Instructions

There are **85 total points**. When asked to provide your answer within a figure or table, be careful to not exceed box boundaries. Bubbles must be filled out completely: ● is correct, ✓ ⊙ ⊗ are incorrect All answers must be given within the provided circles, answer boxes, figures or tables. Write your full name in the box to acknowledge the instructions.

Nick Feamster

Privacy Law and Regulation

1. [4 points]: According to the Fair Information Practice Principles (FIPPs) from the 1973 HEW report, which of the following are among the five core principles? (Select all that apply.)

● No secret record-keeping by organizations

● Right to know what data is collected and how it's used

● Prevention of secondary use without explicit consent

○ Mandatory encryption of all personal data

● Requirement for reasonable security precautions

2. [3 points]: What is a significant limitation of the FIPPs model when applied to modern data systems?

○ FIPPs require too much computational power to implement

● FIPPs assume static databases and don't address queries, machine learning, and probabilistic inference

○ FIPPs only apply to government databases, not private companies

○ FIPPs mandate opt-in consent, which is too restrictive for modern services

3. [4 points]: Which of the following are characteristics of the US privacy law approach? (Select all that apply.)

● Sectoral approach with laws specific to data holder types

● No comprehensive federal privacy statute

○ Single omnibus privacy law covering all sectors

● FTC acting as de facto privacy regulator through Section 5 of the FTC Act

● State-level breach notification laws in all 50 states

4. [3 points]: Explain why the notice and consent model has been criticized as insufficient for protecting user privacy in practice.

(Answer inside the box)

Solution: The notice and consent model has several significant limitations: privacy policies are often long, complex, and unreadable for most users; users rarely read policies before accepting them; implementation often diverges from what is stated in policies; and users have little meaningful choice when services are essential or have network effects. Additionally, the model assumes users can make informed decisions about future uses of their data, which is increasingly unrealistic with machine learning and secondary uses.

GDPR and International Privacy

5. [3 points]: Some critics argue that GDPR's complexity and compliance costs create competitive advantages for large incumbent companies like Google and Facebook. Explain one reason why GDPR might strengthen the market position of these incumbents.

(Answer inside the box)

Solution: GDPR's complex requirements and high compliance costs create barriers to entry that disproportionately affect smaller companies and startups. Large incumbents like Google and Facebook have the resources to hire compliance teams, implement sophisticated consent mechanisms, and navigate regulatory requirements. Additionally, users are more likely to grant consent to established platforms they already use and trust, making it harder for new entrants to compete. The compliance burden can also limit competition by making it expensive for small companies to handle user data at scale.

6. [4 points]: Which of the following are key requirements or provisions of GDPR? (Select all that apply.)

○ Companies can collect any data as long as they notify users

● Consent must be opt-in only, not opt-out

● Right to be forgotten (data erasure)

○ Maximum fine of $1 million for violations

● Restrictions on cross-border data transfers

CCPA/CPRA and Automated Compliance

Example: A website's opt-out page requires users to fill out a form with their full name, email address, phone number, and mailing address to opt out of data sales. The website claims this information is needed to "verify identity and process the request."

7. [4 points]: What type of dark pattern does this example primarily represent?

○ Confirm shaming

● Obstruction

○ Misdirection

○ Interface interference

8. [2 points]: What is the Global Privacy Control (GPC) and how does it relate to CPRA compliance?

(Answer inside the box)

Solution: The Global Privacy Control (GPC) is a browser setting that automatically sends an opt-out signal to websites, indicating the user's preference not to have their personal information sold or shared. Under CPRA, websites must honor GPC signals as valid opt-out requests. GPC is implemented in browsers like Firefox and Chrome, though most users are unaware of this feature. It provides an automated alternative to manual opt-out processes.

Initials: ____

9. [3 points]: Explain what the "spillover effect" (or "California effect") refers to in the context of privacy regulation compliance.

(Answer inside the box)

Solution: The spillover or California effect refers to the phenomenon where companies extend privacy protections beyond the jurisdictions where they are legally required. Research has shown that websites in states without privacy laws (like Illinois) adopted privacy protections similar to those required in California, suggesting that companies found it easier to implement privacy controls uniformly rather than maintaining different systems for different states. This demonstrates how strong state regulations can influence practices nationwide.

10. [4 points]: Studies of CCPA/CPRA compliance found that approximately 30% of covered entities failed to implement an opt-out link or mechanism on their website. Which of the following are valid reasons for this? (Select all that apply.)

● The sites don't actually sell data, so no opt-out link is required

● Implementation errors where links are hidden by JavaScript

● Opt-out functionality is in the privacy policy but not in a visible header/footer link

○ All covered entities intentionally violate the law

● Links are dynamically loaded and absent from HTML source

Dark Patterns

11. [3 points]: Explain why CPRA (California Privacy Rights Act) explicitly prohibits the use of dark patterns in the opt-out process.

(Answer inside the box)

Solution: CPRA prohibits dark patterns because they undermine the fundamental purpose of the opt-out right by manipulating or deceiving users into not exercising their privacy rights. Dark patterns make it difficult, confusing, or frustrating for users to opt out, which defeats the legislative intent of giving users meaningful control over their personal information. By banning dark patterns, CPRA aims to ensure that opt-out processes are straightforward and respect user autonomy.

12. [4 points]: Which of the following are examples of the "obstruction" category of dark patterns? (Select all that apply.)

● Requiring CAPTCHA completion to access privacy portal

● Mutually exclusive choices like "opt out of sale OR sharing, but not both"

● Asymmetry where opting in takes 1 click but opting out takes 2 clicks

○ Using confusing language like "Allow sale" on an opt-out button

● Identity verification that fails if not completed within 48 hours

Case Study: A website's privacy opt-out page displays a toggle switch with no text labels. The switch is currently in the "off" position. Above the switch, text reads "Manage your privacy preferences." It's unclear whether toggling the switch to "on" will enable data selling or enable privacy protection. At the bottom of the page, there is a hyperlink labeled "this form" that is difficult to identify as clickable because it appears in the same color as regular text.

13. [2 points]: Identify which category (or categories) of dark patterns are present in this case study and explain your reasoning.

(Answer inside the box)

Solution: This case contains both "misdirection" and "interface interference" dark patterns. Misdirection is present because the toggle switch has unclear/conflicting instructions—it's ambiguous whether toggling "on" opts out or opts in to data selling. Interface interference is present in the hidden hyperlink labeled "this form" which is hard to identify as clickable due to poor visual design. These patterns make it difficult for users to understand and complete the opt-out process.

14. [3 points]: Explain why the use of OneTrust or similar third-party compliance services does not guarantee actual compliance with privacy regulations.

(Answer inside the box)

Solution: While OneTrust and similar services provide compliance tools, simply deploying these services doesn't guarantee compliance. Companies may deploy the software without proper configuration, fail to test the implementation thoroughly, have corner cases that the default settings don't handle, or misunderstand the statute requirements. Research has shown that sites using OneTrust are frequently out of compliance, suggesting that the tools must be properly configured and tested rather than deployed as turnkey solutions.

AI and Privacy

15. [4 points]: Which of the following represent AI-specific privacy risks that **differ from traditional** privacy risks? (Select all that apply.)

● Model memorization of training data

● Human-like interactions encouraging progressive disclosure

○ Data breaches and unauthorized access

● Prompt injection attacks extracting sensitive information

○ Use or sale of data to third parties

16. [3 points]: What is the "interdependent privacy" problem in the context of LLM usage?

○ LLMs require multiple users to verify their identities

○ Privacy settings must be configured by both the user and the LLM provider

● Users' privacy decisions affect others when they share information about colleagues, clients, or family

○ LLMs cannot function without sharing data across multiple servers

17. [3 points]: Explain how anthropomorphization of LLMs can lead to increased privacy risks for users.

(Answer inside the box)

Solution: Anthropomorphization occurs when users perceive LLMs as human-like conversational partners, which encourages trust and intimacy. This leads to progressive disclosure where users share increasingly sensitive information, especially when the LLM prompts with responses like "Tell me more?" The conversational interface and positive feedback create comfort that may cause users to share personal information they wouldn't share with a traditional software interface. This can lead to disclosure of sensitive personal, medical, financial, or professional information.

Initials: ____

Scenario: An LLM interface displays a prompt: "Do you want me to remember our chats to provide better context and more personalized responses in future conversations?" with buttons for "Yes, remember" and "No, don't remember."

18. [4 points]: Analyze this interface design for potential dark patterns. Does it exhibit characteristics of obstruction, interface interference, or misdirection? Explain your answer.

(Answer inside the box)

Solution: This exhibits misdirection because the framing emphasizes the benefit to the user ("better context and more personalized responses") while downplaying or omitting information about data collection, storage, and potential privacy implications. The phrasing makes it sound like a helpful feature rather than a data collection decision. A more transparent approach would explicitly mention that chat history will be stored and potentially used for training, along with associated privacy implications.

19. [4 points]: Which of the following are strategies users employ to protect privacy when using LLMs? (Select all that apply.)

● Accepting risks as a privacy-convenience tradeoff similar to other tech services

● Avoiding certain uses like financial data with real numbers

● Rewriting prompts to reduce identifiable information about gender, race, or location

○ Only using LLMs through VPNs to hide IP addresses

● Not using LLMs for work-related queries due to company policies

Copyright and Fair Use

20. [4 points]: Which of the following can be protected by copyright? (Select all that apply.)

○ Mathematical formulas

● Literary works

○ Ideas and concepts

● Software code

○ Historical facts

21. [4 points]: By default, copyright grants the creator exclusive rights including making copies, distributing works, creating derivatives, and public performance. ● Yes ○ No

22. [4 points]: According to fair use doctrine, which of the following are among the four factors courts consider when determining whether use of copyrighted material qualifies as fair use? (Select all that apply.)

○ Whether the user obtained permission from the copyright holder

● Purpose and character of use (transformative use)

● Nature of the copyrighted work

● Amount and substantiality of portion used

● Effect on market value of the original work

Case Study: An AI company trains a large language model (LLM) using millions of copyrighted books, articles, and creative works without obtaining permission from or compensating the authors. The company argues that this constitutes fair use. Authors have filed lawsuits claiming copyright infringement, arguing that AI-generated content competes with their work and threatens their livelihoods. The case is expected to reach the Supreme Court.

23. [3 points]: Using the "transformative use" test from fair use doctrine, construct an argument IN FAVOR of the AI company's fair use defense.

(Answer inside the box)

Solution: The AI company's use is transformative because the LLM doesn't merely reproduce or redistribute the original copyrighted works. Instead, it learns patterns, structures, and relationships from the training data to generate entirely new content. The model creates novel outputs that are fundamentally different from any individual training text. This is analogous to how sampling music can be transformative when creating something new. The AI's outputs don't supersede the originals—users seeking a specific book won't use an LLM as a substitute for reading that book.

24. [3 points]: Using the "effect on market value" test from fair use doctrine, construct an argument AGAINST the AI company's fair use defense.

(Answer inside the box)

Solution: The AI company's use harms the market for the original works because AI-generated content directly competes with human creators. Romance novelists, journalists, and other writers argue that instant AI generation damages their sales by providing a cheaper, faster alternative to hiring human writers. Artists claim AI-generated art replaces the market for their commissioned work. The business model of these AI companies is built on providing content that would otherwise be created by the copyright holders, effectively putting authors and creators out of business and destroying the commercial value of their work.

Feedback

25. [1 point]: Interest (1=Boring!; 10=Amazing!): **8** Difficulty (1=Too easy; 10=Too hard): **7**

26. [2 points]: 1. One topic you found most interesting this semester. 2. One suggestion for improvement:

(Answer inside the box)

Solution: The intersection of AI and privacy law was fascinating. More case studies would be helpful.