of Science in Data Science program's website such that the results of my research are available for all to read. My signature on the document signifies my responsibility to fulfill the archiving standards set by Mississippi State University.

_____

STUDENT NAME

[The faculty advisor may change this statement prior to submitting this report.]

[STATEMENT]

This Capstone Project *Technical Documentation* is authored by a Bachelor of Science in Data Science student and has been reviewed and approved by:

_____

FACULTY ADVISOR NAME

_____

FACULTY ADVISOR NAME

_____