Data Intensive (DI) Requirement

We live in a world that is increasingly data and information-rich and the fields for which we are preparing our students are increasingly data-driven. Data are being collected in all areas of the human experience, and massive collections of this data are now readily available online and elsewhere. In aggregate, these data collections are providing unprecedented opportunities for growth in human understanding. Concomitantly, data analytics and interpretation are emerging as fundamental components of daily life (e.g. social media, online commerce, search engines, web media, infrastructure optimization and management, electronic medicine, etc.). Data analysis and modeling are becoming ubiquitous throughout science, technology, engineering, medicine, the humanities, arts, and social sciences, and business. Students will almost certainly encounter data, data analysis, and the results of data analysis in their careers and daily lives, regardless of their major and career path. As data consumers, they will at a minimum need to understand how to use data interpretation and data analytics for decision making. Many will be creators of data analysis solutions in the technological workforce, and others will play decision-making roles where they will have to leverage data resources, or manage their creation and maintenance. Rensselaer has taken a leadership position in defining a well-educated student in the new data driven economy and world by making sure that all students, regardless of major, have data awareness and data dexterity.

Through its *Data Intensive*, or DI, requirement, Rensselaer aims to develop graduates who are sophisticated consumers, producers, and users of data. Our graduates will be able to interpret and/or transform data into actionable insight and solutions to compelling problems and then communicate associated findings effectively to diverse audiences. In order to best prepare our students to operate in this rapidly evolving environment, provide the global workforce with thought leaders, and create breakthroughs that will become tomorrow's new products and solutions, Rensselaer has made data dexterity a core part of the undergraduate experience.

Through a university-wide effort, Rensselaer has defined a distinctive integrated interdisciplinary undergraduate requirement in data dexterity. The data dexterity requirement begins with the premise that all Rensselaer students will acquire basic data awareness and literacy skills as part of their core undergraduate experience. To accomplish this goal, the Institute is creating a sequence of courses that culminates in students developing capabilities in the following data dexterity objectives:

- understanding of how modeling and analysis of data can be used to contribute to the solution of real world problems
- applying quantitative algorithms and techniques and/or qualitative approaches to diverse data and interpreting the results, including an understanding of uncertainty, probability and statistics
- communicating effectively the results and insights gained from data analysis to diverse audiences through oral, written, and multi-media presentations
- developing an awareness of the importance of data stewardship and documentation across the data life cycle
- understanding the ethical use and impact of data on society

This document specifies criteria for DI courses and procedures for proposing a course as DI. Analogous to the current Communication Intensive (CI) requirement, all Rensselaer undergraduates must complete a two course *DI* sequence. The first part of the requirement, termed *DI-1*, is fulfilled by successfully completing a *DI-1* designated introductory course that incorporates one or more instructor-led illustrations, or modules, in which data are used to answer an important question or solve an important problem. The second part of the requirement, termed *DI-2*, is fulfilled by completing a more advanced project-oriented and/or hands-on course in which the students delve more deeply in one or more of the data intensive learning outcomes listed below that is appropriate and relevant to their major discipline.

DI-1 A DI-1 course should have one or more modules or units (at least 4 – 8 contact hours) incorporated into an existing or new introductory course (in any field) that address the use of data in that field – how they are collected, managed, manipulated, analyzed, interpreted, and used to answer questions. For example, a DI-1 course instructor might pose a question or problem whose answer or solution might be provided by collecting and analyzing data. The instructor could describe the data collection process, explain the analytic tool(s), and show how the analysis provides an answer or solution. The instructor might also discuss ethical issues that may be associated with data collection, analysis, reporting the results, misusing the data, and/or

archiving the data. Privacy and security issues could also be discussed. Finally, the instructor could illustrate techniques typically used to organize, analyze, or visualize the data (and results) if that is appropriate. This should lead to a final interpretation of the results and how they answer the original question or lead to a solution of an important problem. All schools should plan to offer multiple DI-1 courses to ensure that every student has the opportunity to take at least one.

DI-2 Each program will define a course or menu of courses (existing or new) within its disciplinary curriculum that develops data dexterity in a more focused discipline-specific manner. DI-2 courses are intended to be those in which students acquire the skills, knowledge, and/or understanding that is relevant to the role of data in that field. The data-focused content should be larger in scope and more in-depth than that covered in the DI-1 courses, and should involve projects, research, and/or hands-on activities. For example, a DI-2 course instructor might pose a problem and ask the students to think about what data are needed, how it can be analyzed and interpreted, whether there are any ethical issues to consider, privacy and security issues that might arise, and how best to visualize and communicate what the data are saying. The students might actually collect data, or select, curate, and validate data from existing datasets, analyze it, and work to develop an answer or solution to an original question or problem.

Students who successfully complete data intensive courses will be able to meet one or more of the following learning outcomes (instructors may also propose other suitable DI course specific learning outcomes):

- Identify different types of data, information and evidence within the relevant discipline, and be able to discuss issues of data quality, curation, validation, and uncertainty.

- Identify appropriate problems to which data can be applied, and discuss limitations, biases, assumptions and interpretations.

- Determine appropriate analytical tools and effectively use them with relevant data types to formulate, analyze, interpret, and/or solve real-world problems.

- Effectively communicate about problems/issues in this field in which data is a relevant tool, including writing about, presenting on, and visualizing data.

- Discuss the ethical issues surrounding data in this field, including, but not limited to, responsible conduct of research, privacy, provenance, privatization, monetization, and social implications.

Faculty who want to have a course designated data intensive should submit the Submission Form for Data Intensive (DI) Courses and the course syllabus as an attachment to the DI Review Committee at: DI_Review@rpi.edu. All proposed courses will be reviewed by a committee that is appointed by the Faculty Senate Curriculum Committee comprising the Associate Deans from all five schools. This committee will:

Determine whether syllabi for proposed courses meet the goals of the Data-Intensive requirements.

- Notify the Registrar's Office of approvals and course status changes;

- Prepare and deliver a status report to the FSCC annually;

- Review all DI courses at three-year intervals.

As of January 2019, the DI review committee consists of the following faculty:

- School of Architecture: David Bell

- School of Engineering: Kurt Anderson, Matt Oehlschlaeger

- School of Humanities and Social Sciences: Brett Fajen

- School of Management: Chanaka Edirisinghe

- School of Science: Lee Ligon

- Ex-officio: Kristin Bennett

Submission Form for Data Intensive (DI) Courses

The requirements for data intensive courses are listed below. In the space following each requirement, **please include information from your course syllabus that shows how your course meets the requirement. Please also attach a copy of the full syllabus to your application.**

Course Subject/Number   **Course Title**

Instructor   **Email**

1. What DI level is this course proposed to meet?
- ☐ DI-1
- ☐ DI-2

2. Data Intensive Learning Outcomes
Students who successfully complete data intensive courses will be able to meet one or more of the following learning outcomes. Check the learning outcomes that will be met by this course.

- ☐ 1. Identify different types of data, information and evidence within the relevant discipline, and be able to discuss issues of data quality, curation, validation, and uncertainty.
- ☐ 2. Identify appropriate problems to which data can be applied, and discuss limitations, biases, assumptions and interpretations.
- ☐ 3. Determine appropriate analytical tools and effectively use them with diverse data types to formulate, analyze, interpret, and/or solve real-world problems.
- ☐ 4. Effectively communicate about problems/issues in this field in which data is a relevant tool, including writing about, presenting on, and visualizing data.
- ☐ 5. Discuss the ethical issues surrounding data in this field, including, but not limited to, responsible conduct of research, privacy, provenance, privatization, monetization, and social implications.
- ☐ 6. Instructor proposed DI learning outcome._____

Describe below how this course will meet the learning outcomes checked above. For DI-1 courses, include the estimated number of contact hours for the DI unit. For DI-2, explain how the course prepares students to be competent in the use of data in their discipline. If only a portion of the course will focus on DI material, include a justification explaining why that is sufficient to achieve the goal of data competence in the discipline.

3. Data Intensive Assignments and Assessments
How will you assess whether students have met the learning outcomes listed above? What assignments/examinations will be used to test students?