

Course Number/Title: COMM 408: Big Data and Communication Tools
University of Michigan-Ann Arbor
Department of Communication and Media
Winter 2023
Time/Dates: Tu/TH 5:30-7:00 PM In-Person
Location: G444A Mason/Angell Hall
Instructor Name: Corinne Jones
Instructor Email: [REDACTED] or Canvas Messenger (preferred)
Office Hours: 4:15-5:15 PM Tu/Th
Office Location: North Quad 5427, or Zoom Office hours (please schedule in advance)

****Please note, this syllabus is subject to change at the discretion of the instructor. Changes will be announced in class and as announcements on Canvas.***

Many people assume “big data” sets are neutral purveyors of truth, and students in humanities and social sciences communication fields increasingly use freely available tools to track website “hits” and to scrape “public” information from social media sites. Often, they use this data to create data visualizations. However, is this data really an accurate representation of the digital landscape, and what are the ethical implications of it? Students in this course learn to use some simple and freely available tools to collect and analyze “public” data alongside critical analytic frameworks to communicate about this data accurately and ethically. In the first half of the class, students compose two smaller papers about their analyses. In the second half of the class, students write a position paper about the limits of big data.

*Some of this course draws from the work of Salter, A., Stanfill, M., Giroux, A. L., Vie, S., deWinter, J., Flores, L., Blodgett, B., Steele, C. K., & Kane, L. (2020). Understanding Digital Culture. Understanding Digital Culture. <https://understandingdigitalculture.hcommons.org/institute-overview/>

*While we will talk about how to use some basic tools in this class, since tools are always changing, the purpose of this class is to give you critical thinking skills that you can use as the tools change.

Course Outcomes

By the end of this course, students will:

- (1) Define big data and the ways it is used outside of this class and impacts our everyday lives and communication
- (2) Explain some of the challenges big data poses for communications researchers.
- (3) Use basic data tools to collect data and create simple data visualizations to accurately and ethically represent that data.
- (4) Critically evaluate and analyze big data and data visualizations and add qualitative analysis to compose well-supported arguments.
- (5) Develop critical and reflexive ethical principles for the use of big data.

Required Texts

- All reading material will be available through U-M’s library or PDFs through Canvas.

Other Required Materials

- Computer, Internet access, and familiarity with computer programs and technology
- A word processing program (e.g., Microsoft Word)

- Some technical skill and a basic understanding of Excel.
- A PDF reader (e.g., Adobe Reader)
- All papers must use MLA or APA formatting (See the Purdue OWL site <http://owl.english.purdue.edu/owl/>.)
- Having your own laptop or computer will be helpful for some assignments, but I have designed the class so that it can be completed without your own computer.

Grade Breakdown:

Participation* (24 required classes x 0.5 point each), 12%	12 points
Define “big data,” uses, and implications (week 1, in-class), C/I, 2%	2 points
(Re)define “big data,” uses, and implications (week 12, in-class), C/I 2%	2 points
Reading Responses, (16 Due x 1.5 points), C/I, 24%	24 points
Peer and Self Review, (3 Due x 2 points each), C/I 6%	6 points
Website Tracking Paper, 15%	15 points
Social Media Paper 15%	15 points
Final Paper, 16%	16 points
Group Presentation, 8 %	8 points
Final Grade out of 100 points	

* Participation is based on in-class work. Since this class is very technological and workshop focused, “participation” is crucial in this class. However, I recognize there are multiple forms of participation, and I understand that attendance must be flexible. If you foresee any issues, please let me know as soon as possible so that we can make other arrangements.

* “C/I” = “Complete/Incomplete”

Grading Scale:

A	94-100	B-	80-83	D+	67-69
A-	90-93	C+	77-79	D	64-66
B+	87-89	C	74-76	D-	60-63
B	84-86	C-	70-73	F	59 and below

If the grades are between cutoff percentages at the end of the semester, I will round the final grade up or down. If the grade is a 93.45%, I round that down to 93% for an A-. If the grade is 93.55%, I round that up to 94% for an A.

Assignment Descriptions

Participation: “Participation” is based on in-class work, and it is crucial since we have so many workshops this semester. However, I recognize there are multiple forms of participation, and I understand that attendance must be flexible. If you foresee any issues, please let me know as soon as possible so that we can make other arrangements. You can miss up to three classes this semester, no questions asked, with no penalty to your grade. **You will lose points for being excessively late.** However, you will be responsible for what we do in class, so I highly suggest that you contact me.

Define and (Re)Define “big data,” uses, and implications, Complete/Incomplete: Since one of the outcomes of this class is to define “big data,” its uses, and implications outside of this class, we will define “big data” and look at some of the possible uses of big data in the beginning of this class, and then we will revisit those definitions, uses, and implications at different

intervals throughout the class. This assignment is completed in class. The purpose is to scaffold you into the final assignment, in which you will define big data based on our course readings and your own experiences.

Reading Responses (1.5 point each, 24 points total), Complete/Incomplete: These are structured responses that you will complete for readings this semester to help scaffold you into major assignments. You will respond to specific questions (listed in Canvas and in the syllabus below). They are **at least 200** words long and they are graded on completion. They should answer all parts of the question for full points.

Peer Reviews, (2 points each, 6 points total) Complete/Incomplete. In class, we will complete peer reviews for our papers. They are graded based on completion. If you complete all parts, you receive full credit. Please turn in a draft (complete or incomplete) of your on Canvas **before class**. Peer reviews are completed in class. After you receive feedback from your peers, please write a brief revision plan and show me before you leave.

Website Tracking Paper: Based on the things that you might use big data research for in your lives outside of this class (for your jobs or otherwise), your **main task** in this report is to make a **concrete, specific, and arguable claim** that is backed up by **evidence** (including some sort of data visualization, and **two** scholarly, peer-reviewed sources from the library) that answers ***How do search engines and web results affect how people can communicate and find information? Who is impacted, and why does this matter?*** The paper will be **750-1000** words, and it will use information from either Semrush or Google Trends, considering both the limitations of the data and the ethical questions that come up with this data. The paper will use an academic writing style and prose. They will have a clear thesis, relevant background information, a research question, methods, evidence, analysis, and a conclusion.

Social Media Paper: Based on the things that you might use big data research for in your lives outside of this class (for your jobs or otherwise), your **main task** in this report is to make a **concrete, specific, and arguable claim** that is backed up by **evidence** (including some sort of data visualization, and **two** scholarly, peer-reviewed sources from the library) that answers ***How do trends and search results on social media (specifically, TikTok and YouTube) affect how people can communicate and find information? Who is impacted, and why does this matter?*** The paper will be **750-1000** words, and it will use information from either TikTok or YouTube, considering both the limitations of the data and the ethical questions that come up with this data. The paper will use an academic writing style and prose. They will have a clear thesis, relevant background information, a research question, methods, evidence, analysis, and a conclusion.

Final Paper: All semester we have been reflecting on the definitions, implications, and ethical principles around “big data.” This is your chance to clearly articulate what you learned in this class, why it matters, and how it will inform your actions moving forward. Your **main task** is to write a double spaced **750-1,000 word** essay that answers ***What is “big data”? What are the implications of “big data” in our everyday lives? What ethical principles do you believe that companies and individuals should employ when using “big data”?*** To answer this question, you will be backed up by **at least four scholarly sources** as testimonial evidence. You may also draw from your own experiences in this class. The paper will use an academic writing style and

prose. They will have a clear thesis, relevant background information, a research question, methods, evidence, analysis, and a conclusion.

Group Presentations: For these presentations you will work in six teams to define the six major challenges of big data, according to boyd and Crawford. Your **main tasks** are to (1) summarize and define the challenges about which you have chosen to present by using some sort of digital visual aid (Powerpoint, etc.) (2) come up with some examples of how these challenges applied to your research during the semester, (3) come up with some examples of how these challenges may apply to the types of things you might use big data for in your daily lives, and (4) answer any questions your peers may have.

Syllabus Statements [altered for institutional context]

Calendar

To view the full academic calendar, please visit:

https://ro.umich.edu/sites/default/files/calendar/pdfs/Cal_2022-2023.pdf

Tentative Schedule: Tentative Schedule:

Week	Readings (Due by class time)	Guiding Questions and In-Class work	Assignments (Due by class time, except where noted)
Unit 1: Introduction and Website Tracking/SEO			
Week 1: Jan 5: Th.		<p>Introduction to course.</p> <p>Go over Claims and Evidence handouts and look at assignments.</p> <p>How might you use “big data” in in communication and media outside of this class? Since people use "big data" for so many different things, thinking critically about what links you look at will help you make the most of this class.</p> <ul style="list-style-type: none"> • Tell Stories with Data • Data Analytics: A New Form of Communication • Big Data and Marketing • Big Data and Media Reporting • Big Data and Politics • Big Data and Risk/Health communication 	<p>Participation 1</p> <p>Define “big data” and briefly outline how you might use “big data” in-class survey. Graded on completion.</p>

		<ul style="list-style-type: none"> • Big Data and Sports Communication • Big data and Non-Profit Communication • Big Data and Journalism 	
Week 2: Jan 10: Tu.	danah boyd & Kate Crawford (2012) Critical Questions for Big Data, Information, Communication & Society , 15:5, 662-679, DOI: 10.1080/1369118X.2012.678878	What is big data? Get into 6 groups for Presentation. Your group will also give an informal presentation defining 1 of 6 key phrases next time to give you a chance to get acquainted.	Participation 2 Response 1: How do boyd and Crawford define “big data”? How is their definition different or the same from how you defined it?
Week 2: Jan. 12 Th.	Hepworth, K. (2017) Big Data Visualization: Promises and Pitfalls , <i>Communication Design Quarterly</i> 4(4), 7-19, doi: 10.1145/3071088.3071090 PiktoChart. (n.d). The 10 Essential Types of graphs and when to use them. , https://piktochart.com/blog/types-of-graphs/ (also in Canvas as Piktochart_Graphs).	How do we (ethically) represent big data? What can different kinds of visualizations tell us?	Participation 3 Response 2: Please define data visualization and tell me why it is impactful and important for big data. Please explain what a line chart, pie/donut chart, and bar graph show and how they are different? (since we will see some of those types of charts in this class).
Week 3: Jan 17 Tu.	Read for definitions: <ul style="list-style-type: none"> • SproutSocial on CPM • Semrush on CPC and PPC 	Where does “Big data” come from? In groups, define: <ul style="list-style-type: none"> • CPM 	Participation 4 Response 3:

	<ul style="list-style-type: none"> • SproutSocial on CTR and Semrush on CTR • Semrush on SOV • Semrush on “keywords” • “negative keywords.” <p>(All in Canvas as PDFs as Keywords_DigitalEconomy1-6).</p> <p>Watch: The Rise of Surveillance Capitalism by Mark Weinstein</p>	<ul style="list-style-type: none"> • CPC/PPC • CTR • SOV • “Keywords” • “Negative keywords” <p>Work through the first part of Semrush_Handout in Canvas</p>	<p>Please define: CPM, CPC/PPC, CTR, SOV, and “Keywords,” and “negative keywords.”</p> <p>To the best of your ability, please explain how you understand the digital economy, and how “big data” is generated.</p>
<p>Week 3:</p> <p>Jan 19 Th.</p>	<p>Andrejevic, M., & Gates, K. (2014). Big data surveillance: introduction. <i>Surveillance & Society</i>, 12(2), 185–196. https://doi.org/10.24908/ss.v12i2.5242</p>	<p>What is big data?</p> <p>What do algorithms have to do with big data, especially in terms of recommendation systems?</p> <p>Work through the second part of Semrush_Handout in Canvas</p>	<p>Participation 5</p> <p>Response 4: How do Andrejevic and Gates define “big data”? How is their definition different or the same from how boyd and Crawford defined it and how you defined it? Why are algorithms important for how we define and interpret “big data”, especially in terms of recommendation systems?</p>

<p>Week 4:</p> <p>Jan. 24</p> <p>Tu.</p>	<p>Markham, A., Tiidenberg, K., Herman, A. (2018). Ethics as Methods: doing ethics in the era of big data research - introduction" <i>Social media and society</i>, 4(3), p. 1-9.</p> <p>https://doi.org/10.1177/2056305118784502</p>	<p>Ethical Reflection Worksheet completed in class for participation.</p> <p>People as data points?</p> <p>How to include data visualizations in your papers.</p> <p>Looking ahead to Google Trends.</p>	<p>Participation 6</p> <p>Turn in ethical reflection in class during class time for participation points.</p>
<p>Week 4:</p> <p>Jan. 26</p> <p>Th.</p>	<ul style="list-style-type: none"> • Rogers, Simon. 2016. What is Google Trends Data—And what does it mean? <i>Medium</i>. 2016. • Google. N.d. The Homepage Explained. <i>Google</i> • Google. N.d. Trends Top Charts Explained. <i>Google</i>. • Google. N.d. The Homepage Explained. <i>Google</i> • Google. N.d. Google Trends: Understanding the Data. • Google. N.d. Troubleshooting. <i>Google</i>. <p>All VERY short readings.</p> <p>(All in Canvas as Google_Trends_Definitions_1-6)</p>	<p>Work through Google Trends. Handout in Canvas</p>	<p>Participation 7</p> <p>Response 5:</p> <p>Please explain to the best of your ability, how Google calculates the 0-100 numbers on the Google Trends graph. What do those numbers represent? What data is excluded from Google Trends? Please explain, to the best of your ability, the difference between topics and queries. Please explain, to the best of your ability, the definition of “rising” terms and “top” terms.</p>

<p>Week 5:</p> <p>Jan. 31 Tu.</p>	<p>Safiya Noble (2018). <i>Algorithms of Oppression</i>. Chapter 1: A Society Searching</p>	<p>First part of class: Talk about Safiya Noble's arguments.</p> <p>Second part of class: Work through Google Trends. Handout in Canvas</p>	<p>Participation 8</p> <p>Response 6: To the best of your ability, please explain why the results that appear that Google's recommends are not necessarily the "best," according to Noble. How do algorithms and "big data" play a part in this process? What are the social impacts/who do these results prioritize?</p>
<p>Week 5:</p> <p>Feb. 2 Th.</p>		<p>Peer Review</p>	<p>Participation 9</p> <p>Turn in a draft of your paper on Canvas and have a draft of the paper accessible to you during class (on paper, or on a computer that you are willing to share). Show me your revision plan before you leave class for points.</p>
<p>Week 6:</p> <p>Feb. 7 Tu.</p>	<p>Gitelman, L. (2016) Raw data is an oxymoron. (PDF in Canvas as "Gitelman RawData")</p>	<p>First part of class:</p> <p>Talk about Gitelman, and why "raw data is an oxymoron."</p>	<p>Participation 10</p> <p>Response 7: What does Gitelman mean</p>

		<p>What does this mean for how we interpret our data?</p> <p>Second part of class:</p> <p>What kind of data visualization are you using in your paper? Why is that specific visualization appropriate?</p> <p>Go over citational practices, and talk about how to include figure labels on the final paper.</p>	<p>that “raw data is an oxymoron”?</p> <p>What can her point reveal about the types of “myths” we hold about “big data” (from boyd and Crawford)?</p>
<p>Week 6:</p> <p>Feb. 9 Th.</p>		<p>Looking ahead to Paper 2 and in-class work toward final paper.</p>	<p>Participation 11</p> <p>Website Tracking Paper Due by class time</p>
Unit 2: Social Media Paper			
<p>Week 7:</p> <p>Feb. 14 Tu.</p>	<p>Jordan Reichel. (2020). Social Media and Participatory Culture. <i>Medium</i>.</p> <p>(In Canvas as “Reichel_ParticipatoryCulture”)</p> <p>Kalantzis and Cope. (n.d.) Jenkins on Participatory Culture. (In Canvas as “Kalantzis and Cope_ParticipatoryCulture”)</p> <p>Watch: TEDxNYED – Henry Jenkins – 03/06/10</p>	<p>First part of Class:</p> <p>How is social media “big data” different than “big data” about websites and Google Trends? What is “participatory culture”?</p> <p>Second Part of Class:</p> <p>Begin working through TikTok Handout.</p>	<p>Participation 12</p> <p>Response 8:</p> <p>How is social media “big data” different than “big data” about websites and Google Trends? What is “participatory culture” and what do these authors say are its benefits?</p>
<p>Week 7:</p>	<p>Christian Fuchs, Social Media: A Critical</p>	<p>First part of Class:</p>	<p>Participation 13</p>

Feb. 16 Th.	<u>Introduction</u> . Chapter 3 (In Canvas as “Fuchs_ParticipatoryC ulture”)	<p>What are the critiques of “big data” and “participatory culture”?</p> <p>Who “owns” big data? (Quick write in class)</p> <p>Second Part of Class:</p> <p>Work through second part of TikTok Handout.</p>	Response 9: What does Fuchs says about “participatory culture”? Why does Fuchs disagree with Jenkins and the previous authors? Briefly, who “owns” data? Is it you? People who reuse your content? Companies? Governments?
Week 8: Feb. 21 Tu.	Catherine D'Ignazio and Lauren F. Klein. <u>Data Feminism:</u> Introduction.	<p>First Part of Class: Talk about data feminism and what it is.</p> <p>Ethical Reflection Worksheet completed in class for participation.</p> <p>Second part of class: Go over YouTube Data handout in Canvas that walks us through how to use:</p> <p>Digital Methods Initiative. (n.d.). <i>DMI Tools</i>. Digital Methods.Net. https://wiki.digitalmethods.net/Dmi/ToolDatabase</p> <p>Work through handout in Canvas (“NetVizz_YouTube_Handout”)</p> <p>Supplementary Handout: NEH YouTube Tutorial here.</p> <p>What are your selection criteria? Why are you selecting these search terms over others?</p>	Participation 14 Response 10: Please define data feminism. How do D’Ignazio and Klein show that data can be used for good? For bad?
Week 8:			Participation 15

Feb. 23 Th.	Ribeiro, M. et al. (2020) Auditing Radicalization Pathways on YouTube . Proceedings of the 2020 Conference on Fairness, Accountability, and Transparency, p. 131 – 141, DOI: 10.1145/3351095.3372879	Second part of class: Go over YouTube Data handout in Canvas that walks us through how to use: Digital Methods Initiative. (n.d.). <i>DMI Tools</i> . Digital Methods.Net. https://wiki.digitalmethods.net/Dmi/ToolDatabase Work through handout in Canvas (“NetVizz_YouTube_Handout”) Supplementary Handout: NEH YouTube Tutorial here . What are your selection criteria? Why are you selecting	
Week 9: Feb. 28 Tu. Spring Break (Feb. 25 – Mar. 6)	Spring Break (Feb. 25 – Mar. 6)		
Week 9: Mar. 2 Th. Spring Break (Feb. 25 – Mar. 6)	Spring Break (Feb. 25 – Mar. 6)		
Week 10: Mar. 7 Tu.	Savolainen, Laura. “The Shadow Banning Controversy: Perceived Governance and Algorithmic Folklore.” <i>Media, Culture & Society</i> , vol. 44, no. 6, 2022, pp. 1091–109, https://doi.org/10.1177/01634437221077174 .	First part of class: Discuss Savolainen. Second part of class: In-class practice making Word Clouds and counting words. We will be using Voyant Tools .	Participation 16 Response 11: Even if we do not know how algorithms make sense of big data, why are algorithmic beliefs/”folklore”, such as those

		(See “Voyant_WordCloud_handout” in Canvas)	<p>around “shadowbanning” important and how do they connect to boyd and Crawford’s point that big data are also “myth”?</p> <p>Additionally, how might “shadowbanning” (real or imagined) impact the types of data you are seeing in your social media papers?</p>
<p>Week 10: Mar. 9 Th.</p>	<p>Feisler, C. & Proferes N. (2018). “Participant” Perceptions of Twitter Research Ethics. <i>Social Media + Society</i>, 4 (1), 1-14. Doi: https://doi.org/10.1177/2056305118763366</p>	<p>First part of class: Discuss Feisler and Proferes.</p> <p>Second part of class:</p> <p>In-class practice using concordance analysis to look at words in context.</p> <p>We will be using Voyant Tools. (See “Voyant_WordCloud_handout” in Canvas)</p>	<p>Participation 17</p> <p>Response 12: Would you be okay with a researcher using your social media posts? Under what circumstances? How do your beliefs about data ownership (see response from Feb. 16) affect your answer?</p>
<p>Week 11: Mar. 14 Tu.</p>		Peer review	<p>Participation 18</p> <p>Turn in a draft of your paper on Canvas and have a draft of</p>

			<p>the paper accessible to you during class (on paper, or on a computer that you are willing to share). Show me your revision plan before you leave class for points.</p>
<p>Week 11:</p> <p>Mar. 16 Th.</p>		<p>Now that we have practiced using big data, return to questions and links from first class about how you might use big data.</p> <ul style="list-style-type: none"> • Tell Stories with Data • Data Analytics: A New Form of Communication • Big Data and Marketing • Big Data and Media Reporting • Big Data and Politics • Big Data and Risk/Health communication • Big Data and Sports Communication • Big data and Non-Profit Communication • Big Data and Journalism <p>Use class time to search the university library and other sources find examples of how you might use big data outside of this class, or to find examples of how big data affects you.</p>	<p>Participation 19</p> <p>Mid-Term Definition of Big Data and Implications.</p> <p>Please turn in a link for an article or an example of how you might use big data after this class. Completed in class.</p>

Week 12: Mar. 21 Tu.		Paper due Come to class with the examples of how you might use big data in your life outside of this class to prepare for final paper.	Participation 20 Social Media Paper due by class time
Unit 3: (Re)defining Big data, uses, and implications			
Week 12: Mar. 23 Th.	<p>Read: Crawford et al. (2014). Critiquing Big Data: Politics, Ethics, Epistemology, <i>International Journal of Communication</i>.</p> <p>Revisit: Andrejevic, M., & Gates, K. (2014). Big data surveillance: introduction. <i>Surveillance & Society</i>, 12(2), 185–196. https://doi.org/10.24908/ss.v12i2.5242</p> <p>danah boyd & Kate Crawford (2012) Critical Questions for Big Data, <i>Information, Communication & Society</i>, 15:5, 662-679, DOI: 10.1080/1369118X.2012.678878</p>	<p>(Re)Defining Big Data</p> <p>How did you define “big data” in the beginning of this class? How do you define it now? Has your definition changed, or have you added to it? Why or why not?</p>	<p>Participation 21</p> <p>Response 13: Please revisit boyd and Crawford (2012) and Andrejevic and Gates (2014). (in Canvas). How does Crawford et al.’s (2014) definition differ or add to existing definitions. Reflect on the trajectory of this class, what you have learned, and their definitions and try to come up with your own definitions of big data.</p>
Week 13: Mar. 28 Tu.		Group Presentations for the six issues boyd and Crawford define with big data.	Participation 22
Week 13: Mar. 30 Th.	<p>Please look through AoIR’s IRE 1.0 IRE 2.0 IRE 3.0</p>	<p>Developing ethical principles.</p> <p>Revisit Markham et al., D’Ignazio and Klein, and Feisler and Proferes.</p> <p>Discuss AoIR and ethical principles that you found.</p>	<p>Participation 23</p> <p>Response 14: Which two ethical principles piqued your interest from</p>

	Find <i>two</i> ethical principles that are interesting to you, and come to class prepared to talk about them.		these examples of internet research ethics? Do you agree with them or disagree with them? Why?
Week 14: Apr. 4 Tu.	Edwards, Dustin. (2021). Critical infrastructure literacies and/as ways of relating in big data ecologies. <i>Computers and Composition</i> https://doi.org/10.1016/j.compcom.2021.102653 Watch: The Hidden Pollution of the Internet	How does big data affect the environment? Who can take action responsible?	Participation 24 Response 15: What are some of the ways that “big data” affects the physical world and environment? Who can take action? People? Companies? Governments?
Week 14: Apr. 6 Th.	Hummel et al. (2021). Data Sovereignty: A Review . <i>Big Data and Society</i> 8(1). 1-17. doi:10.1177/2053951720982012.		Participation 25 Response 16: To the best of your ability, please define data sovereignty. How would this concept apply outside of this class?
Week 15: Apr. 11 Tu.		Peer Review	Participation 26 Turn in a draft of your paper on Canvas and have a draft of the paper accessible to you during class (on paper, or on a computer that

			you are willing to share). Show me your revision plan before you leave class for points.
Week 15: Apr. 13 Th.		Work/Draft Day	Participation 27 Final papers due by the end of class time.
Week 16: Apr. 18 Tu.	Classes End		Participation 28
Apr. 19-23	Study Days		
Apr. 24-27	Final Examinations		