

ENG 395

Data and Society

Spring 2021

MW 6:00 - 7:15 pm

Online only

Course site: <https://lindsaythomas.net/eng395s21/>

Professor Lindsay Thomas

lindsaythomas@miami.edu, lxt308@miami.edu

Virtual office hours: MW 1:30 pm – 3:00 pm, and by appointment

- Please email me to schedule a time to meet. The best time to meet with me is during the above time slot, and slots will be reserved on a first-come, first-served basis. But if you can't make my virtual office hours, just let me know and we'll set up an appointment that works for both of us.

Course Description

Data is often considered the domain of scientists and statisticians. But the proliferation of data and databases across nearly all aspects of daily life – powering everything from the targeted advertisements you see when you go online to the fake news circulating on Facebook to the next financial recession – has made the study and understanding of the concept of data a vital everyday concern. This course provides an introduction to the meanings, uses, and politics of data today. Readings are drawn from literary and cultural studies, media studies, science and technology studies, sociology, information science, and the digital humanities. We will focus in unit 1 on the concept of data, examining its definition, history and relationship to quantification itself. In unit 2, we will explore how researchers in the humanities use data to study culture, and we will analyze existing datasets and create our own. Finally, unit 3 will center on algorithmic processes and what data does in the world. You will complete an algorithm audit for your final project, which will ask you to select and investigate a specific algorithmic process (i.e., Google search auto-completes, or autocorrects in iOS, etc.) in order to understand how it functions conceptually, the data it uses (or might use, if that information is unavailable), and what you see as its most important social consequences.

Student Learning Outcomes

- Summarize important features of the concept and uses of data in contemporary culture using real-world examples.
- Examine existing datasets to distinguish their most important features and articulate how they are used in specific applications and/or for specific purposes.
- Apply concepts and techniques in data modeling to design and construct a small dataset.
- Describe how data is used to study literature, film, images, and other objects of study in the humanities.
- Explain the concept of an algorithm and identify the uses, functions, and significance of specific algorithmic processes in contemporary culture.
- Demonstrate knowledge of the significance of concepts learned throughout the course by formulating original arguments in writing about the role and meaning of data in contemporary culture.

Required Course Materials

- All of our course reading will be available in our class Google drive folder or via our course website.

We will refer to the assigned reading every single day in class, and we will often read passages aloud together. Bring the assigned reading with you every day to class.

How This Course Will Work

This course will be delivered fully online. There are 2 ways to take this course: synchronously, and asynchronously. Everyone who is physically able to take the course synchronously (i.e., residing in a time zone where this is possible) will take the course synchronously. Each synchronous class session will be conducted via Zoom. Synchronous class sessions will be recorded for later viewing by those who are taking the course asynchronously. Each synchronous class session's chat will also be saved as another record of class discussion.

Our class is scheduled to meet two times per week, on Mondays and Wednesdays, and we will meet synchronously over Zoom each of those class days. If our semester is interrupted in some way, we will re-evaluate this schedule and some of our class sessions may become asynchronous.

A) If you are taking the course synchronously:

On Mondays and Wednesdays, you will sign in to our class Zoom session for our class discussion. I will make each day's slides or other class materials available in our shared Google drive folder. When we have group discussions in breakout rooms, sometimes your group will record your discussion/complete your discussion activities in a shared class Google doc. Sometimes designated reporters will report out to the rest of the class, and sometimes I will call on people from each group to report.

B) If you are taking the course asynchronously:

After each synchronous class meeting, you will have until our next synchronous class meeting to view the recording of our class and complete the activities announced during the recording. Often, these activities will be the same as those people did during class, but sometimes they will be different.

Grading

- Participation: 20%
- Response papers: 15%
- Dataset analysis: 15%
- Dataset creation: 15%
- Algorithm audit: 35%
 - Plan part 1: 5%
 - Plan part 2: 5%
 - Final draft: 25%

Additional details on all course assignments can be found via the "Assignments" page on our course website.

Course Digital Infrastructure

We will make use of multiple online systems and programs in this course: a course site, Zoom, Google drive, and Blackboard. The “Passwords and Login Instructions” doc in our shared class folder on Gdrive will contain the passwords and other information you need to use our course digital systems.

Course site:

We will use our course site to manage course information and our schedule. You will find an online version of our course calendar there (including the most up-to-date version of reading assignments and due dates), as well as a copy of our course syllabus. You will also find all course assignment sheets there.

Zoom:

All of our synchronous class sessions will take place over Zoom. If you don't have a Zoom account already, you will need to sign up for one via UMIT (<https://www.it.miami.edu/a-z-listing/zoom/>). You will need a laptop or a smartphone that can run Zoom, with a working microphone and camera.

You will receive our class Zoom call info via your UM email address, and it will also be listed in the “Passwords and Login Instructions” doc in our shared class folder (you will also be able to access Zoom information via Blackboard and our course site). We will use this Zoom call link and password for all class sessions. We will use a separate Zoom session link for office hours.

To access recordings of class sessions:

1. Select “Zoom Class Meeting Info” from the left-hand menu on our Blackboard site.
2. You will be taken to a screen with a link to our class Zoom meeting. This is how you will access this link throughout the semester.
3. Select “Cloud Recordings”. There, you will see a list of all of the class session recordings from the past 30 days. This is also how you access the class chat from each session. A class session recording is automatically deleted after 30 days.

Google drive:

We will have a class Google drive folder, and at the beginning of the semester I will share you into this folder. We will use this folder to store slides and other materials from our class sessions. We will also use the class Google drive folder to distribute course readings. To protect your privacy, you will need to sign in to Google drive using your UM CaneID and password, and to use this account when working on materials related to this class. You will be shared into our class Gdrive folder via your UM email address.

Blackboard:

You will submit your assignments via Blackboard. I will also use Grade Center to record your grades on course assignments, and you will also be able to access recordings of class sessions and chats via Blackboard.

Class Participation

Participation makes up 20% of your grade in this class. It is assessed on the following metrics:

1. **Attendance.** This is a discussion-based class whose success depends on your consistent presence. When you're not here, it disrupts our collective work. This is only more important under the current circumstances, when we all need to rely on each other more than ever. Attendance includes completing asynchronous class activities. You may miss up to 4 class sessions/asynchronous activities for any reason without penalty, and you don't need to inform me of these absences in advance. Barring extended illness or emergency, any absence beyond this will lead to a drop not only in your participation grade, but also in your overall grade for the course. I will take attendance every day. If you would like to check on your attendance record, just ask. In general, if something comes up and you will need to be absent from class for an extended period, just communicate with me. We will work it out. Finally, I realize that wifi connection issues may occasionally interfere with your ability to attend class on a particular day. If this happens to you, you may attend the course asynchronously that day by viewing the synchronous class session and completing the asynchronous class activities for that day. You may also simply elect to miss class that day, and use it as one of your 4 allowed absences.
2. **Preparedness.** Sign in to class having completed the reading and ready to discuss it. Conducting classes over Zoom requires that I call on people during class discussion. Therefore, you should come to class expecting to be called on and ready to contribute.
3. **Contributions to class discussion.** These contributions can take multiple forms, including participation in small-group discussions, contributions to whole-class discussions, contributions to the class chat during class, and meeting with me in virtual office hours. It's perfectly fine to be shy (I am also very shy!), and I understand that Zoom often rewards the gregarious (this is also why I will call on people during discussion). But if you are shy, make sure to take full advantage of the many written and asynchronous forms of participation available to you in this class.
4. **Thoughtful and respectful engagement with all members of the class community.** This means giving the class your full attention while it is happening, and setting aside distractions, especially online distractions. I realize this is hard to do while you are attending a class online, but it is an important way of showing respect for others. This also means listening actively to everyone, acknowledging and interacting with the ideas of others, and speaking to others with respect and dignity. It means refraining from interfering with or dominating class conversation. Finally, it means refraining from posting confidential recordings or transcripts of class sessions on public forums (see below for more on that).

This semester is going to be hard and taxing for all of us. I recognize that many of you simply want things to go back to normal. I do, too -- desperately. Unfortunately, this is not possible right now. If there is something happening in your life that is affecting your work in this class and you feel comfortable talking to me about it, please do so. I am on your side and I want to work with you, not against you. I will extend the benefit of the doubt to you, and I ask that you please extend it to others in the class in turn.

Zoom Etiquette

Our class sessions will be conducted over Zoom. During class sessions, I ask that you please mute your microphone unless you are speaking. The university recommends that I require you to turn on your video during class, but I find that requirement distasteful. While I strongly encourage you to turn your

video on if possible because I would like to be able to see your face, I do not feel it is my place to require it, and wifi connection issues can sometimes make it impossible.

Please do not attend class sessions while you are laying down in bed or on the couch.

I encourage you to participate in discussion via Zoom's chat feature during class. You should feel free to ask and answer questions about the class or what we are discussing there. You should also use the chat feature to place yourself on stack if you would like to speak during class discussion. Please remember that the class Zoom chat will be automatically saved after every class session for future reference, so do not post anything there that you are not comfortable with everyone being able to read after the fact.

Finally, I ask that you refrain from using Zoom's private message feature to try to message me during class. Because I'll be busy leading discussion, it is unlikely that I will see your message. Instead, email me. (In general, just be careful with Zoom's private message feature. When Zoom chats are active, I find it can be difficult to tell which messages are private and which are public. Please be careful and respectful.)

Discussion Guidelines: On trust, on difficulty and on being wrong

Class discussion is a process of creating knowledge together. This is not a lecture course, and the success of this class does not depend on me delivering information to you as you sit there passively. Rather, it depends on your continual engagement and participation in a collective project of knowledge making. This is why your participation in class is so vital. This is hard work -- especially online -- and it *only* works if we all commit to respecting each member of this class through our words and our actions. This doesn't mean that you agree with everyone else, or with me, or with the author of the piece we are reading. It does mean that we all commit to supporting and trusting one another.

As you know, synchronous class discussion will be recorded for viewing by those who will be taking this class asynchronously. You will break our collective trust if you share recordings of class conversations with people outside of this class, or on public forums. Doing so will result in a failing grade in this class. Do not do this.

The things we read in this class will sometimes be difficult or boring or unfamiliar. It's ok to not understand them when you first read them! It's ok that you've never taken a class like this before, and/or that you feel frustrated by the reading. Figuring difficult things out together as a group is what class discussion is for. We will all be entering into uncharted territory in this class, and some of the ideas we discuss may confuse you, or make you uncomfortable and angry, or both. We will all experience moments of panic, of flailing, and of error. I ask that you accept this chance to make mistakes, and that you extend this acceptance to your classmates as well.

Reading Expectations for This Class

The reading in this class will require your time and attention. One of the most common comments I receive on my teaching evaluations at the end of every semester is that I assign "a lot" of reading (but students also generally say that doing all of the reading is "worth it"). You'll be expected to do somewhere in the neighborhood of 45 minutes of reading a night (including the weekend!) for this class. (The estimated reading speed for this calculation is 138 words a minute, based on some research which I'm glad to discuss with you!) Sometimes, you'll be asked to read a little more than this; sometimes, you'll be asked to read a little less. When readings are more difficult to read or understand, I assign fewer pages.

A few tips for staying on top of the reading: (1) Read every day. Make some leisure time for this, and don't try to cram all the pages into the night or morning before class. (2) Read at odd hours. Read between classes. Read before you fall asleep. Read instead of endlessly scrolling your social media of choice. (3) Read ahead. If you ever find yourself with some extra prep time, do your future self a favor.

Late Work

All assignments should be submitted on the due date and in the format indicated in the assignment sheet. Due dates are important because they allow me to organize the work of responding to your assignments in ways that best contribute to your learning. However, life also sometimes gets in the way (especially this year). If you find that you need an extension on a particular assignment, please contact me as soon as possible to arrange an alternative due date (please try to contact me at least 24 hours before the assignment is due, but the sooner the better).

On a personal level, like everyone else, I dislike being lied to. I am not here to judge you; rather, I want to work with you, and I actually do want you to learn. I would much rather that you take a few extra days to complete an assignment to the best of your ability than you turn in something thrown together because you did not have the time you needed to devote to it. You do not need to concoct elaborate stories if requesting an extension on an assignment, or if your work will be late. Simply be honest with me about needing more time and we will work it out.

Email

All students are required to check their official UM email accounts regularly. I will send course information and announcements through email, and we will all be relying even more than usual on email this semester. I endeavor to respond to all emails that you send me within 24 hours during the week and within 48 hours over the weekend, but please do not send me urgent emails regarding your assignments in the hours before they are due and expect a reply.

If you have more involved questions about course material, assignments, or policies, the best thing to do is to talk to me in a one-on-one appointment over Zoom. In fact, the best thing to do in almost any situation that affects your class work is simply to talk to me about it. I am happy to answer questions about the course via email, but I recommend that more involved questions and conversations be conducted over Zoom. I am also happy to read and discuss advance drafts of your assignments with you over Zoom, but I will not read and comment on drafts of assignments via email before they are due.

Technological Failures Are Not Emergencies

Technological failures and mishaps – file corruption, computer crashes, wifi connection problems, uploading the wrong file to Blackboard – are predictable facts of twenty-first century life. They happen all of the time and are thus NOT emergencies. For this course, for all of your courses, for your career, for the rest of your life on this earth, and especially now that most everything has moved online for the time being, you need to develop strategies that take such failures into account. Start your work early, save it often, and save backup copies of important documents off-site on an external hard drive or in the cloud using services like Dropbox and Google drive. Technological failure or mishap – including uploading the wrong file to Blackboard – is not an excuse for late or unfinished work (although it may very well make it difficult for you to attend class on a given day – I understand this).

Please note that I will grade whatever you upload to Blackboard for grading. It is your responsibility to turn in the correct version of your assignment for grading.

Plagiarism

The principle of academic integrity is taken very seriously and violations, especially plagiarism, are treated gravely. In terms of this course, academic integrity means that when you are responsible for a task, you – and no one else – will perform that task. When you rely on someone else's work in performing an aspect of that work, you will give full credit in the proper, accepted form. Turning in work for this class that you have not done yourself or that you have previously completed for other courses is a violation of academic integrity. The University of Miami's honor code can be found here: <https://doso.studentaffairs.miami.edu/honor-council/honor-code/index.html>. Ignorance of what constitutes academic dishonesty is not an acceptable excuse for academic dishonesty.

Violations of academic integrity constitute grounds for failure of the course and possible expulsion from the university. If you have questions about what constitutes plagiarism, please contact me.

Writing and Tutoring Resources

The Writing Center offers **free**, one-on-one assistance with any aspect of the writing process. I strongly suggest you take advantage of this resource. You can schedule a consultation via their website at <https://english.as.miami.edu/writing-center/index.html>.

The Camner Center for Academic Resources also offers **free** tutoring for UM students. You can learn more and schedule an appointment here: <https://camnercenter.miami.edu/tutoring-services/index.html>.

Counseling Resources

UM offers counseling **free of charge** to students who have already paid the Health and Counseling Center fee. The Counseling Center website is <https://counseling.studentaffairs.miami.edu/index.html>, and you can make an appointment by following the steps outlined here: <https://counseling.studentaffairs.miami.edu/appointments/make-an-appointment/index.html>.

Resources for Students with Disabilities

It is important to me that all learning experiences be as accessible as possible. If you anticipate or experience academic barriers due to your disability (including mental health, learning disorders and chronic medical conditions), please let me know immediately so that we can privately discuss options. To establish reasonable accommodations, you also need to contact the Office of Disability Services, <https://camnercenter.miami.edu/disability-services/index.html>.

Schedule

Readings are due – meaning they should be completed – on the dates indicated. If readings are not linked on our course calendar, they can be found in our class Google drive folder.

The most accurate and up-to-date version of this calendar can be found on our course site. Use the online calendar to check on reading assignments, rather than this print version, since the print version of this syllabus will not be updated throughout the semester.

I reserve the right to change the course calendar as needed; adequate advance notice will always be given of any changes.

Unit 1: What is data?

Week 1

Monday, Jan 25

- Introductions

Wednesday, Jan 27

- Lisa Gitelman and Virginia Jackson, “Introduction” from *“Raw Data” is an Oxymoron* (2013)
- Kate Crawford, “The Hidden Biases in Big Data,” *Harvard Business Review* (April 1, 2013), <https://hbr.org/2013/04/the-hidden-biases-in-big-data>

Week 2

Monday, Feb 1

- Daniel Rosenberg, “Data Before the Fact,” from *“Raw Data” is an Oxymoron* (2013)

Wednesday, Feb 3

- Lauren Klein and Catherine D’Ignazio, “What Gets Counted Counts,” *Data Feminism* (2020), <https://data-feminism.mitpress.mit.edu/pub/h1w0nbqp/release/2>

Friday, Feb 5

- **Response paper 1 due**

Week 3

Monday, Feb 8

- Jacqueline Wernimont, Ch 2 “Counting the Dead,” *Numbered Lives* (2018)

Wednesday, Feb 10

- Jessica Marie Johnson, “Markup Bodies: Black [Life] Studies and Slavery [Death] Studies at the Digital Crossroads,” *Social Text* 36.4 137 (2018): 57-59

Friday, Feb 12

- **Response paper 2 due**

Unit 2: How do you create data and how can you use it to study culture?

Week 4

Monday, Feb 15

- “What is Data?”, University of Minnesota Libraries, <https://www.lib.umn.edu/datamanagement/whatdata>
- Explore *Slave Voyages*, <https://www.slavevoyages.org/>
 - Read David Eltis, “Trans-Atlantic Slave Trade -- Understanding the Database,” <https://www.slavevoyages.org/voyage/about#methodology/introduction/0/en/>
 - Read Introduction, Coverage of the Slave Trade, Nature of Sources, Cases and Variables, Data Variables, Age Categories, Dates, Names, and Skim the rest of the essay (see course site for individual links)
 - Familiarize yourself with the Trans-Atlantic Slave Trade section of the website: explore the maps, glance through the essays, look at the timeline and chronology section, etc.
- Hannah Anderson and Matt Daniels, “Film Dialogue from 2,000 screenplays, Broken Down by Gender and Age,” *The Pudding*, April 2016, <https://pudding.cool/2017/03/film-dialogue/>

Wednesday, Feb 17

- Heather Krause, “Data Biographies: Getting to Know Your Data,” March 27, 2017, <https://gijn.org/2017/03/27/data-biographies-getting-to-know-your-data/>
- Women Writer’s Project: <https://wwp.northeastern.edu/>
 - Editorial principles: https://wwp.northeastern.edu/about/methods/editorial_principles.html

Friday, Feb 19

- **Response paper 3 due**

Week 5

Monday, Feb 22

- Jenn Riley, NISO, “Understanding Metadata,” pgs 1-18, 38-41, https://groups.niso.org/apps/group_public/download.php/17446/Understanding%20Metadata.pdf
- Data from WhatEvery1Says project

Wednesday, Feb 24

- Kate Theimer, “Archives in Context and As Context,” *Journal of Digital Humanities* 1.2 (2012), <http://journalofdigitalhumanities.org/1-2/archives-in-context-and-as-context-by-kate-theimer/>
- Women Writer’s Project: <https://wwp.northeastern.edu/>
 - Read about their schema customization: <https://wwp.northeastern.edu/about/methods/customization.html>

Friday, Feb 26

- **Dataset analysis due**

Week 6

Monday, March 1

- Emily Dreyfuss, “‘ICE is Everywhere’: Using Library Science to Map the Separation Crisis,” *Wired*, June 25, 2018, <https://www.wired.com/story/ice-is-everywhere-using-library-science-to-map-child-separation/>
- Explore Mobilized Humanities, Torn Apart / Separados:
 - [Vol 1](#)
 - [Vol 2](#)

Wednesday, March 3 -- “Wellness Wednesday,” NO CLASS

Week 7

Monday, March 8

- Laura B. McGrath, “Compiling White,” *Los Angeles Review of Books*, January 21, 2019, <https://lareviewofbooks.org/article/compiling-white/>
- Alexander Manshel, Laura B. McGrath, and J. D. Porter, “Who Cares About Literary Prizes?”, *Public Books*, September 3, 2019, <https://www.publicbooks.org/who-cares-about-literary-prizes/>

Wednesday, March 10

- Hanna Wallach, “Big Data, Machine Learning, and the Social Sciences,” *Medium* (2014), <https://medium.com/@hannawallach/big-data-machine-learning-and-the-social-sciences-927a8e20460d>

Friday, March 12

- **Response paper 4 due**

Unit 3: How does data matter?**Week 8**

Monday, March 15

- Tarleton Gillespie “Algorithm,” from *Digital Keywords: A Vocabulary of Information Society and Culture* (2016)
- Kate Crawford and Vladan Joler, *Anatomy of an AI System* (2018), <https://anatomyof.ai/>

Wednesday, March 17

- Dataset creation workshop

Friday, March 19

- **Dataset creation due**

Week 9

Monday, March 22

- Safiya Noble, “Introduction: The Power of Algorithms,” and Ch. 1 “A Society, Searching,” from *Algorithms of Oppression* (2018)

Wednesday, March 24

- Safiya Noble, Ch. 2, “Searching for Black Girls,” *Algorithms of Oppression* (2018)

Week 10

Monday, March 29

- Alex Rosenblat, Tamara Kneese, and danah boy, “Algorithmic Accountability,” Workshop for The Social, Cultural, and Ethical Dimensions of “Big Data,” (2014)
- Nicholas Diakopoulos, “Algorithmic Accountability: Journalistic Investigation of Computational Power Structures,” *Digital Journalism* (2014)

Wednesday, March 31

- Christian Sandvig, Kevin Hamilton, Karrie Karahalios, and Cedric Langbort, “Auditing Algorithms: Research Methods for Detecting Discrimination on Internet Platforms” (2014)

Week 11

Monday, April 5

- Terry Nguyen, “How social justice slideshows took over Instagram,” *Vox*, August 12, 2020, <https://www.vox.com/the-goods/21359098/social-justice-slideshows-instagram-activism>
- “How TikTok recommends videos #ForYou,” *TikTok Newsroom*, June 18, 2020, <https://newsroom.tiktok.com/en-us/how-tiktok-recommends-videos-for-you>

Wednesday, April 7

- Rebecca Heilweil, “Right-wing media thrives on Facebook. Whether it rules is more complicated,” *Vox*, September 9, 2020, <https://www.vox.com/recode/21419328/facebook-conservative-bias-right-wing-crowdtangle-election>
- Simon Shuster and Billy Perrigo, “Like, Share, Recruit: How a White-Supremacist Militia Uses Facebook to Radicalize and Train New Members,” *TIME*, Jan 7, 2021, <https://time.com/5926750/azov-far-right-movement-facebook/>

Friday, April 9

- **Algorithm audit plan part 1 due**

Week 12

Monday, April 12

- Joy Buolamwini and Timnit Gebru, “Gender Shades: Intersectional Accuracy Disparities in Commercial Gender Classification,” The Conference on Fairness, Accountability, and Transparency, March 2018 (~20 min video)

Wednesday, April 14 -- “Wellness Wednesday,” NO CLASS

Week 13

Monday, April 19

- **Algorithm audit plan part 2 draft by class**
- Workshop

Wednesday, April 21

- Workshop

Friday, April 23

- **Revised algorithm plan part 2 due**

Week 14

Monday, April 26

- No class; individual meetings about algorithm audits

Wednesday, April 28

- Wrap-up

Algorithm audit due during finals week