Engineering Privacy in Software (Spring 2020): Semester Project

Over the course of the semester students will work in groups on a privacy-oriented software project, meeting a series of milestones designed to guide progress from understanding specific user concerns to delivering running code. Regular in-class progress presentations will provide opportuntities for students to give and receive feedback as well as discover and address common challenges. Groups will be assigned based on the programming, organization, and writing skills of students and interest in specific topic areas.

Project Topics

There are three topic areas for class projects: intimate partner violence, civic engagement, and refugee assistance. These topics have been chosen as they present timely and complex privacy challenges for users and software engineers.

Intimate Partner Violence (IPV)

> "Intimate partner violence is a serious, preventable public health problem that affects millions of Americans. The term "intimate partner violence" describes physical, sexual, or psychological harm by a current or former partner or spouse. This type of violence can occur among heterosexual or same-sex couples and does not require sexual intimacy." - *U.S. Centers for Disease Control*[1]

While smartphones and other technologies have produced numerous societal benefits, they have also made it difficult for abuse victims to seek help and escape their abusers. One recent academic study found that "survivors of intimate partner violence increasingly report that abusers install spyware on devices to track their location, monitor communications, and cause emotional and physical harm".[2] Likewise, another study concluded that "the sociotechnical and relational factors

[1] https://www.cdc.gov/violenceprevention/intimatepartnerviolence/index.html

[2] Chatterjee, Rahul, et al. "The Spyware Used in Intimate Partner Violence." 2018 IEEE Symposium on Security and Privacy (SP). IEEE, 2018.

that characterize IPV make such attacks both extremely damaging to victims and challenging to counteract, in part because they undermine the predominant threat models under which systems have been designed".[3]

As these studies show, there is a pressing need for privacy engineering approaches to be applied to the topic of IPV. Groups working on this topic will need to design systems which preserve their privacy of IPV victims from the individuals closest to them. Such systems may provide means for victims to seek out information, contact help centers, or other tasks (subject to instructor approval). Ronda Fleming from the Women's Center & Shelter of Greater Pittsburgh will come to class to talk about the challenges of providing assistance to victims of IPV.

Refugee Assistance

> "The majority of refugees are traveling with mobile phones...It is a vital tool and for people who have taken so little with them from their homes, and lost most of what they had along the way, their phones are among their most valued possessions." - *Kate Coyer, Harvard University, Central European University*[4]

There have been several refugee crises in the United States and Europe in recent years, with countless people fleeing violence and economic challenges. Smartphones have proven to be essential tools for refugees as they allow for accessing maps, finding ways to seek safe passage, coordinating among groups, and assisting with resettlement. [5] However, smartphones also allow allow authorities to track the location of users, lead to the spread of misinformation, and may not have reliable internet connections.

Groups working on this topic area may focus on systems which facilitate the dissemination of vetted knowledge to refugees (eg legal resources from an assistance organization), allowing aid organizations to anonymously communicate with refugees, allowing separated individuals to find each other, or a topic of the team's

[3] Freed, Diana, et al. "'A Stalker's Paradise': How Intimate Partner Abusers Exploit Technology." Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems. ACM, 2018.

[4] https://www.cnbc.com/2015/09/11/how-smartphones-are-helping-refugees-in-europe.html

[5] https://www.npr.org/2015/09/19/441754735/a-harrowing-journey-into-europe-aided-by-apps-and-internet-access

choosing (subject to instructor approval). Professor Kate Coyer, who assisted Syrian refugees in Hungary in 2015, will speak to the class about technology use by refugees.[6]

Children's Technology

> "When we talk about teens in the early stages of adolescence, we're talking about a brain that's under construction...it's not so much about how they'll behave online, but whether they are ready for what they're going to encounter." - *David Anderson, Senior Director, Child Mind Institute* [7]

Since the early days of the commercial Internet, the need to protect the privacy of children has been viewed as paramount. This view resulted in the 1998 passage of the Children's Online Privacy Protection Act (COPPA) in the United States. The goal of COPPA is to protect "children's privacy by giving parents tools to control what information is collected from their children online".[8] Despite being law for over 20 years, COPPA has been routinely criticized for failing to provide meaningful protections for younger Internet users[9]. Recognizing this problem, the FTC has recently stepped up enforcement, fining Google's YouTube division $170 million dollars for illegally collecting "personal information from children without their parents' consent".[10]

For this project, groups will create a web app which serves the needs of children and their guardians in a way which is privacy protective. Projects may provide interfaces for parents to provide verifiable consent for service usage by their children (eg by securely uploading identification documents), a service which uses machine learning to mask faces of children in photos (eg by leveraging existing datasets and techniques [11]), or a topic of the team's choosing (subject to instructor approval).

[6] https://cmds.ceu.edu/article/2015-09-15/kate-coyer-launches-grassroots-initiative-help-refugees-access-free-wifi-and

[7] https://www.wsj.com/articles/how-13-became-the-internets-age-of-adulthood-11560850201

[8] https://www.ftc.gov/enforcement/statutes/childrens-online-privacy-protection-act

[9] https://www.wsj.com/articles/how-13-became-the-internets-age-of-adulthood-11560850201

[10] https://www.ftc.gov/news-events/press-releases/2019/09/google-youtube-will-pay-record-170-million-alleged-violations

[11] https://talhassner.github.io/home/publication/2015_CVPR

Jennifer Sydeski Hurd from local start-up Connect Wolf will come speak to the class about challenges in the children's technology space.

Milestones

Each project will be different and evolve over the course of the semester, meeting the following milestones will keep you on track and teach you valuable skills.

Skill Quiz and Project Preferences

- Due: January 15, In Class
- Description: To aid the instructor in creating groups we will have a short in-class quiz to gauge familiarity with web development skills and to signal preferences for project topics. This information will be used solely to assign you to groups and there is no grade attached.

Work Organization Plan

- Due: January 22, 5pm
- Description: Submit a two page detailed summary of the communication tools, task managers, and software stack your team plans on using. Teams should discuss the tools and programming languages they are proficient in and reach a consensus on how they plan on organizing. For example, "We will use email and WhatsApp for communication, Jira for task management, and write our app in Java". The work organization plan should also specify areas of primary focus for each team member, such areas are "coding", "writing and presentation", and "organization and planning". Once instructors provide feedback on your plans, teams should move forward with the the next milestone.

To-Do Web App

- Due: February 5, 5pm
- Description: To ensure you are prepared to develop a custom web application you will first demonstrate your ability to code a simple "to-do" application.

The application must be deployed on a web server and will allow a user to perform "**CRUD**" actions, namely: **C**reate a to-do item in a list, **R**ead a list of existing to-do items, **U**pdate/modify any items, and **D**elete items. The application should follow the MVC framework: data should be stored in SQL and you should have both back-end and front-end code which handles interaction. There are numerous examples of CRUD applications on the web, you are highly encouraged to take inspiration from tutorials, but copying code outright will constitute plagiarism and be penalized. If you are using a web application framework you are free to leverage built-in functions. Note there is no requirement for user authentication in this step. Students will submit a URL of their system on Canvas.

Authentication fo To-Do Web App

- Due: February 12, 5pm
- Description: For this milestone you will add authentication to your app. You must have a login page which directs users to their own to-do lists. You do not need to implement an account creation system and you may use hard-coded accounts. Students will submit a URL of their system on Canvas along with account credentials.

Project Proposal

- Due: February 19, 5pm
- Description: The group will turn in a two page high-level description of the project which fulfills the criteria listed below. **Note that this document may be subject to significant revision based on what you learn from our speakers and lecture materials, it is a proposal, not a contract.**
  - List of team members and email addresses
  - Background on the topic including at least one academic article, one press article, and one example of an existing project in the space
  - Description of the specific users of your system, the privacy threats they face, and the problem your software solves for them

  - Description of how your system will work and the steps you will take to protect against threats to user privacy
  - High-level details on your chosen technical stack and the affordances it brings to your project

Proposal Presentations

- Due: February 19, In Class
- Description: This presentation should be in 2 parts. The first part will be 5 minutes and will be an elevator pitch. Use this time as if you were pitching the idea to funders, or to upper management in your organization. You will need to explain why this project is important and who will use it. The second part will be a 5 minute presentation for your peers, in which you provide the technical background on the project (as in the project description). You should consider part one to be *why*-focused, and part two to be *how*-focused.

Project Requirements

- Due Date: March 4, 5pm
- Description: Submit a 10-15 page report detailing the requirements you have for your project. This is still high-level, but unlike the proposal you are now committing to a feature set and design. Your requirements report must include the following:
  - Introduction which lays out goals of the system
  - User stories for main types of users
  - Diagrams and wireframes to show how features will work from a UI perspective
  - Diagrams to show how features will be implemented (you can use UML if you like)
  - Database schema
  - Data you will return for Subject Access Requests

  - Description of how you will meet the objectives of the eight privacy design strategies from Jaap-Henk Hoepman's Little Blue Book

Requirements Presentation

- Due Date: March 4, 5pm
- Description: Each group will present a 15 minute presentation on their requirements and participate in an in-class critique.

Project Critiques

- Due Date: March 6, 5pm
- Description: Each individual must read another teams' project requirements and provide 1-2 pages of constructive criticism. Imagine you are a kind and helpful colleague or boss who wants this project to succeed, and will ask insightful questions to help clarify the project before work gets underway.

Revised Project Requirements, Timeline, and Responsibilities

- Due Date: March 18, 5pm
- Description: Based on instructor feedback provide a revised version of your project requirements document. You should also add two new sections:
  - A timeline which includes major and minor milestones, tasks needing to be completed, and time estimates for how long such tasks will take. Use the assignment deadlines as your guide for this section, be as precise as possible, and err on the side of over-estimating the time needed to complete a step.
  - A responsibilities list which provides a break down of which team members will be responsible for which milestones. Note some team members may have meta tasks such as coordinating work and those tasks should be included in this section even if they are not tied to a specific deliverable.

Prototype One

- Due Date: April 8, 5pm
- Description: For this assignment you will submit the first working prototype for your projects which should follow directly from your requirements document and timeline.

  The deliverables are as follows:

  - URL where instructors can view your project, this should be a staging server, not your dev branch
  - Code with inline documentation, detailed enough for a competent programmer to read who does not have familiarity with your language or framework of choice
  - Database schema, initialization script (if applicable), and implementation details
  - Other supporting documents as needed (eg notes on plans for UI, etc)
  - Submit all materials as a zip file

  Requirements are as follows:

  - Account creation and user authentication implemented (subject to revision, hard-wired passwords acceptable)
  - Top-level navigation complete
  - Landing pages for all sections of site, for all users (eg the dashboard and action pages for clients and admins)
  - Placeholders for all functionality, both on front-end and back-end (eg a form on the front-end, and an empty function on the back-end which would produce desired functionality)
  - Relevant notifications to explain to users how their data is used (eg, "By creating an account we will store your chosen user name, password, and phone number. Your user name is used to maintain your account and chat history, your phone number is securely encrypted and used only for sending account reset information.")

Prototype Two

- Due Date: April 27, 5pm
- Description: For this assignment you will submit the first working prototype for your projects which should follow directly from your requirements document and timeline.

  The deliverables are as follows:

  - URL where instructors can view your project, this should be a staging server, not your dev branch
  - Code with inline documentation, detailed enough for a competent programmer to read who does not have familiarity with your language or framework of choice
  - Database schema, initialization script (if applicable), and implementation details
  - Other supporting documents as needed (eg notes on plans for UI, etc)
  - Submit all materials as a zip file

  Requirements are as follows:

  - Ability to create and delete user accounts, fully functional, secure, and private
  - Implementation of all features.
  - No penalties for bugs provided they are explained in supporting documentation and inline in code (eg "Landing page is not displaying most recent status updates, we believe the bug is in the file 'status_updater.py'")

Final Presentation and Tech Demo

- Due Date: April 29, In Class
- Description: Each team will prepare a 15 minute highly-polished presentation which should be treated as a pitch to non-profit organizations to use your product. The presentation must:

  - Outline the motivation for the project, including background on the topic area, challenges faced by users, and challenges faced by non-profit organizations seeking to help users
  - Specify how your site provides ways to solve the problems faced by user and non-profits
  - High-level discussion of how you implemented your solution, this should cover how data is collected, stored, and processed, user on boarding process, and any specific features unique to your solution. You should not describe technical implementation aspects in detail as the audience for this pitch is not presumed to be technical
  - A demo whereby you demonstrate both user and admin-facing features on your live site, the demo should make a non-profit eager to deploy your solution
  - You will have 5 minutes for Q/A. We will be doing all three groups in one class session so time will be strictly kept.

Submission of Final Project

- Due Date: May 01, 5pm
- Description: The following must be submitted:
  - URL for your project, to be made available until you have received your final grade and not modified after May 8th
  - URL will direct to a feature complete site where instructors may create accounts, and engage in all user behaviors
  - Final code, with inline documentation
  - Final report

  The final report must include the following in 7-10 pages :

  - Motivation and objectives of project goals
  - Motivation and objectives of implementation decisions

  - Detailed description of how each PbD strategy from the little blue book has been addressed
  - Detailed description of three biggest implementation challenges
  - The following elements of a Privacy Impact Assessment:
    * Granular description of stakeholders
    * Very specific stakeholder consultation plan
    * Granular analysis of information flows with diagrams
    * Identification of privacy risks and solutions
    * Explanations of tradeoffs made between usability and privacy

Presentations and Deliberation

Throughout the semester each team will present status updates to the class. These updates be 5-10 minute presentations which will cover progress made, challenges faced, and outstanding issues. Presentations will be followed by questions, feedback, and discussion with the class and instructors. Participation grades will be based on both the quality of presentations and feedback given. It is not necessary that all team members speak each week, but by the end of the semester everybody is required to have presented at least once.

Technical Considerations

Your final project will be a system which users may access from a website or mobile app, solves a specific problem, and protects privacy. Beyond these requirements, teams are free to choose any programming languages or frameworks they are comfortable with. Note that teams should discuss and agree on tools *prior* to writing any code. Freedom in choosing tools means documentation must be clear and comprehensive enough that somebody not familiar with the tools you choose will understand your code. This means both high-level technical documentation will be needed as well as extensive code comments.