

Data Science for Social Impact

APSTA-GE 2331 | Fall 2017 | 3 credits

Wednesday & Thursday, 9:00 – 10:15 am

7 East 12th Street, 121



NEW YORK UNIVERSITY

Instructors

Jennifer Hill

Jennifer.hill@nyu.edu

246 Greene Street, 2nd Floor

Office hours by appointment

Laura Norén

laura.noren@nyu.edu

60 5th Avenue, 6th Floor

Office hours Tuesday & Thursday 1:00 – 2:00 PM

Course Description

Course focuses on how analysts use data and quantitative evidence to impact policy and practice. Students will learn how to gather and analyze data to address questions about program efficacy and efficient targeting of resources. Topics will include how to choose organizational partners, implement change, build trust with organizations and civic agencies, satisfy the needs of stakeholders and manage legal, ethical, and logistical constraints. Students will discuss real case studies and appropriate ways to address them.

Course Prerequisites

- **REQUIRED:** Two semesters of Statistics or Data Science (for instance APSTA-GE 2003/2004 or DS-GA 1001/1002 or equivalent as approved by the instructor).

Learning Objectives

Upon completion of the course, students will be able to

- 1) integrate organizational goals and available data into an actionable research plan
- 2) identify and articulate basic legal, ethical, logistical and financial constraints in research involving human subjects and human-generated data
- 3) critically assess and develop applicable frameworks for an ethically sound balance between privacy protection and transparency/reproducibility
- 4) identify the basics of best practice in reproducible research
- 5) articulate how organizations may have to serve the needs of various stakeholders and outline potential pathways to negotiate potentially competing desiderata
- 6) identify and demonstrate how to address complications associated with combining datasets and satisfying human subjects requirements

Course Format

Two 1.25 hour sections of class time per. The course will feature weekly guest lectures on Wednesdays (usually) followed by lead small group discussions and hands-on activities on Thursdays (usually).

Students will be exposed to research partnership issues and solutions through case studies presented by guest lecturers of the following types: researchers who have directly encountered such challenges in the field, practitioners working to promote social good at organizations that hire or partner with data scientists, experts on particular topics (e.g. IRB compliance or reproducibility in research). Students will be expected to be active, respectful, and intellectually additive classroom participants.

Course Requirements

The grade for this course will be determined as follows:

- class participation (15% of grade),
- written/computational assignments (60%),
- and one group final project (25%).

Students will be required to participate in class activities (small or large group) by asking questions or making comments a minimum of 20 times during the course of the semester. Participation will be documented via students sending the text of the questions they ask of the guest lecturer to Profs Hill and Norén. To receive credit, questions must be asked in class OR, if there are too many questions in class, an attempt to ask the question must have been made. This is a requirement in part communication is an important learning goal.

Assignments will be the following:

- 1) interviewing someone who works at the intersection of Data Science and Social Impact and writing a blog post about research-practice partnerships. Due 6 October 2017
- 2) creation of replicable computer code for a program evaluation with messy data and multiple datasets that need to be combined. Due 27 October 2017
- 3) practice dissemination of findings in the form of an Op-Ed and a report to program administrators for the same program. Due 17 November 2017
- 4) complete CITI human subjects certification. Due 1 December 2017
- 5) Final presentations of major project. Due 13 December 2017

In the final group project the group will be presented with a hypothetical situation defined by the available data, resources, organizational goals and constraints and the group will be required to devise a research and dissemination strategy (these case studies will be created in consultation with the speakers recruited each term and thus will vary each year).

The following readings may be supplemented by materials suggested by guest speakers.

COURSE OUTLINE

Week 1 Defining Data Science and Social Impact

6-7 Sept Guest Speaker: Wednesday

Tensie Whelan, Clinical Professor of Business & Society; Director, Center for Sustainable Business; Former President of the Rainforest Alliance where she increased the budget from \$4.5m to \$50m and certified 17% of the world's tea, 14% of the world's cocoa, and 5% of the world's coffee in compliance with Rainforest Alliance labor and environmental standards.

Readings

O'Neil, Kathy and Schutt, Rachel. (2015) Doing Data Science, Chapter 1. O'Reilly.

Pue et al. "Toward a Theory of Social Innovation" (2016) Monk School of Global Affairs White Paper, University of Toronto.

Stuek, W. (2011) "The arguments for and against Vancouver's supervised injection site,"
Globe and Mail

Week 2 Defining impact, standards of evidence: What Works Clearinghouse, Campbell Collaboration, Cochrane Collaboration

13-14 Sept Guest Speaker: Thursday

Chris Weiss, Senior Education Research Scientist, Institute of Education Sciences and former Director of Substance Use and Mental Health Program area at the Vera Institute of Justice.

Readings

Song, M and Herman, R. (2010) "A Practical Guide on Designing and Conducting Impact Studies in Education: Lessons Learned From the What Works Clearinghouse (Phase I)," American Institutes of Research Working Paper.

Taubes, Gary, New York Times Magazine Sept 16, 2007 "Do We Really Know What Makes Us Healthy"

Week 3 Opportunities and challenges in research-practice partnerships

20-21 Sept Guest Speaker: Wednesday

JeanCarlo Bonilla, Director of Engagement and Impact at DataKind who spearheads mission-driven projects with partner organizations and volunteer DataCorps and DataDive programs as well as the new DataKind Labs team of full-time data scientists.

Readings

Brinkerhoff, J. M. (2002) "Assessing and improving partnership relations and outcomes: A proposed framework," *Evaluation and Program Planning*, 25, 215-231.

Coburn, C.E., Penuel, W.R. (2016) "Research-Practice Partnerships in Education: Outcomes, Dynamics and Open Questions," *Educational Researcher*
[plus associated case studies from WT Grant site]

Week 4 (More) Opportunities and challenges in research-practice partnerships

27-28 Sept Guest Speaker: Thursday

Elise Cappella, Associate Professor of Applied Psychology at NYU. Cappella's research integrates education and psychology with the goal to better understand what disrupts, and alternatively, promotes children's positive adaptation in schools and communities. She has participated in a school-randomized control trial of a theory-based program to align parents and teachers around temperament-based strategies to promote children's behavior and learning.

Readings

Capella, Elise; Jackson, Daisy R.; Bilal, Caroline; Hamre, Bridget; and Soulé, Charles. (2011) Bridging Mental Health and Education in Urban Elementary Schools: Participatory Research to Inform Intervention Development. *School Psychology Review*. Vol. 40(4): 486-508.

Capella, Elise; Hamre, Bridget; Kim, Ha Yeon; Henry, David B., Frazier, Stacy L., and

- Atkins, Marc S. (2012) Teacher Consultation and Coaching Within Mental Health Practice: Classroom and Child Effects in Urban Elementary Schools. *American Psychological Association*. Vol. 80(4): 597-610. DOI: 10.1037/a0027725
- Reid, P. T., & Vianna, E. (2001) "Negotiating partnerships in research on poverty with community-based agencies," *Journal of Social Issues*, 57(2), 337-354.
- Rynes, S. (2007) "Let's Create a Tipping Point: What Academics and Practitioners Can Do, Alone and Together," *Academy of Management Journal*, 50(5): 1046-1054.

Week 5 Incorporating qualitative and quantitative data

4-5 Oct Guest Speaker: Wednesday

Dana Burde, Associate Professor and Director of International Education at NYU
Professor Burde focuses on the effects of conflict on education, the efforts of humanitarian organizations to mitigate these effects, and the relationship between education and political violence or peace.

Readings

Bamberger, M. (August 2012) "Introduction to Mixed Methods in Impact Evaluation," *Impact Evaluation Notes*, Interaction (with support from Rockefeller Foundation)

6 Oct **ASSIGNMENT 1 DUE: Blog post based on interview with someone who works at the intersection of data science and social impact (must submit interview audio file).**

Week 6 Best practice in transparent and reproducible programming

11-12 Oct Guest Speaker: Wednesday

Vicky Steeves, Librarian for Research Data Management and Reproducibility, NYU
Ms. Steeves provide instructional and consultation services to the NYU Community in research data management and conducts research on the role of reproducibility in research.

Readings

Smith-Spangler, Crystal M. (2012) "Transparency and Reproducible Research in Modeling: Why We Need It and How to Get There," *Medical Decision Making*, 32 (5): 663-666.

Week 7 Mechanics and politics of heterogeneous data integration

18-19 Oct Guest Speaker: Wednesday

Julia Ingrid Lane, Professor of Public Service at NYU Wagner Graduate School of Public Service and NYU Provostial Fellow for Innovation Analytics. Julia has led many initiatives, including co-founding the UMETRICS and STAR METRICS programs at the National Science Foundation. She conceptualized and established a data enclave at NORC/University of Chicago. This provides a confidential, protected environment within which authorized researchers can access sensitive microdata remotely and provides data producers with a secure dissemination platform.

Readings

Foster, Ian; Ghani, Rayid; Jarmin, Ron S., and Lane, Julia, eds. (2016) *Big Data and Social Science*, Chapter 3: Record Linkage. Chapman & Hall.

Vatsalan, D., Christen, P., O'Keefe, C., and Verykios, Va. (2014) "An Evaluation Framework for Privacy-Preserving Record Linkage," *Journal of Privacy and Confidentiality*: 6(1), Article 3.

Week 8 Understanding and balancing the needs of different stakeholders and target audiences

25-26 Oct Guest Speaker: Wednesday

Jack Buckley, Senior Vice President at American Institutes for Research

He leads AIR's Research and Evaluation area where he oversees projects across our entire range of subject areas including education, health, and the workforce both in the United States and internationally, always with the goal expanding our knowledge about how best to improve people's lives, particularly the disadvantaged. Before joining AIR, he helped lead the redesign of the SAT at the College Board.

Readings

Antonacopoulou, E., & Meric, J. (2005). "From power to knowledge relationships: Stakeholder interactions as learning partnerships," *Stakeholder Theory. A European Perspective*, 125-147.

Harrison, J. S., & Freeman, R. E. (1999) "Stakeholders, social responsibility, and performance: Empirical evidence and theoretical perspectives," *Academy of Management Journal*, 42(5), 479-485.

LeRoux, K. (2009) "Managing stakeholder demands: Balancing responsiveness to clients and funding agents in nonprofit social service organizations," *Administration & Society*, 41 (2), 158-184.

27 Oct **ASSIGNMENT 2 DUE: Code creation of replicable computer code**

Week 9 Disseminating results to multiple stakeholders; handling null or negative results

1-2 Nov Guest Speaker: TBD

Readings

Turcotte-Tremblay, A., Fregonesse, F., Alam, N. (2013) "Ethical Considerations of Dissemination and Restitution of Findings in Global Health Research", *BioéthiqueOnline*, 2/7

Week 10 Translating results into practice

8-9 Nov Guest Speaker: Wednesday

Michael Holland, Executive Director of NYU Center for Urban Science & Progress

Mr. Holland ensures the effectiveness of day-to-day operations and the optimal use of available resources. His role is also to provide leadership and direction for budget and financial planning, and to manage special projects and strategic planning. Mike comes to CUSP with a background in research policy and the oversight of federal research programs. He previously worked with Director Koonin as Senior Advisor and Staff Director in the

Office of the Under Secretary for Science at the US Department of Energy. Before that, Mike oversaw the Department of Energy's Office of Science for a decade as a program examiner in the White House Office of Management & Budget.

Readings

- Carr, K.M and Bradley-Levine, J. (2016) "Translational Research Design: Collaborating with Stakeholders for Program Evaluation," *The Qualitative Report*, 21(1): 44-58
- Daniel, D. B. (2012) "Promising Principles: Translating the Science of Learning to Educational Practice," *Journal of Applied Research in Memory and Cognition*, 1, 251–253
- Glasgow, R. E., & Emmons, K. M. (2007), "How can we increase translation of research into practice? Types of evidence needed," *Annual Review of Public Health*, 28, 413–433.
- Lenfant C. Shattuck lecture—clinical research to clinical practice—lost in translation? *N Engl J Med* 2003; 349: 868–74
- Palinkas, L. A., & Soydan, H. (2012). Translation and implementation of evidence-based practice. New York: Oxford University Press.
- Weyrauch, V. and Díaz Langou, G. (2011) "Sound Expectations: from Impact Evaluations to Policy Change" *International Initiative for Impact Evaluation* working paper

Week 11 Careful use of data and data-driven decision making: Messy data, significance, and confidentiality

15-16 Nov Guest Speaker: Wednesday
Connor Concannon, deputy director for data analytics at the Manhattan DA's office

Readings

- Buhl-Mortensen, Lene and Welin, Stellan. (1998) "The Ethics of Doing Policy Relevant Science: The Precautionary Principle and the Significance of Non-significant Results," *Science and Engineering Ethics*, Vol. 4: 401-412.
- Fienberg, S. (2001) "Statistical perspectives on confidentiality and data access in public health," *Statistics in Medicine*, 20(9-10): 1347-1356.
- Foster, Ian; Ghani, Rayid; Jarmin, Ron S., and Lane, Julia, eds. (2016) *Big Data and Social Science*. Chapter 11: *Privacy and Confidentiality*. Chapman & Hall.

17 Nov. ASSIGNMENT 3 DUE: Write an op-ed and a report to program administrators covering the same findings

Thanksgiving November 22-23

Week 12 Legal framework for data collection, analysis, and sharing (IRB, FERPA, HIPAA)

29-30 Nov Guest speakers: Wednesday
June Ahn, Associate Professor of Educational Communication & Technology
Ethical treatment of data panel
Dr. Michael Cantor, Medical Doctor (Internal medicine) and Director of Clinical Research Informatics

Additional panelist(s) TBD

Readings

NYU human subjects tutorial: <http://www.nyu.edu/ucaihs/tutorial/>

Bailar, John C. "Science, Statistics and Deception." *Research Ethics: A Reader*, Deni Elliott and Judy E. Stern Eds. Hanover: University Press of New England, 1997.

Brydon-Miller, M., & Greenwood, D. (2006) "A re-examination of the relationship between action research and human subjects review processes," *Action Research*, 4(1), 117-128.

Resnick, D. B. (2000) "Statistics, ethics, and research: An agenda for education and reform," *Accountability in Research: Politics and Quality Assurance*, 8: 163:188.

1 Dec. **ASSIGNMENT 4 DUE: CITI training complete, must submit certificate of completion**

Week 13 Scaling social impact

6-7 Dec Guest Speaker: TBD

Readings

Dees, G., Anderson, B.B. and Wei-skillern, Jane (2014) "Scaling Social Impact: Strategies for spreading social innovations," *Stanford Social Innovation Review*.

Week 14 Wrap-up and presentations

13-14 Dec **ASSIGNMENT 5 DUE: Student presentations of major project**

Academic Integrity

All students are responsible for understanding and complying with the New York University Steinhardt School Statement on Academic Integrity. A copy of this statement is available at: http://steinhardt.nyu.edu/policies/academic_integrity.

Students with Disabilities

Students with physical or learning disabilities are required to register with the Moses Center for Students with Disabilities, 726 Broadway, 2nd Floor, (212-998-4980 and online at <http://www.nyu.edu/csd>) and are required to present a letter from the Center to the instructor at the start of the semester in order to be considered for appropriate accommodation.