International Journal of Science and Research (IJSR)
ISSN: 2319-7064
SJIF (2022): 7.942

Ethics in Software Engineering: Privacy

Prasanth Yadla

Department of Computer Science North Carolina State University Raleigh, NC 27695
Email: *pyadla2[at]alumni. ncsu.edu*

Abstract: ***The rise of digital technologies has intensified privacy concerns, as devices and services increasingly expose personal information. From smartphones and computers to smart assistants and web browsing, users face risks of unauthorized recording, data transmission, and tracking. These vulnerabilities are exacerbated by unencrypted networks and unethical software practices. This article underscores the importance of integrating privacy into the software development lifecycle and explores the implications of neglecting it. Real - world cases, including Google's StreetView, highlight privacy breaches and their consequences. We propose actionable design principles to enhance user privacy and evaluate their effectiveness in mitigating risks, emphasizing the ethical responsibility of developers.***

Keywords: Software Engineering, Ethics, Risk management

1. Problem Statement

With the advent of technology and the digital world, privacy is becoming increasingly concerning. Privacy, according to Wikipedia [2] is defined as the ability of an individual or group to seclude themselves or information about themselves, and thereby express themselves selectively. [3] Privacy is interwoven into our daily life in many ways. Almost everyone owns a computer or a smartphone these days. There is no guarantee that our smartphone is not actively listening to use and not sending any data the the manufacturer or third parties. There is no proof that the video camera or microphone in our computer is turned off and is not passively recording. These even motivated many people to cover their video slit in mechanical ways.

With the rise of peer - to - peer devices, different computing devices are more connected than ever. Un - Encrypted network packets floating all around motivated adversaries to sniff them and steal personal information, including but not limited to name, address, credit card information, medical records., etc.

Browsing the web is also a risk to privacy. When visiting websites, there is also tracking conducted passively. Website developers embed 3rd party tracker cookies in browser, so that your activity is monitors within and outside of the website you are visiting.

Smart Assistant technology like Alexa, Siri, Google Home are becoming more prevalent in US households. While they may be useful, there are many privacy risks associated with the same. They are always switched on, and could be recording household conversations. Recent studies have also indicated that they can be invoked by inaudible sound from far - off also.

All the above problems arrive when the manufacturers or developers of the software product or service did not take privacy into their architectural design and requirements.

The purpose of this article is to highlight the necessity of privacy in the context of software engineering and ethics. The far - reaching and inevitable consequences of privacy violations can have it's importance in every aspect of the phases, like requirements gathering, design, implementation and finally deployment and release. The real - world facts and evidence surrounding breaches, it's implications and relation to the unethical practises of software and indirectly the developers who coded the software. We also present some useful design suggestions and **ideas/techniques and options to improve and respect** the privacy of customers. Lastly we evaluate and reason our suggestions/choices in the general context.

We also study in - detail a privacy case, with Google's StreetView feature [4]. This feature displays searchable aerial and street - view photographs of neighborhoods, city blocks, stores, and even individual residences. From the beginning, the privacy concerns were evident, but it didn't take people long to realize that photographs of customers, children, adults and private pictures were made public through streetview. This data was also being used by Burglers and Criminals to plan their offense. There was no process to remove these images initially, but after a lot of complaints and outpouring of media, they finally made a systemtic process to remove the sensitive images.

2. Relevant Facts

To assess the risk and far - reaching impact of privacy, we analyze a few real - world incidents of data breach and legal consequences. [5] Google's StreetView became the center of a new privacy scandal. It had been discovered that Google vehicles doing drive - by photography had been collecting data from unencrypted Wi - Fi networks, including SSID's, device identifiers, passwords and email content. They initially refused to consent that the data was stored, but they eventually agreed that the data was collected and retained as well. Google initially blamed one engineer for the breaches. However, it was later revealed that the engineer communicated with the management regarding this feature. At one point, 12 countries have launched formal investigations of Google for breaches and later U. S states have filed lawsuits with fines over 7 million for data breaches. Google finally acknowledged their culpability for privacy breaches and promised to improve. [6] In the case of twitter, there was a massive data breach in 2018, the plain text passwords of 330 million users were exposed to the twitter internal network. Traditionally, twitter encrypts and masks the passwords of users by using a hashing algorithm to convert the password into a random string of characters.

This random string is stored in the system and various internal logs. However, a bug in the software algorithm caused the passwords to be stored in plain text and anybody with access could view the passwords. Twitter immediately reached out to it's respective users to change their passwords. This is indeed another example of privacy breach. However, Twitter did remain slightly ethical by informing users immediately and publicly.

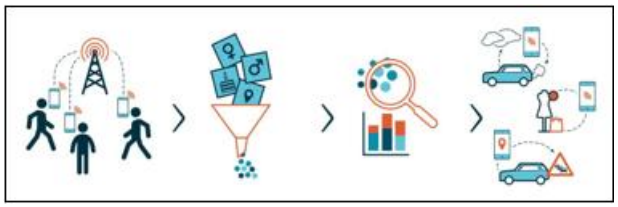

Figure 1: De - Identification Use - Cases

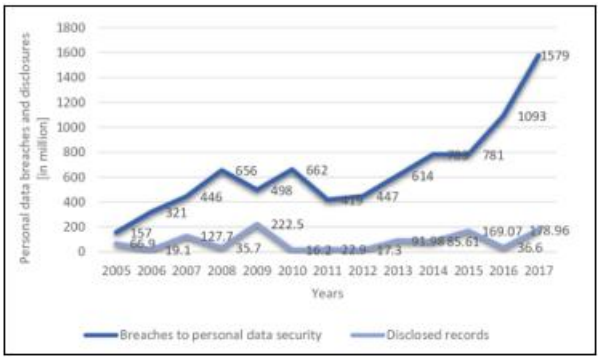

Figure 2: Increase in Data Breaches

[7] The Facebook–Cambridge Analytica data scandal was a major incident where millions of Facebook user's personal data was sold to Cambridge Analytica by Facebook without the consent of the individuals. It is claimed to be the largest known leak in social networking history. The data is supposedly used for political campaigning and targeted advertising during the 2016 elections. Facebook was fined by the UK for exposing the data to a serious risk.

Stakeholders

There are many stakeholders in the context of privacy.

- Most Importantly, the customer and users who are sharing thier data in good faith. They are the most important stakeholders
- The management of the company that provides the service. This includes the CEO, owner., etc. They are eventually held responsible for any legal challenges.
- Investors at company have spent a great deal of money, wouldn't want their deal to go bad.
- Finally, the developers who wrote the software or service. They will be known to write privacy - ignorant code.

There will always be varying levels of importance of each of the mentions stakeholders depending on the context of the problem in hand.

3. Considered Options with Ethics Tests

A. Anonymize the Data

Data anonymization is a type of information sanitization measure whose intent is privacy protection. It is the process of remove personally identifyable information from the data, so that the data can remain anonymous. In the case of Google's StreetView, the faces and sensitive pictures will be **blurred**. It will only show the part of the image relevant to StreetView. The faces of people will not be identifiable.

With regards to the harm test, the aggregate statistics obtained from any kind of analytics might not represent the exact value as there is loss of information is present in the dataset after sanitization. However, the sensitive information will be masked so that even if the adversaries gain hold of the data, they won't be able to target a specific individual.

This decision treats other people with respect, as their data is being anonymized to the best extent possible and they would not be revealed.

For example, if it was my data that was being collected, say, demographic and age is revealed, but my specific information like, height, weight, nationality is hidden, only aggregate information would be revealed, and would make me feel confident with giving access to my data because I am anonymous.

Another example is in the context of web - browsers. Websites hosted by the anonymized network **Tor [**1], uses an onion strategy to obfuscate the IP address of it's users. Tor has been very successful in preserving privacy.

B. Implement Network Security Protocols and Authentication Mechanisms

Use of encryption schemes will help prevent exposing any personal information over the network. Mechanisms like Public Key cryptography will encrypt data, so that in a case that the adversary manages to get hold the data at rest or at transfer, they will not be able to decrypt it. Using these schemes in routers and internet modems public restaurants can prevent the incident in Google StreetView Vehicles, which attempted to collect personal data from un - encrypted Wi - Fi networks, including SSIDs, device identifiers, passwords., etc.

The only harm this option can cause is the **overhead in latency and compute resources**. Some complex algorithms like RSA and Kerberos requires lot of communication before - hand so they might slow up initial setup. However, the **benefits of this options far outweigh the communication overhead** as we are exposing less data. If in a public place, the Wifi required 2 factor authentication, I would be more confident of using it as it is encrypted. The WWW internet is moving towards encrypted data, with average encrypted web traffic at 50% at 2014, to 80% - 90% today.96% of the world's top 100 sites default to HTTPS.

C. Tell what personal data is being collected and for what purpose

Users should naturally have the privilege of knowing whether their personal data is being collected by the service they are utilizing. Companies might be requesting users for sensitive information like SSN, DOB, age, medical information., etc. Users should have the right to also know how this collected data will be used and for what purpose.

In the case of Google StreetView, Google Vehicles should release publicly the places the vehicle is collecting data and for what purpose it is collecting it for. This would induce more transparency. This option does not harm, as users are gaining access to information about the collection and usage of their data. This is in - fact a fundamental right of the constitution of the United States. If I knew where and what personal information of mine is being collected, my peace of mind will definitely be improved.

D. Give users the right to view/delete or restrict information collection

If the users personal data has been collected, users should be provided with the opportunity to control their information which is stored by the company. For example, the user can view the data the company collected and can choose to delete it permanently if he wishes to.

They can also restrict information collection at any point in time. In the case of GoogleStreet View, Google has to provide the opportunity to users requesting sensitive images deleted. Google did not initially have a streamlined process for this. This option definitely does not provide any harm, as we are getting access to our own data being collected. We have control over our data and hence exercising our fundamental rights. I would feel more confident if I have control of what I can view/delete/restrict my personal data collected.

E. Inform users in case of a data breach

In case of a data breach, the company responsible should not delay the notice to relevant users. They should immediately notify the customer what information was breached, ideally withing 72 hours of breach. This would ensure minimal consequences on users after breach.

In the case of Google StreetView, they did not provide notice of sensitive photographs, until an outpouring of complaints and media stories. They also did not acknowledge the fact that they were storing data collected by Google Vehicles. Timely notice of a data breach does more good than harm. For example when NITRO company informed of a database leak of it's PDF software, employees at Google, Apple quickly deleted their information to prevent further repercussions. I would be most satisfied if I was informed of any leak/breach of my data at the earliest.

4. Conclusion

Using a few of the measures suggested above can significantly improve the privacy guarantees of users. In - fact, Apple mentions that Privacy is a fundamental right and they design their products to protect and give more control over your information.

On iOS 14, it is have observed that all these above options have been implemented in the iPhone. Studies have also shown that customers are more pleased with privacy guarantees.

Following the given options would substantially reduce the impact of Google's StreetView Data Leak Scandal. Ranging from blurred faces, encrypted Wifi - networks, to data breach notice, this will be a major step in the frontier of Privacy and Ethics in Software Engineering.

Most of these options have also been compiled and legally established privacy acts such as HIPAA (for medical records), COPAA (for children's data), CCPA (California's Privacy Rights Act.).

References

[1] Tor: Anonymized Browser, https: //www.torproject. org/
[2] Privacy https: //en. wikipedia. org/wiki/Privacy
[3] Privacy in the Contexts of Everyday Life https: //citeseerx. ist. psu. edu/
[4] Ethics Privacy http: //courses. ics. hawaii. edu/ReviewICS314/
[5] Street View. https: //epic. org/privacy/streetview/
[6] Twitter Security Flaw. https: //www.theverge. com/2018/5/3/17316684/
[7] Facebook–Cambridge Analytica data scandal. https: //en. wikipedia. org/wiki/
[8] Apple Privacy. https: //www.apple. com/privacy/