



Privacy, Security, and Ethics for Big Data

INST408Y
Spring 2019

Course Description

The increasing number of networked information technologies—including internet of things (IoT), wearables, ubiquitous sensing, social sharing platforms, and other AI-driven systems—are generating a tremendous amount of data about individuals, companies, and societies. These technologies offer enormous benefits but also create enormous risks to individual privacy and national security. Further, the ease with which data can be collected from online sources, analyzed, and inferences drawn about individual users raises a wide range of ethical questions about these technologies, their creators, and their users.

In this course, students will evaluate major privacy and security questions raised by big data and related technologies. Students will learn about the history of research ethics and consider how ethical frameworks can and should be applied to digital data. They will work through case studies from real world scenarios to understand the complex interactions between data security, privacy, and ethics in modern businesses. This course is required for the InfoSci Cybersecurity and Privacy Specialization.

Learning Outcomes

After successfully completing this course, you will be able to:

- Identify and explain basic ethical and policy-based frameworks for working with big data and apply these frameworks to real-world cases.
- Explain differences between and shared values across data, ethics, and society.
- Identify situations where data is sensitive, assess the risks, and describe how various stakeholders could respond to those risks.
- Describe how to minimize privacy/security compromises through the data lifecycle (from collection through dissemination).
- Implement good security and privacy practices in personal data storage, use, and reporting.

Required Resources

Course website: elms.umd.edu

I use Canvas for all course readings, slide decks, study guides (when relevant), and communication. View the current week's Module for upcoming readings and assignments. I also post class-wide announcements through ELMS using Announcements. Make sure you have ELMS set up to forward Announcements to your email and/or regularly check your account to ensure you don't miss any class-related information.

You are required to download the TurningPoint "virtual clicker" app and set up an account to be used in class. For more information, see

<https://it.umd.edu/news/2017/clickers>

Dr. Jessica Vitak

Pronounced VEE-tack
she/her/hers
jvitak@umd.edu

Priya Kumar (TA)

she/her/hers
pkumar12@umd.edu

Class Meets

TuTh 9:30-10:45am

Office Hours

Vitak: TuTh 12pm-1pm
HBK #2117G and by
appointment

Kumar: Wednesdays by
appointment:

<https://calbird.com/priya/students>

Prerequisites

INST201 or permission of
instructor

Course Communication

Time-sensitive
announcements will be
posted on ELMS.

- Contact me via ELMS or by email (be sure to include "INST408" in the subject).
- Contact Priya regarding assignments
- Email correspondence should be like face-to-face correspondence. Include a salutation (Dear Dr./Prof Vitak) and end with your name.

I will reply to emails within 48 hours; if for any reason you have not heard from me after that time, please send me a follow-up email.

Campus Policies

It is our shared responsibility to know and abide by the University of Maryland's policies that relate to all courses, which include topics like:

- Academic integrity
- Student and instructor conduct
- Accessibility and accommodations
- Attendance and excused absences
- Grades and appeals
- Copyright and intellectual property

Please visit www.ugst.umd.edu/courserelatedpolicies.html for the Office of Undergraduate Studies' full list of campus-wide policies and follow up with me if you have questions.

Activities, Learning Assessments, & Expectations for Students

Course Activities

Your final grade will be based on the following components:

Learning Assessments	#	Points Each	Category Total
In class activities	10	2.5	25
Case study evaluations	5	7.5	37.5
Ethics Training	1	5	5
Managing Your Data	1	7.5	7.5
Terms of Service/Privacy Policy Analysis	1	10	10
Final Exam	1	15	15
Total Points:			100

In-Class Activities (2.5% each, 25% total): While attendance will not be taken in this course, students will regularly participate in class-based activities around that day's lecture. This may include individual or group work and/or may require participation using the TurningPoint mobile app for polling/interaction.

Because life happens, I recognize you may have to miss class from time to time; therefore, I will only count the top 10 activities toward your final grade. So if we have 13 in-class activities, the three lowest scores (including zeros) will be dropped. If you have an excused absence (e.g., religious holiday, doctor's note), I will allow you to make up the activity outside of class, as long as it is submitted within one week of returning to class.

Case Studies (7.5% each; 37.5% total): In this class, we'll be using case studies from real-world examples that highlight concepts from the class at the intersection of big data, ethics, and privacy/security. For each case study, you will be assigned a reading that overviews the case study, then you will complete a 700-1000 word write-up on the case study based on guiding questions I provide. I will provide six case study assignments throughout the semester and you will be asked to complete five (or you may complete all six and the lowest grade will be dropped).

Ethics Training (5%): To understand how research is conducted in university settings (and most organizations that have an internal review board), you will need to complete the basic CITI training in Social and Behavioral Responsible Conduct of Research (CITI stands for Collaborative Institutional Training Initiative). This involves reading and completing quizzes for several modules on data collection and management. You need a cumulative score of 80% across all quizzes to receive credit.

To receive credit, you must upload the certificate you receive for successfully completing training. Instructions on how to complete CITI training are found here (pdf): https://research.umd.edu/sites/default/files/documents/irb-forms/CITI_Training_Completion_Guide.pdf

Managing Your Data Activity (7.5%): One of the goals of this class is to make you more aware of and informed about how your data is used, and to empower you to take more control over your digital footprint. For this activity, you will explore your use of apps, websites, and other digital services, consider the various privacy tradeoffs you make to use them, and make decisions based on that information. The assignment will include a set of tasks to complete, including going through your mobile phone’s location settings, reviewing social media sharing and public/private settings, and Googling your name). Then, you will write up a 600-800 word reflection on the activity, what you discovered, and whether this activity changed your opinions about the data you share online.

Terms of Service/Privacy Policy Analysis (10%): Companies are required to provide consumers with details regarding the terms of service, which include end-user rights, as well as details on what the company can do with any data the collect from consumers. The vast majority of people never engage with these overly-long, hard-to-understand documents. But perhaps we should, so we can make more informed decisions about our data.

In this assignment, you will select a company you use, read its ToS and privacy policy documents, and critically evaluate the information they contain using the FTC’s five Fair Information Practice Principles (FIPPs): Transparency, Choice, Information Review and Correction, Information Protection, and Accountability (for a review: <https://ethics.berkeley.edu/privacy/fipps>). Address areas where the site’s policies are strong, where they fall short, and make suggestions for how the company could improve its privacy practices. Finally, consider the ethical and legal ramifications of the policy as it currently stands.

Papers should be 750-1000 words and should cover the following:

- Include name of company and URLs to relevant policies
- Brief (one paragraph) overview of the organization being evaluated and its history (especially anything that relates to privacy and security of user information)
- Analysis of how the site’s privacy policy meets (or falls short) for each of the FTC Fair Information Practice Principles
- Evaluate potential legal and ethical issues related to the company’s privacy policy and offer recommendations for how to address these issues.

Final Exam (15%): The goal of the final exam is to assess whether you have successfully met the learning outcomes of this class. The exam will require you to complete 3-4 essays (750-1000 words each) on topics related to content covered in the course. The exam questions will be assigned on the final day of classes and you will have until the scheduled final exam date and time to submit your responses electronically. There will be no physical exam during the scheduled exam time.

Final Grades: Final letter grades are assigned based on the percentage of total assessment points earned. To be fair to everyone, I have to establish clear standards and apply them consistently, so please understand that being close to a cutoff is not the same this as making the cut (**89.99 ≠ 90.00**). It would be unethical to make exceptions for some and not others.

Final Grade Cutoffs									
+	97.00%	+	87.00%	+	77.00%	+	67.00%		
A	94.00%	B	84.00%	C	74.00%	D	64.00%	F	<60.0%
-	90.00%	-	80.00%	-	70.00%	-	60.00%		

Course-Specific Policies

Use of computers, phones or tablet devices during our class meetings. The structure of this class is highly interactive. You will be working in groups frequently on questions, case studies, or scenarios I present. Sometimes these activities will require (or will benefit from) technology use. Therefore, I will not ban technology use in class. However, because we are a relatively small class and because participation and interaction are such important components of class, I expect any technology use to be related to class content. If I observe you engaging in technology use that may be distracting to their classmates, I will ask you to either stop or to leave for the day.

I expect you to make the responsible and respectful decision to refrain from using your cellphone in class. If you have critical communication to attend to, please excuse yourself and return when you are ready.

Missing Class. If you are going to miss class, you should speak with a classmate to get notes. You do not need to contact me if you are going to miss a single class; however, if you miss a speaker or are going to miss more than one class in a row, email me so we can make sure you do not fall behind. Get some extra credit for reading the syllabus – if you are reading this, simply email me and the TA before Class 4 a picture of a meme that directly relates to class material. But don't tell anyone else – let them earn the points. Repeated absences throughout the course of the semester will likely have a negative impact on final grade, as you will not be able to participate in class activities and discussions.

Policy on Academic Misconduct. Cases of academic misconduct will be referred to the Office of Student Conduct irrespective of scope and circumstances, as required by university rules and regulations. It is crucial to understand that the instructors do not have a choice of following other courses of actions in handling these cases. There are severe consequences of academic misconduct, some of which are permanent and reflected on the student's transcript. For details about procedures governing such referrals and possible consequences for the student please visit <http://osc.umd.edu/OSC/Default.aspx>

It is very important that you complete your own assignments, and do not share any files or other work. The best course of action to take when a student is having problems with an assignment question is to contact the instructor. The instructor will be happy to work with students while they work on the assignments.

University of Maryland Code of Academic Integrity. The University of Maryland, College Park has a nationally recognized Code of Academic Integrity, administered by the Student Honor Council. This Code sets standards for academic integrity at Maryland for all undergraduate and graduate students. As a student you are responsible for upholding these standards for this course. It is very important for you to be aware of the consequences of cheating, fabrication, facilitation, and plagiarism. For more information on the Code of Academic Integrity or the Student Honor Council, please visit <http://shc.umd.edu/SHC/Default.aspx>

Note on Turnitin Originality Checker and Plagiarism. For this course, some of your assignments may be collected via Turnitin on our course ELMS page. I have chosen to use this tool because it can help you improve your scholarly writing and help me verify the integrity of student work. For information about Turnitin, how it works, and the feedback reports you may have access to, visit [Turnitin Originality Checker for Students](#).

Even for assignments that do not require you to submit through Turnitin, you are responsible for confirming that anything you submit does not contain plagiarism. This includes not including proper formatting to indicate when you directly quote or paraphrase another source. If you do not know how to correctly cite material, refer to the UMD Library resources, available at <https://lib.guides.umd.edu/c.php?g=327184&p=2588295>.

If you are found to have committed plagiarism, you will receive a zero on that assignment. You can resubmit the assignment within one week with a 50% penalty. If you commit plagiarism more than once, you will receive a zero in the class and will be referred to the [Student Honor Council](#).

Special Needs. Students with disabilities should inform the instructor of their needs at the beginning of the semester. Please also contact the Disability Support Services (301-314-7682 or <http://www.counseling.umd.edu/DSS/>). DSS will make arrangements with the student and the instructor to determine and implement appropriate academic accommodations. Students encountering psychological problems that hamper their course work are referred to the Counseling Center (301-314-7651 or <http://www.counseling.umd.edu/>) for expert help.

Get Some Help!

You are expected to take personal responsibility for your own learning. This includes acknowledging when your performance does not match your goals and doing something about it. Everyone can benefit from some expert guidance on time management, note taking, and exam preparation, so I encourage you to consider visiting <http://ter.ps/learn> and schedule an appointment with an academic coach. Sharpen your communication skills (and improve your grade) by visiting <http://ter.ps/writing> and schedule an appointment with the campus Writing Center. Finally, if you just need someone to talk to, visit <http://www.counseling.umd.edu>.

Everything is free because you have already paid for it, and **everyone needs help**... all you have to do is ask for it.

Names/Pronouns and Self Identifications

The University of Maryland recognizes the importance of a diverse student body, and we are committed to fostering equitable classroom environments. I invite you, if you wish, to tell us how you want to be referred to both in terms of your name and your pronouns (he/him, she/her, they/them, etc.). The pronouns someone indicates are not necessarily indicative of their gender identity. Visit trans.umd.edu to learn more.

Additionally, how you identify in terms of your gender, race, class, sexuality, religion, and dis/ability, among all aspects of your identity, is your choice whether to disclose (e.g., should it come up in classroom conversation about our experiences and perspectives) and should be self-identified, not presumed or imposed. I will do my best to address and refer to all students accordingly, and I ask you to do the same for all of your fellow Terps.

Course Schedule

Note: readings may be added/changed throughout semester. Check ELMS Modules for up-to-date list.

Week	TOPIC & READINGS	WHAT'S DUE?
Part 1: Core Concepts		
1 Jan 29 Jan 31	Situating Data: Data in the Digital Age Big Data <ul style="list-style-type: none"> boyd & Crawford (2012), "Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon" 	Read the syllabus! Complete class survey
2 Feb 5 Feb 7	Ethical Frameworks for Evaluating Big Data <ul style="list-style-type: none"> Ethics Codes: History, Context, and Challenges. <i>Data & Society</i>. The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research [pdf] ACM's Code of Ethics: https://www.acm.org/code-of-ethics 	

Week	TOPIC & READINGS	WHAT'S DUE?
3 Feb 12 Feb 14	<p>Applying Ethics to Digital Data: Challenges & Opportunities</p> <ul style="list-style-type: none"> Engaging the Ethics of Data Science in Practice. Communications of the ACM: https://cacm.acm.org/magazines/2017/11/222176-engaging-the-ethics-of-data-science-in-practice/fulltext Anonymity and the Netflix Dataset: https://www.schneier.com/blog/archives/2007/12/anonymity_and_t_2.html <p>Case Study Readings:</p> <ul style="list-style-type: none"> Case study: No Encore for Encore? Ethical questions for web-based censorship measurement. Browse the original article: Burnett, S., & Feamster, N. (2015). Encore: Lightweight measurement of web censorship with cross-origin requests. In <i>ACM SIGCOMM Computer Communication Review</i> (Vol. 45, No. 4, pp. 653-667). ACM. <ul style="list-style-type: none"> Or watch a video of the paper being presented: https://www.youtube.com/watch?v=cMUTU0OrOAc <p>Recommended:</p> <ul style="list-style-type: none"> Robust De-anonymization of Large Sparse Datasets: https://www.cs.utexas.edu/~shmat/shmat_oak08netflix.pdf 	<p>Case Study: Encore</p>
4 Feb 19 Feb 21	<p>Privacy and Big Data</p> <ul style="list-style-type: none"> Tavani (2012): Privacy and Cyberspace. boyd (2015): Which students get to have privacy? (Medium) The Right to be Forgotten or the Duty to be Remembered? Twitter data reuse and implications for user privacy. <i>The Council for Big Data, Ethics, and Society</i>. The Real Privacy Problem. MIT Technology Review (2013). https://www.technologyreview.com/s/520426/the-real-privacy-problem/ 	<p>Managing Your Data</p>
5 Feb 26 Feb 28	<p>Security Challenges with Big Data</p> <ul style="list-style-type: none"> Should you fear the internet of things? https://bluenotes.anz.com/posts/2017/06/longread-should-you-fear-the-internet-of-things Zhang, 2018: Big Data Security and Privacy Protection (pdf) <p>Case Study Readings:</p> <ul style="list-style-type: none"> Read EPIC's backgrounder on the case: https://www.epic.org/amicus/crypto/apple/ Apple vs. the FBI Is Really, Really Complicated. Harvard Business Review: https://hbr.org/2016/02/apple-vs-the-fbi-is-really-really-complicated Ethics Unwrapped The FBI & Apple Security vs. Privacy: http://ethicsunwrapped.utexas.edu/wp- 	<p>Case Study: Apple vs. DOJ/FBI</p>

Week	TOPIC & READINGS	WHAT'S DUE?
	content/uploads/2016/10/21-The-FBI-Apple-Security-vs.-Privacy.pdf	
Part 2: Examining the Lifecycle of Data: From Collection to Dissemination		
6 Mar 5 March 7	<p>Issues from data collection/data mining</p> <ul style="list-style-type: none"> Willson, M., & Leaver, T. (2015). Zynga's FarmVille, social games, and the ethics of big data mining. <i>Communication and Research Practice</i>, 1(2), 147-158. Hill, K. (2012, February 16). How Target figured out a teen girl was pregnant before her father did. <i>Forbes</i>. Available at http://www.forbes.com/sites/kashmirhill/2012/02/16/how-target-figured-out-a-teengirl-was-pregnant-before-her-father-did/#7089a99434c6 Duus, R., & Cooray, M. (2015, December 27). Wearable fitness trackers: the dark side. <i>The Independent</i>. Available at http://www.independent.co.uk/life-style/health-and-families/health-news/wearable-fitness-trackers-the-dark-side-a6787171.html <p>Case Study Readings:</p> <ul style="list-style-type: none"> The mind-bending messiness of the Ashley Madison data dump (The Verge): https://www.theverge.com/2015/8/19/9178855/ashley-madison-data-breach-implications Philosophers On The Ashley Madison Hack: http://dailynous.com/2015/08/24/philosophers-on-the-ashley-madison-hack 	<p>Case Study: Ashley Madison data breach</p>
7 Mar 12 Mar 14	<p>Issues with data exploration/analysis</p> <ul style="list-style-type: none"> Lerman, J. (2013). Big data and its exclusions. <i>Stanford Law Review Online</i>, 66, 55-63. Crawford, K. (2013). The hidden biases in big data. <i>Harvard Business Review</i>. https://hbr.org/2013/04/the-hidden-biases-in-big-data 	
UMD Spring Break March 17-24 NO CLASS		
8 Mar 26 Mar 28	<p>Issues with algorithms & automated systems</p> <ul style="list-style-type: none"> Gillespie, T. (2014). The relevance of algorithms. <i>Media technologies: Essays on communication, materiality, and society</i>. Tufekci, Zeynep. "Machine intelligence makes human morals more important." presented at the TED, Banff, Alberta, 2016. https://www.youtube.com/watch?v=hSSmmlridUM. Jeong, S. (2016, March 25). How to make a bot that isn't racist. <i>Motherboard</i>. http://motherboard.vice.com/read/how-to-make-a-not-racist-bot 	
9 Apr 2 Apr 4	<p>Issues with dissemination/evaluation of data</p> <ul style="list-style-type: none"> Distrust Your Data: Jacob Harris on Six Ways to Make Mistakes with Data. https://source.opennews.org/articles/distrust-your-data/ The deception that lurks in our data-driven world (Splinter): https://splinternews.com/the-deception-that-lurks-in-our-data-driven-world-1793851547 	<p>Privacy Policy Analysis</p>

Week	TOPIC & READINGS	WHAT'S DUE?
	<ul style="list-style-type: none"> Weight Loss On Your Wrist? Fitness Trackers May Not Help (NPR): https://www.npr.org/sections/health-shots/2016/09/20/494631423/weight-loss-on-your-wrist-fitness-trackers-may-not-help 	
Part 3: Digging Into Developing Domains		
10 Apr 9 Apr 11	<p>Learning About You Through Your Social Network Data</p> <ul style="list-style-type: none"> Were OkCupid’s and Facebook’s Experiments Unethical? Harvard Business Review: https://hbr.org/2014/07/were-okcupids-and-facebooks-experiments-unethical What We Can Learn From The Epic Failure Of Google Flu Trends (WIRED): https://www.wired.com/2015/10/can-learn-epic-failure-google-flu-trends/ Your Apps Know Where You Were Last Night, and They’re Not Keeping It Secret (NY Times): https://www.nytimes.com/interactive/2018/12/10/business/location-data-privacy-apps.html <p>Case Study Readings:</p> <ul style="list-style-type: none"> WATCH: The Cambridge Analytica Scandal: Understanding Facebook’s data privacy debacle (The Verge): https://www.theverge.com/2018/4/10/17165130/facebook-cambridge-analytica-scandal Cambridge Analytica controversy must spur researchers to update data ethics (Nature): https://www.nature.com/articles/d41586-018-03856-4 The Murky Ethics of Data Gathering in a Post-Cambridge Analytica World (Medium): https://medium.com/ama-marketing-news/the-murky-ethics-of-data-gathering-in-a-post-cambridge-analytica-world-33848084bc4a <p>Recommended:</p> <ul style="list-style-type: none"> Ethical Considerations When Companies Study – And Fail to Study – Their Customers (The Cambridge Handbook of Consumer Privacy) 	<p>Case Study: Cambridge Analytica</p>
11 Apr 16 Apr 18	<p>Privacy as a Luxury Good: Ethical Research With Marginalized Populations</p> <ul style="list-style-type: none"> Making “Safe”: Community-Centered Practices in a Virtual World Dedicated to Children with Autism (2015 CSCW) Privacy, Poverty, and Big Data: A Matrix of Vulnerabilities for Poor Americans. <i>Washington University Law Review</i> (2017) <p>Case Study Readings</p> <ul style="list-style-type: none"> Deep neural networks are more accurate than humans at detecting sexual orientation from facial images. <i>Journal of Personality and Social Psychology</i> (2018) 	<p>Case Study: “Gaydar” Research</p>

Week	TOPIC & READINGS	WHAT'S DUE?
	<ul style="list-style-type: none"> AI 'Gaydar' And How The Future Of AI Will Be Exempt From Ethical Review (Forbes): https://www.forbes.com/sites/kalevleetaru/2017/09/16/ai-gaydar-and-how-the-future-of-ai-will-be-exempt-from-ethical-review/#319598e32c09 	
12 Apr 23 Apr 25	Internet of Things <ul style="list-style-type: none"> DDoS Attacks Evolve To Conscript Devices Onto The IoT: https://www.forbes.com/sites/centurylink/2018/02/04/ddos-attacks-evolve-to-conscript-devices-onto-the-iot/#385972946aaa Privacy and the Internet of Things: Emerging Frameworks for Policy and Design: https://cltc.berkeley.edu/IoTprivacy/ Amazon-owned Ring has reportedly been spying on customer camera feeds. BGR: https://bgr.com/2019/01/10/ring-camera-customer-feeds-accessed-creepy-privacy-violation/ 	
13 Apr 30 May 2	Search Engines & Recommender Systems <ul style="list-style-type: none"> As Google Fights Fake News, Voices on the Margins Raise Alarm (NY Times): https://www.nytimes.com/2017/09/26/technology/google-search-bias-claims.html We Followed YouTube's Recommendation Algorithm Down The Rabbit Hole (Buzzfeed News): https://www.buzzfeednews.com/article/carolineodonovan/down-youtubes-recommendation-rabbithole Case Study Readings: <ul style="list-style-type: none"> "It was a matter of life and death": A YouTube Engineer's Decision to Alter Data in the 'It Gets Better Project'. <i>Data & Society</i>. How an ex-YouTube insider investigated its secret algorithm (The Guardian): https://www.theguardian.com/technology/2018/feb/02/youtube-algorithm-election-clinton-trump-guillaume-chaslot Recommended: <ul style="list-style-type: none"> Selecting Effective Means to Any End: Futures and Ethics of Persuasion Profiling. <i>International Conference on Persuasive Technology</i>. 	Case Study: YouTube "It Gets Better"
14 May 7 May 9	Law & Policy in the U.S. and Abroad <ul style="list-style-type: none"> Warren, S., & Brandeis, L. (1890). The right to privacy. <i>Harvard Law Review</i>. Tech companies are open to privacy regulations. Congress should act. <i>Washington Post</i>: https://www.washingtonpost.com/opinions/facebook-is-failing-its-own-test-time-for-national-privacy-rules-for-tech/2018/09/28/bc605fac-c1d5-11e8-a1f0-a4051b6ad114_story.html?utm_term=.d54442a91d58 	

Week	TOPIC & READINGS	WHAT'S DUE?
	<ul style="list-style-type: none"> • Three Years of the Right to Be Forgotten: https://elie.net/static/files/three-years-of-the-right-to-be-forgotten/three-years-of-the-right-to-be-forgotten-paper.pdf • What the GDPR means for Facebook, the EU and you: https://www.cnet.com/how-to/what-gdpr-means-for-facebook-google-the-eu-us-and-you/ 	
15 May 14	Course Wrap-Up	
Final Exam: Due Date TBD		