



SCHOOL OF INFORMATICS AND COMPUTING

INDIANA UNIVERSITY
Department of Information and Library Science
Bloomington

Z604: The Social and Organizational Informatics of Big Data Fall 2015

Online

Instructors:

Dr. Howard Rosenbaum <hrosenba@indiana.edu>
Dr. Pnina Fichman <fichman@indiana.edu>
Ashley Dainas (Associate Instructor) <ardainas@iu.edu>

INTRODUCTION

The world is awash in digital data from the web, a wide range of sensor technologies, a wider range of information and communication technologies including email and social media, and an even wider range of public and private sector organizations. Royster (2013, 9) provides an overview of the national labor market need.

A major impediment to the widespread use of big data is the lack of workers with the appropriate training and skills. Big data work can require not only knowledge of statistical analysis and computer systems, but experience in the relevant field or industry, such as healthcare or physics.

Manyika et.al (2011, 110) describe three types of skilled big data employees, who will be in demand in the next five years, one of which is the type of informational professional this course is intended to prepare; these are “the managers and analysts who know how to operate companies by using insights from big data.”

This course provides a survey of the organizational, legal, political, and social issues that surround the creation, dissemination and use of big data from the perspective of social and organizational informatics. It is part of the data science specialization in the MIS and MLS degree programs and in the decision maker track in the Data Science program and focuses on the social and organizational impacts of big data in complex organizations. The course addresses a need in these programs to provide you with an understanding of the ways in which the integration of big data is changing organizational structure, culture, and work practices. The goal of the course is to help you understand the complexity of the organizational, legal, political, and social issues that surround the creation, dissemination and use of big data. You will have the opportunity to develop your decision-making skills as you focus on the impacts of big data on a domain of your choice.

Sources

1. Manyika, J., Chui, M., Brown, B., Bughin, J., Dobbs, R., Roxburgh, C. and Byers, A.H. (2011). Big data: The next frontier for innovation, competitiveness and productivity. McKinsey Global Institute.
2. Royster, S. (2013). Working with Big Data. Occupational Outlook Quarterly. Bureau of Labor Statistics

COURSE OBJECTIVES

By the end of this course, you will:

- Understand the role of big data in organizational decision making
- Understand the ways in which the integration of big data is changing the culture, structure, and work practices of complex organizations
- Understand the social and organizational impacts of big data

These learning objectives will be assessed through a series of in-class exercises, homework assignments, and a group final project.

COURSE REQUIREMENTS

This course is provided in a completely online format and uses pre-recorded video as the main mode of information delivery. The Powerpoint slides and all of the readings and course materials will be available through a site in Canvas, the University's online learning management system. The course is structured in a series of modules that are topic-based and designed to correspond to a fifteen-week semester. Each module has between seven and twelve videos where the readings and topics for that module are discussed. The first video in each module described the entire module and there is a Powerpoint slide with the outline for that module.

Because this is a fully online course, you will be taking a much more active role in your learning. In a face-to-face course, we see you at least weekly and can therefore play a more active part in directing your learning. What we have done here is to design a designing a digital space that allows you learn at your own pace, at your convenience, at times that are convenient for you (with one important exception, discussed below). Since you have a greater degree of control, you need to be disciplined enough to stay up to date with the course. We hope that you find the following suggestions useful:

1. Regular Canvas login: Log in to Canvas at least every other day to see if there are new announcements, discussion posts and replies to your posts. You may need to log in one or more times a day when there are group discussions happening, when there will be an interactive meeting, or when group submissions are due.
2. Manage your time: We'll be dealing with complex materials and academic readings. Working with this information online adds a level of complexity beyond what you might be used to in face-to-face courses. Handling the requirements of the class requires good planning, which allows you to meet unexpected challenges, such as illness, travel, family obligations, net outages etc. Allow for 8-10 hours per week on this course, or about two hours a day. This is the equivalent time commitment for a face-to-face course in our program, where you'd meet for about three hours a week and spend roughly three hours outside of classes for every hour you are in class.
3. Keep track of due dates: Use the calendar tool on Canvas to keep track of due dates. Print out the schedule at the end of the syllabus, but look for updates on Canvas; depending on our schedules and commitments, we may change assignment due dates during the semester.

Being a good citizen in the course

In online social interactions, there are ways of acting and norms to follow that make communication more enjoyable and productive. The following tips for interacting online in e-mail and/or Discussion Board messages are adapted from guidelines originally compiled by Von Rospach and Spafford.

- a. Remember that the person receiving your message is someone like you, someone who deserves and appreciates courtesy and respect.
- b. Be brief. Succinct, thoughtful messages have the greatest impact.
- c. Your messages reflect on YOU. Take time to make sure that you are proud of their form and content.

- d. Use descriptive subject headings in e-mail messages.
- e. Think about your audience and the relevance of your messages.
- f. Be careful with humor and sarcasm; without the voice inflections and body language of face-to-face communication, Internet messages can be easily misinterpreted.
- g. When making follow-up comments, summarize the parts of the message to which you are responding.
- h. Avoid repeating what has already been said. Needless repetition is ineffective communication.

Giving Feedback

This course makes use of collaborative learning. Therefore, it is important that we all understand how to provide quality feedback to each other. Here are a few tips for providing, positive, constructive, and useful feedback to peers:

- a. Be empathetic and remember that this environment is a safe place for making mistakes.
- b. Use nonjudgmental language and phrases that do not attack an individual. One way of doing this is to ask the individual to discuss his/her process for making the final decision.
- c. Use specific questions, examples, and references to research as a way of making your point.
- d. Make your feedback useful by providing suggestions that the individual can understand and use to improve her/his work

What you can expect from us

We are committed to providing a quality learning experience through thoughtful planning, implementation, and assessment of course activities. We are also committed to being readily available to students throughout the semester, by having synchronous office hour sessions, asynchronous discussions and chats, by replying to questions in the forum/email within one to two business days and to returning graded course work with feedback within five to ten business days of each assignment's due date (this is because of the very large size of the class).

When using the Canvas interactions tools, it is our job to initiate thoughtful, on-topic discussions, encourage student-to-student communication, and mediate when necessary. Therefore, it is not our responsibility to respond to every post, but encourage students to take ownership of the learning process by responding to each other. Either the Instructors or the Associate Instructor assigned to the class will respond to your questions about the class materials, assignments, and lectures and we will schedule interactive sessions where you can communicate with us synchronously.

Other requirements

To receive a passing grade in this course, you must turn in all of the assignments and the term project and do your presentations. All assignments will be submitted through Canvas. You cannot pass this course without doing all of the assigned work, however, turning in all of the work is not a guarantee that you will pass the course. Grades of <I> (Incomplete) may be assigned in this course after discussion with the instructors, but, depending on the circumstances, there will be a penalty applied at the discretion of the instructors.

All papers and assignments must be submitted on the dates specified in this syllabus. If you cannot submit an assignment or cannot deliver a presentation on the date it is due, it is your responsibility to discuss your situation with the instructors, preferably in advance. Given that your reasons or problems are legitimate,

arrangements for the completion of the outstanding work can be made; this will occur, however, at the discretion of the instructors. There will be a penalty for work turned in after the assigned date, and this will also be applied at the discretion of the instructors.

Your written and oral work will be evaluated according to four criteria; it must:

- Be clearly written or presented;
- Demonstrate a degree of insight into the concepts, issues, and trends in both the areas you investigate in the assignments and in the course content;
- Demonstrate a degree of originality in your assignments and the term project; and
- Display some familiarity with the appropriate current and/or classic literatures.

Borderline grades will be decided (up or down) on the basis of class contributions and participation throughout the semester.

Academic dishonesty

There is extensive documentation and discussion of the issue of academic dishonesty in the Indiana University "Code of Student Rights, Responsibilities and Conduct;" this is available at <http://www.indiana.edu/~code/>. Of particular relevance is the section on plagiarism:

3. Plagiarism

Plagiarism is defined as presenting someone else's work, including the work of other students, as one's own. Any ideas or materials taken from another source for either written or oral use must be fully acknowledged, unless the information is common knowledge. What is considered "common knowledge" may differ from course to course.

- a. A student must not adopt or reproduce ideas, opinions, theories, formulas, graphics, or pictures of another person without acknowledgment.
- b. A student must give credit to the originality of others and acknowledge indebtedness whenever:
 1. Directly quoting another person's actual words, whether oral or written;
 2. Using another person's ideas, opinions, or theories;
 3. Paraphrasing the words, ideas, opinions, or theories of others, whether oral or written;
 4. Borrowing facts, statistics, or illustrative material; or
 5. Offering materials assembled or collected by others in the form of projects or collections without acknowledgment.

From: Part II: Student Responsibilities - G. Uphold and maintain academic and professional honesty and integrity - 3. Plagiarism

<http://www.indiana.edu/~code/code/responsibilities/academic/index.shtml>

Plagiarism is the use of someone else's ideas, words, or opinions without attribution. Any assignment that contains plagiarized material or indicates any other form of academic dishonesty will receive a grade of "F". A second instance will result in an automatic grade of "F" for the course. Penalties may be harsher depending upon the severity of the offense. See Indiana University's "Code of Student Rights, Responsibilities and Conduct" (link above).

There is more to avoiding plagiarism than simply citing a reference. To aid students both in recognizing plagiarism and in avoiding the appearance of plagiarism, Indiana University's Writing Tutorial Services has prepared a short guide entitled *Plagiarism: What it is and how to recognize and avoid it*. For example, here are some strategies for avoiding plagiarism:

1. Put in **quotations** everything that comes directly from the text especially when taking notes.

2. **Paraphrase**, but be sure you are not just rearranging or replacing a few words. Instead, read over what you want to paraphrase carefully; cover up the text with your hand, or close the text so you can't see any of it (and so aren't tempted to use the text as a "guide"). Write out the idea in your own words without peeking.

3. **Check your paraphrase** against the original text to be sure you have not accidentally used the same phrases or words, and that the information is accurate.

From: <http://www.indiana.edu/~wts/pamphlets/plagiarism.shtml>

This guide is available here: <http://www.indiana.edu/~wts/pamphlets/plagiarism.shtml>. It provides explicit examples of plagiarism and offers strategies for avoiding it. Each student should be familiar with this document and use it as a guide when completing assignments. In fact, there are many pamphlets at Writing Tutorial Services that you might find useful as you begin your graduate work. They are listed here:

<http://www.indiana.edu/~wts/pamphlets.shtml>.

Indiana University and School of Library and Information Science policies on academic dishonesty will be followed. Students found to be engaging in plagiarism, cheating, and other types of dishonesty will receive an F for the assignment and additional penalties applied at the discretion of the instructors. As a rule of thumb, when in doubt, cite the source!

Grades in the Department of Information and Library Science

The following definitions of letter grades have been defined by student and faculty members of the Committee on Improvement of Instruction and have been approved by the faculty (November 11, 1996) as an aid in evaluation of academic performance and to assist students by giving them an understanding of the grading standards of the Department of Information and Library Science:

Grade	GPA	Meaning
A	4.0	Outstanding achievement. Student performance demonstrates full command of the course materials and evinces a high level of originality and/or creativity that far surpasses course expectations
A-	3.7	Excellent achievement. Student performance demonstrates thorough knowledge of the course materials and exceeds course expectations by completing all requirements in a superior manner
B+	3.3	Very good work. Student performance demonstrates above-average comprehension of the course materials and exceeds course expectations on all tasks as defined in the course syllabus
B	3.0	Good work. Student performance meets designated course expectations, demonstrates understanding of the course materials and is at an acceptable level
B-	2.7	Marginal work. Student performance demonstrates incomplete understanding of course materials.
C+	2.3	Unsatisfactory work. Student performance demonstrates incomplete and inadequate understanding of course materials
C	2.0	
C-	1.7	Unacceptable work. Coursework performed at this level will not count toward the MLS or MIS degree. For the course to count towards the degree, the student must repeat the course with a passing grade.
D+	1.3	
D	1.0	
D-	.07	
F	0.0	Failing. Student may continue in program only with permission of the Dean.

OTHER IMPORTANT INFORMATION

There are three ways you can get in touch with us outside of class:

1. Rosenbaum will be the primary faculty contact person. His office phone number is 812-855-3250 and his office hours are Monday, 1:00-2:30 PM EST. He can also make other arrangements to talk with you if these hours are not convenient. He has voice mail, so you can always leave us a message.
2. Rosenbaum's email address is <hrosenba@indiana.edu>. Please use Z604 in the subject line of your email if you wish to contact us about class.
3. There will be periodic synchronous "Conferences" held within Canvas. The Assistant Instructor will manage these sessions and will be announced no less than a week in advance. Rosenbaum will also use "Conferences" for online office hours.

ASSIGNMENTS

You will have four assignments in this class. These assignments are described below, and will be discussed in greater detail in class. They include:

1. In class exercises: Periodically, you will complete a series of in class exercises based on the readings that will accompany the various modules that make up the content of the class. Credit will be given for completing the exercise. Successful completion of these exercises will be worth **10%** of the final grade.
2. Critical reviews: You will complete one critical review of a reading selected from the syllabus reading list. This review will be worth **20%** of the final grade. You will select and critically review an article from the readings for the module for which you have signed up. You will email the article choice(s) to the instructor for approval at any time, but no later than the week before your article review is due.

As you read each article, try to engage with the key issues in the article. As you analyze the article, you should try to answer the following questions:

- What is the main point of the article?
- How does the author develop a persuasive argument to convince the reader of the importance of this point?
- Is this argument persuasive? Why or why not?
- What types of evidence are offered?
- What background is required to understand and make use of this research?
- What is its significance to the field or to data science as a whole?
- What critiques have been offered (or could be offered) regarding its approach?

Your completed essays will be ~1,000 words and will be uploaded to the Canvas "Assignments" section. When you sign up for a module, you will be assigned due date for the essays.

3. Leading group discussion: This assignment includes three components. You will be assigned to a small group of 7-8 people. In this group you will lead one group session that focuses on an article that you have selected that is not included in the syllabus but that is relevant to one of the class sessions. You must have

the article approved by the instructor or Associate Instructor two weeks before the date when you will upload your presentation.

You will complete this assignment by

A. Preparing a PowerPoint presentation of about 8-10 slides that will be posted to your group's space in Canvas and to the Canvas "Assignments" section. Follow the same format as you did for the critical review assignment and develop the presentation around the themes you noticed in the article. Address the following questions:

- What are the main themes of this article?
- What do they tell us about the aspect of data science the author is addressing?
- How well or poorly did they present their arguments?

When you post your presentation, include a slide with at least two questions that you think will encourage discussion among your group members.

B. You will monitor and participate in the online discussion that will follow in the three days after the slides are posted.

C. When others in your group take their turn leading the discussion, you will comment at least once on each of their presentations.

This assignment will be worth **20%** of the final grade.

4. Final paper: You have been asked by your supervisor to assess the impact of big data on your organization (this organization is in a specific domain of your choosing). Your investigation will take the form of a white paper where some aspect of the social and organizational impacts of big data will be investigated in some depth. You can make use of academic and trade resources. The audience for your white paper will be your supervisor and other high level decision makers in your organization.

Your white paper will have the following sections*:

- Title: Choose a title that clearly outlines the problem or issue you are addressing and is appropriate for your target audience.
- Abstract or executive summary: Provide a one-paragraph summary of the main points and the main conclusion of your paper. This is a way to grab and hold the reader's attention. Provide enough detail to satisfy a busy executive while encouraging her to continue to read the paper.
- Introduction: Define the problem or issue and provide background discussion. Building credibility by referencing relevant sources.
- Problem or issue definition: Describe the nature of the problem or issue based on your analysis of available literature. Explain why the audience should care about the problem or issue and how it is likely to have impacts on your organization. This section should be written from the perspective of the target audience.
- High-level solution: Describe the relevant solution at a high level, including the technologies and organizational changes involved. Support your arguments with data, if possible; this can include tables, charts, and graphics. Quote industry experts as needed to bolster your positions. You are educating your target audience on the current state of the art, as well as where your solution fits.

- **Solution details:** Having explained the problem and the general approach to solving it, describe your solution in more detail. Describe how your solution works and why it is superior to the competition. Remember your audience and use appropriate language and level of detail. This is the heart of your white paper and you can use case studies or customer testimonials (if you have been able to find any) to support your arguments.

- **Business benefits and risks:** This is an important section. Discuss the relevant benefits that you see coming to your organization, such as return on investment (ROI), usability, adherence to standards, and speed of implementation, supported by relevant data where possible. Explain what the risks might be as well (and how you suggest that they be handled).

- **Summary:** In several paragraphs, emphasize the benefits of your solution as well as the risks to readers who decide not to follow your recommendations. Many readers may skip the entire document and read this section only, so write this section as if it were an entirely standalone document summarizing the main selling points about your solution. Conclude with the most important point that you want the reader to remember.

*Adapted from: <http://www.yourwhitepaper.com/whitepaperstructure.htm>

Your white paper will be ~1500 words and will be **45%** of the final grade. It will be due at the end of the semester.

GRADING AND DUE DATES

This table shows the assignments you have to do and the percentage of the final grade each is worth.

Assignment	% of Final Grade	Date due
In-class exercises	10%	Throughout
Critical article review	20%	As assigned
Leading class discussion	20%	As assigned
Final paper	45%	December 1
Participation	5%	Throughout

Note that there is a portion of the overall grade that has been allocated for class participation. Participation will be determined in three main ways. One will involve a demonstration of your effort and interest in class. Since this class is online, participation in this sense is defined primarily as contributing to class discussion. A second involves other ways to demonstrate that you are engaged in the class, such as participating in the online synchronous sessions to discuss your work or ask questions. The final way is to engage in email discussions with the instructors and the graduate assistants about class materials and assignments.

REQUIRED TEXTS

There are no required texts for this course. Readings are available through the Canvas web site, the Library's databases, or on the web.

NOTE: The URLs for the readings were last checked on August 31, 2015

TOPIC OUTLINE, READING SCHEDULE AND ASSIGNMENT DUE DATES

SECTION I: SOCIETY

Class 1: Course Overview – Introduction to big data

Readings

Agrawal, D., Bernstein, P., Bertino, E., Davidson, S., Dayal, U. (2011). Challenges and Opportunities with Big Data 2011-1. Cyber Center Technical Reports. Purdue University. Purdue e-Pubs. 1-19.

Bughin, J., Chui, M., and Manyika, J. (2010). Clouds, big data, and smart assets: Ten tech-enabled business trends to watch. McKinsey Quarterly. 1-14.

Cukier, K. and Mayer-Schoenberger, V, (2013). The Rise of Big Data. Foreign Affairs. 92(3). 28-40.

Diebold, F.X. (2012). A personal perspective on the origin(s) and development of “big data”: The phenomenon, the term, and the discipline. Second Version (November 26, 2012). PIER Working Paper No. 13-003. 1-7
<http://ssrn.com/abstract=2202843>

Fisher, D., DeLine, R., Czerwinski, M., and Drucker, S. (2012). Interactions with Big Data Analytics. Interactions. May-June, 50-59.

King, J. Magoulas, R. (2014). Data Science Salary Survey: Tools, Trends, What Pays (and What Doesn't) for Data Professionals. O'Reilly Media. 1-35.

Winter, R. (2011). Big Data: Business Opportunities, Requirements and Oracle's Approach. Winter Corp. 1-8.

Class 2: Social and cultural impacts of big data

Readings

Brown, B., Chui, M., and Manyika, J. (2011). Are you ready for the era of big data? McKinsey Quarterly. 1-12.

Federal Trade Commission. (2014). Big data: A tool for inclusion or exclusion. 4-27.

Gillespie, T. (forthcoming). The relevance of algorithms. in Media Technologies, ed. Tarleton Gillespie, T., Boczkowski, P., and Foot, K.. Cambridge, MA: MIT Press. 1-32

Ross, D. (2013). Big data and social organizations: A state of the art review. Nominet Trust. 6-24.

Shilton, K. (2012). Participatory personal data: An emerging research challenge for the information sciences. Preprint - Journal of the American Society for Information Science and Technology. 1-35.

Class 3: Big data and policy

Readings

Crawford, K., and Schultz, J. (2013). Big data and due process: Toward a framework to redress predictive privacy harms. Public Law & Legal Theory Research Paper Series Working Paper No. 13-64. Law & Economics Research Paper Series Working Paper No. 13-36. New York University School of Law. 92-128.

Hemerly, J. (2013, June). Public policy considerations for data-driven innovation. *Computer*. 25-31.

Nguyen, T-V.T., and Fornara, N. (2014). Policy enforcement for the web of data. IAT4SIS '14, August 18, Prague, Czech Republic. 1-6.

Prinsloo, P., and Slade, S. (2013). An evaluation of policy frameworks for addressing ethical considerations in learning analytics. LAK '13, April 08 - 12 2013, Leuven, Belgium, 240-244.

Tallon, P.P. (2013). Corporate Governance of Big Data: Perspectives on Value, Risk, and Cost. *Computer*. 32-38.

Tene, O., and Polonetsky, J. (2013). Big data for all: Privacy and user control in the age of analytics. *Northwestern Journal of Technology and Intellectual Property*, 11(5). 240-273.

Class 4: Ethical and legal concerns of big data

Readings

Diakopoulos, N. (2013). Algorithmic accountability reporting: On the investigation of black boxes. *Columbia Journalism School*. 1-37.

Goodman, K.W. and Meslin, E.M. (2014). Ethics, Information Technology, and Public Health: Duties and Challenges in Computational Epidemiology. In Magnuson, J.A. and Fu, P.C. (eds.), *Public Health Informatics and Information Systems*, 191 Health Informatics, Springer-Verlag London, 191-209.

Ioannidis, J.P.A. (2013). Informed consent, big data, and the oxymoron of research that is not research. *The American Journal of Bioethics*. 40-42.

Nunan, D. and Di Domenico, M. (2013). Market research and the ethics of big data. *International Journal of Market Research*, 55(4), 1-13.

Richards, N.M. and King, J.H. (2013). Three paradoxes of big data. *66 Stanford Law Review Online* 41-46.

Sweeney, L. (2014). Transparency establishes trust. *Federal Trade Commission*. 1-13.

SECTION II: ORGANIZATIONS

Class 5: Big data and decision making

Readings

Court, D. (2015). Getting big impact from big data. *McKinsey Quarterly*. 1-8.

Davenport, T.H., Barth, P., and Bean, R. (2012). How “big data” is different” *MIT Sloan Management Review*, July 30, 1-6.

Lehikoinen, J. and Koistinen, V. (2014). In big data we trust? *Interactions*. September-October. 38-41.

Monroe, M. (2014). Seeing is believing. *XRDS*, 21(2), 44-47.

Pitt, J., Bourazeri, A., Nowak, A., Roszczynska-Kurasinska, M., Rychwalska, A., Santiago, I.R., Sánchez, M.L., Florea, M., and Sanduleac, M. (2013, June). Transforming big data into collective awareness. *Computer*. 40-45.

Provost, F. and Fawcett, T. (2013). Data science and its relationship to big data and data-driven decision making. *Big Data* 1(1), 51-59.

Tien, J.M. (2013). Big Data: Unleashing information. *Journal of Systems Science Systems Engineering*, 22(2), 127-151.

Class 6: Using big data for innovation

Bloem, J., van Doorn, M., Duivesteyn, S., van Manen, T., and Ommeren, E. (2012). Big social: Predicting behavior with big data. *The Sogeti Trend Lab. VINT.* 4-36.

Bloem, J., van Doorn, M., Duivesteyn, S., van Manen, T., and Ommeren, E. (2012). Your big data potential: The art of the possible. *The Sogeti Trend Lab. VINT.* 4-43.

Desouza, K.C. and Smith, K.L. (2014). Big data for social innovation. *Stanford Social Innovation Review*. 38-43.

Hilbert, M. (2013). Big data for development: From information – to knowledge societies. *SSRN Scholarly Paper No. ID 2205145*). 1-39.

Electronic copy available at: <http://ssrn.com/abstract=2205145>

Manyika, J., Chui, M., Groves, P., Farrell, D., Van Kuiken, S., and Doshi, E.A. (2013). Open data: Unlocking innovation and performance with liquid information. Ch. 3: Consumer products. *McKinsey Global Institute, McKinsey Center for Government, McKinsey Business Technology Office.* 43-54.

World Economic Forum. (2012). Big data, big impact: New possibilities for international development. 1-10.

Class 7: Data curation and technical big data work

Readings

Borgman, C.L. (2012). The conundrum of sharing research data. *Journal of the American Society for Information Science*, 63(6):1059–1078.

Chaudhuri, S. (2012). What next? A half-dozen data management research goals for big data and the cloud. *PODS '12, May 21–23, 2012, Scottsdale, Arizona, USA.* 1-4.

Jaakkola, H., Mäkinen, T., Eteläaho, A. (2014). Open Data – Opportunities and Challenges. *International Conference on Computer Systems and Technologies - CompSysTech'14.* 25-39.

Kaisler, S., Armour, F., Espinosa, A.J., and Money, W. (2013). Big data: Issues and challenges moving forward. *2013 46th Hawaii International Conference on System Sciences*, 995-1004.

Kervin, K.E., Cook, R.B., and Michener, W.K. (2014). The backstage work of data sharing. *GROUP '14, November 9-12, Sansibel Island, FL, USA.* 152-156.

Lynch, C. (2008). How do your data grow: Commentary. *Nature* (455). 28-29.

Rosenthal, D.S.H., Rosenthal, D.C., Miller, E.L., Adams, I.F., Storer, M.W., Zadok, E. (n.d.). The economics of long-term digital storage. 1-8.

Class 8: Critiques of big data

Readings

- boyd, b. and Crawford, K. (2011). Six provocations for big data. Paper presented at Oxford Internet Institute's "A Decade in Internet Time: Symposium on the Dynamics of the Internet and Society" on September 21. 1-17.
- boyd, d. and Crawford, K. (2012). Critical questions for big data. *Information, Communication & Society*, 15(5), 662-679.
- Bryant, A. and Raja, U. (2014). In the realm of big data. *First Monday*, Volume 19, Number 2-3. 1-22.
- Buhl, H.U.B., Röglinger, M., Moser, F., Heidemann, J. (2013). Big data: A fashionable topic with(out) sustainable relevance for research and practice? *BISE*, 65-69.
- Ekbia, H., Mattioli, M., Kouper, I., Arave, G., Ghazinejad, A., Bowman, T., Suri, V.R., Tsou, A., Weingart, S., and Sugimoto, C.R. (2014). Big data, bigger dilemmas: A critical review. *Journal of the Association for Information Science and Technology*, 1523-1545.
- Lazer, D., Kennedy, R., King, G., and Vespignani, A. (2014). The parable of Google flu: Traps in big data analysis. *Science* 343(6176) (March 14), 1-7.
- Lesk. M. (2013). Big Data, Big Brother, Big Money. *IEEE Security & Privacy*. 85-59.
- Wigan, M.R. and Clarke, R. (2013, June). Big data's big unintended consequences. *Computer*. 46-53.

SECTION III: DOMAINS

Class 9: Big data and business

Readings

- Bloem, J., van Doorn, M., Duivesteyn, S., van Manen, T., and Ommeren, E. (2012). Creating clarity with big data. *The Sogeti Trend Lab. VINT*. 1-32.
- Gopalkrishnan, V., Steier, D., Lewis, H., Guszczka, J. (2012). Big Data, Big Business: Bridging the Gap. *BigMine'12*, August 12, 2012, Beijing, China. 7-11.
- MasterCard. (2013). Around the world in 5 personas: How global consumers think about their data online. 1-10.
- Neamtu, R., Ahsan, R., Stokes, J., Hoxha, A., Bao, J., Gvozdenovic, S., Meyer, T., Patel, N., Rangan, R., Wang, Y., Zhang, D., and Rundensteiner, E.A. (2014). Taming big data: Integrating diverse public data sources for economic competitiveness analytics. *Data4U '14*, 25-28.
- Schull, N.D. (2013). The touch-point collective: Crowd contouring on the casino floor. *Limn* (2), 1-7.
<http://limn.it/the-touch-point-collective-crowd-contouring-on-the-casino-floor/>

Class 10: Big data and business analytics

Readings

- Ayankoya, K., Calitz, A., and Greyling, J. (2014). Intrinsic relations between data science, big data, business analytics and datafication. SAICSIT2014, Centurion, South Africa. 192-198.
- Chen, H., Chaiang, R.H.L., and Storey, V.C. (2012). Business intelligence and analytics: From big data to big impact. *MIS Quarterly*, 36(4), 1165-1188.
- Davenport, T.H. (2013). Analytics 3.0. *Harvard Business Review*. December. 1-9.
- Provost, F. and Fawcett, T. (2013). *Data Science for Business*. Ch. 1: Introduction: Data-Analytic thinking. O'Reilly Publishing. 1-17.
- Waller, M.A. and Fawcett, S.E. (2013). Data science, predictive analytics, and big data: A revolution that will transform supply chain design and management. *Journal of Business Logistics*, 34(2), 77–84.

Class 11: Big data and education

Readings

- Long, P.D. and Siemens, G. (2011). Penetrating the Fog: Analytics in Learning and Education. *Educes Review*. 1-6.
- Marsh, J.A., Pane, J.F., and Hamilton, L.S. (2012). Making Sense of Data-Driven Decision Making in Education: Evidence from Recent RAND Research. *RAND Education*. 1-18.
- Picciano, A.G. (2012). The Evolution of Big Data and Learning Analytics in American Higher Education. *Journal of Asynchronous Learning Networks*, 16(3), 9-20.
- Siemens, G. (2013). Learning analytics: The emergence of a discipline. *American Behavioral Scientist*, 57(10) 1380–1400.
- West, D.M. (2013). Big data for education: Data mining, data analytics, and web dashboards. *Governance Studies at Brookings*. 1-11.

Class 12: Big data and government

Readings

- Ansolabehere, S. & Hersh, E. (2012). Validation: What big data reveal about survey misreporting and the real electorate. *Political Analysis*, 20(4), 1-31.
- Bertot, J.C. and Choi, H. (2013). Big data and e-Government: Issues, policies, and recommendations. *The Proceedings of the 14th Annual International Conference on Digital Government Research*. 1-10.
- Bertot, J.C., Butler, B.S., and Travis, D.M. (2014). Local big data: The role of libraries in building community data infrastructures. *Digital Government '14*. 1-7.
- Donovan, A., Finn R., and Wadhwa, K. (editor), Oruc, S., Werker, C., Cunningham, S.W., Vega Gorgojo, G., Soylu, A., Roman, D., Akerkar, R., Garcia, J.M., Lammertant, H., Galetta, A., De Hert, P., Vrije, Grumbach, S., Faravelon, A., and Ramirez, A. (2014). Deliverable D2.1: Report on legal, economic, social, ethical and political issues. 1-102.

Harrison, T.M. (2014). Using Big Data for Digital Government Research. *Digital Government '14*. 309-310.

Kim, G.H., Trimi, S., and Chung, J.H. (2014). Big-data applications in the government sector. *Communications of the ACM*. 57(3). 78-85.

O'Hara, K. (2012). Transparency, open data and trust in government: Shaping the infosphere. *WebSci '12*. 223-232.

Class 13: Big data and science

Readings

Callebaut, W. (2012). Scientific perspectivism: A philosopher of science's response to the challenge of big data biology. *Studies in History and Philosophy of Biological and Biomedical Sciences* 43 (2012) 69–80.

Demchenko, Y., Grosso, P., de Laat, C., and Membrey, P. (2013). Addressing big data issues in scientific data infrastructure. 2013 International Conference on Collaboration Technologies and Systems (CTS), 48-55.

Fricke, M. (2015). Big data and its epistemology. *Journal of the Association for Information Science and Technology*, 66(4), 651–661.

Tufekci, Z. (2013). Big data: Pitfalls, methods, and concepts for an emergent field. 1-24. Draft paper available <http://ssrn.com/abstract=2229952>

Zaslavsky, A., Perera, C., and Georgakopoulos, D. (2013). Sensing as a service and big data. 1-8.

Class 14: Big data and health

Readings

Chawla, N.V. and Davis, D.A. (2013). Bringing big data to personalized healthcare: A patient-centered framework. *Journal of General Internal Medicine* 28(Suppl 3):660–665.

Ewing, E.T., Gad, S. and Ramakrishnan, N. (2013). Gaining insights into epidemics by mining historical newspapers. *Computer*. 68-72.

Groves, P., Kayyali, B., Knott, D., and Van Kuiken, S. (2013). The 'big data' revolution in health care: Accelerating value and innovation. McKinsey and Company. 1-22.

Hay, S.I., George, D.B., Moyes, C.L., and Brownstein, J.S. (2013). Big data opportunities for global infectious disease surveillance. *PLOS Medicine*, 10(4), 1-4.

Manyika, J., Chui, M., Groves, P., Farrell, D., Van Kuiken, S., and Doshi, E.A. (2013). Open data: Unlocking innovation and performance with liquid information. Ch. 6: Health. McKinsey Global Institute, McKinsey Center for Government, McKinsey Business Technology Office. 79-90.

McGregor, C. (2013, June). Big data in neonatal intensive care. *Computer*. 54-59.