Privacy - Overview

INFR11158/11230 Usable Security and Privacy

Dr. Jingjie Li

08/03/2024
Overview

• Warm-up
• A taxonomy of privacy
• Privacy by design
• Contextual integrity
• Take-home
Contact tracing
Have you used contact tracing? How’s your experience?
What is privacy?
Defining privacy

• There are many definitions
  • The right to be let alone
  • The right to control one’s own data

• Many common security goals overlap with privacy ones
  • Confidentiality
  • Access control of information
  • Protection from unwanted intrusions
A taxonomy of privacy (by Daniel J. Solove)
A TAXONOMY OF PRIVACY

INFORMATION PROCESSING

AGGREGATION
- Combining of various pieces of personal information
- A credit bureau combining an individual’s payment history from multiple creditors

SECONDARY USE
- Using personal information for a purpose other than the purpose for which it was collected
- The U.S. Government using census data collected for the purpose of apportioning Congressional districts to identify and inter those of Japanese descent in WW2

EXCLUSION
- Failing to let an individual know about the information that others have about them and participate in its handling or use
- A company using customer call history without the customer’s knowledge, to shift their order in a queue (i.e. “Your call will be answered in this order (HOT) received”)

INSECURITY
- Failing to protect information
- An ecommerce website allowing others to view an individual’s purchase history by changing the URL (e.g. enterprivacy.com/NID=123)

IDENTIFICATION
- Linking of information to an individual. [Sometimes called ‘singing out’]
- A researcher linking medical files to the Governor of a state using only date of birth, zip code and gender

INFORMATION DISSEMINATION

DISCLOSURE
- Revealing truthful information about a person that impacts their security or the way others judge their character
- A government agency revealing an individual’s address to a stalker, resulting in the individual’s murder

EXPOSURE
- Revealing a person’s nudity, grief, or bodily functions
- A store forcing a customer to remove clothing revealing a colostomy bag

BREACH OF CONFIDENTIALITY
- Breaking a promise to keep a person’s information confidential
- A doctor revealing patient information to friends on a social media website

INCREASED ACCESSIBILITY
- Amplifying the accessibility of personal information
- A court making proceedings searchable on the Internet without redacting personal information

APPROPRIATION
- Using an individual’s identity to serve the aims and interests of another
- A social media site using customer’s images in advertising

DISTORTION
- Disseminating false or misleading information about a person
- A creditor reporting a paid bill as unpaid to a credit bureau

COLLECTION

SURVEILLANCE
- Watching, listening to, or recording of a person’s activities
- A website monitoring cursor movements of a visitor while visiting the website

INTERROGATION
- Questioning or probing for personal information
- An interviewee asking an inappropriate question, such as marital status during an employment interview

INVASION

INTRUSION
- Disturbing a person’s tranquility or solitude
- An augmented reality game directing players onto private residential property

DECISIONAL INTERFERENCE
- Intruding into a person’s decision making regarding their private affairs
- A payment processor declining transactions for contraceptives

https://www.jstor.org/stable/40041279
What are the privacy risks of contact tracing apps given the framework?
Privacy by design
Privacy by design – definition

Framework for building privacy proactively into new systems, proposed in 2009. Widely accepted as an international standard for good privacy engineering. GDPR also basis some of its principles on Privacy by Design.

• Proactive not Reactive; Preventative not Remedial
• Privacy as the Default
• Privacy Embedded into Design
• Full Functionality – Positive-Sum, not Zero-Sum
• End-to-End Security – Lifecycle Protection
• Visibility and Transparency
• Respect for User Privacy
Privacy by design – strategies
Privacy by design – Minimize

• Definition
  • Limit as much as possible the processing of personal data.

• Tactics
  • Select only relevant people and relevant attributes for processing.
  • Exclude people or attributes in advance of processing it, or better delete it.
  • Strip away (remove) data as soon as it is no longer needed. Also, auto-delete after it is no longer needed.
  • Destroy data that is no longer needed. Build systems that support complete destruction of data and do not leave it in unexpected parts of the system.

• Example
  • “Google announced a revised log retention policy, saying ‘we’ll anonymize IP addresses on our server logs after 9 months,’ instead of the previous 18-24 months.”
Privacy by design – Separate

• Definition
  • Separate the processing of personal data as much as possible.
• Tactics
  • **Isolate.** Collect and process data in different databases or applications.
  • **Distribute** the collection and processing over different entities. Use the equipment of the user as much as possible.
• Example
  • Tor’s Onion Routing structure along with the many organizations that host exit nodes ensures that no one entity has visibility over the whole network. Data is distributed across many nodes, knowledge of who sent the data and where it is going is also strictly distributed between nodes.
Privacy by design – Abstract

• Definition
  • Limit as much as possible the detail in which personal data is processed.

• Tactics
  • **Summarise** detailed attributes into more coarse-grained, general attributes. For example, use age categories instead of birthdate.
  • **Group.** Aggregate information about a group of people instead of processing data individually. Present data as averages.
  • **Perturb** data values to create an approximisation, for example, by adding random noise.

• Examples
  • Pubs are required to check that patrons are above the legal drinking age. But they don’t need to know the exact birthdate of the person, just if they are above that age or not.
Privacy by design – Hide

• Definition
  • Protect personal data or make it unlinkable or unobservable. Make sure it does not become public or known.

• Tactics
  • **Restrict** access to personal data. Setup a strict access-control policy.
  • **Obfuscate.** Use tools like encryption, hashes, and pseudonym’s to ensure that only people with the ability to decrypt can get the data.
  • **Dissociate.** Break the link between events, persons, and data.
  • **Mix** data into larger sets to ensure that data is not easy to re-connect.

• Examples
  • Most user studies promise to disassociate participants’ names from their data: “you will be assigned a random participant number, your name will be stored separately from the data we collect.”
Privacy by design – Inform

• **Definition**
  • Inform data subjects about the processing of their personal data in a timely and adequate manner.

• **Tactics**
  • **Supply** resources on the processing of personal data including, policies, processes, and risks. Provide information about *which* personal data, *how* processed, and *why* processed.
  • **Explain** clearly why data needs to be processed.
  • **Notify** users when their data is being processed, shared with third parties, or after a data leak.

• **Examples**
  • Apple shows an icon whenever location data is being accessed by an app.
Privacy by design – Control

• Definition
  • Provide data subjects adequate control over the processing of their personal data.

• Tactics
  • **Consent.** Ask users for their explicit consent to data processing.
  • **Choose.** Offer users a real choice with basic functionality available to those who opt-out.
  • **Update.** Give users the means to review and update their personal data.
  • **Retract.** Allow users to retract (or delete) their personal information.

• Examples
  • Cookies allow users to consent (cookie pop-up) and delete the cookie (retract).
Privacy by design – Enforce

• Definition
  Commit to processing personal data in a privacy-friendly way, and adequately enforce this.

• Tactics
  • **Create.** Make a Privacy Policy, and assign resources to execute
  • **Maintain.** Uphold policy and ensure all technical and org. controls. Apply to 3rd parties
  • **Uphold.** Verify policy regularly and adjust implementation when necessary.

• Examples
  • Potential approach is to implement privacy management system like the plan-do-check-act cycle from Information Security Management Standard (ISO 27001)
Privacy by design – Demonstrate

• Definition
  • Demonstrate you are processing personal data in a privacy-friendly way.

• Tactics
  • **Record.** Document all (important steps). Record decisions and motivate them.
  • **Audit.** Regularly audit and review org. processes and how personal data is processed
  • **Report.** Provide results of audits to Data Protection Authority (DPA).

• Examples
  • Certified against internationally recognised standards for privacy friendliness like TRUSTe or EuroPriSe.
Privacy by design in development cycle
Contextual integrity
**Contextual integrity**

- Privacy is defined by how **information flows**
- Information flow is appropriate when it conforms with **contextual privacy norms**
- A contextual norm can be described by (at least) five parameters
  - data type (what sort of information is being shared)
  - data subject (who/what the information is about)
  - sender (who/what is sharing the data)
  - recipient (who/what is getting the data)
  - transmission principle (the constraints imposed on the flow/how), e.g., with one’s consent.
- New norms and flows are evaluated through their context

Contextual integrity

https://www.dli.tech.cornell.edu/post/privacy-policies-as-contextual-integrity-beyond-rules-compliance
Stop the Spread: A Contextual Integrity Perspective on the Appropriateness of COVID-19 Vaccination Certificates

Shikun Zhang
Carnegie Mellon University
Pittsburgh, PA, USA
shikunz@cs.cmu.edu

Yan Shvatzshnaider
York University
Toronto, Canada
yansh@yorku.ca

Yuanyuan Feng
University of Vermont
Burlington, VT, USA
yuanyuan.feng@uvm.edu

Helen Nissenbaum
Cornell Tech
New York, NY, USA
hn288@cornell.edu

Norman Sadeh
Carnegie Mellon University
Pittsburgh, PA, USA
sadeh@cs.cmu.edu

ABSTRACT

We present an empirical study exploring how privacy influences the acceptance of vaccination certificate (VC) deployments across different realistic usage scenarios. The study employed the privacy framework of Contextual Integrity, which has been shown to be particularly effective in capturing people’s privacy expectations across different contexts. We use a vignette methodology, where we selectively manipulate salient contextual parameters to learn whether and how they affect people’s attitudes towards VCs. We surveyed 890 participants from a demographically-stratified sample of the US population to gauge the acceptance and overall attitudes towards possible VC deployments to enforce vaccination mandates and the different information flows VCs might entail. Analysis of results collected as part of this study is used to derive general normative observations about different possible VC practices and to provide guidance for the possible deployments of VCs in different contexts.

1 INTRODUCTION

The prolonged and devastating COVID-19 pandemic has affected every aspect of people’s lives as well as the global economy. In an attempt to curb the spread of highly contagious variants, governments around the world have contemplated or adopted vaccination mandates (VMs) and vaccination certificates (or passports) (VCs) in schools, hospitals, public transportation, and other social contexts [15, 27, 42, 43, 50, 53, 62]. COVID VMs and VCs challenge established societal norms and conventions. While vaccination mandates and certificates are not new (e.g., vaccination mandates for children attending schools, “yellow cards” for travel to or from a country with a high risk of diseases such as yellow fever [55]), the sudden and unprecedented requirement to show proof of vaccination to gain access to public venues or engage in a range of daily activities has triggered a fierce global debate on the appropriateness of COVID-19 VMs and VCs in light of established societal norms and conventions, perceived privacy harms, and civil liberty expectations [9, 34, 36, 61, 69].
Take-home


- (Blog) WIRED - ‘AI Girlfriends’ Are a Privacy Nightmare