

SocGen188/IS180 Bio/Data: The Social and Political Consequences of Big Data in Biology and Medicine

Thursdays, 9:30-11:50AM

Public Affairs 2317

Instructor: Christopher Kelty (ckelty@ucla.edu) and Emmanuel Didier (edidier@ucla.edu)

Syllabus

Catalog Description

An elective class designed to introduce students to technical, social and political issues concerning "big data" in biology and the health sciences. Includes contemporary and historical cases. Topics may include the history of statistics and the uses of data in public health and medicine; insurance data; the relationship of quantification, classification and individuality; disease monitoring and 'sentinels'; statistics and stratification; the politics of surveillance; participatory data production; monitoring and self-monitoring; huge data in genomics. Students will learn basic concepts in statistics, data analysis and data mining.

Introduction

Data seem to be everywhere nowadays, especially "big" data. This course approaches this issue with special attention to health and medicine, and primarily from a social and cultural perspective. Health, both individual and collective, is now the subject of huge data capture and increasingly complex tools for computing and analyzing such data. But what are these data? How are they gathered and organized? Do they change the way the specialists look at patients, the way we are cured and cared for? Do they change how public health is conceived, how disease surveillance or intervention are practices? What lasting effects will these new data have? Are they used to set up new public health policies, or change individual health behavior through increasingly intimate means? Students will be asked to read a mix of popular and academic writing; to learn some of the basic concepts and technologies involved in data analysis, statistics and computation; and to engage in research on the forms and practices of data collection in health and medicine today.

Course structure

This course meets once a week for 2.5 hours. Each week, the course will be divided into two sections of roughly 1 hour each. The first hour will include lecture, discussion and debate conducted in a seminar-style. The second hour will focus on more practical questions, especially as they relate to the quarter-long research project.

Course goals

- ❖ Understand the basics of the relationship amongst data, statistical technique and computation/technology.
- ❖ Understand the political uses of statistics and data: statistics as both an instrument of truth-telling and of the governance of populations.

- ❖ Understand basic concepts in statistics, data analysis and computation, how they are used, and their implications for health and medicine
- ❖ Understand how data has developed historically and what constitutes the major shifts in quality, quantity and analysis of data.
- ❖ Understand the difference between individual and collective uses of data; from preventive and personalized medicine to public health and disease surveillance.
- ❖ Understand how statistics and big data are playing into contemporary aspects of market capitalism and policy-making in governments, in the domains of health, medicine and biology.

Grading Structure

Your grade in this course is divided evenly:

- ❖ 50% in class discussion, reading response papers (weekly) and participation.
 - Discussion and participation are mandatory.
 - Reading response papers will be assigned each week and due the day before class (Wednesdays by 5pm); class members will be selected to start off discussion each week
- ❖ 50% Research Project (see below for breakdown)

Required Texts:

All required texts will be available on the CCLE site for the class, or via the UCLA Library's electronic resources.

Honor

Please don't plagiarize. Plagiarizing means using anyone else's work and pretending it is your own. Don't do it, cause it sucks, and when we find out, the Dean of Students will make you write a humiliating letter apologizing to everyone involved and you will probably fail the course. Learn how to keep track of your sources and cite them properly, so that you can indicate where different ideas and evidence come from.

Statement on Disability

In compliance with the Federal Rehabilitation Act of 1973, as amended (Public Law 93-112) and the Americans with Disabilities Act of 1990 (Public Law 101-336), University of California policy prohibits unlawful discrimination on the basis of disability in its programs, services, and activities. If you wish to request accommodations for a disability, please contact the Office for Students with Disabilities (OSD) as soon as possible at A255 Murphy Hall, 310.825.1501 or 310.206.6083 (telephone device for the deaf). Website: www.osd.ucla.edu. In addition to registering with OSD, please feel free to contact the professors privately.

Technology in the Classroom

During the lecture and discussion portion of the classes you are asked to participate fully, which means no devices unless you must absolutely use them to access your reading or notes. We prefer that you print out your readings and notes for this part of the course. For the lab and assignment prep portion of the course (second half of each meeting) devices are allowed for research purposes, data collection/analysis, note-taking etc. Leave your friends and your various addictions at home, please.

Assignments

Weekly reading responses/discussion questions

Each week you are required to produce a 1-page (around 300-400 words) paper that analyzes some aspect of the reading for the week. Response papers should do the following three things: (1) Demonstrate that you grasped the reading by summarizing the main points quickly and in a few sentences; (2) zero in on some aspect of the reading that interests you and explain it in more detail, and raise questions about it—what do you understand and what seems mysterious? Do you agree or disagree with the arguments? Does the reading spark new ideas or new questions for you? (3) Connect the reading to questions you have about the role of data, statistics and computation in health and medicine, in particular, even if that is not the topic of the readings. Here you should refer to other readings in this course, or another course you are taking, or perhaps to issues in the news media or in your own experience.

These response papers are due by Wednesday at 5pm each week (uploaded to the CCLE website). Late papers will be graded accordingly.

Research project

A substantial portion of your grade (50%) will come from a quarter-long research project designed to allow you to investigate the use of data in the domains of medicine, health, behavior or biology. In most weeks, we will set aside part of class to work in groups on the project. You will be expected to share your findings in a small group, talk through ideas and methods for investigating these issues, and potentially ways to find, create, collect and analyze your own data.

A more detailed description of the Assignment will be distributed in class and posted on the CCLE Website.

Readings and Schedule

NB: This schedule is tentative and likely to change. Check the CCLE Website for the most recent schedule and readings.

Week 1: Introduction, Orientation...the “Big Data” hype

An introduction to the course and lectures on the basic topics and concepts to be covered. Introduction to the Research project and first assignment. On day one, we will view two videos (total of about 2.5 hrs)—one of them is a generally uncritical portrait of the promises of Big Data, the other is more critical, sometimes paranoid, portrait. “Big Data” will be our point of departure—because it is seemingly everywhere today, but we will try to quickly move beyond it and into more complex and challenging terrain.

To be viewed in class:

BBC Horizon’s “The Age of Big Data”

<http://gsis.mediacore.tv/media/bbc-horizon-the-age-of-big-data>

“Terms and Conditions May Apply”

<http://tacma.net/>

Optional videos and articles:

You may also be interested in these other examples of “Big Data” promotion and/or critique.

Hans Rosling’s Ted Talks:

http://www.ted.com/speakers/hans_rosling

Kenneth Cukier’s talks:

<http://www.cukier.com/>

Janet Vertesi. 2014. “My Experiment Opting Out of Big Data Made Me Look Like a Criminal” *Time Magazine*, May 1. <http://time.com/83200/privacy-internet-big-data-opt-out/> See also:

<http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teen-girl-was-pregnant-before-her-father-did/>

Week 2: Data, Statistics, Computation

In this week we dive into some of the core concepts at issue in understanding data: what do we mean by statistics, by data, or by computation? What kinds of techniques, technologies and concepts do we need in order to understand these issues in more depth?

Required reading:

Mayer-Schönberger, Viktor, and Kenneth. Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Boston: Houghton Mifflin Harcourt. Introduction and Chapter 5 “Datafication” pp. 73-98 (PDF on CCLE)

Anderson, Chris. 2008. "The end of theory." *Wired magazine* 16.7: 16-07.
http://archive.wired.com/science/discoveries/magazine/16-07/pb_theory

Boyd, Danah, and Kate Crawford. 2012. "Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon." *Information, Communication & Society* 15.5: 662-679.
http://www.tandfonline.com/doi/abs/10.1080/1369118X.2012.678878#.VKd_eYrF_Iw

Lazer, D., R. Kennedy, G. King, and A. Vespignani. 2014. "The Parable of Google Flu: Traps in Big Data Analysis." *Science* 343 (6176) (March 14): 1203–1205.
<http://www.sciencemag.org/content/343/6176/1203>

Gina Neff. 2013. "Why Big Data Won't Cure us" *Big Data*. Sept; 1(3): 117–123.
<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC4114418/>

Vivien Marx. 2013. "Biology: The big challenges of big data," *Nature*, 498, 255–260 (13 June) <http://www.nature.com/nature/journal/v498/n7453/full/498255a.html>

Optional:

Liran Einav and Jonathan Levin. 2014. "Economics in the age of big data" *Science* 7 November 2014: 346 (6210), 1243089 [DOI:10.1126/science.1243089]

Kate Crawford. 2013. "The Hidden Biases in Big Data," *Harvard Business Review* blog, <https://hbr.org/2013/04/the-hidden-biases-in-big-data>

Week 3: From the Avalanche of numbers to the Data Deluge.

The question of "big" data—the role that the amount of data plays—has existed ever since data were first collected. It is interesting to notice that each time a new statistical technique appears it raises the question again, formulated another way. We will see how the question of quantity of data was formulated as early the 17th century attempts to understand disease, again in the 19th century establishment of administrative offices, and once more in the 1930s with the establishment of sampling procedures

Andrea Rusnak, 2002. *Vital Accounts, Quantifying Health and Population in Eighteenth-Century England and France*. Cambridge University Press, 2002. Intro and Chapter 1. (PDF on CCLE)

Hacking Ian. 1982. "Bio-power and the avalanche of printed number," *Humanities in Society*, 5/(3&4), pp. 279-295. (PDF on CCLE)

E Didier. 2011. "Counting on Relief: Industrializing the Statistical Interviewer during the New Deal", *Science in Context*, 24(2), pp. 281-310.
<http://dx.doi.org/10.1017/S026988971100010X>

Optional:

Daniel Rosenberg. 2012. "Data Before the Fact," in "Raw Data" is an Oxymoron, (Lisa Gitelman, ed.), pp. 15-40, The MIT Press, 2012. (PDF on CCLE)

Stephan, Frederick F. 1948. "History of the Uses of Modern Sampling Procedures." *Journal of the American Statistical Association* 43(241):12–39.
<http://www.jstor.org/stable/2280064>

S Leonelli. 2014. "What difference does quantity make? On the epistemology of Big Data in biology," *Big Data & Society* April-June 1: 2053951714534395, first published on June 3, 2014. doi:10.1177/2053951714534395
<http://bds.sagepub.com/content/1/1/2053951714534395.full>

Week 4: Aggregation. collections, maps, databases, documents.

How is data collected, where does it come from, and how is it stored? In more recent times, data are gathered in very different ensembles that have different characteristics and uses. We might consider the difference between collections, maps, and databases as overarching forms aggregating different kinds of data.

Strasser BJ. 2011. "The Experimenter's Museum: GenBank, Natural History, and the Moral Economies of Biomedicine." *Isis; an International Review Devoted to the History of Science and Its Cultural Influences* 102, no. 1 (2011): 60–96.
<http://www.jstor.org/stable/10.1086/658657>

Kelty, Christopher M. 2012. "This is not an article: Model organism newsletters and the question of 'open science.'" *Biosocieties* 7(2):140-168. <http://www.palgrave-journals.com/biosoc/journal/v7/n2/full/biosoc20128a.html>

Adam Bostanci. 2004. "Sequencing the human genome," Gaudillière, Jean-Paul, and Hans-Jörg Rheinberger, eds. *From Molecular Genetics to Genomics: The Mapping Cultures of Twentieth-Century Genetics*. Routledge Studies in the History of Science, Technology, and Medicine 20. London ; New York: Routledge, pp. 158 – 179. (PDF on CCLE)

Week 5: Labor. New producers of data

How is data produced? Humans play a key role both in the collection of data and in the ways that "automated" data capture is structured, through algorithms, machine learning or other techniques. Who are the new elite scientists and engineers of data? And who are the hidden laboring classes?

Howe, Doug, Maria Costanzo, Petra Fey, Takashi Gojobori, Linda Hannick, Winston Hide, David P. Hill, et al. 2008. "Big Data: The Future of Biocuration." *Nature* 455, no. 7209 (September 4): 47–50. doi:10.1038/455047a.
<http://www.nature.com/nature/journal/v455/n7209/full/455047a.html>

Harris Harlan, Murphy Sean, Vaisman Marck. 2013. *Analyzing the Analyzers, An Introspective Survey of Data Scientists and Their Work*, O'Reilly Press. (PDF on CCLE)

Carlo Caduff. 2014. Sick Weather Ahead: On Data-Mining, Crowd-Sourcing and White Noise, *Cambridge Anthropology*, Volume 32, Number 1: 32-46(15).
<http://dx.doi.org/10.3167/ca.2014.320104>

Irani, Lilly C., and M. Six Silberman. 2013. "Turkopticon." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems - CHI '13*. New York, New York, USA: ACM Press, P. 611
<http://dl.acm.org/citation.cfm?id=2470654.2470742> (Accessed December 18, 2014).

Mayer-Schönberger, Viktor, and Kenneth. Cukier. 2013. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Boston: Houghton Mifflin Harcourt. Introduction and Chapter 7 "Implications" pp. 244-296 (PDF on CCLE)

Week 6: Tracking individuals

In the next two weeks we look at the differences between information for tracking individuals and for tracking populations, and think about the differences between them. First we turn to some issues related to individuals—self-tracking and tracking of others health and wealth.

Lauer, Josh. 2008. "From rumor to written record: Credit reporting and the invention of financial identity in nineteenth-century America." *Technology and Culture* 49.2: 301-324.

https://muse.jhu.edu/journals/technology_and_culture/v049/49.2.lauer.pdf

Nafus, Dawn, and Jamie Sherman. 2014. "Big Data, Big Questions| This One Does Not Go Up To 11: The Quantified Self Movement as an Alternative Big Data Practice," *International Journal of Communication* 8 (2014): 11.

<http://ijoc.org/index.php/ijoc/article/view/2170>

Melanie Swan. 2012. "Sensor Mania! The Internet of Things, Wearable Computing, Objective Metrics, and the Quantified Self 2.0" *JSAN* 2012, 1(3), 217-253; doi:10.3390/jsan1030217, <http://www.mdpi.com/2224-2708/1/3/217>

Eubanks, Virginia. 2014. "Want to Predict the Future of Surveillance? Ask Poor Communities. January 15, 2014. <http://prospect.org/article/want-predict-future-surveillance-ask-poor-communities>

Stop Talking Start Doing - NYPD - Powered by IBM, Video 2009,

<https://www.youtube.com/watch?v=QrpGvmmYh1U>

Optional:

Leon, Pedro, et al. 2012. "Why Johnny can't opt out: A usability evaluation of tools to limit online behavioral advertising." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. ACM.

Week 7: Tracking populations

In this week we look to techniques and practices aimed at identifying populations and dealing with them as objects in their own right, specifically for the purposes of tracking disease.

Fearnley, Lyle. 2010. "Epidemic intelligence. Langmuir and the birth of disease surveillance." *Behemoth A Journal on Civilisation* 3.3: 36-56. <http://ojs.ub.uni-freiburg.de/behemoth/article/view/683>

Andre Lakoff. 2014. "Real-time biopolitics: the actuary and the sentinel in global public health," *Economy and Society* Volume 44 Number 1 Month 2014: 1– 19. (PDF on CCLE).

Brownstein, J. S., Freifeld, C. C., & Madoff, L. C. 2009. Digital Disease Detection — Harnessing the Web for Public Health Surveillance. *New England Journal of Medicine*, 360(21), 2153–2157. doi:10.1056/NEJMp0900702
<http://www.nejm.org/doi/full/10.1056/NEJMp0900702>

Hay SI, George DB, Moyes CL, Brownstein JS. 2013. Big Data Opportunities for Global Infectious Disease Surveillance. *PLoS Med* 10(4): e1001413. doi: 10.1371/journal.pmed.1001413
<http://www.plosmedicine.org/article/info%3Adoi%2F10.1371%2Fjournal.pmed.1001413>

Optional:

Trifirò G, Coloma PM, Rijnbeek PR, et.al. 2014. Combining multiple healthcare databases for postmarketing drug and vaccine safety surveillance: why and how?. (Review) *J Intern Med* 2014; 275: 551–561.

Week 8: Knowing populations in order to know individuals: personalized medicine.

Paradoxically, big data in medicine has given rise to something called "personalized medicine". The more data is available, the more specific its application may be. But science is supposed to be general, valid for every one. How a very concept of personalized medicine might even be conceivable?

Collins, Francis S. *The Language of Life: DNA and the Revolution in Personalized Medicine*. New York: Harper, 2010. Chapter 3. (PDF on CCLE).

Rabeharisoa, Vololona, Michel Callon, Angela Marques Filipe, João Arriscado Nunes, Florence Paterson, and Frédéric Vergnaud. 2014. "From 'politics of numbers' to 'politics of singularisation': Patients' Activism and Engagement in Research on Rare Diseases in France and Portugal." *BioSocieties* 9, no. 2: 194–217. doi:10.1057/biosoc.2014.4. <http://www.palgrave-journals.com/biosoc/journal/v9/n2/abs/biosoc20144a.html>

Wicks, Paul et al. 2010. "Sharing Health Data for Better Outcomes on PatientsLikeMe." Ed. Gunther Eysenbach. *Journal of Medical Internet Research* 12.2: e19. PMC. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2956230/>

Optional:

Greshake, Bastian et al. "openSNP—A Crowdsourced Web Resource for Personal Genomics." Ed. Tricia A. Thornton-Wells. *PLoS ONE* 9.3 (2014): e89204. PMC. Web. 18 Dec. 2014. <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3960092/>

Week 9: Data and Policy

The way information about society is gathered is more or less congruent with certain public policies. As you have policies that are related, for example, to liberalism, to planning, to

Keynesinism, some methods of data collection are related to such general kinds of policies. What are the policies that are linked to big data ?

Schadt, Eric E. 2012. "The changing privacy landscape in the era of big data." *Molecular Systems Biology* 8.1: 612.

<http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3472686/>

Evans, Barbara J, Michael O Dorschner, Wylie Burke, and Gail P Jarvik. 2014.

"Regulatory changes raise troubling questions for genomic testing." *Genet Med* 16:799–803. <http://dx.doi.org/10.1038/gim.2014.127>

Roberts, J. Scott, Dana Dolinoy, and Beth A. Tarini. 2014. "Emerging Issues in Public Health Genomics." *Annual Review of Genomics and Human Genetics* 15.1:461-80.

<http://www.ncbi.nlm.nih.gov/pubmed/25184533>

Alain Desrosières. 2003. "Managing the Economy: the State, the Market and Statistics", in Porter T. et Ross D. (eds.), *The Cambridge History of Science*, vol.7: Modern Social and Behavioral Sciences, Cambridge University Press. (PDF on CCLE)

Rouvroy, Antoinette. 2008. "Human Genes and Neoliberal Governance. A Foucauldian Critique. (NEW PAPERBACK EDITION, SEPT. 2009)." *Routledge*, Chapter 3 "Policy implications : discourses of genetic enlightenment as new disciplinary device." (PDF on CCLE)

Optional:

Zeynep Tufekci, 2014. Engineering the public: Big data, surveillance and computational politics, *First Monday*, Volume 19, Number 7 - 7 July 2014

<http://firstmonday.org/ojs/index.php/fm/article/view/4901/4097>

Week 10: Data and Aesthetics

Last, but not least, a major impact of data and statistics is visual and rhetorical (i.e. persuasive), and this is accomplished through the presentation of data in compelling and simple visual forms. We explore here at the end of the class some of the ways data are presented, and some of the ways artists find to critique the uses and abuses of data.

Tufte, Edward R., and P. R. Graves-Morris. 1983. *The visual display of quantitative information*. Vol. 2. Cheshire, CT: Graphics press. Selections. (PDF on CCLE).

"Safety in Numbers? Seven Artists on Big data," *Frieze Magazine*, March 2014,

<http://www.frieze.com/issue/article/safety-in-numbers/>

Paul Flicek and Ewan Birney. 2009. "Visualizing the Epigenome," in Ferguson-Smith, John M. Grealley, and Robert A. Martienssen, *Epigenomics*, p. 55–66. Dordrecht: Springer Netherlands. (PDF on CCLE).

Julien Prévieux, *Anomalies construites*, Video, 2011, 7'41":

<http://www.previeux.net/html/videos/AnomaliesFr.html>

"What Only Artists Can Teach Us About Technology, Data, and Surveillance." Accessed December 16, 2014. <http://www.thestranger.com/seattle/what-only-artists-can-teach-us-about-technology-data-and-surveillance/Content?oid=21162466>

Optional:

Rosenberg, Daniel, and Anthony Grafton. *Cartographies of time: A history of the timeline*. Princeton Architectural Press, 2013.

Rhizome Artbase: <http://rhizome.org/artbase/?ref=header>

Ars Electronica archive: <http://archive.aec.at/prix/>

Museum-goers Are Big Data's Latest Targets, *artnet.com*, Coline Milliard, Monday, December 15, 2014 <http://news.artnet.com/art-world/museum-goers-are-big-datas-latest-targets-196936>

Socgen 188/IS180 BioData

Research Assignment, Winter 2015

The goal of this research assignment is to challenge you to explore a particular use of data, statistics and computation and how it relates to a problem in the domain of medicine and biology. By doing so you will demonstrate your understanding of the core concepts and social/political implications of “big data” as it is emerging today, as well as how it relates to the uses of data and statistics in the past.

The final report will be a 10-15 page analysis of your chosen topic which will present an original, empirically-based, analysis of a problem involving the use of data, statistics and computation. This is an individual project, but you will have ample time in class to work in groups, discuss your progress and contribute to other students' projects as well.

Each week we will focus on different aspects of your research projects and use the time in class to think about where and how to find relevant information, ranging from data/statistics themselves to scholarly articles, interviews or other sources of information and analysis.

Keep in mind that this is an exercise in *research*. Your goal is to discover something you do not know, and preferably something that other people do not know or understand either. You may very well not succeed—but that does not mean you will fail the course, as long as your final report clearly outlines what is known, what you have and have not been able to discover, and discusses the origin, history, and implications of the topic you have chosen.

Jan 8: Overview of the project

Assignment (Due Jan. 15th): As you go through your week, note all cases where you think data might be collected about you. List all the apps on your phone or tablet that are collecting data; list the websites you think might keep data on you. If you go somewhere, are there sensors, cameras, or people collecting and creating data? Don't restrict it only to yourself—think about what groups you belong to: family, neighborhood, city, state; race, ethnicity, religion; consumer demographic, social network, etc. You may also look into what people around you are doing, but don't be creepy.

This data need not all be related to biology, health or medicine, but in your list, make special note of those that are (put them in bold or give them a star or something).

Jan 15th: Share your list, refine your interest in a problem

In class we will make a long list of all the data sources that people have identified. Alongside this we will try to identify whether that data is actually used to solve a problem, intended to be used, or is simply collected because it can be.

Assignment (Due 1/22): Define the topic. From your list, you must identify both a kind of data, and the problem that data is intended to solve. Describe the problem and the data, as well as what you do or do not know about it, what you think you can find out, and how you will go about doing so (**1-2 page**). Your goal is to identify a phenomenon—a problem, a set of techniques and some kind of data that all seem to come together. Moreover you will want to make sure that it is something you think you can research for the remaining 7 weeks—not too specific and not too general!

Jan 22nd: Debate your proposal

In class we will meet in small groups and share proposals with each other. In groups of 4 to 5, you will share your idea and make suggestions for others.

Assignment (Due 1/29): Refine your assignment in light of discussion. Now start researching.

Jan 29th TBD (Refining research questions)

Feb 5th, Feb 12th and Feb 19th: Research workshops

Over the next three weeks, in class, you will share what you've found with your group, and try to help others find resources and generate ideas for further research.

For each of these three assignments you may choose to do one or more of the following things:

1. Hunt for legal, medical, or scientific literature about your problem. Keep track of your sources and take good notes!
2. Find people who know more than you and ask them about your project. You may choose to do interviews—but if so, make sure you alert the instructors about your plans.
3. Find data about data—what kinds of data or statistics can you imagine creating, downloading or otherwise obtaining? What would you do with it if you had it? What are you learning about data that you could apply to your own collection, analysis and use of data?

Assignment (Due Feb 5th): Get technical. What technologies, tools or organizations are involved in your case. How is the data generated, who has access to it (can you get access to it?), where is it stored, etc. Is it automated or produced by hand?

Assignment (Due Feb 12th): Get legal. Who owns the data, who regulates it and who makes money on it? Are their local or federal laws related to it? Court cases?

Assignment (Due Feb 19th): Get historical. Research the history of the problem: how has it been approached in the past; what solutions have been attempted and has “data” “statistics” or “computation” played a central role? What kinds of people, institutions, or cultural contexts help make sense of the problem?

Feb 26th: Have an argument.

By this point you should know a great deal about your topic. With a **new** group you will attempt to make an argument or thesis concerning your chosen topic. Other members of the group (presuming that they do not know as much about your project) will ask you hard questions. You will try to answer them and defend your thesis. Take good notes and then think about revising your thesis and laying it out in your final report.

March 5th, March 11th: Drafting and Revising of report.

In class time will be used to workshop ideas and sections of the papers, as necessary.

Final Assignment (Due 3/21): Integrate it all into a report. What are the ethical, social, or political implications of this data—beyond the attempt to solve a particular problem? Are there unintended or uncertain effects? Of what sort, and for whom are they an issue?