syllabus for COMP 198 *"***AI for Everyone***"*

Instructor:	Mark LeBlanc (**mleblanc**@wheatoncollege.edu)	**Open Hours** to meet in my office: by appointment
Office:	DC-1322	*or* **TW 10-11** *or* **TW 2 – 3:30**
Phone:	286-3970 (on campus: x3970)	Meeting: **TR** 12:30-1:50 pm, **DC 1315 ("csLab")**

Texts (*selected online sites and free tools*) …

Content:

This introductory AI course, open to all, covers basic data carpentry reasoning skills and the workflows that drive our algorithmic, data-centric world. Students will explore AI's benefits, ethical challenges, and societal impacts: Just how do we fight the instinct to trust a "human-sounding" machine? Particular emphasis will be placed on converting and mapping data from its raw, messy form into a clean, structured format for analysis and storytelling, as well as the foundations of data ethics (e.g., being able to articulate how AI is an extractive industry, understanding sampling and representation bias, and ensuring the reproducibility of results).

AI is everywhere ... well, not *everywhere* ... but you know, like "everywhere" … being used by "everyone". This course focuses your attention on AI: how did we get here? What are today's most modern tools and models? What are the skills needed for AI work, and what is the probability of doom that AI and bots bring to our world? Humans are and will be responsible for the types of bots we build and how we interact with them ... (*them*? should I use pronouns? Aren't those for humans?)

The short history and progress of Artificial Intelligence (AI) have featured more hype than reality, yet we now find ourselves faced with systems and devices that exceed many of our expectations. New tools and careers are emerging at a rapid rate and will require some degree of computational skills, along with the ability to think critically about how Machine Learning (ML) can and should be used.

This course bridges the gap between curiosity and practical understanding, offering a hands-on exploration of AI. We will work as a small group to increase levels of competence in data manipulation, computational and visualization tools, and the algorithmic workflows in ML applications. Conversations center on crucial AI concerns: ethics, algorithm biases, distinguishing between intrusive and helpful AI applications, considering AI as an extractive industry, and asking who is paid for the data used to build models?

By course-end, students develop a foundational grasp of AI concepts and practical skills in data management, analysis, effective communication, especially in writing, and critical evaluation of information. Equipped with these skills, you'll navigate the ethical and practical aspects of AI, recognizing its potential for both risk and positive change. Empowered with essential computing skills and the ability to think critically, students craft a final project reflecting their personal interests.

Student Learning Outcomes

Upon completion of this course, you should meet the following learning objectives:

Technical Writing and Speech

- Increase proficiency when communicating about computing topics in writing and in conversation
- Understand the importance of reducing ambiguity by applying professional notation in multiple areas
- Learn writing tips and techniques when drawing, annotating, and presenting diagrams of computational systems and graphs representing results

Collaboration

- Professionally participate in "meetings" (class), practice note-taking, and use of the whiteboard
- Take initiative with leadership opportunities and practice teamwork
- Engage with peers in discussion on the ethical issues surrounding the use of LLMs
- "Control the room" when presenting orally to others

Five (5) Data Reasoning Skills

1. **Managing your Digital Workspace**: understanding local *vs.* cloud-based storage, knowing and describing where your data is stored (file paths) for access, reproducibility, privacy, and intellectual independence.

2. **Data Wrangling and Tidying**: Before you can analyze, visualize, or tell a story about data, one often needs to manipulate data to organize it in ways that are systematic and standardized to make it useful in future workflows. This requires the skills of "cleaning" the raw material: converting and mapping data from its raw, messy form into a clean, structured format for analysis and knowing how to determine if data sets are reliable.

3. **AI Literacy**: A paradigm shift has occurred in the way we not only access information, but in the ways we read, write, code, and share information. Teaching and learning will never be the same, and thus, wise practices require increasing levels of expertise to guide us as we embrace the potential, understand the limits, and acknowledge the costs of AI.

4. **Foundational Data Ethics**: Foundational Data Ethics includes a set of practical guidelines to ensure data work is accurate, secure, and fair. This is no longer an abstract philosophical debate; an ethical grounding to AI facilitates a practical "measure twice, cut once" grounding in data work to avoid disastrous, real-world mistakes. It's about professional responsibility, not just theory

5. **Data Visualization and Interpretation**: Before you can analyze, visualize, or tell a story about data, one often needs to manipulate data to organize it in ways that are systematic and standardized to make it useful in future workflows. This requires the skills of "cleaning" the raw material and converting and mapping data from its raw, messy form into a clean, structured format for analysis.

CORE SKILLS for each Data Carpentry area

1. **Managing your Digital Workspace Core Skills:**

   - **Understanding where your files are stored** (on your laptop, in the cloud). Understanding file path notation in order to communicate that to others.

   - **Employing good organizational and naming conventions** for your data (e.g., how organized is your Google Drive?).

   - **Fluency with file sizes**; Understanding digital file sizes in terms of what range of sizes is appropriate for the file type.

   - Knowing of **metadata** (e.g., timestamps, author information, and watermarks).

   - Demonstrating **problem solving proficiency** in workflows requiring the use of data across multiple apps (e.g., Google Sheets to Docs to Slides).

2. **Data Wrangling and Tidying Core Skills:**

   - Opening, formatting, and analyzing data in a **spreadsheet** (e.g., Google Sheets), including the use of formulas, conditional formatting, *etc.* **Percentage calculations and basic aggregations (sum, average, count).**

   - **Recognizing "untidy" data:** Knowing how to use your spreadsheet to spot issues like a single cell containing multiple values (e.g., "New York, NY"), inconsistent naming ("USA," "U.S.A.", "United States"), mixed data types in a column (numbers and text), or detecting duplicate lines.

   - **Handling missing values:** Knowing what to do when data cells are blank.

   - **Data validation**: Checking that data makes sense (e.g., ages can't be negative, dates should be in valid ranges); noting some areas of potential bias in data samples; anonymizing data.

   - Identifying and **Automating Repetitive Tasks** such as *recognizing a pattern* that can be automated.

   - **Structuring data for use in other tools:** Understanding the basic principle that each row should be an observation and each column is typically a variable or feature of the data. To work properly, this is essential for almost every data tool (e.g., Google Sheets), as well as the workflows required for machine learning.

3. **AI Literacy Core Skills:**

   - Knowing the current landscape of **leading language and image models**

   - Understanding the **limitations and costs of AI**, including the possibility of hallucinations, biased data, and viewing AI as an extractive industry

   - Facilitating effective human-computer interaction: crafting productive prompts for AI tools

   - Experiencing AI teaching and conversational agents, e.g., practicing for an interview or pitch with a bot

   - Understanding the tenets of machine learning, including model-building stages of training and testing

   - Debating when to use AI *vs.* when human judgment is essential

4. Foundational Data Ethics Core Skills:

- **Understanding Sampling and Representation Bias:** Practically, this means asking critically: "Does my data actually represent the group I'm talking about?" A model built on biased data is a faulty tool that will fail in the real world, like a hiring algorithm trained predominantly on resumes from men or training a facial recognition system with an inadequate representation of people of color.

- **Ensuring Reproducibility:** The ability to document your steps (in a simple text file or document) so that someone else (or your future self) can follow your process and get the exact same result. This builds trust and makes work verifiable.

5. Data Visualization and Interpretation Core Skills:

- **Create and interpret** a simple bar chart, a line chart, and a histogram in a tool like Google Sheets or Excel.

- **Choosing the Appropriate Visualization.** W*hy* is a bar chart used for comparisons, a line chart for trends over time, and a scatter plot to show relationships?

- **Writing a 1-Sentence Assertion.** Can the student look at a simple chart and write a single, declarative sentence that summarizes the main point? This is the *real* "data storytelling" skill: interpretation, not just formatting. This language is useful because it connects to discussions about bias.

- Knowing good conventions for **formatting Tables and Figures**, including rich annotations and captions.

About your Professor:

Why am I teaching this course? I'm an organized geek who loves computing, and I love to teach and motivate college students to leverage computing, no matter their interests.

How do I define successful learning? A willingness to study, practice, and demonstrate *your* competencies while sharing what you've learned in a spirit of collaboration.

What you can expect from me? I will be organized and relentlessly striving to help you map your new computing skills onto your passions. I strive to make requirements transparent and will grade accordingly. I am very open to any discussion or concern about a previous grade.

Goals & Sequence (note: our Canvas site will be modified often and regularly, as needed)

Week	Topics	Data Reasoning Competencies and Badges Earned
Weeks **1-4**:	Large Language Models (LLMs): Gaining expertise with chatGPT, Gemini, Copilot, Claude, *et al.* – Levels of AI, History of AI, AI as an "extractive industry"	AI Literacy Managing your Digital Workspace
Weeks **5-7**:	Data Scraping and Data Cleaning: investigating the past and current use and ethics of "scraping", working with spreadsheets and data in `.csv` files, detecting bias in data	Data Wrangling and Tidying Data Visualization and Interpretation
Week **8**:	Spring Break	
Week **9**:	Intro to Machine Learning: training and testing Applied Machine Learning: Cluster Analysis of written texts	AI Literacy Foundational Data Ethics
Week **10**:	More Machine Learning: classification	
Weeks **11-13**:	Open Work on Final Project Design and Implementation	
Weeks **14-15**:	Final Project Presentations	

Expectations

1. **Honor Code Revisited**: It goes without saying that all submitted work will be the student's own, in keeping with the Wheaton Honor Code. For homework, all work must be your own from beginning to end. See below for our AI policy.

2. **Many homework solutions** must be converted to .pdf format and uploaded to our course Canvas site. Printed, stapled hardcopies are also required in some assignments. (staples are not optional :)

3. **Class attendance and participation is required**. In class, you will often work with others to solve problems and practice the terminology of AI (in small groups and/or at the whiteboard). You should consider class meetings like a "briefing" at a real job, thus you must be involved, participate, question, and engage with the material during class time.

4. **You will earn micro-credentials or "badges"** by passing assessments for the Data Reasoning skills. You will contribute to our collective construction of the graphic designs for badge icons as well as contributing to the construction of assessment instruments, e.g., by designing Google Form questions that verify if someone is "AI literate". As noted above, you'll need to **be in class and take notes** to develop good questions, as well as learn the answers as you go for the test :)

5. You "earn" your grade in this course **by showing mastery of the learning objectives/skills** as they appear within and are applied to current topics in computing. In addition to numerous opportunities to "play" in class, your level of mastery is directly dependent on your time on task outside of class.

6. The analogies of an athlete or musician are appropriate. You know when you watch a game or listen to a concert, much practice has occurred prior to the game or performance. In fact, most performers would *never* attempt to perform without hours and hours of practice. Is practice glorious? Does anyone cheer? Nope. Yet, returning to the topic of your own work, **it is sometimes the case that a student thinks that one or two attempts at a problem are sufficient for submission; why is that?** *Think about it.* The spectators in the audience *know* when they see a poorly prepared performer. Your professor (*a.k.a.* your future boss) is in the audience. Be disciplined. Avoid ambiguity. Impress your professor.

Course AI Policy

This course will follow Wheaton's "Unrestricted Use of AI with Acknowledgment" policy.

In this course, I encourage you to use all the tools at your disposal. However, as with any other resource you use to aid your work... *you* must acknowledge the AI tools that you use in the development of your work. We will also agree to **include citations and acknowledgments of AI use** in our reports.

And to be fair, from my faculty point of view, I will include a clear statement in any class materials or assessments that involve AI use.

Assessments & Grades

This course has been designed in compliance with [Wheaton College's credit hour policy](#).

Any and all assigned work will be due at the beginning of class on the assigned date. Thus, if a homework is due on a Thursday, you must bring your professionally formatted and printed/stapled hardcopy to class that day. Submissions will be collected at the beginning of the class, prior to beginning that day's topic.

Class **attendance** and **participation** is required.

Writing: You will be asked to write up (document) your experimental workflows, summaries, and examples of using tools, including how to format your work and apply professional notation and writing tips and techniques when drawing and annotating diagrams of computational systems. Your submissions will resemble professional documents.

Computational lab skills will be assessed in class. For example, we might practice a new topic with spreadsheets on Tuesday. You will then be asked, in class on Thursday, to replicate that type of work. We will view these in-class, hands-on sessions as your demonstration that you are proficient. While you will finish most "labs" during our class time together, our emphasis will be on demonstrating the proficiency of the skill, not the speed of a skill. Thus, you will be allowed to repeat and/or complete unmet proficiencies. The overall goal will be for you to be exposed to and show a level of mastery of a set of computational skills. Grading rubrics will vary depending on the type of work, but you can be assured that your Prof will always be open to your questions on how a certain in-class "lab" might be graded.

Final Projects: The grading rubric will include points for level of difficulty and originality.

Your grade:

In-class participation	10%	attendance and participation required each day
Technical Writing and Speech	45%	continual throughout the semester (1) close reading of "AI as an Extractive Industry" (2) spreadsheets (3) machine learning experiment
Data Reasoning Badges	25%	Q/A contributions to certification tests
Lab proficiencies	5%	continual throughout the semester
Final Project	15%	Weeks 11-13: work++ … preparation, rehearsals Weeks 14-15: presentation

Accessibility Statement

Wheaton College is committed to providing equitable access and supportive services for all students to fully access and thrive in the academic, residential and social aspects of student life at Wheaton College. Affirmatively, Wheaton provides appropriate accommodations for eligible students with documented disabilities to afford equal access to educational programs and services. Individuals with disabilities and other access concerns requiring accommodations or information on accessibility should reach out to Accessibility Services at the Filene Center, either via email at accessibility@wheatoncollege.edu or via phone at (508) 286-3794.

Academic Honesty Statement

Established by students for the purpose of self-governance in 1921, The Honor Code is a commitment to the ideals of academic excellence and individual responsibility: *As members of the Wheaton community, we commit ourselves to act honestly, responsibly, and above all, with honor and integrity in all areas of campus life. We are accountable for all that we say and write. We are responsible for the academic integrity of our work. We pledge that we will not misrepresent our work nor give or receive unauthorized aid. We commit ourselves to behave in a manner that demonstrates concern for the personal dignity, rights and freedoms of all members of the community. We are respectful of college property and the property of others. We will not tolerate a lack of respect for these values.*

In short, you signed the Honor Code; thus, you've declared that you will not cheat. I believe you.