Privacy, Security, and Ethics for Big Data

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.

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

TA: [removed]

Class Meets (Zoom)
TUESDAYS 1-2:15pm

Office Hours (Zoom)
THURSDAYS 1-2:15pm

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 “INST366-0102” in the subject).
• Contact TA 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](http://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.

Class Structure
This is a “flipped class.” Weekly content will become available Mondays at 9am ET. This includes readings, assignments, and pre-recorded lectures. All assignments for that week are due the following Sunday at 11:59pm ET.

Because lectures are pre-recorded, you can watch them at any point during the week. However, we’ll meet once a week during our regularly scheduled time (Tuesdays at 1pm ET) to talk about the prior week’s content and engage in virtual activities and discussions. Some of these activities will count toward your grade, so you should plan to attend the Tuesday sessions. I will also record and post each Tuesday’s Zoom session to Canvas. I will hold my office hours during the Thursday time slot; you are welcome to drop into the Zoom call during that time, but it is not required.

I also want to acknowledge that the last year has been stressful for many people, and many of you are balancing work and school and/or living at home. If you are experiencing problems keeping up with class material, please reach out to me and we can work on a plan to help you make it through the semester. I am also offering each student a “get out of jail free” card for the semester: for any assignment (except the final), you may submit it up to one week late without a penalty. To use this card, you must add a note when you submit on Canvas that you want to use it. Without a note, your assignment will be graded as late.

Activities, Learning Assessments, & Expectations for Students

Course Activities
Your final grade will be based on the following components:

<table>
<thead>
<tr>
<th>Learning Assessments</th>
<th>#</th>
<th>Points Each</th>
<th>Category Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Exit tickets</td>
<td>12</td>
<td>1</td>
<td>12</td>
</tr>
<tr>
<td>Zoom activities (during Tuesday sessions)</td>
<td>8</td>
<td>1.5</td>
<td>12</td>
</tr>
<tr>
<td>Ethics Training</td>
<td>1</td>
<td>5</td>
<td>5</td>
</tr>
<tr>
<td>Case Study Evaluations</td>
<td>5</td>
<td>6</td>
<td>30</td>
</tr>
<tr>
<td>Managing Your Data</td>
<td>1</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Ethics of Data Mining Activity</td>
<td>1</td>
<td>8</td>
<td>8</td>
</tr>
<tr>
<td>Terms of Service/Privacy Policy Analysis</td>
<td>1</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Final Exam</td>
<td>1</td>
<td>15</td>
<td>15</td>
</tr>
<tr>
<td><strong>Total Points:</strong></td>
<td></td>
<td><strong>100</strong></td>
<td></td>
</tr>
</tbody>
</table>
Exit tickets (1% each, 12% total): For each week’s readings, you need to briefly summarize the two most important and/or interesting things you learned from the materials. You must also provide either a question that came up based on the week’s content or a relevant news article you found on the topic. I will only count the top 12 scores for this assignment; anything beyond that is dropped.

Zoom Activities (1.5% each, 12% total): While attendance will not be taken in this course, students will regularly participate in class-based activities around that day’s lecture. Sometimes, this will be done via polls in Zoom; other times, students may complete tasks or work in breakout groups. I will only count the top eight scores for this assignment; any activities beyond those eight will be dropped from your final grade. If you have an excused absence (e.g., religious holiday), 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 (6% each; 36% 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 600-800 word write-up on the case study based on guiding questions I provide. I will provide seven case study assignments throughout the semester and you will be asked to complete six (or you may complete all seven and the lowest grade will be dropped).

Ethics Training (5%): When university researchers in the US want to conduct research with human subjects, they must first complete ethics training for social and behavioral research through CITI (Collaborative Institutional Training Initiative). You might have already completed this for another class. If so, great! If not, you need to do this first. The training 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. Regardless of whether you have completed the basic training, you also need to complete an elective module on big data (title: Human Subjects Considerations and Big Data Research; ID 19126) and one of internet-based research (ID: 510). To receive credit, you must upload the PDF certificate you receive for successfully completing training. Make sure you review the instructions on Canvas on how to save a certificate showing completion of all modules.

Managing Your Data Activity (8%): 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 an 600-800 word reflection on the activity, what you discovered, and whether this activity changed your opinions about the data you share online.

Ethics of Data Mining (8%): In Part 2 of the class, we consider some of the ethical considerations for large-scale data collection and analysis. To help you understand some of the concerns, you’ll be using an existing tool (TAGS) to collect tweets related to a specific set of keywords or hashtags. You’ll review the data and respond to a series of prompts listed on the assignment page in 750-1000 words.

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 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.

**Extra Credit Opportunities (1-3%, max of 6% during semester).** There may be extra credit opportunities throughout the semester. If there are, I will notify the class via Announcements and provide details on what is required and when it is due. A student cannot receive more than six points in extra credit throughout the semester.

**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.

<table>
<thead>
<tr>
<th>Final Grade Cutoffs</th>
</tr>
</thead>
<tbody>
<tr>
<td>+ 97.00%  + 87.00%  + 77.00%  + 67.00%</td>
</tr>
<tr>
<td>A 93.00%  B 83.00%  C 73.00%  D 63.00%  F &lt;60.0%</td>
</tr>
<tr>
<td>- 90.00%  - 80.00%  - 70.00%  - 60.00%</td>
</tr>
</tbody>
</table>

**Late Assignments**
A late penalty will be applied to any assignment that is submitted late (starting one minute beyond the due date and time). Late assignments will be graded according to the following formula:

\[
\text{[grade on assignment]} \times 0.9^{\text{[number of days late]}}
\]

For example, if you turn in a 10-point assignment 1 day late, and earned a 100% on that assignment, your final grade for that assignment would be \((1.00)(0.9^1)\), or 90%. If you turned in the same assignment 3 days late, and earned an 80%, your grade would be \((0.80)(0.9^3)\), or 58%. Therefore, although your grade will take a serious hit if the assignment is late, in this class it is always better to turn in an assignment late than to not do it at all. Note that late assignments may not be graded quickly.
Please prepare in advance so that you will not encounter technical difficulties that will result in your work receiving a late penalty. **Technical difficulties are not an excuse for late assignments — if you are having trouble submitting an assignment on Canvas, e-mail it to me before the deadline** to avoid a penalty. If you have a conflict with the due date, assignments can always be submitted early. Generally speaking, illnesses are not an excuse for late assignments because you will receive the assignments at least one week before they are due.

Also remember that you also have two “get out of jail free” cards to use at any time during the semester (note: these cannot be used on the midterm or team project assignments).

**Course-Specific Policies**

**Zoom Meetings.** During class Zoom calls, you should be respectful in any discussions, both when we meet as a full class and when we breakout into groups. Here are some of my rules and expectations for Zoom classes.

- I encourage you to have your camera on but I will not require it. I think it’s especially important to turn on the camera during breakouts.
- If you have your camera on, be “professional”: have a shirt on and have the camera showing your face. Don’t smoke, drink alcohol, or get drunk during class.
- Keep your mic muted unless you’re talking. Avoid interrupting or talking over others.
- If you need to clarify something while in a breakout room, click the Ask for Help button and I will join the room to offer help.
- Do not share pornographic or other adult content. This will result in being kicked out of Zoom immediately.

**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 by February 5 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](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://she.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](http://shc.umd.edu/SHC/Default.aspx).

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](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 violate Academic Integrity more than once, you will receive a zero in the class and will be referred to the [Student Honor Council](http://shc.umd.edu/SHC/Default.aspx).

**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](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](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](http://www.counseling.umd.edu).

**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](http://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.*
<table>
<thead>
<tr>
<th>Week/Module</th>
<th>TOPIC &amp; READINGS</th>
<th>WHAT’S DUE BY THE END OF WEEK?</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Part 1: Core Concepts</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
| 1 Week of Jan 24 | **Situating “Big Data”: Data in the Digital Age**  
- Human insights missing from big data [TED Talk] | Read the syllabus!  
Complete class survey |
| 2 Week of Jan 31 | **Ethical Frameworks for Evaluating (Big) Data**  
- AoIR’s most recent ethics guidelines (only read Part 3)  
- Giving algorithms a sense of uncertainty could make them more ethical. *MIT Technology Review.* [url]  
**Recommended (not required):**  
- The Menlo Report: Ethical Principles Guiding Information and Communication Technology Research [pdf]  
- RadioLab podcast episode, “The Bad Show” [url]  
- Engaging the Ethics of Data Science in Practice. *Communications of the ACM* [url]  
| 3 Week of Feb 7 | **Privacy and Big Data**  
- Apple's App 'Privacy Labels' Are Here—and They’re a Big Step Forward. WIRED. [https://www.wired.com/story/apple-app-privacy-labels/](https://www.wired.com/story/apple-app-privacy-labels/)  
| | **Recommended (not required):**  
- The Right to be Forgotten or the Duty to be Remembered? Twitter data reuse and implications for user privacy. *The Council for Big Data, Ethics, and Society.*  
<table>
<thead>
<tr>
<th>Week/Module</th>
<th>Topic &amp; Readings</th>
<th>What’s Due by the End of Week?</th>
</tr>
</thead>
</table>
| 4 Week of Feb 14 | **Security and Big Data**  
- Russia’s hacking frenzy is a reckoning. WIRED. [https://www.wired.com/story/russia-hack-supply-chain-reckoning/](https://www.wired.com/story/russia-hack-supply-chain-reckoning/)  
**Recommended (not required)**  
| 5 Week of Feb 21 | **Privacy, Security, Big Data, & Legislation**  
- Tech companies are open to privacy regulations. Congress should act. *Washington Post* [url]  
- One Clear Message From Voters This Election? More Privacy. WIRED [url]  
- Five Years of the Right to Be Forgotten [url]  
- What the GDPR means for Facebook, the EU and you. CNET. [url] | Ethics Training |
| 6 Week of Feb 28 | **Issues from data collection/data mining**  
<p>| 7 | <strong>Issues with data exploration/analysis and dissemination</strong> | |</p>
<table>
<thead>
<tr>
<th>Week/Module</th>
<th>Topic &amp; Readings</th>
<th>What’s Due by the End of Week?</th>
</tr>
</thead>
</table>
• Distrust Your Data: Jacob Harris on Six Ways to Make Mistakes with Data. [https://source.opennews.org/articles/distrust-your-data/](https://source.opennews.org/articles/distrust-your-data/)  
| Week of Mar 21 | **Issues with algorithms & automated systems**  
• Tufekci, Zeynep. “Machine intelligence makes human morals more important.” TED, Banff, Alberta, 2016. [https://www.youtube.com/watch?v=hSSmmlRidUM](https://www.youtube.com/watch?v=hSSmmlRidUM)  
Recommended:  
• Bias detectives: the researchers striving to make algorithms fair [https://www.nature.com/articles/d41586-018-05469-3](https://www.nature.com/articles/d41586-018-05469-3)  
• Can we open the black box of AI? [Nature]  
| Week of Mar 28 | **Part 3: Digging Into Developing Domains**  
**Learning About You Through Your Social Network Data**  
• Amazon, Facebook and Google don’t need to spy on your conversations to know what you’re talking about [The Conversation]  
• Were OkCupid’s and Facebook’s Experiments Unethical? *Harvard Business Review*: [https://hbr.org/2014/07/were-okcupids-and-facebooksexperiments-unethical](https://hbr.org/2014/07/were-okcupids-and-facebooksexperiments-unethical)  
Recommended: | *Case Study: Cambridge Analytica* |
<table>
<thead>
<tr>
<th>Week/Module</th>
<th><strong>TOPIC &amp; READINGS</strong></th>
<th><strong>WHAT’S DUE BY THE END OF WEEK?</strong></th>
</tr>
</thead>
</table>
| 10 Week of Apr 4 | • Ethical Considerations When Companies Study – And Fail to Study – Their Customers (The Cambridge Handbook of Consumer Privacy)  
• 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/](https://www.wired.com/2015/10/can-learn-epic-failure-google-flu-trends/)  
**Privacy as a Luxury Good**  
• The Disparate Impact of Surveillance: [https://tcf.org/content/report/disparate-impact-surveillance/?agreed=1](https://tcf.org/content/report/disparate-impact-surveillance/?agreed=1)  
• Predictive policing algorithms are racist. They need to be dismantled. [MIT Technology Review]  
**Recommended:**  
| | **Case Study: Golden State Killer** |
| 11 Week of Apr 11 | • What is the Internet of Things? WIRED explains. [https://www.wired.co.uk/article/internet-of-things-what-is-explained-iot](https://www.wired.co.uk/article/internet-of-things-what-is-explained-iot)  
• What your smart devices know (and share) about you [TED Talk]  
• Sidewalk Labs’ vision and your data privacy: A guide to the saga on Toronto’s waterfront [Globe and Mail]  
**Recommended**  
• The Big Hack. NY Mag [url]  
| **The Internet of Things, Smart Homes, and Smart Cities** |
| 12 Week of Apr 18 | • Who’s the Boss? "The Elf on the Shelf" and the normalization of surveillance: [https://www.policyalternatives.ca/publications/commentary/who-s-boss](https://www.policyalternatives.ca/publications/commentary/who-s-boss)  
| **Surveillance at Home and Work**  
<p>| <strong>Privacy Policy Analysis</strong> |</p>
<table>
<thead>
<tr>
<th>Week/Module</th>
<th>Topic &amp; Readings</th>
<th>What’s Due by the End of Week?</th>
</tr>
</thead>
</table>
| 13 Week of Apr 25 | **Big Data, Better Health?**  
  - Data privacy vs. sharing **TBD**  
  - Wearables article **TBD**  
| 14 Week of May 2 | **Looking to the Future**  
  - Rise of the machines: has technology evolved beyond our control? [The Guardian](https://www.theguardian.com/)  
  - **TBD**                                                                                                                                                  |                                |
| 15 Week of May 9 | **Course Wrap Up** (final Zoom meeting is May 11)                                                                                                                                                                   |                                |
|              | **Final Exam (take home): Due Date TBD**                                                                                                                                                                           |                                |