Spring 25: Human Centered Data Science
Course Number (CRN): 72773
Time: Monday, 1:00 - 3:50 pm
Instructor: Dong Wang (dwang24@illinois.edu)
Prerequisite: Comfortable with Python programming
Grading: Grading is mainly based on course projects and class presentations – There are no exams or quizzes.

Description: This course introduces the principles of human-centered data science (HCDS), a discipline that places people at the core of data-driven decision-making processes. In today's digital age, industries and public organizations harness vast amounts of data from social media and online platforms to tailor marketing strategies, predict consumer behaviors, improve public health, and more. In addition, academic institutions and researchers use various human centered data sources such as census data, in-depth interviews, and surveys to gain insights into social patterns, behavioral motivation, and cultural norms. However, while this abundance of data can uncover trends and preferences, overlooking human-centered considerations can lead to the development of biased algorithms and discriminatory practices. Notable examples include the use of racially biased facial recognition software and socioeconomic biases in crime prediction algorithms. Given this background, this course aims to adopt a *human-centered approach* in data science. Specifically, students will get an opportunity to learn practical skills in data science while deepening their theoretical understanding, particularly focusing on making informed decisions throughout the data science workflow—from data collection, data preparation and exploration, data analysis, to result communication. Additionally, this course will explore various human-centered case studies, demonstrating the application of these principles across different sectors such as healthcare, information systems, traffic management, disaster response, and finance.
Some **sample topics** that will be covered include:

- Reliable Data Collection and Filtering Techniques from Noisy Human Data Sources
- Enhancing Decision-Making through Human-LLM Collaboration
- Explainable AI for Transparency
- Promoting Equity through Fairness in Algorithm Design
- Privacy-Preserving Techniques in Data Science
- Time Series Analysis for Predictive Insights
- Advancing AI Systems through Multimodal Data Fusion
- Open-end topics on Human-centered Data for the Course Project