MACS 30000: Perspectives on Computational Analysis

Autumn 2024

Fabricio Vasselai
vasselai@uchicago.edu
Office: Room 213-A (1155 E. 60th St.)
Office Hours: by appointment.
You can sign up through: this appointment page. *Please email me if you can't find a good slot.*

Course Overview

Massive digital traces of human behavior and ubiquitous computation have both extended and altered classical social science inquiry. This course surveys successful social science applications of computational approaches to the representation of complex data, information visualization, and model construction and estimation . We will reexamine the scientific method in the social sciences in context of both theory development and testing, exploring how computation and digital data enables new answers to classic investigations, the posing of novel questions, and new ethical challenges and opportunities. Students will review fundamental research designs such as observational studies and experiments, statistical summaries, visualization of data, and how computational opportunities can enhance them. The focus of the course is on exploring the wide range of contemporary approaches to computational social science.

Learning objectives

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

- Explain how digital data and computational approaches complement and expand upon traditional research designs in the social sciences
- Analyze the strengths and weaknesses of competing computationally-enhanced research designs for a question of interest, as well as key limitations of current computational approaches
- Evaluate the ethical implications of digital research in the social sciences
- Produce their own computationally-enhanced research design to answer a social scientific question of interest

Course Structure

Class sessions will consist of a mix of lecture and discussion, giving you the opportunity to learn about computational research design strategies and critique them with your peers.

All course content and readings are organized on Canvas under "Files" on the left-hand side of your screen. You should read all of the readings listed for a given class session ***ahead of class time*** and be prepared for in-class discussion.

Our textbook for this course is (note that be clicking on the link you can read it online for free):

> Salganik, Matthew J., *Bit by Bit: Social Research in the Digital Age*, Princeton University Press, 2018.

In the course calendar below we cite *Bit by Bit* chapters that correspond to the lesson for each week.

Evaluation

Midterm Paper 1: Assessing computationally-enhanced research designs (20%)

Due Date: 11/2/2024, 11:59pm CT
Description: You will write a paper of around 1800-2200 words comparing and contrasting the usage of an observational study vs. a digital experiment to answer a research question. Note that your paper is to be divided into two parts. **Specific instructions for the assignment will be provided on the Canvas course site.**

Midterm Paper 2: Ethics in Computational Social Science (20%)

Due Date: 11/16/2024, 11:59pm CT
Description: You will write a paper of around 1500-2000 words addressing a specific issue related to CSS and ethics. **Specific instructions for the assignment will be provided on the Canvas course site.**

Final Paper: Research proposal (40%)

Due Date: 12/13/2024, 11:59pm CT
For the final assignment, you will write an original research proposal for a social scientific question of your own interest and develop a computationally-enhanced research design to answer the question.
Specific instructions for the assignment will be provided on the Canvas course site.
NB: You need to submit your CITI certificate along with your final paper.

Participation (20%)

A big part of this class involves in-class discussion of the assigned readings, which will drive you to engage deeper with specific computational social science research designs and allow you to practice taking part in scholarly discussions.

Plagiarism and Academic Honesty

In our first class meeting, we will discuss what constitutes plagiarism and how to avoid it. Academic honesty is an extremely important principle in academia and at the University of Chicago.

- Writing assignments must put in quotes and cite any excerpts taken from another work.
- If the cited work is the particular paper referenced in the Assignment, no works cited or references are necessary at the end of the composition.
- If the cited work is not the particular paper referenced in the Assignment, you MUST include a works cited or references section at the end of the composition.
- Any copying of others' work will result in a zero grade and potential further academic discipline.
- In this course, we will be developing skills and knowledge that are important to discover and practice on your own. Because use of AI tools inhibits development of these skills and knowledge, students are not allowed to use any AI tools, such as ChatGPT for the initial draft of any of the assignments in this course. Some students find it helpful to use AI tools, however, for polishing their writing and suggesting ways to improve their overall English writing fluency after writing a paper draft on their own. This form of AI assistance is permitted, provided that students submit both the original draft of the assignment, as well as the polished, AI-assisted draft for the assignment. Students should also submit a complete log of the prompts that were used to generate

the polished version of the draft and properly cite the use of the AI tool in their paper. If you are unclear if something is an AI tool or how to cite it, please check with your instructor. Using AI tools for any other purposes in this course will violate the University's academic integrity policy.

- When in doubt, please ask! It is far better to check with us prior to submitting an assignment than waiting

Submitting Late Assignments

All assignments should be submitted by the deadline listed on Canvas. The only exceptions are for religious observances, injury/illness, and similarly exceptional situations. If you are unable to submit an assignment on-time and you know in advance, you must contact your instructor prior to the assignment deadline to arrange an extension. If you fall ill or are unable to ask in advance, please contact your instructor as soon as possible. We are generally reasonable in these situations, but you need to let us know.

Revising and Resubmitting Midterm Papers

If you receive a grade below 90 you can resubmit within seven days for re-grading, and earn 50% of the remaining points (up to 90) deducted in your first attempt, so long as you made a good-faith effort to write the paper in your first submission.[1]Your revisions must be submitted within 7 days of grades being released for the assignment in question. With your revisions, you should include a cover letter detailing what you changed in your submission and why your revisions address all of the issues with your original paper.

Grading

Grades are not curved in this class or, at least, not in the traditional sense. We use a standard set of grade boundaries:

- 95-100: A
- 90-95: A-
- 85-90: B+
- 80-85: B
- 75-80: B-
- 70-75: C+
- <70: Dealt on a case-by-case basis

We curve only to the extent we might lower the boundaries for one or more letter grades, depending on the distribution of the raw scores across sections of this course. We will not raise the boundaries in response to the distribution.

So, for example, if you have a total score of 82 in the course, you are guaranteed to get, at least, a B (but may potentially get a higher grade if the boundary for a B+ is lowered).

[1] Example: you get 80 in your first attempt, and do everything right the second time, you get 80 + (90 - 80)/2 = 85

If you would like to be graded on a Pass/Fail (P/F) basis, send a message to your instructor before the Final Research Proposal is due. A total score of 75 and above in the class will qualify for a "P" in the class. Note that MACSS and CSS Certificate students must take this course for a letter grade.

Statement on diversity, inclusion, and disability

The University of Chicago is committed to diversity and rigorous inquiry from multiple perspectives. The MAPSS, CIR, and Computation programs share this commitment and seek to foster productive learning environments based upon inclusion, open communication, and mutual respect for a diverse range of identities, experiences, and positions.

The University of Chicago is committed to ensuring equitable access to our academic programs and services. Students with disabilities who have been approved for the use of academic accommodations by Student Disability Services (SDS) and need a reasonable accommodation(s) to participate fully in this course should follow the procedures established by SDS for using accommodations. Timely notifications are required in order to ensure that your accommodations can be implemented. Please meet with your instructor to discuss your access needs in this class after you have completed the SDS procedures for requesting accommodations.

- Email: disabilities@uchicago.edu
- Phone: 773-702-6000

Course Schedule

Note: *Schedule is subject to change. Check on Canvas for updates as the course progresses.*

Week	Date	Class and Readings
1	**Oct. 1st & 3rd**	**CLASS:** *What is computational social science research?* ● Introduction to: ○ Class Format ○ Computational Social Science ○ Social Science Research Questions, Constructs, and Operationalization ● How to read a social science paper.
		READINGS: ● **[handout]** "How to read a (Computational) Social Science Paper" ● Watts, D. J. (2007). A twenty-first century science. ● Lazer et. al. (2009) Computational Social Science. ● Evans, J. (2020). Social computing unhinged. ● *Bit by Bit* - chapter 1

		● Bhattacherjee, A. (2012). Social science research: Principles, methods, and practices. (Chapters 1-4; Extremely useful for students without a strong background in the social sciences) ● Chapter 2&4 - Doing honest work in college. ● Handout: How to read a (Computational) Social Science Paper
2	**Oct. 8th & 10th**	**CLASS:** *Observational Studies I* ● Digital Trace (Big) Data ● Big Data Research Strategies: Counting (and Measuring) Things **READINGS:** ● *Bit by Bit* - 2.1-2.4.1 ● Chan, H. F., et al. (2021). Can psychological traits explain mobility behavior during the COVID-19 pandemic? *Social Psychological and Personality Science,* 12(6), 1018–1029. ● Shi, F., Shi, Y., Dokshin, F. A., Evans, J. A., & Macy, M. W. (2017). Millions of online book co-purchases reveal partisan differences in the consumption of science. *Nature Human Behaviour*, 1(4), 0079.
3	**Oct. 15th & 17th**	**CLASS:** *Observational Studies II* ● Big Data Research Strategies: ○ Forecasting ○ Approximating Experiments **READINGS:** ● *Bit by Bit* – 2.4.2 ○ Salganik et al. (2020). Measuring the predictability of life outcomes with a scientific mass collaboration. PNAS, 117 (15), 8398-8403. ○ Shi, F. and Evans, J. (2023). Surprising combinations of research contents and contexts are related to impact and emerge with scientific outsiders from distant disciplines. Nature Communications, 14, 1641. ● *Bit by Bit* – 2.4.3 ○ Hersh, E. D. (2013). Long-term effect of September 11 on the political behavior of victims' families and neighbors. ○ Legewie, J. (2016). Racial profiling and use of force in police stops: How local events trigger periods of increased discrimination. **Optional** ● [Nice overview] Keele L. The Statistics of Causal Inference: A View from Political Methodology. Political Analysis. 2015;23(3):313-335. ● [applied prediction exercise] Hegre, Håvard, Paola Vesco, and Michael Colaresi. 2022. "Lessons from an Escalation Prediction Competition." International Interactions 48 (4): 521–54.
4	**Oct. 22th & 24th**	**CLASS:** *Asking Questions* ● Fundamentals of (Digital) Survey Design ● Digital Enhancements to Traditional Survey Design

		READINGS: ● *Bit by Bit* – 3.1-3.4 ○ Wang, Wei, David Rothschild, Sharad Goel, and Andrew Gelman, "Forecasting Elections with Non-Representative Polls," International Journal of Forecasting, 31:3 (2015) pp. 980-991. ● *Bit by Bit* – 3.5-3.7 ○ York Cornwell, E., & Cagney, K. A. (2017). Aging in activity space: results from smartphone-based GPS-tracking of urban seniors. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *72*(5), 864-875. ○ Blumenstock, J., Cadamuro, G., & On, R. (2015). Predicting poverty and wealth from mobile phone metadata. *Science*, *350*(6264), 1073-1076. ○ Kosinski, M., Stillwell, D., & Graepel, T. (2013). Private traits and attributes are predictable from digital records of human behavior. *Proceedings of the national academy of sciences*, *110*(15), 5802-5805.
5	**Oct. 29th & 31st**	**CLASS:** *Experiments* ● Digital Experimental Design ● Running Digital Experiments
		READINGS: ● *Bit by Bit* – chapter 4.1-4.4 ○ Bond, R. M., Fariss, C. J., Jones, J. J., Kramer, A. D., Marlow, C., Settle, J. E., & Fowler, J. H. (2012). A 61-million-person experiment in social influence and political mobilization. ● *Bit by Bit* – 4.5-4.7 ○ Van de Rijt, A., Kang, S. M., Restivo, M., & Patil, A. (2014). Field experiments of success-breeds-success dynamics. *Proceedings of the National Academy of Sciences*, *111*(19), 6934-6939. ○ Vijaykumar, S., Jin, Y., Rogerson, D., Lu, X., Sharma, S., Maughan, A., ... & Morris, D. (2021). How shades of truth and age affect responses to COVID-19 (Mis) information: randomized survey experiment among WhatsApp users in UK and Brazil. *Humanities and Social Sciences Communications*, *8*(1), 1-12.
NB: Item Due **Nov 2 11:59pm CT**		*Midterm Paper 1: Assessing computationally-enhanced research designs*
6	**Nov. 5th & 7th**	**CLASS:** *Ethics* ● Principles of Ethical Digital Research ● Evaluating the Ethics of Digital Research
		READINGS: ● *Bit by Bit* – chapter 6 ○ Navigating the IRB Process (Slides) ● Kramer, A. D., Guillory, J. E., & Hancock, J. T. (2014). Experimental evidence of massive-scale emotional contagion through social networks. ● Watts, D. J. (2014). Stop complaining about the Facebook study. It's a golden age for research.

		• Zimmer, M. (2010). "But the data is already public": on the ethics of research in Facebook. • Milkman, K. L., Akinola, M., & Chugh, D. (2012). Temporal distance and discrimination: An audit study in academia. Psychological science, 23(7), 710-717.
NB: Item Due **December 13th 11:59pm CT**		*CITI Training certificate (see canvas for more information)* *Note that this is not due now, but you will need it before submitting your last assignment.*
7	**Nov. 12th & 14th**	**CLASS:** *Simulations* • Growing Artificial Societies • "Explaining" Data via Simulation
		READINGS: **Agent-based modeling** • Epstein and Axtell 1996 *Growing Artificial Societies: Social Science from the Bottom Up (p. 1-53)* • Page, S. 2015. "What sociologists should know about complexity." Annual Review of Sociology 41: 21–41. • Dean et al. 2000 "Understanding Anasazi Culture Change through Agent-based Modeling" • Janssen 2009 "Understanding Artificial Anasazi" • Stonedahl and Wilensky 2010 "Evolutionary Robustness Checking in the Artificial Anasazi Model" **Generative AI:** • Bail, Christopher A. "Can Generative AI improve social science?." Proceedings of the National Academy of Sciences 121, no. 21 (2024). • Irving, Geoffrey, and Amanda Askell. "AI safety needs social scientists." Distill 4, no. 2 (2019). • Optional (for fun): An interactive simulation of neighborhood segregation (based on Schelling's famous 1971 study, mentioned in the Epstein and Axtell reading) • Optional: If you download/install NetLogo, you can play around with model parameters for yourself (AA and Sugarscape are a sample models included with the installation) • Optional (using the same approach in another domain): Axtell 2018 "Endogenous Firm Dynamics and Labor Flows via Heterogeneous Agents.
NB: Item Due **November 16th 11:59pm CT**		*Midterm Paper 2: Ethics in Computational Social Science*
8	**Nov. 19th**	**CLASS:** *Mass Collaboration + Writing your own Research Question*

	& 21st	**READINGS:** ● *Bit by Bit* – chapter 5 ● Bell, R. M., Koren, Y., & Volinsky, C. (2010). All together now: A perspective on the netflix prize. ● Optional: Salesses, P., Schechtner, K., & Hidalgo, C. A. (2013). The collaborative image of the city: mapping the inequality of urban perception. ● Booth et al (2016). The Craft of Research. Chapters 3&4 ● Peng & Matsui (2007) The Art of Data Science. Chapter 2: Epicycles of Analysis ● Peng & Matsui (2007) The Art of Data Science. Chapter 3: Stating and Refining the Question ● Sheppard (2020) Research Methods for the Social Sciences: An Introduction. Section 3.3
	**Nov 26th 28th**	**Thanksgiving Break**; No Class *(Suggested: Apple Cider Donuts)*
9	**Dec 3rd & 5th**	**CLASS:** *Designing your own CSS Research* ● From Research Question to Well-Researched Proposal ● Proposal Workshop
		READINGS: ● Booth et al (2016). The Craft of Research. Chapters 5&6 ● What are Citation Management Tools? ● Which Tool is Best for me?
NB: Item Due **December 13th 11:59pm CT**		*Final Paper: Research proposal*