Last update: 12/01/2022

15.572 Analytics Lab

Action Learning Seminar on Analytics and Machine Learning

Instructors: Professor Sinan Aral
sinan@mit.edu; URL: http://www.sinanaral.io
Office hours by appointment

Professor Abdullah Almaatouq
amaatouq@mit.edu; URL: http://www.amaatouq.com/
Office hours by appointment

Class Times: Thursday 4:00-5:30pm, E62-276
Special Sessions (Pitch Day and Final Presentations):
Pitch Day: Friday, September 9, 3-6:30 pm, Virtual (Link TBD)
Final Presentations: Friday, December 2, 3-7pm, Virtual (Link TBD)

Recitation Times: Thursday 5:30-6:30pm, E62-276
Friday 12:00-1:00pm, E51-345

TAs: Hong-Yi TuYe hytuye@mit.edu Hirotaka Miura hmiura@mit.edu

Course support: Aileen Menounos aileenrm@mit.edu Carrie Reynolds carrie1@mit.edu

Summary and Objectives

The growth in big data, analytics and machine learning is transforming management decision-making, operations, marketing, finance, and product innovation. Businesses across the world are wrestling with these challenges and opportunities. We are on the cusp of a second machine age – an era where machines can automate or augment more and more of the mental tasks that previously only humans could do.

The purpose of the Analytics Lab (A-Lab) is to match student teams with leading-edge projects involving analytics and machine learning as they apply to business questions and problems. The primary focus of the projects is on the technical and analytical aspects, but business relevance provides the context and strategy.

Course Principles and Expectations

The primary criterion for projects is to provide a rich learning experience for the students. In addition, the projects should be of high relevance and interest to the supporting organization and senior managers and professionals in it.

Project teams of four students are expected to work independent of regular class meetings. Project sponsoring organizations will cover costs of travel and lodging, if any. Each project team will have an MIT-associated mentor to provide guidance and assistance and a link to outside project sponsors on an as-needed basis. The ultimate decision making and responsibility for project direction and completion rests with the team members themselves.

Two special sessions are scheduled: **Pitch Day on September 9 and Final Presentations on December 2. Attendance at both sessions is required.** Please arrange your schedule accordingly.

Notes on Class Activities and Due Dates:

- 9/9: On Pitch Day, we will meet jointly with the representatives from project proposing companies. Each will briefly describe their project as proposed, and students will have an opportunity to meet and informally mix with them and fellow students. The session will be followed by a reception. The chief aim of this session is to help inform student team formation and project selection.

- 9/16, 11:59pm: **DUE: Project Ranks**. After Pitch Day, students should start forming teams of 4 with team requirements in mind (e.g. 3 technical to 1 non technical ratio). 9/15's networking session will be another opportunity for team formation but we encourage students to start conversations as early as possible. In the following days, faculty, mentors, and the course support team will work out assignments of projects to students/teams, subject to review by the proposing company.

- 9/18: Final team-project pairings will be communicated to students. MIT and every proposing company have executed a jointly signed NDA. Each student team member will be required to review and sign an acknowledgment stating that all will abide by the terms agreed to in the NDA. Additional information will follow from Ellen Baum.

- 10/12, 11:59pm: **DUE**: **Project plan**. Each team should submit one document to their mentor (via email) and to the TAs (via Canvas). The project plan is to be developed by the team, reviewed by the team's mentor, and endorsed by the project sponsor before the deadline. It should be thought of as a working document, used by the team and mentor to assess progress and adjust and adapt through the semester. Here is a suggested outline:
    - Purpose and Scope: The project purpose and scope should serve as a compass that guides the team throughout the duration of the project. It should reflect the company proposal, but be more focused. Remember, the project is intended to be a rich learning experience for your team. This project is *not* a consulting engagement with the project sponsor. Bear this in mind when drafting your project purpose and scope. We welcome both creativity and practicality.
    - Objectives: Break the project down into high-level objectives that you intend to achieve. It is wise to define a clear "minimum viable product" as an initial goal, leaving room for expanding on that in a modular fashion as time and resources permit, including a "stretch goal" if all goes well.
    - Tasks: For each objective, list one or more granular tasks. For each task, define the following: a) due date, b) specific deliverables, c) who is responsible, d) current status. Revisit the objectives and tasks at least weekly to see if you're on track. Are there new opportunities or unanticipated barriers? Feel free to use a simple shared spreadsheet or more elaborate project management tool of your choice to track progress.

- 10/13, 10:00am: **DUE: Mid-term presentation slides**. Each team should submit their slides to their mentor (via email) and the TAs (via Canvas).

- 10/13 & 10/20, teams will deliver 3-minute presentations on their project work to date and potential lines of future analysis. The chief aim of these sessions is to help illuminate issues common across teams in order to foster discussion and collaboration. This is a great chance to get feedback and suggestions. Teams should be open and honest about their progress and challenges. Class participants should be supportive, helpful and generous with comments and advice.

- 12/1, 11:59pm: **DELIVERABLES DUE on Canvas (for TAs) and via email (for mentors).**

    - **Final report** (10 pages maximum, 3000 words, not including figures or references, not including executive summary).
        - **Summary of findings** (one-page executive summary); summary should include a high-level statement of the challenge addressed by the project and the key insights the team generated during the semester. This can include a graphic, but should fit on one page. What's interesting, useful and surprising about your work? Please attach this at the top of the final report PDF when submitting via Canvas.
    - **Final presentation slides.**
    - **[OPTIONAL UNDER EXCUSED ABSENCE] Final presentation video**. Each team who is excused should submit their 4-minute final presentation video to their mentor and upload to the shared drive [LINK TBD]. Feel free to upload them earlier!

- 12/2: During the Final Presentations session, each team will present their project work to an audience of experts, entrepreneurs, and executives, including representatives from project sponsoring organizations, as well as three outside judges. Teams will have 4 minutes to present their project work, plus 2 minutes for Q&A and judge remarks (6 minutes total per team). See the "Grading" section below for the judges' evaluation criteria.

N.B.: that teams are required to share your final report, summary of findings, and final presentation slides with project sponsors with enough lead time for them to review for inadvertent disclosure of Confidential Information.

Grading:

- 30% Final presentation content and delivery – team-wide; presentations will be evaluated according to the following criteria:
    - Technical and Analytical: How creative or advanced were the techniques used? Were they appropriate to the task and correctly applied?
    - Effort and Contribution: How much improvement was delivered? What alternative techniques were attempted before the team selected the one(s) that seemed best?
    - Business Impact: Beyond the data analytics described, how clearly did the team convey the bottom-line, real-world impact of their findings? What are the managerial or strategic implications? Can the potential benefits be quantified? Why do these results matter for the business or organization? What more general lessons can be learned? What are the next steps?
    - Presentation: How clear, informative and interesting was the presentation itself and how well was it delivered? Was it fun to see and hear? How did the team handle questions?
- 30% Final report + 15% Summary of Findings/Executive Summary – team-wide, using the same criteria as the presentation. Be sure to carefully and fully document all your references and data sources.
- 15% Contribution to class discussions and team project enablement – individual. Students in the class are co-producers of class discussions and collective learning. Your contributions to this learning process all semester long will be appraised in addition to the specific content that you contribute.
    - Independently evaluated by instructor, mentors, and team members.
- 10% Mid-point presentation content and delivery – team-wide.

Team Formation Requirements:

- If possible, students should form teams of 4.
- To simulate a real world data science project, teams should be composed of technical and business oriented members. As such, at least one team member should be business oriented (i.e. not MBAn) in a group of 4 (1:3 ratio).
- If in a group of 3, the minimum ratio drops to 1:2.
- TAs may rebalance teams as needed.

Required Books:

Data Science for Business: What You Need to Know About Data Mining and Data-Analytic Thinking, Foster Provost and Tom Fawcett. 2013. O'Reilly Media Inc. (Online access available at https://mit.primo.exlibrisgroup.com/discovery/openurl?institution=01MIT_INST&vid=01MIT_INST:MIT&rft.mms_id=9935050662306761&u.ignore_date_coverage=true

Data Destruction:

The following states the MIT Action Learning Office's policies on data destruction:

Project sponsors share confidential and proprietary information to student teams doing Action Learning projects. MIT Sloan has an obligation to destroy that data at the end of the project so that it does not inadvertently get disclosed to unauthorized people and it is not used for any other purpose than the project.

MIT Sloan depends on the student teams for destroying the data in a timely and appropriate manner. Please note that destruction of data is a requisite step for the completion of course requirements.

What data is required to be destroyed?
Any information supplied by the company in any format- emails, notes from a phone meeting, worksheets, records, company documents, any kind of company data. This includes data that is marked confidential and unmarked data. If the company supplied it, it must be destroyed at the end of the project.

What data is NOT required to be destroyed?
Students can keep their final paper and other derivative work that does not include company proprietary or confidential information. If there is any doubt, ask for help to discern what needs to be destroyed.

What are acceptable destruction methods?

- Printed Materials: Documents should be recycled in MIT approved secure recycle bins. Each academic area and many program offices have these bins.
- Digital Data Controlled by Students: If students have the data in Dropbox or on their computer, they must delete the data using appropriate tools.
- Digital Data Controlled by Sloan Technology Services: STS will destroy the data according to MIT Sloan IT policies.

If there are any issues or questions on this issue, please contact Ellen Baum, Contract Administration, at 3-5617 at ebaum@mit.edu.

Class Schedule:

	Date	Time	Session	Lecturer
S1	9/8	4:00-5:30	Welcome & Legal considerations and guidelines	Sinan Aral & Abdullah Almaatouq, Ellen Baum & Rajiv Shridhar
	9/9	2:00-6:00	**Pitch Day**	
S2	9/15	4:00-5:30	Social Analytics	Sinan Aral
NW	9/15	5:30-6:30	Networking and Team Formation	Hiro/Hong-Yi
S3	9/22	4:00-5:30	Guest Lecture	Claudia Perlich (Confirmed)
OSI	9/22	5:30-6:30	Intro to HPC: Cluster Computing	Hirotaka Miura
OSA	9/23	12:00-1:00	No REC - Student Holiday	
S4	9/29	4:00-5:30	Data Science for Social Good	Abdullah Almaatouq
OSI	9/29	5:30-6:30	Skill Seminar: Data Wrangling in R	Emma Van Inwegen
OSA	9/30	12:00-1:00	Intro to HPC: Parallel Computing	Hirotaka Miura
S5	10/6	4:00-5:30	Guest Lecture	Martin Tingley (confirmed)
OSI	10/6	5:30-6:30	Skill Seminar: Data Wrangling and Visualization in Python	Hong-Yi TuYe
OSA	10/7	12:00-1:00	Skill Seminar: Methods for observational causal inference (Intermediate)	Alex Moehring
S6	10/13	4:00-5:30	**Mid-Point Presentations I**	
OSI	10/13	5:30-6:30	Skill Seminar: Methods for observational causal inference (Advanced)	Alex Moehring
OSA	10/14	12:00-1:00	Skill Seminar: Intro to NLP	Mohammed Alsobay
S7	10/20	4:00-5:30	**Mid-Point Presentations II (and - CODE@MIT)**	
OSI	10/20	5:30-6:30	Skill Seminar: Deep Learning Foundations and Applications (and - CODE@MIT) (Intermediate)	Hong-Yi TuYe
OSA	10/21	12:00-1:00	No REC, Attend CODE@MIT	CODE@MIT
	10/27	4:00-5:30	SIP Week (No Class, No REC)	
S8	11/3	4:00-5:30	Guest Lecture	Gary King (Confirmed)
OSI	11/3	5:30-6:30	Project Office Hours	Hong-Yi TuYe, Hirotaka Miura
OSA	11/4	12:00-1:00	Skill Seminar: Deep Learning & Recommendation Systems	Madhav Kumar
S9	11/10	4:00-5:30	Guest Lecture	Gary Marcus (Confirmed)

OSI	11/10	5:30-6:30	Project Office Hours	Hong-Yi TuYe, Hirotaka Miura
OSA	11/11	12:00-1:00	Veteran's Day, No REC	
S10	11/17	4:00-5:30	Guest Lecture	Manish Raghavan (confirmed)
OSI	11/17	5:30-6:30	MLOps and Gradio	Abubakar Abid
OSA	11/18	12:00-1:00	TBD Skill Workshop or OHs	Hong-Yi TuYe, Hirotaka Miura
	11/24	4:00-5:30	Thanksgiving (No Class, No REC)	
S11	12/1	4:00-5:30	Class Devoted to Project Time	
	12/1	5:30-6:30	NA	
S12	12/2 **Friday**	2:00-7:00	**Final Presentations Session**	

Reading List (Check for Updates as the Semester Progresses**):**

Session 1: *Welcome – Intro to Analytics* (Sinan Aral)

1. Review all project proposals and the A-lab syllabus.
2. "*Big Data: The Management Revolution*" Brynjolfsson, E. and McAfee, A. 2012. *Harvard Business Review*, 90(10); October: 60-68; http://libproxy.mit.edu/login?url=http://search.ebscohost.com/login.aspx?direct=true&db=bth&AN=79996279&site=ehost-live.
3. "*The Business of AI*" Brynjolfsson, E. and McAfee, A. 2017. *Harvard Business Review*, July; https://hbr.org/cover-story/2017/07/the-business-of-artificial-intelligence

Recommended Readings

4. "*Chapter 1: Introduction: Data Analytic Thinking*" Provost, F. and Fawcett T. 2013. *Data Science for Business*, O'Reilly Media Inc.; http://library.mit.edu/item/002221893.
5. *"The Unreasonable Effectiveness of Data"* Halevy, A., Norvig, P. and Pereira, F. 2009. *IEEE Intelligent Systems;* http://static.googleusercontent.com/media/research.google.com/en//pubs/archive/35179.pdf
6. *"The Rapid Adoption of Data-Driven Decision-Making."* Brynjolfsson, E. and McElheran, K. 2016. *American Economic Review*, 106(5): 133-39. https://www.aeaweb.org/articles?id=10.1257/aer.p20161016
7. "*Big Data: New Tricks for Econometrics*" Varian, H. 2014. *Journal of Economic Perspectives*, 28(2): 3-28; https://www.aeaweb.org/articles.php?doi=10.1257/jep.28.2.3.
8. "*Lectures on Machine Learning*" Athey, S. and Imbens, G. 2015. NBER; http://conference.nber.org/confer/2015/SI2015/ML/syllabus.pdf.
9. "*The Future of Prediction: How Google Searches Foreshadow Housing Prices and Sales*" Wu, L. and Brynjolfsson, E. 2014. *Economics of Digitization* (A. Goldfarb, S. Greenstein, and C. Tucker, eds.), Univ. of Chicago Press; http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2022293.

Session 2a: *Social Analytics* (Sinan Aral)

10. *"Chapter 3: Hype Machine"* Sinan Aral. *The Hype Machine*. 2020.
11. *"Chapter 6: Personalized Mass Persuasion"* Sinan Aral. *The Hype Machine*. 2020.

12. *"Chapter 8: Strategies for a hyper-socialized world"* Sinan Aral. *The Hype Machine*. 2020.

Recommended Readings

13. "Social Influence Bias: A Randomized Experiment" Muchnik, L., Aral, S., and Taylor, S. 2013. Science, 341(6146); August 9: 647-651; http://www.sciencemag.org/content/341/6146/647.full.
14. 20. "Identifying Influential and Susceptible Members of Social Networks" Aral, S. and Walker, D. 2012. Science, 337(6092); July 20: 337-341; http://www.sciencemag.org/content/337/6092/337.full.
15. "Creating Social Contagion through Viral Product Design: A Randomized Trial of Peer Influence in Networks" Aral, S. and Walker, D. 2011. Management Science, 57(9); September: 1623-1639; http://pubsonline.informs.org/doi/abs/10.1287/mnsc.1110.1421.
16. "Distinguishing Influence Based Contagion from Homophily Driven Diffusionin Dynamic Networks"Aral, S., Muchnik, L., and Sundararajan, A. 2009. Proceedings of the National Academy of Sciences (PNAS), 106(51); December 22: 21544-21549; http://www.pnas.org/content/106/51/21544.full.

Session 2b: *Covid-19 Analytics* (Sinan Aral)

17. *"Interdependence and the Cost of Uncoordinated Responses to COVID-19."* Holtz, David, Michael Zhao, Seth G. Benzell, Cathy Y. Cao, Mohammad Amin Rahimian, Jeremy Yang, Jennifer Allen, Avinash Collis, Alex Moehring, Tara Sowrirajan, Dipayan Ghosh, Yunhao Zhang, Paramveer S. Dhillon, Christos Nicolaides, Dean Eckles, and Sinan Aral. Proceedings of the National Academy of Sciences Vol. 117, No. 33 (2020): 19837-43. https://www.pnas.org/content/117/33/19837
18. *"Interdependent Program Evaluation: Geographic and Social Spillovers in COVID-19 Closures and Reopenings in the United States."* Zhao, Michael, David Holtz, and Sinan K. Aral. Science Advances Vol. 7, No. 31 (2021): eabe7733. https://www.science.org/doi/10.1126/sciadv.abe7733
19. *"Surfacing Norms to Increase Vaccine Acceptance."* Moehring, Alex, Avinash Collis, Kiran Garimella, M. Amin Rahimian, Sinan Aral, and Dean Eckles. 2021. PsyArXiv. February 8. doi:10.31234/osf.io/srv6t. https://psyarxiv.com/srv6t/

Session 3: Data Science for Social Good (Abdullah Almaatouq)

20. "Computational social science." Lazer, David, et al. 2009. Science. 323(5915), 721-723 https://gking.harvard.edu/files/LazPenAda09.pdf
21. "Predicting poverty and wealth from mobile phone metadata." Blumenstock, Joshua, Gabriel Cadamuro, and Robert On. Science 350.6264 (2015): 1073-1076. https://www.unhcr.org/innovation/wp-content/uploads/2016/11/blumenstock-science-2015.pdf
22. "Unique in the shopping mall: On the reidentifiability of credit card metadata." De Montjoye, Yves-Alexandre, Laura Radaelli, and Vivek Kumar Singh. Science 347.6221 (2015): 536-539. https://science.sciencemag.org/content/347/6221/536.full.pdf
23. "Measuring the predictability of life outcomes with a scientific mass collaboration." Salganik, Matthew et al. 2020. Proceedings of the National Academy of Sciences (PNAS)
24. "The world's technological capacity to store, communicate, and compute information" Hilbert, Martin, and Priscila López. Science (2011). http://www.uvm.edu/pdodds/files/papers/others/2011/hilbert2011a.pdf

Recommended Readings

25. Chapters 1-3 "Bit by bit: Social research in the digital age". Salganik, Matthew. Princeton University Press, 2019.
26. "Mobile communication signatures of unemployment." Almaatouq, Abdullah, Francisco Prieto-Castrillo, and Alex Pentland. International conference on social informatics. Springer, Cham, 2016. https://arxiv.org/pdf/1609.01778.pdf
27. "Unique in the crowd: The privacy bounds of human mobility." De Montjoye, Yves-Alexandre, et al. Scientific reports 3 (2013): 1376. https://www.nature.com/articles/srep01376

Session 4: Guest Lecture, TBA

Readings TBA

Session 5: Guest Lecture, Gary Marcus
Readings TBA

Session 8: Guest Lecture, Gary King
Readings TBA

Session 9: Guest Lecture, Martin Tingley
Readings TBA

Session 10: Guest Lecture, Manish Raghavan
Readings TBA

Session 11: Guest Lecture, TBA
Readings TBA

Optional Skill Seminar: *Data Wrangling in Python* (Jaeyun Song)

Recommended Readings
28. Wickham, H. (2014). Tidy Data. Journal of Statistical Software, 59(10), 1 - 23: https://www.jstatsoft.org/index.php/jss/article/view/v059i10/v59i10.pdf
29. "10 minutes to pandas": https://pandas.pydata.org/docs/user_guide/10min.html#min
30. Chapters 5 and 7-10 of McKinney, W. (2017). Python for data analysis: Data wrangling with Pandas, NumPy, and IPython.
31. "Regular Expression HOWTO": https://docs.python.org/3/howto/regex.html

Optional Skill Seminar: *Data Wrangling and Visualization in R* (Emma Van Inwegen)

Recommended Readings
32. Wickham, H. (2014). Tidy Data. Journal of Statistical Software, 59(10), 1 - 23: https://www.jstatsoft.org/index.php/jss/article/view/v059i10/v59i10.pdf
33. 'R for Data Science' Chapter 5 (other chapters are very good as well): https://r4ds.had.co.nz/ (online version)
34. Advanced dplyr verbs:
    - join (https://dplyr.tidyverse.org/reference/join.html)
    - mutate_all (https://dplyr.tidyverse.org/reference/mutate_all.html)
    - summarize_all (https://dplyr.tidyverse.org/reference/summarise_all.html)
    - complete (https://tidyr.tidyverse.org/reference/complete.html)
    - pivot_wider (https://tidyr.tidyverse.org/reference/pivot_wider.html)
    - pivot_longer (https://tidyr.tidyverse.org/reference/pivot_longer.html)
35. 'ggplot2: elegant graphics for data analysis': https://ggplot2-book.org/ (online version)
36. Tufte, E. R. 2007. The Visual Display of Quantitative Information., Cheshire, Conn. : Graphics Press, c2007. https://search.ebscohost.com/login.aspx?direct=true&db=cat00916a&AN=mit.001657880&site=eds-live&scope=site.

37. Pipe Operators (Chaining): https://magrittr.tidyverse.org/

Optional Skill Seminar: *Intro to NLP* (Zanele Munyikwa)

38. Section 2 of Text Mining with R: Sentiment Analysis with tidy data https://www.tidytextmining.com/sentiment.html
39. Read First Three Sections of Chapter 10: Representing and Mining Text (Introduction, Why Text is Important, Why Text is Difficult) From Provost, F. and Fawcett T. 2013. Data Science for Business, O'Reilly Media Inc.; http://library.mit.edu/item/002221893
40. Golder, Scott A., and Michael W. Macy. "Diurnal and seasonal mood vary with work, sleep, and daylength across diverse cultures." Science 333, no. 6051 (2011): 1878-1881.

Recommended Readings

41. Hovy, Dirk, and Shannon L. Spruit. "The social impact of natural language processing." In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers), pp. 591-598. 2016. https://www.aclweb.org/anthology/P16-2096.pdf
42. Kiritchenko, Svetlana, and Saif M. Mohammad. "Examining gender and race bias in two hundred sentiment analysis systems." arXiv preprint arXiv:1805.04508 (2018).

Optional Skill Seminar: *Intro to Deep Learning and Machine Vision* (Hong-Yi TuYe)

Recommended Readings

43. CS231n Notes Module 1: Neural Networks. **CS231n: Convolutional Neural Networks for Visual Recognition**
44. CS231n Notes Module 2: Convolutional Neural Networks **CS231n: Convolutional Neural Networks for Visual Recognition**

Optional Skill Seminar: *Deep-Learning Based Recommendation Systems* (Madhav Kumar)

45. Kumar, M., Eckles, D., & Aral, S. (2020). Scalable bundling via dense product embeddings. *arXiv preprint arXiv:2002.00100*.
46. Shuai Zhang, Lina Yao, Aixin Sun, and Yi Tay. 2018. Deep Learning based Recommender System: A Survey and New Perspectives. ACM Comput. Surv. 1, 1, Article 1 (July 2018)
47. Stanley, J. 2017. 3 Million Instacart Orders, Open Sourced. https://tech.instacart.com/3-million-instacart-orders-open-sourced-d40d29ead6f2

SPEAKER BIOS

Claudia Perlich

Claudia Perlich is currently the SVP of Data Science at Two Sigma. Prior to this, she was the Chief Scientist at Dstillery where she designed, developed, analyzed, and optimized machine learning that drives digital advertising. An active industry speaker and frequent contributor to industry publications, Claudia enjoys serving as a guide in the world of data. In 2013, she was named the winner of the Advertising Research Foundation's (ARF) Grand Innovation Award and was selected as a member of the Crain's NY annual 40 Under 40 list. Additionally, she has been published in over 30 scientific journals and holds multiple patents in machine learning. She has won many data mining competitions, including the prestigious KDD Cup three times for her work on movie ratings in 2007, breast cancer detection in 2008, and churn and propensity predictions for telecom customers in 2009, as well as the KDD best paper award for data leakage in 2011 and bid optimization in 2012. She started her career in Data Science at the IBM T.J. Watson Research Center, concentrating on research in data analytics and machine learning for complex real-world domains and applications. She received her PhD in Information Systems from the NYU Stern School of Business, and holds a Master of Computer Science from Colorado University. Claudia takes active interest in the making of the next generation of data scientists and is teaching "Data Mining for Business Intelligence" in the NYU Stern MBA program.

Gary Marcus

Gary Marcus is a scientist, best-selling author, and entrepreneur. He is Founder and CEO of Robust.AI, and was Founder and CEO of Geometric Intelligence, a machine learning company acquired by Uber in 2016. He is the author of five books, including The Algebraic Mind, Kluge, The Birth of the Mind, and The New York Times best seller Guitar Zero, as well as editor of The Future of the Brain and The Norton Psychology Reader. Marcus attended Hampshire College, where he designed his own major, cognitive science, working on human reasoning. He continued on to graduate school at Massachusetts Institute of Technology, where his advisor was the experimental psychologist Steven Pinker. He received his Ph.D. in 1993.

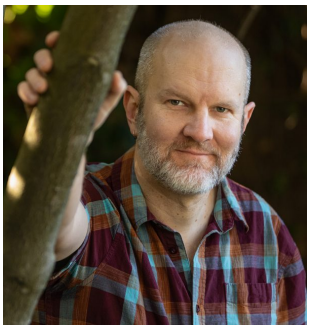
Gary King

Gary King is the Albert J. Weatherhead III University Professor at Harvard University -- one of 25 with Harvard's most distinguished faculty title -- and Director of the Institute for Quantitative Social Science. King develops and applies empirical methods in many areas of social science, focusing on innovations that span the range from statistical theory to practical application. King is a proud graduate of SUNY New Paltz (B.A., 1980) and the University of Wisconsin-Madison (M.A., Ph.D., 1984). He taught at NYU for three years before coming to Harvard in 1987.

Martin Tingley

Martin Tingley leads product experimentation research at Netflix. Before joining Netflix, he held positions at Insurance Australia Group and served on the faculty of Penn State University. Tingley is a graduate of the University of Toronto (B.Sc., 2003) and Harvard University (M.A., Ph.D., 2009).

Travis Brooks



Travis Brooks is a Product Manager of Experimentation Platform at Netflix. Before joining Netflix, he held positions at Yelp and SLAC National Accelerator Laboratory. Brooks is a graduate of Case Western Reserve University (B.S., 1997) and Stanford University (M.S., 2001).

Manish Raghavan

Manish Raghavan is an Assistant Professor of Informational Technology at the MIT Sloan School of Management. Manish was most recently a postdoctoral fellow at the Harvard Center for Research on Computation and Society (CRCS), working with Cynthia Dwork. He completed his PhD at the computer science department at Cornell University, advised by Jon Kleinberg. His research studies the impacts of computational tools on society with a focus on decision-making, behavioral economics, and hiring algorithms.