Data Ethics and Privacy

Privacy, Security, and Policy in the Age of the Internet

Presented by Dr. Clifford Neuman
Director, USC Center for Computer Systems Security
11 January 2022

Computer Security is Hard

- You can’t make a system totally secure.
- The consequences of failure may be devastating.
- Someone wants you to fail.
- The bad actors may be beyond our reach.

We make the problem harder

- You can’t make a system totally secure.
  - We make our systems bigger.
- The consequences of failure may be devastating.
  - We can’t clearly state what constitutes failure.
- Someone wants you to fail.
- The bad actors may be beyond our reach.
  - We give them remote access to our systems.

Defining the problem

Secure – A system is secure if it correctly enforces a correctly stated policy for a system.

- A system can only be secure with respect to a particular set of policies and under a set of stated assumptions.
- There is no system that is absolutely secure.

What do we want to achieve

Confidentiality
- Keep data out of the wrong hands.

Integrity
- Prevent unauthorized modification of data.

Availability
- Keep the system running and reachable.
- Prevent destruction of systems and data.

- The rest is policy and mechanism.

Mechanism

- A security mechanism is a method or tool used by the system to enforce a policy (recall our definition of secure).
  - Mechanism can support prevention, detection, and reaction.
  - e.g. access control, cryptography, availability, detection.

- A second role for security mechanisms is to enforce the underlying assumptions of the system.
  - e.g. containment, identity management.

Policy (specification)

- Policy defines what is allowed (or not) and how the system and security mechanisms should act.
  - It tells us what we mean by confidentiality, integrity, and availability.
  - Policies are enforced by mechanisms.

- Authorization (final goal of security) tells the system when an operation is allowed.
  - It depends upon policy, possibly authentication, other characteristics.
  - The answer is often simply a yes or a no.

Policy (public, ethics, agreement)

- Public policy includes laws/regulations imposed on data processors pertaining to allowed “processing” of data.
  - Often legal consequences for not properly implementing (civil or criminal).
- Agreed security policy terms include contractual requirements such as service level agreements and privacy policies.
  - Enforced in their breach through litigation (Civil), arbitration, or other means of redress.
- Ethical processing of data is based on concern for the customer or data subject.
  - Ethical, social, and societal norms may differ from location to location
    - Especially w.r.t. Privacy.
  - My require consideration of ethical issues in underlying system design.
  - Enforcement in their breach may based on “public policy” or reputation.

Privacy

- It is primarily a policy issue
  - What is permitted to be done with data
- Privacy concerns Personally Identifiable Information (PII)
  - It is the linking of data that associates it with a subject
- Security and Privacy are interrelated
  - Security is a necessary mechanisms to enforce a privacy policy, but
  - Accountability and attribution may create links to subjects that limit privacy
- Some aspects of privacy depend on user education
  - So that users understand how “their” data may be used
- Privacy might need to be balanced with other rights, such as that of free expression.

Permitted Processing of PII

- Public Policy on Commercial collection and use
  - GDPR (EU), CCPA and CPRA [Prop 24] (CA), and PIPL (CN)
  - Basic goals of accountability, control, and transparency, with some fundamental differences
- Governmental Processing of PII
  - Varies significantly from country to country
  - In the US, constitutional rights dictate procedure for access to information: predominantly 4th and 5th, but also 1st amendment.
  - In some cases, government use of data is more limited than public commercial use.
  - In some countries, government access is presumed to be unlimited.
- The Internet crosses jurisdictions – some rules are extraterritorial.
  - Creating conflicting rules for some multinational corporations

Ethical Data Processing

Transparency is a key goal of most privacy regulations.

- Required disclosure (and in some cases consent) of the purpose for which information is collected and processed.
- But what if the rules or purpose changes?

A worthwhile principle of ethical data processing then might be to take the steps necessary to ensure that data can only be used for the purposes for which disclosure was made.

- Data minimization.
- Data protection – i.e. computer security.
- Technologies limiting linkability (e.g. k-anonymity, differential privacy)

Think through potential for mis-use of your system or technology.

- By criminals, corporations, governments, and users.
- Are their biases that are present in the technology or the data sets it processes.
- Can you prevent these possibilities.

Before you start

You should determine what you must protect.
That begins with an assessment.

- Classifying what's important – Inventory of Data
  - What's important to you?
  - What do you have that's important to others?
  - What are the consequences of loss or disclosure?
  - What can a third party do with this data?

- How they get at it
  - Where is the data stored?
  - What systems do you have?
  - What software do you run (including your apps)?
  - What vulnerabilities exist in your system?

Managing Risk

- No system is absolutely secure, so the goal is to minimize risk

  Risk = *consequence* X *likelihood*

- Our calculation for risk will only be as accurate as the values we enter for *consequence* and *likelihood*.
  - For security breaches, these values are often no better than guesses.

- So, why then is risk management useful?
  - Because we know of steps that reduce *likelihood* and reduce *consequence*.

Reducing Consequence

- Consider the consequence for the data from your inventory.
  - What happens if you lose access?
  - What happens if someone else gains access?
  - It's not **your** data.

- Reducing the consequences
  - Effective backups – to protect against loss of access.
  - Don't keep the most sensitive data on connected systems.
  - Reduce what can be accessed from your systems.

Treat Personally Identifiable Information like toxic waste

Reducing Likelihood

- Risk is reduced when we reduce likelihood

  Risk = *consequence* x (likelihood = *threat* x *vulnerability*)

- Still imprecise and based on guesses.
  - Vulnerability - A weakness in a system that may be exploited.
  - Threat – The capability of an adversary to exploit a vulnerability.
  - Often just low, medium, and high.

- Reducing Threat and Vulnerability
  - Reducing “attack surface” – Blocking access – Minimizing programs and function.
  - Patching Systems – Reducing Bugs.
  - Training.

For Security Practitioners – Tools of the Trade

- Configuration Management
  - Tracks systems, data, and software
- Containment
  - Firewalls and Routers
  - Subnetworks
  - Encryption
  - Disconnected disks
- Access and Identity Management
  - Authentication – including 2nd factor
  - File access policies
- Detection and Response
  - Antivirus and Internet Security Suites
  - Security Incident Event Management Systems (Detection)
- Minimization
  - Get rid of data
  - Get rid of apps
  - Make data less accessible (move it)

Some guidelines to remember

- A system is secure if it correctly enforces a correctly stated policy.
- Policies describe rules for confidentiality, integrity, and availability.
- You can’t make a system absolutely secure.
- Privacy is a statement of policy.
- Data should only be processed for the purpose for which it was collected.
- Security and privacy require managing risk.
- The most effective and simplest steps involve minimizing our existing systems: “Less is more”.
  - Fewer programs.
  - Less access.
  - Less data.