ILS 103: Introduction to Data Lifecycle Management
CRN 20215 | updated 8/16/2024

Fall 2024 Semester - Purdue University
Monday 12:30-1:20 p.m. WALC 1087

Professor [Michael Witt](#)
Cell phone: 765-494-8703
Email: mwitt@purdue.edu
Office hours: Mondays 10:30-11:30 a.m.
WALC 3053T or Discord or other times by request

Course Description

The Habits of Mind (Costa & Kallick, 2008) are an identified set of problem-solving, life-related patterns of behavior that are necessary to effectively operate in society and promote strategic reasoning, insightfulness, perseverance, creativity, and craftsmanship. Introduction to Data Lifecycle Management is an introductory, one-credit, seminar course that exposes students to an additional six "Habits of Mind for Data" that are necessary for engineers to navigate in the world of data. Students will be introduced to concepts and begin to develop their own habits for 1) being aware of types of data and their lifecycles; 2) finding and evaluating data purposefully; 3) creating and sharing data for reuse, accountability, and enhancement; 4) making decisions and communicating using data, including basic data analysis and visualization; 4) protecting and archiving data; and 6) using data ethically and responsibly.

Students will engage these topics through readings, code examples, videos, discussion, lectures, guest speakers, and hands-on activities in class. As a final project, students will demonstrate application of the material by planning, finding and reusing, collecting, analyzing, presenting, and preparing data to be archived.

Enrollment in this section is limited to students accepted into and placed with the [Engineering in the World of Data](#) Learning Community. This one credit-hour course is a prerequisite for ENGR 103 Computational Methods of Data Science for Engineers in the spring semester.

Learning Resources, Technology & Texts

- No textbook is required.
- All required reading / viewing / listening will be provided via Brightspace.
- We will be using [Google Colab](#) for code examples.

Learning Outcomes

At the end of this course, students will be able to:

- describe the data lifecycle and types of data that are used in their discipline or profession,
- locate and evaluate datasets for reuse that are relevant to their research or work,
- apply a variety of introductory data collection, analysis, and visualization techniques,
- discuss the ethical and responsible use of data,and
- understand the basics of securing, archiving, and sharing data.

Course Structure

The six Habits of Mind for Data will be explored in six modules. Students will <u>prepare</u> for class by reading or viewing an article, video, or other content that is assigned in Brightspace. They will <u>reflect</u> on the content by responding to an instructor's prompt in an online discussion board and commenting on a classmate's response for that week's module in Brightspace. The class will meet for a lecture or guest speaker who will <u>concentrate</u> on the module and relate their experience and expertise to it. The class will <u>recap</u> the topic by revisiting main points, answering questions, and relating the module to the final project. Students will break into teams to <u>engage</u> the module through a hands-on activity in class and take a short quiz to <u>assess</u> their knowledge learned from the module.

Each of the 6 modules typically follows this pattern:

- *Prepare*: Assigned reading, code example, video, or other content to read
- *Reflect*: Post a response to instructor and comment to another student in Brightspace online discussions
- *Concentrate*: Attend and participate in lecture in class
- *Recap*: Revisit main points in class and how they relate to data science
- *Engage*: Hands-on activity in class
- *Assess*: Quiz

Depending on the calendar, one or two modules may span multiple weeks. While we will not be learning how to program in this class, examples of code will be shared to begin to familiarize students with python. Students should bring a laptop computer to class and be prepared to work in small teams, in person, and online. The primary modality for the class is face-to-face.

Engineering in the World of Data Learning Community

All students enrolled in this section of ILS 103 belong to the Engineering in the World of Data Learning Community that will offer a variety of social and educational activities throughout the year. The learning community strives to increase students' awareness of data-related behaviors and ethics in the fields of engineering through combined assignments, learning and social activities, and other shared experiences.

Class Schedule

This is a *tentative* schedule for the semester that is subject to change, for example, to accommodate guest speakers or to address relevant, emerging issues and current events. The current schedule will be updated and maintained in the calendar in Brightspace.

Dates	Topics
August 19	**Syllabus review / learning community introduction** **Module 1: Be aware of data and its lifecycle**
August 26	**Module 1: Be aware of data and its lifecycle** Quiz Module #1
September 2	**Labor Day - No class**
September 9	**Python with Pythons**
September 16	**Module 2: Find and evaluate data purposefully**
September 23	**Module 2: Find and evaluate data purposefully** Quiz Module #2 **Module 3: Create and share data for reuse and accountability**
September 30	**Module 3: Create and share data for reuse and accountability** Quiz Module #3
October 7	**Fall break - No class**
October 14	**Module 4: Make decisions & communicate using data**
October 21	**Module 4: Make decisions & communicate using data**
October 28	**Module 5: Mitigate risk and archive data** Quiz Module #5 Final project assigned
November 4	**Module 6: Use data responsibly and ethically**
November 11	**Guest: Dr. Mohamed Yakout, Senior Product Manager, Product Knowledge Team, Amazon.com**

November 18	**Module 6: Use data responsibly and ethically** **Guest: Ad Liberation**
November 25	**Final project work session** (optional but encouraged)
December 2	**Year-end class party / wrap-up** Course evaluations

Assignments and Grading

Grades are based on 100 points available through participation (in class and online discussions), attendance of learning community co-curricular/extracurricular activities, quizzes, and the final project. Late assignments will not be accepted without prior approval of an instructor.

Quizzes	6 quizzes (5 points each)	30
Participation	Discussion Board reflections and in-class participation (4 points each)	24
Learning community	Attend 2 Learning Community events (8 points each)	16
Final project	Data management and curation project	30
Total		100

Quizzes are given in class unless otherwise instructed. There will be one 5-point quiz per module for a total of 30 points.

For each of the six modules, 4 participation points are available for a total of 24 points. For each module, students will read materials, write an answer to an instructor prompt in the online discussion boards in Brightspace (2 points), and respond to one other student's answer (1 point). An additional point is earned by class attendance (online or in person), active listening, and participation. To summarize, for each module:

- Class attendance, listening, and participation: 1 point
- Complete answer to the instructor's online discussion prompt: 2 points
- <u>Meaningful</u> response to another student's post: 1 point

The rubric and assignment for the final project (30 points total) will be distributed in class. The following grading scale will be used:

Grade	Value
A	94-100 points
A-	90-93 point
B+	87-89 points
B	84-86 points
B-	80-83 points
C+	77-79 points
C	74-76 points
C-	70-73 points
D	60-69 points
F	Below 60 points

Communication

Brightspace will be the primary course website; all assignments, discussion board posts and quizzes will be administered and turned in using Brightspace. Instructors may be contacted via Brightspace or by the email addresses listed above. They are available online during posted office hours or by appointment to meet online or in person at other times.

You are expected to read and respond to your @purdue.edu email on a frequent basis.

A Discord server has been established for the Engineering in the World of Data Learning Community that can be used for informal communication among faculty and students, e.g., information about related seminars and events, social activities, current news and articles related to data, forming study groups, etc. Discord does not act as an official course website; no coursework should be submitted via Discord, and no course-related or assignment related posts will be posted in Discord that are not also posted in Brightspace.

Professional Expectations

Everyone in our learning environment helps shape the environment to be positive and productive for everyone. Behaving professionally includes arriving for class on time and being prepared; focusing during class on ILS 103; controlling your behavior to minimize distractions to those around you; and engaging with others in a respectful and professional manner. Instructors can deduct points from your semester total for behavior that is disruptive to your class or to your team's dynamics and performance.

Attendance

Students are expected to be present for every meeting of the classes in which they are enrolled. Only an instructor can excuse a student from a course requirement or responsibility. If you are not able to attend a class during the regularly scheduled time due to illness or quarantine, you are required to contact the instructors for an accommodation to replace your participation points with an additional, short writing assignment or to meet and discuss the module with the instructors during office hours. There will be no penalty for students who are following Protect Purdue and COVID-19 protocols.

Academic Integrity and Artificial Intelligence (AI)

Students are expected to follow the [Purdue Honor Pledge](#) and University policies related to plagiarism and academic honesty. AI is a tool, much like a shovel or a calculator, and like any tool, it can be used properly or improperly. The proper uses of AI assist (but do not replace) the process of learning, original thinking and work. We will be considerate of what data has been used to train generative models that may belong to others, be cautious of inaccuracies that may occur from AI-assisted results, and be sure to clearly attribute what work has been done using AI in assignments. Undisclosed use of AI on assignments and the use of AI on quizzes and tests will be treated as academic dishonesty.

Important University Policies and Resources

Students are encouraged and expected to familiarize themselves with information and follow policies from the university that are linked in Brightspace relating to Academic Integrity, Nondiscrimination, Accessibility, Mental Health and Wellness, Basic Needs Security, and Emergency Preparedness.