CSC 533: Privacy in the Digital Age (Fall 2023)
Home Assignment #5
Assigned: Nov 10, 2023, Due: Nov 21, 2023 **(No extension possible)**

Instruction: Completed homework should be typed (e.g., using LaTeX or word document) or hand-written clearly and scanned and uploaded into Moodle. You can discuss about how to use certain tools for data collection and analysis, but no collaboration is permitted to solve the problems.

1. **Learning objective:** Designing a more usable privacy notice.

   Develop a mockup of a usable alternative privacy notice (**remember this is draft notice content and you can customize contents, e.g., the type of data required, based on the specific app functionality you envision to provide**). You may choose an existing app or website and remodel its real privacy notice to a **consumer-friendly** format. Your mockup may be hand drawn, drawn using PowerPoint or your favorite drawing tools, or designed using a rapid-prototyping tool (e.g., Wireframe). If the privacy notice takes up multiple screens or includes interactive features (like expansion upon clicking a button), show all the screens, pop-ups, etc.

   a) Write a paragraph explaining the rationale behind your major design decisions. [10 points]
   b) Write a paragraph explaining where/when in the course of selecting, downloading, installing, or using the app users will have the opportunity to see this privacy notice. Why do you recommend making the notice available in this way? [10 points]
   c) Include a mockup/prototype design in your submission. [10 points]

2. **Learning objective:** Determining policy compliance using ChatGPT

   Let's assume the following 8 checkpoints determine to what extent privacy policies are complete and compliant with different regulations.

   Data collection and purpose*:" if the privacy policy mentions how and why a service provider collects user information.",*

   Data sharing*: " if the privacy policy mentions how user information may be shared with or collected by third parties.",*

   User Control*:" if the privacy policy mentions choices and control options for the user.",*

   User Right*:" if the privacy policy mentions how users may access, edit, or delete their information",*

   Data Retention*:" if the privacy policy mentions how long user information is stored",*

   Data protection*:" if the privacy policy mentions how user information is protected",*

   Policy Change Notification*:" if the privacy policy mentions how users will be informed about changes to the privacy policy",*

   Restriction for specific audience:" if the privacy policy mentions practices that pertain only to a specific group of users, such as children, Europeans, or California residents",

   a) Select any **3** apps from Google play store. Then compare the privacy policies with the **8-checklist** provided above for compliance issues. Construct a **table** for each app listing the various privacy policy segments (like sentences or paragraphs) that address your checklist items. If not found for a given checkpoint, state **'not found'**. [15 points]

      Table format example

Policy segment	Checkpoint
An operator of a Web site or online service shall retain personal information collected online from a child for only as long as is reasonably necessary to fulfill the purpose for which the information was collected.	Data retention policy

CSC 533: Privacy in the Digital Age (Fall 2023)
Home Assignment #5
Assigned: Nov 10, 2023, Due: Nov 21, 2023 **(No extension possible)**

b) Rate the apps in terms of grade A, B, C and D depending on how **compliant** they are with your checklist. And specify the **rubric** you use for such ratings (i.e., reasons for providing such grading). [10 points]

c) Construct a simple word-vector matching (you can use python wordnet) algorithm to prototype the policy compliance system. So, you would extract **key words** from **each checkpoint** and create a bag of words (including synonyms or similar words) and then **mark/identify sentences or segments** related to that checkpoint inside the policy (this would reduce the manual task of reading the full policy).

Evaluate how **accurate** your prototype is in correctly finding all the relevant segments in the privacy policies of the **three selected** apps in part (a), compared to your manual analysis. **Remember your uploaded code should list/print the privacy policy segments/sentences that match your checklist items**. Similar output as part (a) but using the automated approach this time. [15 points]

Table format example

Policy segment **(automated detection)**	Checkpoint
An operator of a Web site or online service shall retain personal information collected online from a child for only as long as is reasonably necessary to fulfill the purpose for which the information was collected.	Data retention policy (key words: retain, retention, kept)

d) Now repeat part (c) but use a large-language model like ChatGPT (using the free web console version) and ask it to find statements that might match the 8 checkpoints. You will need to check by chunking the policy in parts as the whole text might not fit with the character limit. Produce similar outputs as part (c) for each checkpoint and the corresponding text that talks about a checkpoint, if any. Note you can prompt ChatGPT to not only determine if a policy text contains information about a checkpoint but also highlight the part that is relevant. Such instructions can be passed through the portal. [15 points]

Table format example

Policy segment **(ChatGPT detection)**	Checkpoint
An operator of a Web site or online service shall retain personal information collected online from a child for only as long as is reasonably necessary to fulfill the purpose for which the information was collected.	Data retention policy (prompt: *"Determine if the following text mentions how long user information is stored; also highlight corresponding text, if found")*

e) Using the same rubric as defined in part (b), compare the grading of the apps for the different approaches? Which approach performed better and why? [15 points]

Report on PDF

Policy/App	Grade using manual approach	Grade using automated approach	Grade using ChatGPT
App1			
App2			
App3			

Submission:

You have to submit three files:

1. Merge all the written parts into a single pdf file named <your unity id>_HW5.**pdf**.
2. Rename the program file you used for as <your unity id>_HW5_QX.extension (e.g., .c/.cpp/.java/.py for question X).
3. Add a README file regarding how to run your code.

Zip all files into <your unity id>_HW5.zip and submit the zip file on Moodle.