Columbia University
School of International and Public Affairs

Artificial Intelligence Institutions

INAF 6548U, Spring 2025

Instructor: **Daniel Björkegren**
dan.bjork@columbia.edu

Meetings: **Thursdays 2.10-4pm** in IAB XXX

Office Hours: A variety of times are available, both virtual and in person. Sign up for a slot using the link on my website (http://dan.bjorkegren.com). If none of the listed times work for you, you may email about other potential meeting times.

New technologies afford—and may require—different ways of organizing society. We will consider historical episodes of technological change and our current era, following how shifts in technology can shift the economy and society. We will first use this course itself as a laboratory to explore the impacts of AI on education. We will then consider how AI may reshape other sectors, including governance and transportation; and the cross-cutting questions it raises about values, economic wellbeing, and purpose.

As the topics in this course are dynamic, this course is experimental.

Prerequisites are quantitative analysis II (SIPA U6501) and microeconomics (SIPA U6300 or U6400 or equivalent). Students will benefit from coding experience but software development is not a strict prerequisite.

Course Components

Class readings will consist of online writing, research papers, and book excerpts. We will ground our discussions based on first principles, and, where possible, data. Given the dynamic nature of these topics, readings are subject to change. Students are encouraged to suggest readings they find thought provoking.

Participation
Student discussion is the core of this course. You will be expected to have read any readings, to come to class with questions and observations of your own, and to actively engage with questions and comments. Students are expected to attend all classes unless they have an excused absence, which may include severe medical issues, family emergencies, or significant career-related activities.

Assignments
In the pre-ChatGPT era, I would have asked you to submit 4 memos and a term paper. But AI has expanded what you are capable of creating. So our first task in this class is a meta task: determining what work output would help you learn the most, and be most useful for your careers and impact on the world.

First assignment
Your first assignment, which will be due two days before our second class meeting, is to submit a half page proposal for the remainder of assignments of the course. In our second meeting, we will discuss selected proposals and the class will collectively decide on a single series of assignments that will contribute the most to your learning and demonstrate mastery. We'll use this discussion as a springboard to evaluate the impact of AI on how and what our societies teach.

I anticipate that these proposals will include several regular assignments, such as, for example, writing memos, critiquing memos or arguments, or producing demos to test hypotheses. There are placeholder due dates in the schedule below, though the number and timing of these will be determined during our discussion.

Although you are expected to have conceptual familiarity with microeconomics, statistics, and computation, technical analysis or complex math will be optional for assignments.

Final Project
In addition to regular assignments, you are expected to produce a final artifact. This may take several forms, such as a product demo, business plan, policy proposal, or research paper.

In the final week of the course each group will have the opportunity to present their artifact. These presentations will allow the rest of the class to learn about your work as well as provide ideas and feedback. A final written document will be due after reading period.

Grades are assembled from these components:

Assignments	50% total (each weighted equally)
Final project	30%
Class participation	20%

After the discussion in the second class, the grading and assignment plan will be finalized by the morning of January 31. If you are concerned about the uncertainty, note that that day is also SIPA's add/drop deadline.

Office Hours are a great learning opportunity. Please come to office hours with questions on the material covered in class, comments on the course, or if you want to talk about ideas. Please do not use office hours to talk about grades.

Policies

Artificial Intelligence
AI tools, including LLMs and interfaces like ChatGPT and GitHub Copilot, can improve productivity (Peng et al. 2023). However, AI tools make mistakes that are difficult to detect unless one understands the underlying principles, and using AI tools prematurely in this course can interfere with learning those principles (Bastani et al. 2024). For that reason, there may be assignments in this course for which AI assistance will not be allowed, which will be appropriately designated. If you are unsure if a certain tool is allowed, ask.

When I communicate with you, I value your time and select each word to write to you by hand. Please respect my time by not sending communication written with AI tools.

Regrades
Requests for reconsideration of grades are not encouraged, and will be accepted only in writing, with a clear statement of what has been misgraded, within one week of receiving the graded assignment. Please submit your full assignment so grading on all questions can be reconsidered.

Academic Integrity
The School of International & Public Affairs does not tolerate cheating or plagiarism in any form. Students who violate the Code of Academic & Professional Conduct will be subject to the Dean's Disciplinary Procedures. Familiarize yourself with proper citation methods using resources available online: http://bulletin.columbia.edu/sipa/academic-policies/. Report violations to the Associate Dean for Student Affairs.

Disability Statement
SIPA is committed to providing reasonable accommodations for students registered with Columbia University's Disability Services (DS). Students with DS-certified accommodation letters should discuss their accommodations with their instructor. Contact SIPA's DS liaison for additional questions: disability@sipa.columbia.edu or 212-854-8690.

Class Schedule

This schedule is subject to change based on your interests and advances in the space. Readings for each course are shown in indent.

Class 1. January 23.
Primer on Artificial Intelligence
What is artificial intelligence? How does it relate to statistics?

Explainers:
https://www.understandingai.org/p/large-language-models-explained-with
https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/

LLM Visualizer https://bbycroft.net/llm

Videos
3Blue1Brown neural network explainer
https://www.youtube.com/playlist?list=PLZHQObOWTQDNU6R1_67000Dx_ZCJB-3pi
Karpathy's video "Let's build GPT from scratch"
Neural networks from scratch https://nnfs.io

☐ **First Assignment Due** (by 11am January 28)
Proposal for assignments for this course, to spur class discussion

Class 2. January 30.
Education
What are the opportunities and risks of AI in education? What does that mean for this course, your career—and the world?

Meeks, F., Agrawal, A., Tana, E., 2024. *How to go from -1 to 0*. South Park Commons. https://blog.southparkcommons.com/how-to-go-from-minus-1-to-0/
- *Read specifically the contrast between 'demo' and 'memo':*

Excerpts from Sal Khan, *Brave New Words.*

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., Mariman, R., 2024. *Generative AI Can Harm Learning*. https://doi.org/10.2139/ssrn.4895486

Class 3. February 6.
Capabilities
What is AI capable of? How might its capabilities evolve?

Scaling

Kaplan, J., McCandlish, S., Henighan, T., Brown, T.B., Chess, B., Child, R., Gray, S., Radford, A., Wu, J., Amodei, D., 2020. *Scaling Laws for Neural Language Models*. https://doi.org/10.48550/arXiv.2001.08361

- *Focus on understanding Figures 1 and 2.*

Jack Clark's GPT 2 retrospective

Mitchell, M., Krakauer, D.C., 2023. *The Debate Over Understanding in AI's Large Language Models*. https://doi.org/10.48550/arXiv.2210.13966

Stanford HAI, 2024. *AI Index Report*. URL https://aiindex.stanford.edu/report/

Narayanan, A., Kapoor, S., 2024. *AI Snake Oil: What Artificial Intelligence Can Do, What It Can't, and How to Tell the Difference*. Princeton University Press.

No class. February 13.

☐ **Tentative Second Assignment Due** (by 8pm February 19)
To be announced

Class 4. February 20.
Precedents
Do technology shifts lead to institution shifts?

Acemoglu, D., Johnson, S., 2023. *Power and Progress: Our Thousand-Year Struggle Over Technology and Prosperity*, 1st edition. ed. PublicAffairs, New York.

Rosenberg, N., Trajtenberg, M., 2004. *A General-Purpose Technology at Work: The Corliss Steam Engine in the Late-Nineteenth-Century United States*. The Journal of Economic History 64, 61–99. https://doi.org/10.1017/S0022050704002608

Kremer, M., Leino, J., Miguel, E., Zwane, A.P., 2011. *Spring Cleaning: Rural Water Impacts, Valuation, and Property Rights Institutions*. The Quarterly Journal of Economics 126, 145–205. https://doi.org/10.1093/qje/qjq010

Class 5. February 27.

Artificial Intelligence Institutions

Work
How might AI affect work? What will humans do?

Ford, H., 1939. *Machines as Ministers to Man*. The New York Times 59, 89.

Impacts on existing jobs

Eloundou, T., Manning, S., Mishkin, P., Rock, D., 2023. *GPTs are GPTs: An Early Look at the Labor Market Impact Potential of Large Language Models*. https://doi.org/10.48550/arXiv.2303.10130

Brynjolfsson, E., Li, D., Raymond, L.R., 2023. *Generative AI at Work*. Working Paper Series. https://doi.org/10.3386/w31161

Peng, S., Kalliamvakou, E., Cihon, P., Demirer, M., 2023. *The Impact of AI on Developer Productivity: Evidence from GitHub Copilot*. https://doi.org/10.48550/arXiv.2302.06590

Work and meaning

Hussam, R., Kelley, E.M., Lane, G., Zahra, F., 2022. *The Psychosocial Value of Employment: Evidence from a Refugee Camp*. American Economic Review 112, 3694–3724. https://doi.org/10.1257/aer.20211616

Universal Basic Income

Bartik, A.W., Rhodes, E., Broockman, D.E., Krause, P.K., Miller, S., Vivalt, E., 2024. *The Impact of Unconditional Cash Transfers on Consumption and Household Balance Sheets: Experimental Evidence from Two US States*. Working Paper Series. https://doi.org/10.3386/w32784

Miller, S., Rhodes, E., Bartik, A.W., Broockman, D.E., Krause, P.K., Vivalt, E., 2024. *Does Income Affect Health? Evidence from a Randomized Controlled Trial of a Guaranteed Income*. Working Paper Series. https://doi.org/10.3386/w32711

Vivalt, E., Rhodes, E., Bartik, A.W., Broockman, D.E., Miller, S., 2024. *The Employment Effects of a Guaranteed Income: Experimental Evidence from Two U.S. States*. Working Paper Series. https://doi.org/10.3386/w32719

Crosta, T., Karlan, D., Ong, F., Rüschenpöhler, J., Udry, C.R., 2024. *Unconditional Cash Transfers: A Bayesian Meta-Analysis of Randomized Evaluations in Low and Middle Income Countries*. Working Paper Series. https://doi.org/10.3386/w32779

☐ **Tentative Third Assignment Due** (by 8pm March 5)
To be announced

Class 6. March 6.

Fakes, Signaling, and Safeguards

How might AI affect the ability to signal? What new institutions might be needed?

Gans, J.S., 2024. *How will Generative AI impact communication?* Economics Letters 242, 111872.

Costello, T.H., Pennycook, G., Rand, D.G., 2024. *Durably reducing conspiracy beliefs through dialogues with AI*. Science 385, eadq1814. https://doi.org/10.1126/science.adq1814

Manipulation and brittleness

Cao, S., Jiang, W., Yang, B., Zhang, A.L., 2023. *How to Talk When a Machine Is Listening: Corporate Disclosure in the Age of AI*. The Review of Financial Studies 36, 3603–3642. https://doi.org/10.1093/rfs/hhad021

Eykholt, K., Evtimov, I., Fernandes, E., Li, B., Rahmati, A., Xiao, C., Prakash, A., Kohno, T., Song, D., 2018. *Robust Physical-World Attacks on Deep Learning Visual Classification*. Presented at the Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 1625–1634.

Björkegren, D., Blumenstock, J.E., Knight, S., 2024. *Manipulation-Robust Prediction.*

Artificial Intelligence Institutions

Class 7. March 13.

Emerging Markets

How do opportunities and risks of AI differ among low-income populations?

Björkegren, D., 2023. *Artificial Intelligence for the Poor*. Foreign Affairs.

Machine Learning

Aiken, E., Bellue, S., Karlan, D., Udry, C., Blumenstock, J.E., 2022. *Machine learning and phone data can improve targeting of humanitarian aid.* Nature 603, 864–870. https://doi.org/10.1038/s41586-022-04484-9

Björkegren, D., Grissen, D., 2018. *The Potential of Digital Credit to Bank the Poor*. American Economic Association Papers and Proceedings.

Artificial Intelligence

Otis, N., Clarke, R., Delecourt, S., Holtz, D., Koning, R., 2024. *The Uneven Impact of Generative AI on Entrepreneurial Performance*. https://doi.org/10.2139/ssrn.4671369

Choi, J.H., Garrod, O., Atherton, P., Joyce-Gibbons, A., Mason-Sesay, M., Björkegren, D., 2024. *Are LLMs Useful in the Poorest Schools? TheTeacher.AI in Sierra Leone*. https://doi.org/10.48550/arXiv.2310.02982

McPeak, G., Sautmann, A., George, O., Hallal, A., Simal, E.A., Schwartz, A.L., Abaluck, J., Ravi, N., Pless, R., 2024. *An LLM's Medical Testing Recommendations in a Nigerian Clinic: Potential and Limits of Prompt Engineering for Clinical Decision Support,* in: 2024 IEEE 12th International Conference on Healthcare Informatics (ICHI). Presented at the 2024 IEEE 12th International Conference on Healthcare Informatics (ICHI), pp. 586–591. *https://doi.org/10.1109/ICHI61247.2024.00094*

No class. March 20. Spring Recess.

☐ **Tentative Fourth Assignment Due** (by 8pm March 26)
To be announced

Artificial Intelligence Institutions

Class 8. March 27.
Alignment
What values should guide automated systems?

Kleinberg, J., Ludwig, J., Mullainathan, S., Sunstein, C.R., 2020. *Algorithms as discrimination detectors*. PNAS. https://doi.org/10.1073/pnas.1912790117

Kasy, M., 2023. *The political economy of AI: Towards democratic control of the means of prediction.*

Abebe, R., Kasy, M., 2021. *The Means of Prediction*. Boston Review.

Class 9. April 3.
Teaching the machines
Who will teach the machines? How should they be paid?

Returns to data: scaling laws revisited

Gans, J.S., 2024. *Copyright policy options for generative artificial intelligence.* National Bureau of Economic Research.

Ghorbani, A., Zou, J., 2019. *Data Shapley: Equitable Valuation of Data for Machine Learning*, in: Proceedings of the 36th International Conference on Machine Learning. Presented at the International Conference on Machine Learning, PMLR, pp. 2242–2251.

☐ **Tentative Project Proposal Due** (by 8pm April 9)
To be announced

Class 10. April 10.
Digital governance
Are new forms of governance possible—or necessary?

Weyl, G., Tang, A., 2024. *Plurality: The Future of Collaborative Technology and Democracy.*

The Digitalist Papers. https://www.digitalistpapers.com

Pol.is

https://communitynotes.x.com/guide/en/about/introduction

Class 11. April 17.
Markets and planning
How can data and algorithms reshape markets?

Hayek, F.A., 1945. *The Use of Knowledge in Society*. The American Economic Review 35, 519–530.

Farronato, C., Fradkin, A., Larsen, B.J., Brynjolfsson, E., 2024. *Consumer Protection in an Online World: An Analysis of Occupational Licensing*. American Economic Journal: Applied Economics 16, 549–79. https://doi.org/10.1257/app.20210716

Luca, M., 2016. *Reviews, Reputation, and Revenue: The Case of Yelp.Com*. https://doi.org/10.2139/ssrn.1928601

Class 12. April 24.
Transportation
What should cities look like when cars drive themselves?

Ostrovsky, M., Schwarz, M., 2019. *Carpooling and the Economics of Self-Driving Cars*, in: Proceedings of the 2019 ACM Conference on Economics and Computation, EC '19. Association for Computing Machinery, New York, NY, USA, pp. 581–582. https://doi.org/10.1145/3328526.3329625

Björkegren, D., Duhaut, A., Nagpal, G., Tsivanidis, N. 2024. *Public and Private Transit: Evidence from Lagos.*

Class 13. May 1.
Project presentations
To be announced

☐ **Tentative Final Project Due** (by 8pm May 9)
To be announced

Version 18 November 2024