Big Data in Society

Academic Year	2017-18, Hilary Term
Day and Time	Wednesdays, Weeks 1-9, 09:15-11:15
Location	Seminar Room, Oxford Internet Institute, 1 St Giles, Oxford, OX1 3JS
Course Providers	Prof. Eric T Meyer, Oxford Internet Institute, eric.meyer@oii.ox.ac.uk Prof. Ralph Schroeder, Oxford Internet Institute, ralph.schroeder@oii.ox.ac.uk
Prerequisites	None

Background

'Big data' has been touted as the next big breakthrough in research, business, and government. Beyond this hype around big data, there are unprecedented opportunities in some areas, but serious limitations in others. This course will move beyond the hype, and use concrete examples from the empirical literature to examine the challenges and opportunities inherent in big data approaches to advancing knowledge in a variety of domains.

The course will begin by defining 'data' and introducing the social implications of big data research. Then it will survey research in various areas in which big data has been used to address previously unanswerable questions, focusing in particular on social media such as Twitter, mobile phone data and collaboration platforms like Wikipedia. It will also look at a range of empirical studies that allow us to compare the approaches taken in a number of disciplines in which big data research has been undertaken, including economics, literature, and development studies. Further, it will examine the methods used in big data approaches, the claims made, and the limitations in terms of representativeness and reliability for various approaches. The aim of the course is to arrive at a critical understanding of big data research, its possibilities and current limitations.

Course Objectives

The course will provide an overview of current empirical research and new directions in the area of big data. The aim is to equip students with a critical understanding of big data research, to recognize the challenges and opportunities within this area, and to plan and assess this kind of research with an awareness of the theoretical and practical issues involved. Students will also gain an appreciation of different disciplinary perspectives on big

data and an awareness of privacy, anonymity and data sharing issues. They will be able to assess the strengths and weaknesses of different approaches to analysing big data, and the need to ground this research in theoretical questions and to connect it with other, existing findings and approaches.

Learning Outcomes

At the end of the course, students will:
- be able to identify the benefits and drawbacks of various big data approaches;
- have a thorough grasp of the social implications of the ethical, legal and social issues in this area of research;
- be familiar with the various approaches and findings using big data, especially in the social sciences but also in cognate areas;
- be able to design research with big data to advance social science research and relate this research to solving practical problems.

Teaching Arrangements

The course will consist of 8 two hour weekly sessions, each being divided into an hour's lecture with the remainder discussion and presentations of student work.

Students will be engaged in collaborative research over the course of the sessions, which will be presented on a weekly basis (this will not be formally assessed, but involve ongoing presentations on a question set at the beginning of the course).

Assessment

Students will be assessed through a final essay that is no longer than 5000 words which must be submitted via Weblearn by 12 noon of Monday of Week 1 of Trinity term (23 April). The essay will cover one of the course topics, and students will choose a topic in consultation with the course tutors in advance. The essay should be clearly related to the topics of the course.

Formative Assessment

Each student will also be required to give a presentation on a specific aspect of the session topic or to review a big data study under the additional readings for each session topic. Details of these presentations will be agreed in Week 1. Students will also be required to write one short (advised length: 1500-3000 words) formative essay on any of the 8 weekly topics. This essay will provide a means for students to obtain feedback on the progress they have achieved.

Submission of Summative Assignment

The summative assignment for this course is due on Monday of Trinity Term Week 1 (23 April) by 12.00pm and should be submitted electronically via the Assignment Submission WebLearn Site. The assignment should also be submitted electronically by 5:00 pm on the same day to teaching@oii.ox.ac.uk. If anything goes wrong with your submission, email

teaching@oii.ox.ac.uk immediately. In cases where a technical fault that is later determined to be a fault of the Weblearn system (and not a fault of your computer) prevents your submitting the assessment on time, having a time stamped email message will help the Proctors determine if your assessment will be accepted. Please note that you should not wait until the last minute to submit materials since Weblearn can run slowly at peak submission times and this is not considered a technical fault.

Full instructions on using WebLearn for electronic submissions can be found on Plato under General Information. There is also an FAQ page on the Assignment Submission WebLearn Site.

Please note that work submitted after the deadline will be processed in the standard manner and, in addition, the late submission will be reported to the Proctors' Office. If a student is concerned that they will not meet the deadline they must contact their college office or examinations school for advice. For details on the regulations for late and non-submissions please refer to the Proctors website at https://www.admin.ox.ac.uk/proctors/examinations/candidates/.

Any student failing this assessment will need to follow the rules set out in the OII Examining Conventions regarding re-submitting failed work.

Topics

> Week 1. Defining Big Data
> Week 2. Big Data and Politics: Access and Algorithms
> Week 3. Big Data and the Economy: New Ways to Target Consumers
> Week 4. Large-Scale Collaboration: A Case Study in Wikipedia
> *Week 5. Reading week – no class*
> Week 6. Big Data in Academia: Digital Humanities
> Week 7. Big Data and Geography: The Cases of Crisis Response and Development
> Week 8. Big Data and Ethics: From Facebook to the NHS
> Week 9. Social Implications and Outlook for Big Data Research

Key to Readings
A reading list is given below for each class. Those items marked with an asterisk (*) are essential reading and MUST be read by all students in preparation for the class. The rest of the items which are not marked with an asterisk are additional readings which need only be consulted in the preparation of student presentations or for essays.

General Readings

Eagle, N. & Greene, K. 2014. *Reality Mining: Using Big Data to Engineer a Better World.* Cambridge MA: MIT Press.
Mayer-Schonberger, Viktor; Cukier, Kenneth. 2013. *Big Data: A Revolution That Will Transform How We Live, Work and Think.* London: John Murray.
Borgman, Christine. 2014. *Big Data, Little Data and Beyond.* Cambridge, MA: MIT Press.
Stephens-Davidowitz, Seth. 2017. *Everybody Lies: What the Internet Can Tell Us About Who We Really Are.* New York: Harper Collins.

Schroeder, R. 2016. Big Data and Communication Research. *Oxford Research Encyclopedia of Communication.*
https://doi.org/10.1093/acrefore/9780190228613.013.276

Golder, Scott and Macy, Michael. 2014. Digital Footprints: Opportunities and Challenges for Online Social Research. *Annual Review of Sociology 40*:6.1–6.24
https://doi.org/10.1146/annurev-soc-071913-043145

Week 1. Defining Big Data

Big data approaches are transforming how we understand society and how society understands itself. Yet, much remains unknown and contested. In order to critically interrogate the application of big data to academic and social issues, we must first consider what constitutes big data conceptually and analytically. Doing so raises a number of key questions: How to define big data? What does such data represent? Is big data merely different from existing data in quantity, or does the scale of this data make a qualitative difference? Who can (or cannot) access this data and who governs this accessibility? By addressing these questions we begin to grapple with the issues and cases introduced in the subsequent lectures.

* Lazer, D., Pentland, A. S., Adamic, L., Aral, S., Barabasi, A. L., Brewer, D., … others. (2009). Computational social science. *Science 323*(5915): 721.
https://dx.doi.org/10.1126/science.1167742
* boyd, D and Crawford, K (2012) Critical Questions for big data: Provocations for a cultural, technological and scholarly phenomenon. *Information, Communication and Society, 15(5)*: 662-79. http://dx.doi.org/10.1080/1369118X.2012.678878

Mayer-Schoenberger, V & Cukier, K (2013) Big Data: A Revolution that will transform how we live, work and think. London: John Murray. [Chapters 4 & 5]
Porter, T (2008) Statistics and Statistical Methods. In T. Porter and D. Ross (eds.) *The Modern Social Sciences*. Cambridge: Cambridge University Press, pp.238-50.
Ekbia, H. et al. (2015). 'Big Data, Bigger Dilemmas: A Critical Review', *Journal of the Association for Information Science and Technology*. 66(8): 1523-45.
Collins, R (1994) Why the Social Sciences Won't Become High-Consensus, Rapid-Discovery Science', Sociological Forum, 9(2): 155-77.
Silver, Nate. (2012). The Signal and the Noise: The Art and Science of Prediction. London: Allen Lane.

Week 2. Big Data and Politics: Access and Algorithms

In the past ten years, we have seen a veritable gold rush of data mined from Twitter, as well as some surprising yet powerful findings from experimental research on Facebook. We have also seen a shift in data accessibility. What do social media reveal about publics, or about voters and constituents and supporters? How are these data being used by elites? And are the data representative, or do they introduce new biases? The capacity to replicate or reproduce this research has come into question, and so raises questions about the scientific claims based on this work. And some research can only be done with access to proprietary social media data, but it has also been argued that publicly funded survey data still outweighs the importance of online data. What rights should citizens and the public have to this data? Control and access to data is clearly becoming a central issue in politics.

* Bakshy, E., Messing, S., & Adamic, L. (2015). Exposure to ideologically diverse news and opinion on Facebook. *Science*, *348*(6239), 1130–1132.
https://dx.doi.org/10.1126/science.aaa1160
* Bradshaw, S. & Howard, P.N. (2017). Troops, Trolls and Troublemakers: A Global Inventory of Organized Social Media Manipulation. Comprop Working Paper 2017.12.

Oxford, UK: Project on Computational Propaganda. http://comprop.oii.ox.ac.uk/wp-content/uploads/sites/89/2017/07/Troops-Trolls-and-Troublemakers.pdf

* Dutton, W., Reisdorf, B.C., Dubois, E., Blank, G. (2017). Search and Politics: The Uses and Impacts of Search in Britain, France, Germany, Italy, Poland, Spain, and the United States. Quello Center Working Paper No. 5-1-17. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2960697

Karpf, D. (2016). Analytic Activism: Digital Listening and the New Political Strategy. Oxford: Oxford University Press, chapter 1,2.

Hersh, E. (2015). Hacking the Electorate: How Campaigns Perceive Voters. Cambridge: Cambridge University Press.

Pariser, E. (2011). "Chapter 4. The You Loop". *The Filter Bubble: What the Internet Is Hiding from You*. New York, NY: The Penguin Press.

Backstrom, L, Boldi, P, Rosa, M, Ugander, J, & Vigna, S (2012) Four degrees of separation, in Proceedings of the 3rd Annual ACM Web Science Conference (WebSci '12). ACM, New York, NY, USA, pp.33-42.

Morstatter, F., Pfeffer, J., Liu, H., & Carley, K. (2013). Is the sample good enough? Comparing data from Twitter's streaming API with Twitter's firehose. *Proceedings of ICWSM*, 400–408. doi:10.1007/978-3-319-05579-4_10

González-Bailón S., Wang, N., Rivero, A., Borge-Holthoefer, J., & Moreno Y. (2014). Assessing the bias in samples of large online networks. Social Networks, 38, 16–27.

Cha, M., Haddadi, H., Benevenuto, F., & Gummadi, K. P. (2010). Measuring user influence in twitter: The million follower fallacy. In *4th International AAAI Conference on Weblogs and Social Media (ICWSM)*. Retrieved from http://www.aaai.org/ocs/index.php/ICWSM/ICWSM10/paper/download/1538/1826

Week 3. Big Data and the Economy: New Ways to Target Consumers

Before 'big data' came into fashion, Savage and Burrows considered the large scale collection of data to have ushered in a period of 'knowing capitalism'. A decade later, big data is used to track work and inventories, personalize advertisements, and a shift from marketing based on guesswork about audiences to marketing based on data science. There are also attempts to predict the stock market using Twitter and use search engine behavior to predict prices and employment. The use of such data has led some to suggest 'the end of theory' in favor of continuous reaction to high-fidelity market signals. Do these approaches conflate causation and correlation? Do they emphasize individuals as informational consumers rather than citizens? How can such approaches be regulated when even the European Commission, for example, cannot examine Google's code to determine if there has been foul play in their ranking algorithm?

* Moat, H. S., Curme, C., Avakian, A., Kenett, D. Y., Stanley, H. E., & Preis, T. (2013) Quantifying wikipedia usage patterns before stock market moves. *Scientific Reports*, 3, Article number 1801. http://dx.doi.org/10.1038/srep01801

* Bollen, J., Mao, H., & Zeng, X. (2011). Twitter mood predicts the stock market. *Journal of Computational Science 2(1)*: 1-8. https://doi.org/10.1016/j.jocs.2010.12.007

Savage, M., & Burrows, R. (2007). The Coming Crisis of Empirical Sociology. *Sociology*, *41*(5), 885–899. doi:10.1177/0038038507080443

Einav, L. & Levin, J. D. (2013). The Data Revolution and Economic Analysis. *NBER Working Paper Series No. 19035*. Retrieved from http://www.nber.org/papers/w19035

Choi, H. & Varian, H. (2012). Predicting the Present with Google Trends. *Economic Record, 88*(S1), 2-9.

Cavallo, A. (2011). *Scraped Data and Sticky Prices*. Paper presented at the American Economic Association Annual Conference, Denver, CO. http://www.aeaweb.org/aea/2011conference/program/retrieve.php?pdfid=403

Manyika, J. et al. (2011). 'Big data: the next frontier for innovation, competition and productivity', McKinsey Global Institute, available at: http://www.mckinsey.com/insights/mgi/research/technology_and_innovation/big_data_the_next_frontier_for_innovation (last accessed August 29, 2012).

Tancer, B. (2009). Click: What Millions of People are Doing Online and Why It Matters. New York: Harper Collins, 2009.

Turow, J. 2017. The Aisles Have Eyes: How Retailers Track Your Shopping, Strip Your Privacy, and Define Your Power. New Haven: Yale University Press.

Week 4. Large-Scale Collaboration: A Case Study in Wikipedia

It has been said Wikipedia does not work in theory, but it does work in practice. Part of the reason for its success has been the way Wikimedia as well as external academics have used data from Wikipedia to generate insights about the site. Alongside Twitter, Wikipedia data is one of the most popular sources of data for analysis because so much of it is openly available. Insights from this data have shown how subtle changes in the interface can make a difference to collaboration as well as demonstrate how editing patterns appear to reinforce existing global inequalities.

Through our examination of Wikipedia we will explore these issues as well as questions such as: how do different language versions represent the same topic? How do the locations and demographics of users skew what constitutes legitimate encyclopedic knowledge? How does Wikipedia function as a large-scale distributed collaboration platform? And apart from how Wikipedia is produced, a key issue for using Wikipedia as a research object is: who reads or uses it?

* Yasseri, T., Sumi, R., Rung, A., Kornai, A., & Kertész, J. (2012). Dynamics of conflicts in Wikipedia. *PloS ONE*, 7(6): e38869. https://doi.org/10.1371/journal.pone.0038869
* Mestyán, M., Yasseri, T., & Kertész, J. (2013). Early Prediction of Movie Box Office Success based on Wikipedia Activity Big Data. *PLoS ONE* 8 (8) e71226. https://doi.org/10.1371/journal.pone.0071226

Bao, P., Hecht, B., Carton, S., Quaderi, M., Horn, M., & Gergle, D. (2012). 'Omnipedia: Bridging the wikipedia language gap', in Proceedings of the 2012 ACM annual conference on Human Factors in Computing Systems, May, pp. 1075-1084.
Heilman, J. & West, A. 2015. Wikipedia and Medicine: Quantifying Readership, Editors, and the Significance of Natural Language, Journal of Medical Internet Research, 17(3):e62.
Reagle, J.M. (2010) *Good Faith Collaboration: The Culture of Wikipedia*. Cambridge MA: MIT Press.
West, R., Weber, I., Castillo, C. (2012) 'Drawing a Data-Driven Portrait of Wikipedia Editors', *Proceedings of WikiSym '12*, August 27–29, 2012, Linz, Austria.
Zhang, X. & Zhu, F. (2011) 'Group Size and Incentives to Contribute: A Natural Experiment at Chinese Wikipedia', *American Economic Review*, 101, pp. 1601–1615.
Waller, V. 2011b. The Search Queries that took Australian Internet Users to Wikipedia. Information Research, 16 (2), online at http://search.ebscohost.com/login.aspx?direct=true&db=lxh&AN=62852994&site=ehost-live
Taylor, L. and Schroeder, R. 'Big Data and Wikipedia Research: Social Science Knowledge across Disciplinary Divides', *Information, Communication and Society*, 18(9): 1039–1056.

Week 5. Reading Week – no class meeting

Week 6. Big Data in Academia: Digital Humanities
Guest Lecturer: Kathryn Eccles

Big data analyses, especially of text, have led to new approaches to literature and the study of culture. What new insights can be derived from patterns in the word frequency in thousands of novels, as in the work of Morretti and Jockers? And what about patterns of culture than can be detected by using Google Books, or 'culturomics'? In the humanities, there has been a backlash against scientific approaches to the study of literature, history and culture. What are strengths in analysis of patterns of text, and how text be linked to patterns of thought, belief, and culture? How closely do understandings of, for example, literary and artistic reputation that can be gained from the analysis of author and artist names fit with what we know about reputation from other research? This session will focus on humanities research, but link this research to enduring questions in the social sciences.

* Michelet, J. al. (2010). Quantitative Analysis of Culture Using Millions of Digitized Books. *Science 331*(6014): 176-182. https://doi.org/10.1126/science.1199644
* Moretti, F. (2000). Conjectures on World Literature. *New Left Review 1:* 54-68. https://newleftreview.org/II/1/franco-moretti-conjectures-on-world-literature

Schroeder, Ralph (2014). 'Big Data: Towards a More Scientific Social Science and Humanities' in Mark Graham and William H Dutton (eds.), Society and the Internet: How Networks of Information are Changing our Lives, pp.164-76.

Aiden, Erez and Michel, Jean-Baptiste. 2014.Uncharted: Big Data as a Lens on Human Culture.New York: Riverhead Books.

Duguid, P. (2007). 'Inheritance and loss? A brief survey of Google Books'. *FirstMonday* 12(8). Online http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/view/1972/1847

Heuser, R. and Le-Khac, L. (2011), 'Learning to Read Data: Bringing out the Humanistic in the Digital Humanities', *Victorian Studies* 54.1: 79-86.

Moretti, F. (2013). Distant Reading. London: Verso.

Stauffer, A. (2011). 'Introduction: Searching Engines, Readings Machines'.*Victorian Studies* 54.1, 63-68.

Jockers, Matthew. 2013. Macroanalysis: Digital Methods and Literary History. Urbana. University of Illinois Press.

Week 7. Big Data and Geography: The Cases of Crisis Response and Development

The use of geographic big data has been especially useful in understanding patterns of unequal development and in crisis response. While some of this data is volunteered geographic information (such as geotagged Wikipedia articles), this data also can come from movements of populations, trace data from mobile phones and sensor networks.

This data can be used to map patterns of mobility, diffusion of diseases and economic flows that were heretofore not possible. In the context of low income countries, a key question is about the visibility or invisibility of different parts of the population. Interesting contrasts can also be made between a 'bird's eye' perspective on population movements as against a perspective 'from the ground up', including logging the activity of individual users on mobile devices. This session will be devoted to examining the reliability of different studies, the gaps and the advantages of using digital sources, as well as the policy implications of different ways of quantifying and measuring populations and their movements.

* Bengtsson L, Lu X, Thorson A, Garfield R, von Schreeb J. (2011) Improved Response to Disasters and Outbreaks by Tracking Population Movements with Mobile Phone Network Data: A Post-Earthquake Geospatial Study in Haiti. *PLoS Med 8(8)*: e1001083. https://doi.org/10.1371/journal.pmed.1001083

* Blumenstock, J. E. (2012). Inferring patterns of internal migration from mobile phone call records: Evidence from Rwanda. *Information Technology for Development, 18(2)*: 107-125. http://dx.doi.org/10.1080/02681102.2011.643209

Schroeder, Ralph, & Taylor, Linnet (2014). 'Is bigger better? The emergence of big data as a tool for international development policy.' GeoJournal. DOI 10.1007/s10708-014-9603-5

Sutton, J., Spiro, E. S., Johnson, B., Fitzhugh, S., Gibson, B., & Butts, C. T. (2014). Warning tweets: serial transmission of messages during the warning phase of a disaster event. *Information, Communication & Society*, *17*(6), 765–787. doi:10.1080/1369118X.2013.862561

Hale, S. A. (2012). Net increase? Cross-lingual linking in the blogosphere. *Journal of Computer-Mediated Communication*, *17*(2), 135–151. doi:10.1111/j.1083-6101.2011.01568.x

Mao, H., Shuai, X., Ahn, Y. Y., & Bollen, J. (2015). Quantifying socio-economic indicators in developing countries from mobile phone communication data: applications to Côte d'Ivoire. *EPJ Data Science*, *4*(1), 1–16. doi:10.1140/epjds/s13688-015-0053-1

Boase, J. and Ling, R. (2013), Measuring Mobile Phone Use: Self-Report Versus Log Data. Journal of Computer-Mediated Communication, 18: 508–519. doi: 10.1111/jcc4.12021

de Montjoye, Y. A., Hidalgo, C. A., Verleysen, M., & Blondel, V. D. (2013). Unique in the Crowd: The privacy bounds of human mobility. *Scientific reports,* 3.

Week 8. Big Data and Ethics: From Facebook to the NHS

It is obvious that the collection, storage and analysis of big data comes with huge risks to privacy. However, there are additional ethical risks that might not be obvious from the outset. This week will begin with a review of big data ethics and then explore the case study of networks and big data. This week follows on directly from the preceding week's discussion of

access and politics and will be focused some specific case, like Facebook data and NHS data.

* Guillory, J. E., Verma, I. M., & Hancock, J. T. (2014). Editorial Expression of Concern: Experimental evidence of massive scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences*, *111*(29), 10779–10779. https://dx.doi.org/10.1073/pnas.1412469111
    * Also skim original article: Kramer et al. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences 111(24)*: 8788-8790. http://www.pnas.org/content/111/24/8788.short
* Schroeder, R. (2014). Big Data and the brave new world of social media research. *Big Data and Society*, July-December: 1-11. http://journals.sagepub.com/doi/abs/10.1177/2053951714563194
* Hawkes, N. (2016). NHS data sharing deal with Google prompts concern. *BMJ 353*: i2573. https://search.proquest.com/openview/5b5a2ee79362679e57629e6761dd2fbb/1
    * Also skim: Hodson, H. (2016). Google Knows Your Ills. *New Scientist 3072 (7 May 2016)*. https://www.newscientist.com/article/2086454-revealed-google-ai-has-access-to-huge-haul-of-nhs-patient-data/

Zimmer, M (2010) "But the Data is Already Public": On the Ethics of Research in Facebook, Ethics and Information Technology, 12 (4): 313-325.
Pasquale, F. (2005). The Black Box Society: The Secret Algorithms that Control Money and Information. Cambridge MA: Harvard University Press.
Rule, J (2007) Privacy in Peril: How We are Sacrificing a Fundamental Right in Exchange for Security and Convenience. New York: Oxford University Press, part IV.
Lewis, K., Kaufman, J., Gonzalez, M., Wimmer, A., & Christakis, N. (2008). Tastes, ties, and time: A new social network dataset using Facebook. com. *Social Networks*, *30*(4), 330–342.
Bond, R, Farris, C, Jones, J, Kramer, A, Marlow, C, Settle, J & Fowler, J (2012) A 61-million-person experiment in social influence and political mobilization, Nature, 489: 295–298.

Week 9. Outlook for Big Data Research

Privacy and anonymity are only some of the issues raised in relation to big data. One task of this session will be to establish to what extent these can be separated from the epistemological and methodological issues that will have been discussed on the course. Another will be to examine legal questions: are these unique to big data, and how do they relate, for example, to different national contexts? Finally, are there social issues that are neither ethical nor legal, but rather concerned with the role of new technology in society, and its power to influence our behavior? We will examine solutions proposed for these questions, current debates, and the outlook for policy and research. Looking forward, we will discuss how the public responds to big data being used, to predict and influence behavior.

* Mayer-Schoenberger, V & Cukier, K (2013) Big Data: A Revolution that will transform how we live, work and think. London: John Murray. [Chapters 8 and 9]
* Lazer, David, Ryan Kennedy, Gary King, and Alessandro Vespignani. 2014. The Parable of Google Flu: Traps in Big Data Analysis. *Science 343, no. 14* March: 1203-1205. https://dx.doi.org/10.1126/science.1248506

Axelsson, Anne-Sofie, and Ralph Schroeder."Making it Open and Keeping it Safe: e-Enabled Data Sharing in Sweden." Acta Sociologica 2009. 52 (3):213-226.
Greenleaf, G. (2012) Global data privacy laws: 89 countries, and accelerating. Queen Mary University of London, School of Law Legal Studies Research Paper No. 98/2012
Lane, J., Stodden, V. (2014). Privacy, Big Data, and the Public Good: Frameworks for Engagement. Cambridge: Cambridge University Press.
O'Neil, Cathy. (2016). Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy. London: Allen Lane.