CS/INFO 5306 Course Readings

- Mon 8/22: Lecture 1: **Overview & Introduction**
- Wed 8/24: Lecture 2: **Overview & Introduction**
- Mon 8/29: Lecture 3: **Wikipedia**
  - "Beyond opening up the black box: Investigating the role of algorithmic systems in Wikipedian organizational culture" (2017)
  - "The Evolution and Consequences of Peer Producing Wikipedia's Rules" (2017).
  - Some suggestions on papers you can consult if you need some help on how to read research papers:
    - Michael Mitzenmacher's "How to read a research paper"
    - Srinivasan Keshav's "How to Read a Paper"
    - Jennifer Raff's "How to read and understand a scientific article"
- Wed 8/31: Lecture 4: **Other Examples of Collaborative Creation**
  - "Distribution of labor, productivity and innovation in collaborative science" (2022)
  - "The dynamics of collective social behavior in a crowd controlled game" (2019)
  - "Latent structure in collaboration: the case of Reddit R/place" (2018)
- Wed 9/7: Lecture 5: **Project 1**
  - Review the list of papers at the end of the Project 1 description document. Class will be focused on brainstorming about project ideas
- Mon 9/12: Lecture 6: **Open Source Software**
  - "The Shifting Sands of Motivation: Revisiting What Drives Contributors in Open Source" (2021)
- Wed 9/14: Lecture 7: **Social Media "Wisdom of Crowds"**
  - "Predicting the Future With Social Media" (2010)
  - "Widespread worry and the stock market" (2010)
  - "How (Not) To Predict Elections" (2011)
- Mon 9/19: Lecture 8: **Mining Search Query Data**
  - "Detecting influenza epidemics using search engine query data" (2009)
  - "The parable of Google Flu: traps in big data analysis" (2014)
- Wed 9/21: Lecture 9: **Other Examples of Mining Online User Behavior I**
  - "The scaling laws of human travel" (2006)
  - "Mining complaints for traffic-jam estimation: A social sensor application" (2015)
  - "Using Social Media to Detect and Locate Wildfires" (2016)
- Mon 9/26: Lecture 10: **Other Examples of Mining Online User Behavior II**
  - "Guess who's not coming to dinner? Evaluating online restaurant reservations for disease surveillance" (2014)
  - "Leveraging Mobility Flows from Location Technology Platforms" (2020)
  - "Mammon and the Archer" by O. Henry. In *The Four Million*, by O. Henry, 1906. **[Note: This is a story written in 1906 and you should expect it to be highly dated.]**
- Wed 9/28: Lecture 11: **Human computation: Introduction**
  - "AI gets a brain: New technology allows software to tap real human intelligence" (2006)
  - "Cheap and fast–but is it good? evaluating non-expert annotations for natural language tasks" (2008)
  - "Get another label? improving data quality and data mining using multiple, noisy labelers" (2008)
  - "Heads in the Cloud" (2010)
- Mon 10/3: Lecture 12: **Project 1 Summaries, Project 2 Introduction**

- o The first part of the class will be each group's 1-minute overview of what they did for Project 1.
  - o The second part of the class will be a discussion of Project 2 and brainstorming about possible project ideas.
- Wed 10/5: Lecture 13: **Microlabor, Pay, and Motivation**
  - o "Financial incentives and the "performance of crowds"" (2009)
- Wed 10/12: Lecture 14: **Cognitive Biases**
  - o Please look over the Wikipedia article on cognitive biases at https://en.wikipedia.org/wiki/Cognitive_bias
  - o Near the top the article has a figure, the Cognitive Bias Codex - which, strangely, isn't actually discussed in the article. This was an image created a few years ago by John Manoogian III and Buster Benson to give structure to what we know about cognitive biases. (The original is at https://www.teachthought.com/critical-thinking/cognitive-biases/.) Please review the image, too - it probably makes the most sense to go to it after reading the Wikipedia article.
  - o Drill down on at least two of the cognitive biases that are in the Wikipedia article and/or Codex to get a more grounded idea about cognitive biases.
- Mon 10/17: Lecture 15: **Pay Rates**
  - o "Worker Demographics and Earnings on Amazon Mechanical Turk: An Exploratory Analysis" (2019)
- Wed 10/19: Lecture 16: **Worker Experience I**
  - o "Being a Turker" (2014)
  - o "Turk-Life in India" (2014)
- Mon 10/24: Lecture 17: **Worker Experience II**
  - o "Web Workers, Unite! Addressing Challenges of Online Laborers" (2011)
  - o "Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk" (2013)
  - o "Taking a HIT: Designing around Rejection, Mistrust, Risk, and Workers' Experiences in Amazon Mechanical Turk" (2016)
- Wed 10/26: Lecture 18: **Worker Experience III**
  - o "The Communication Network Within the Crowd" (2016)
  - o "Becoming the Super Turker: Increasing Wages via a Strategy from High Earning Workers" (2020)
  - o "Goal-Setting Behavior of Workers on Crowdsourcing Platforms: An Exploratory Study on MTurk and Prolific" (2022)
- Mon 10/31: Lecture 19: **Data Labeling I**
  - o "Iterative Quality Control Strategies for Expert Medical Image Labeling" (2021)
  - o "Comparing Experts and Novices for AI Data Work: Insights on Allocating Human Intelligence to Design a Conversational Agent" (2022)
- Wed 11/2: Lecture 20: **Data Labeling II**
  - o "Eliciting and Learning with Soft Labels from Every Annotator" (2022)
  - o "Enhancing Image Classification Capabilities of Crowdsourcing-Based Methods through Expanded Input Elicitation" (2021)
- Mon 11/7: Lecture 21: **Data Labeling III**
  - o "A Checklist to Combat Cognitive Biases in Crowdsourcing" (2021)
  - o "More Data Can Lead Us Astray: Active Data Acquisition in the Presence of Label Bias" (2022)
- Wed 11/9: Lecture 22: **Privacy/Safety/Manipulation**
  - o "Information Extraction and Manipulation Threats in Crowd-Powered Systems" (2014)

- o "Pterodactyl: Two-Step Redaction of Images for Robust Face Deidentification" (2021)
- o "Fast, Accurate, and Healthier: Interactive Blurring Helps Moderators Reduce Exposure to Harmful Content" (2020)
- Mon 11/14: Lecture 23: **Programming Human Computation**
  - o "TurKit: Human Computation Algorithms on Mechanical Turk" (2010)
  - o "CrowdScape: Interactively Visualizing User Behavior and Output" (2012)
  - o "TaskLint: Automated Detection of Ambiguities in Task Instructions" (2022)
- Wed 11/16: Lecture 24: **Citizen Science**
  - o "From Conservation to Crowdsourcing: A Typology of Citizen Science" (2011)
  - o "Collective self-experimentation in patient-led research: How online health communities foster innovation" (2019)
  - o "Monitoring the world's bird populations with community science data" (2020)
  - o "Twelve years of Galaxy Zoo" (2020)
- Mon 11/21: Lecture 25: **Games with a Purpose**
  - o "Anda's Game" by Cory Doctorow (Warning: This is a work of fiction that uses what some might view as coarse language.)
  - o "Designing Games with a Purpose" (2008)
  - o "Repurposing Citizen Science Games as Software Tools for Professional Scientists" (2018)
- Wed 11/23: **Thanksgiving break**
  - o (Optional, just for fun - this is a 1972 science fiction story about teleportation and flash mobs) "Flash Crowd" by Larry Niven
- Mon 11/28: Lecture 26: **Prediction Markets**
  - o "Prediction Markets in Theory and Practice" (2006)
  - o "Market Design, Manipulation, and Accuracy in Political Prediction Markets: Lessons from the Iowa Electronic Markets" (2014)
  - o "Are markets more accurate than polls? The surprising informational value of "just asking" " (2019)
  - o (Optional: Documents DARPA's failed attempt to establish a prediction market for predicting political developments in the Middle East - Senators testified on the Senate floor calling it "morally repugnant and grotesque", making front page news in the New York Times) "The Policy Analysis Market: A Thwarted Experiment in the Use of Prediction Markets for Public Policy" (2007)
- Wed 11/30: Lecture 27: **Contests**
  - o "Test-Driving the Future: How Design Competitions Are Changing Innovation" (2014) **(just pages 14-24)**
  - o "All Together Now: A Perspective on the Netflix Prize" (2010)
  - o "Reflecting on the DARPA Red Balloon Challenge" (2011)
  - o "Crowdsourcing Data Science: A Qualitative Analysis of Organizations' Usage of Kaggle Competitions" (2020)
- Mon 12/5: Lecture 28: **Crowdfunding**
  - o "Financing by and for the Masses: An Introduction to the Special Issue on Crowdfunding" (2016)
  - o Pick two of these:
    - ▪ "Speaking the same language: the power of words in crowdfunding success and failure" (2021)
    - ▪ "GoFundMe as a Medical Plan: Ecological Study of Crowdfunding Insulin" (2022)
    - ▪ "Crowdfunding a monthly income: an analysis of the membership platform Patreon" (2021)