Problem 1 (3 points)
In some cases sharing personally identifiable information in a raw form can be illegal, but sharing hashed versions is allowed. Discuss the differences between raw and hashed identifiers including the related privacy implications.

Problem 2 (3 points)
How do ad blocking tools like AdBlock Plus, Ghostery, and Disconnect mitigate propagation of personal information to advertisers via RTB?

Problem 3 (9 points)
In this problem, you will investigate and evaluate various privacy-enhancing technologies.

(a) Describe what Chrome’s Incognito browsing mode provides. What types of tracking techniques are blocked in this mode? What types of tracking still work?

(b) Cover Your Tracks is a service that allows you to evaluate the uniqueness of your browser fingerprint. What parts of your browser fingerprint provide the most identifying information (please describe your configuration: OS, browser make and model, etc.)? What can you do to reduce the identifiability of your fingerprint?

(c) Recent research has shown that browser fingerprints change frequently from normal use. Say that on average, the fingerprint of your browser changes every 5 days. With this knowledge, along with not storing any cookies, are you safe from tracking? Why or why not?

Problem 4 (3 points)
Compare and contrast Do Not Track and Global Privacy Control. What are the differences in enforcement? Do you think GPC will be more successful than DNT and, if so, why?

Problem 5 (3 points)
Brave is a new, privacy-focused Web browser. How does Brave support behavioral advertising while claiming it preserves users’ privacy?
Problem 6 (3 points)
Describe all methods for cross-device tracking mentioned in the Forbrukerrådet report. For each method, list all identifiers used for linkage.

Problem 7 (6 points)
Pick any two mobile apps mentioned in the Forbrukerrådet report.

(a) Read the sections of their privacy policies that cover information sharing with third parties and re-state them using the contextual integrity framework and specifying all CI parameters. If there is not enough information in a policy to fill in some of the parameters, indicate so.

(b) Describe the actual flows of information from these apps to third parties and compare with the policies.

Problem 8 (1 point)
Apple says this about privacy (https://www.apple.com/privacy/):

Privacy is a fundamental human right. At Apple, it’s also one of our core values. Your devices are important to so many parts of your life. What you share from those experiences, and who you share it with, should be up to you. We design Apple products to protect your privacy and give you control over your information. It’s not always easy. But that’s the kind of innovation we believe in.

What concept of privacy does this statement represent?

Problem 9 (5 points)
(a) What follows from claiming that “Privacy is a right”?

(b) What else could it be and why does it make a difference?

(c) What does it mean to declare that we’re looking for a “meaningful definition of privacy”?

Problem 10 (4 points)
We discussed two categories of privacy definition: control and secrecy. From PIC,

(a) cite two privacy theorists in each of these of each categories, and

(b) quote one instance of each

Problem 11 (6 points)
Tom Gerety was concerned that privacy “expands like a gas to fill up available space.”

(a) Explain with examples drawn from your answers to Problem 10
(b) What was his (and others’) approach to addressing this concern?

(c) Based on reading PIC 6 discuss one reason why the basis for this approach runs into problems of its own. (Your answers must reference page numbers in the relevant PIC chapters.)

Problem 12 (7 points)
(a) What is the privacy paradox and why is it referred to as such?

(b) Whose interests in society are best served by promoting the idea of a paradox? Explain how it may be used by these interests. You can earn 2 BONUS points if your answer mentions any of the actors (people, companies, systems) discussed in Prof. Shmatikoff’s lectures.

(c) List two rebuttals to the paradox. (Your answers must reference page numbers in the relevant PIC chapters.)

Problem 13 (7 points)
According to Contextual Integrity, informational (privacy) norms/rules have five parameters.

(a) What are they?

(b) Give us a well formed contextual informational rule/norm from your home background (culture) that you think might surprise us or some of your classmates.

(c) How might the parameters of CI be used in an argument against the so-called privacy paradox?

(d) In 2009, Amazon deleted copies of George Orwell’s 1984 from customers’ Kindle devices. (Yes, they refunded the 99 cents that customers had paid.) There are many perspectives one could take on this action but how might you use CI as a lens through which to see the privacy aspect.

CS ONLY

Problem CS1
OpenWPM is an open-source Web privacy measurement framework. Use it to analyze 100 popular websites from diverse categories (explain your algorithm for selecting these sites).

(a - 6 points) What are the most common trackers on your chosen sites?

(b - 12 points) How many of the trackers include persistent identifiers? What do you consider a persistent identifier? Identify fingerprinting code — use this link for the list of probable fingerprinting commands (try recursive_dump_page_source method).
(c - 12 points) Count the flows of information across different contexts. To do this, assign each site to a particular social context (e.g., news, entertainment, etc.) – it’s Ok to use an automated assignment algorithm. There exists an information flow between context A and context B if the same tracker appears in a page that belongs to context A and a page that belongs to context B.

You can upload a documented Jupyter notebook or report in PDF.

Useful links and helpful hints:
- [OpenWPM tutorial](#)
- [OpenWPM configuration settings available](#)
- You can create your own function in custom_command.py (there’s already an example function there)
  - Might be useful: [Selenium with Python](#) for programmatically accessing webpage elements (this is what the LinkCountingCommand() there does)
- openwpm/command_sequence.py might also have some useful stuff
  - command_sequence.recursive_dump_page_source() will dump the page source to the sources directory
- crawl-data.sqlite contains some information retrieved from the web crawl
  - [This link](#) explains some of the tables it stores

INFO ONLY

Problem INFO1 (6 points)
Privacy skeptics claim that privacy, in fact, is not particularly important, or worth protecting as privacy advocates claim. Briefly (a few sentences each) describe three skeptical arguments. (Your answers must reference page numbers in the relevant PIC chapters.)

Problem INFO2 (6 points)
“It’s vital to sacrifice personal privacy for national security and social welfare.” Do you agree or disagree? Explain your answer comparing how two theorists might comment on this statement. (Hint: PIC Chapter 4 and 6)

Problem INFO3 (10 points)
Recent revelations made by Politico regarding the mental health non-profit Crisis Text Line have raised numerous data-ethics and privacy concerns over the use of personal data.

(a) Whose privacy is violated?

(b) Using van den Hoven, explain how this scenario does or does not amount to informational injustice.

(c) Some critics might say that privacy was violated because the data in question was sensitive.
Problem INFO4 (8 points)
Someone says to you, “If you’ve done nothing wrong (got nothing to hide), you don’t need privacy!” How might you draw on the definition of Contextual Integrity to push back. (Feel free to make up your own examples or draw any from lectures that are useful for your answer.)