Privacy: Introduction

Jorge Cuellar

WS 18-19

What to learn

This year (WS 18-19) the main focus of the course is PETs based on

- Cryptography methods

This set of slides contains an introduction to it

- rather at the end
  - slides are marked with 

Besides that,

- Two other slides (also marked) are very important:
  - Privacy is a set of concrete rights
  - Basic principles of the Fair Information Practices

The rest of these slides includes

- topics that we have discussed in class
- but are to be considered as "for your information"
  - not necessary for the exam

Privacy has different dimensions

Some Dimensions/Aspects of Privacy

- Spatial privacy
  - The right to be let alone
- Informational self-determination
  - "My data, my rules"
- Transparency
  - Who has data about me?
    - What conclusions are drawn?
    - what consequences do they have?
- The need to participate & share information
  - selectively
- Technological Problems
- Surveillance, Legal interception
  - Protection against terrorism and crime

A Privacy Course should define Privacy?

No: we won't

- There is no single definition of privacy

Privacy: Main Definitions

Wikipedia (a plausible start)

Privacy is the ability of an individual (or group)

- to seclude himself, or information about himself
  - and thereby express himself selectively

The boundaries and content of

- what is considered private
  - differ among cultures and individual
- When something is private to a person
  - it usually means that something is
    - inherently special or sensitive to them

Privacy: Main Definitions

Samuel Warren & Louis Brandeis, The Right to Privacy

- 4 Harv. L. Rev. 193 (1890)
  - This foundational article
  - inspired the development of privacy law in 20th century
    - argued that privacy was protected by the common law as
    - "the right to be let alone"

Privacy: Main Definitions

William Prosser, Privacy, 48 Cal. L. Rev. 383 (1960)

- surveyed all the common law privacy tort cases
- identified four central interests to be protected
  - which remain in widespread use today

Privacy: Main Definitions

Prosser: 4 Torts (legal wrongs, injustices)

1. Intrusion upon the plaintiff's seclusion or solitude, or into his private affairs
2. Public disclosure of embarrassing private facts about the plaintiff
3. Appropriation, for the defendant's advantage, of the plaintiff's name or likeness
4. Publicity which places the plaintiff in a false light in the public eye

Privacy: Main Definitions

Alan Westin 1967

- The claim of individuals, groups and institutions
  - to determine for themselves, when, how and to what extent
  - information about them is communicated to others

Privacy: Main Definitions

Solove: Privacy is not a single concept

(See slides below):

- Privacy is not reducible to
  - a singular essence or
  - a single common characteristic
- It is a plurality of different things
  - that do not share one element in common
  - but bear a resemblance to each other

Solove: Privacy Taxonomy

DJ Solove: Let us define Privacy bottom-up:

Starting from a set of Problems

- that we feel are "privacy problems"

Privacy describes a set of "problems"

- draw from a common pool of similar things

Privacy is not a single concept but a set of

- "family resemblances"

Solove: Family Resemblances

DJ Solove: suggests to abandon

- the traditional way of
  - conceptualizing privacy

instead understanding it with

- Wittgenstein's notion of "family resemblances"

No single common characteristic

Privacy is not reducible to

- a singular essence or
- a single common characteristic

- It is a plurality of different things
- that do not share one element in common
  - but that nevertheless
    - bear a resemblance to each other

Solove: Privacy Taxonomy

- Privacy violations consist of a web of
  - related problems that are
    - not connected by a common element, but
    - bear some resemblances to each other
- Privacy should be conceptualized
  - bottom up rather than top down
    - from particular contexts rather than in the abstract

Solove: Privacy Taxonomy

. . . is a bottom- up ongoing project

- As new problems arise
  - taxonomy will be revised
- Whether a particular problem enters this classification
  - it is not as important as whether it is "recognized" as a "privacy problem"
  - The important things are:
    - it is a problem that should be avoided
    - it has some clear resemblance to privacy
- Regardless of whether
  - we label the problem as part of the privacy cluster,
    - it still is a problem, and
  - protecting against it still has a value

Solove: Privacy Taxonomy

- For example, distortion,
  - which involves disseminating
    - false or misleading information about a person
- Some argue that distortion is not really a privacy harm
  - because privacy only involves true information
- Regardless of whether distortion is "recognized"
  - as a privacy problem, it is
    - 1. nevertheless a problem
    - 2. it bears some resemblance to other privacy problems
- Thus, viewing and classifying them together
  - might be helpful in addressing them

Privacy: Main Definitions

IT Privacy Threats

LINNDUM refers to the threats

- Linkability
- Identifiability
- Non-repudiation
- Detectability
- Disclosure of information
- Unawareness
- Non-compliance

IT Privacy Threats

LINNDUM: Linkability

- Possible to link IoI (items of interest) together
  - for instance, events or records
  - which belong to the same data subject,
    - or otherwise related subjects
- Example:
  - if all records have the same pseudonymized ID
    - then it is possible to link them
- Examples of linkable IoIs:
  - anonymous postings written by the same person
  - web page visits by the same user
  - entries in two databases related to the same person
  - people related by a friendship link
  - RFID responses

IT Privacy Threats

LINNDUM: Linkability

- Note that the individual may not be identifiable
  - identifiability is orthogonal to linkability
- The risk with Linkability is
  - it can lead to Identifiability
    - A person borrows a book from the library
    - the book has an RFID . . . (continue!)
  - it can also lead to inference
    - Geo-Data (explain!)

IT Privacy Threats

LINNDUM: Identifiability

- linking identity with information
- possible to correctly assign an IoI (event or record)
  - to an identifiable person
  - with a higher probability than random

IT Privacy Threats

LINNDUM: Non-repudiation

- The inability of a victim to deny a claim
  - The attacker can thus prove a user knows,
    - has done or has said something
    - He has evidence to support this
- Typical non-repudiation examples
  - in the context of anonymous online voting systems
    - (Explain!)
  - and whistle-blowing systems where plausible deniability is a requirement

IT Privacy Threats

LINNDUM: Detectability

- to distinguish whether an entity or an IoI exists
  - by detecting whether an IOI exists
  - one can deduce certain information
    - even without access to that IoI itself
  - by knowing that a celebrity has a health record in a
    - particular hospital, you can deduce something

IT Privacy Threats

LINNDUM: Information Disclosure

- revealing personal information

LINNDUM: Unawareness

- not knowing the consequences of sharing information

LINNDUM: Non-compliance

- not being compliant with legislation, regulations, and policies

IT Privacy Threats

Addressability

- Using some information – say a pseudonym,
  - to target or address a specific individual
  - not necessarily identifiable

Inferences

- possible to draw inferences of some kind
  - from data analysis

What is Privacy?

Privacy seems to be about everything

. . . , and therefore it appears to be nothing

- "Privacy" has proven to be a powerful
  - rhetorical battle cry in a plethora of
    - unrelated contexts
- "Privacy" means many
  - different things to
    - so many different people:
- it has lost any precise
  - legal connotation that it might once had
    - J. Thomas McCarthy

Privacy is an extremely overloaded term

"Privacy is a value so complex,

- so entangled in competing and contradictory dimensions,
- so engorged with various and distinct meanings,
  - that I sometimes despair whether
    - it can be usefully addressed at all"
    - Robert C. Post, Three Concepts of Privacy,

Privacy is sometimes like oxygen

- you really feel its need, as soon as it is gone

What is the course about?

Relation of Privacy and Technology

- 1. IT poses serious threats to privacy
  - We will try to see why those problems arise
  - Motto: "IT makes surveillance easy for everyone"
  - Motto: "code becomes common use, becomes expectations, becomes law"
- 2. There are IT building blocks
  - that help to protect the privacy
    - We will see many of them: PETs

What is the course about?

1. IT-Technology is a menace against Privacy

- IT enables the collection and mining of data on a scale previously unimaginable via
  - Availability of
    - fast and cheap computers
    - massive storage devices
- This opens the door to potential abuse of individuals' information
  - There has been considerable research exploring the
    - tension between utility and privacy

What is the course about?

2. IT-Technologies (PETs) can help to protect Privacy

- Data privacy
  - Techniques for achieving privacy (or trying to)
  - . . . and their limitations
  - Privacy issues in specific settings

Note

- The 2 points do not cover the whole IT-Privacy:
  - 1. Privacy problems of new Technologies
  - 2. PETs

Example of areas that we do not cover in detail:

- Privacy Engineering / Privacy by Design / "Privacy Development Lifecycle"
- Privacy Management in an Enterprise
- "Privacy Maturity Models"
- Privacy as a Process
- Privacy Certification

Relation of Privacy and IT-Technology

- Not only new technology
  - poses serious problems to privacy
    - We will try to see why those problems arise
- But also there are some
  - technological building blocks that
    - help to protect the privacy
  - We will see many PETs

Discuss the following statements

Exercise: Are they true? Should they be true?

1. Privacy is not necessary
   - I have nothing to hide
2. Privacy and confidentiality are essentially the same
   - Explain: when encrypted health records
     - should be additionally pseudonymized/anonymized?
3. Divulging publicly know information is not a problem
4. Companies are punished by the government
   - for treating personal data in unauthorized ways
5. Privacy is a Policy Issue
   - Security a Technology one

Discuss the following statements

Exercise: Are they true? Should they be true?

1. Anonymous data needs no privacy protection
2. On the Internet, you may remain anonymous
   - Recall the cartoon:
     - "On the Internet, nobody knows you're a dog"?
3. Privacy is good for business
4. Privacy is about protecting users from harmful violations
5. Privacy is only about user control

The argument: "I have nothing to hide"

This argument assumes privacy is defending against

- a government surveillance agent
  - very trusted
  - only uses the information for a very special reason
    - say, to fight terrorism
- Most people have something to hide from somebody
- Governments and intelligence agencies are large and
  - agents may get corrupted
  - who can control what they do?

The argument: "I have nothing to hide"

Assumes Privacy is a form of concealment or secrecy

- Privacy thus means concealment of bad things
- If people believe the argument,
  - people learn not to expect privacy for financial, health or location data, preferences, etc

Surveillance generates a chilling effect

- on free speech, free association, and other rights essential for democracy
- See Solove: "I've Got Nothing to Hide", San Diego Law Review, Vol. 44: 745, 2007

"I have nothing to hide": Example

In 2002 media revealed that DoD was

- constructing a data mining project
  - "Total Information Awareness" (TIA)

Vision: gather a

- variety of information about people
  - financial, educational, health, and other data
- which would be analyzed for suspicious behavior patterns
- "...to look for patterns of activity
  - e.g. based on observations from past terrorist attacks

Privacy vs Confidentiality

Personal Data

The individual to which the data refers

- is the Data Subject

Privacy law describes (in essence)

- rights of data subjects and
- obligations of organizations
  - that collect, use, store, process or disclose personal data

Those obligations (more later)

- include: purpose, minimization
  - of personal information

What is personal information? Example

- In the US, when the government gathers data from third parties
  - there is no Fourth Amendment protection because
    - people lack a "reasonable expectation of privacy" in information exposed to others
- In US v. Miller, Supreme Court concluded that there is
  - "no reasonable expectation of privacy in bank records"
  - "because the documents obtained, including financial statements and deposit slips"
    - "contain only information voluntarily conveyed to the banks and exposed to their employees in the ordinary course of business"

What is personal information? Example

In Smith v. Maryland, the Supreme Court held that

- people lack a reasonable expectation of privacy
  - in the phone numbers they dial because they
  - "know that they must convey numerical information to the phone company,"
- and therefore they cannot
  - "harbor any general expectation that the numbers they dial will remain secret"

Privacy is a type of Confidentiality?

1. Privacy is often confused with confidentiality

Some people use the term "privacy" to denote the

- required protection for
- any type of sensitive data
  - for instance information about a company that should not be seen by the competitors

This is simply confidentiality, not privacy, regulated by internal procedures of the company (security policies) but not by goverment regulation

- Privacy refers only to personal data

2. Privacy-as-confidentiality

Guaranteeing the confidentiality of personal information

- is one way of preserving or enhancing privacy

(See next slides)

Privacy-as-Confidentiality

Refers to an approach to IT-Privacy

and to the motto of a set of PETs

- in contraposition to
  - Privacy-as-Control or
  - Privacy-as-Practice

Methods of this approach

- Anonymity
- Cryptographic protocols
- Obfuscation through
  - introduction of "noise"
  - generation of fake data

Privacy-as-Confidentiality

Main Principles of the approach

1. Disclosure of information is by default prevented
   - or information is minimally disclosed
     - in a way that cannot be linked back to the individual
   - Individuals may still disclose information voluntarily to others
2. Systems should avoid a trust model
   - which forces users to rely on a single entity to protect their privacy
3. Code and development of privacy tools are open to public scrutiny

Privacy-as-Confidentiality

Assumes that information sharing by that person puts only that specific individual at risk

- But some personal information refers to a group of people not a single one
  - genetic information, group photographs, etc

Privacy-as-Control

Assumption: information will be anyway disclosed

- via social media, IoT, web services, etc

Privacy is protected by organizations' procedures and mechanisms

- making data collection and processing
  - transparent to data subjects and the general public
- giving data subjects control about
  - when, how, which information about them is communicated to other
    - and in which granularity

Privacy-as-Control

These mechanisms allow users to

- make informed decisions
- have greater control over the collection and flows of their personal information
- detect or mitigate abuse

Privacy as Practice in the Society

Privacy is seen as the (sometime implicit) negotiation of

- social boundaries
  - what is allowed, what not, what is expected, etc
- through activities of users (collectively or individually)
- Privacy is negotiated through collective dynamics

Privacy as Practice in the Society

Engineers, lawers try to design

- principles, mechanisms
  - that help enforce users' privacy expectations
- Transparency and feedback mechanisms for raising awareness, e.g:
  - possible consequences of a user's actions can be made understandable to her
  - information about how many friends have visited a user's profile
  - information about how other friends manage their privacy settings
  - details about a recommender system
  - users are encouraged to review regularly their privacy settings

The reality: difficult to remain anonymous in the Internet

De-anonymization cases

- AOL Search, Netflix, Massachusetts GIC medical DB

Birth date, postcode, gender

- are enough to identify uniquely 87% of U_S population

Web browser information

- is a fingerprint for 94% of 500K users

Quiz: Assume a company maintains your data confidential

Say: securely encrypted

- and ensuring it is not disclosed or leaked
- will that be enough for being privacy-compliant?
- could that still be a privacy breach against you?

The problem of divulging publicly known information

In the US, the Supreme Court has held that

- bank records (say, financial statements, deposit slips), tel numbers dialed
  - are not protected by privacy rules

Addresses, Telephone numbers, number of children, name of spouse, etc

- are "publicly known" (via FB, LinkedIn, home pages, etc)

Exercise:

Using the "taxonomy of Solove" or Torts or "Priv as Confid" slides

- and give examples where divulging publicly know information is clearly a problem

Does privacy regulation work?

Does regulation sanction improper use of personal data?

The General Data Protection Regulation (GDPR) of the EU allows

- to fine companies who do not comply with EU rules with
  - up to 4% of their global annual turnover, or imposing
  - regular periodic DP audits
- those sanctions can be issued
  - without the necessity to demonstrate harm

Does regulations note improper use of personal data?

Difficult to say . . .

Privacy as a "Policy Issue"

Empirical Social Sciences, Political Studies, Sociology

From the perspective of Political and public policy issues

- study the laws, codes, guidelines, conventions, practices
  - that regulate the processing of personal information
  - in a society

Political Debates, Governance, Regulatory Instruments

Globalization and the consequences

- The ability of any country to protect privacy is linked to
  - actions of public and private organizations outside its borders

Machine Readable Languages to express

Privacy Policies

- both from users or organizations

And to negotiate Compromises

Need PETs for "anonymous data"?

Indeed

Recall: Is it not easy to be anonymous in the Internet

De-anonymization, fingerprinting, etc

Privacy is good for business

The argument goes like this:

- Business are better off, if they
  - Safeguard all personal information under their care and control
  - Ensure compliance with privacy legislation
  - Protect its reputation
  - Enhance consumer trust
- This all helps to
  - Improve efficiency
  - Create greater innovation
  - Heighten competitive advantage

"Privacy is good for business" is not true

Example: Data breach of Sony's Playstation Network, 2014

Initially estimated to cost more than $100 million

- Next quarter financial report:
  - Sony statement:costs reduced to $15 million

"Good or bad, don't care; what matters: Publicity!"

Sony explained the attack as an attempt to

- stop the distribution of a controversial movie they
  - produced and released in their entertainment network
    - gaining more subscribers

Privacy is about protecting users from harmful violations

Yes (See "Legal" part)

- But there is much more to privacy than this

Privacy is only about user control

Yes (see ???)

- But there is much more to privacy than this

Privacy is about user control

- Goals of the Approach
- Support User's control over collection and use of personal data
  - Purpose
  - Informed Consent
  - Notice
  - Privacy Settings, Privacy Policies
- Support the Organization to demonstrate data protection compliance
  - prevent / detect the abuse of personal information for unauthorized purposes
  - "accountability"

Legal vs technical Approaches to Privacy

Privacy is a legal term: Privacy is . . .

1. A fundamental right
2. A set of concrete rights
3. A set of Principles (Mostly: business obligations)
4. Protection against certain harmful activities
5. Privacy is the set of obligations organizations have
   - to safeguard all personal information under their care and control
6. Privacy is a Contract btw org and user

Privacy is a Technical term: Information Privacy

- Including a development methodology

1. Privacy is a fundamental right

Human Right

- The protection of [. . .] personal data is a fundamental right
- Citizens have the right to the protection of personal data
  - this is provided by
    - Universal Declaration of Human Rights (Dec 1948), Art 12
    - Charter of Fundamental Rights of the European Union, Art 8
    - Treaty on the Functioning of the European Union (TFEU), Art 16

1. Privacy is a fundamental right

Universal Declaration of Human Rights:

- No one shall be subjected to arbitrary interference
  - with his privacy, family, home or correspondence
  - nor to attacks upon his honour and reputation
- Everyone has the right to the protection of the law
  - against such interference or attacks

International Covenant on Civil and Political Rights (Dec 1966), Art 17:

- No one shall be subjected to arbitrary or unlawful interference
  - with his privacy, family, home or correspondence
  - nor to unlawful attacks on his honour and reputation
- Everyone has the right to the protection of the law against such interference or attacks

2. Privacy is a set of concrete rights

The following list shows how scholars have theorized about privacy:

- 1. The right to be let alone
  - Warren-Brandeis
- 2. Limited access to the self
  - the ability to shield oneself from unwanted access by others
- 3. Control over personal information
  - the ability to exercise control over information about oneself
- 4. Secrecy
  - the concealment of certain matters from others
- 5. Personhood
  - the protection of one's personality, individuality, and dignity
- 6. Intimacy
  - control over, or limited access to, one's intimate relationships or aspects of life
- D.J. Solove: Conceptualizing Privacy, Cal Law Rev, 90(4), 2002

2. Privacy – Concrete Rights: 1. "Right to be let alone"

- This early vague legal concept
  - did not define a broad legal privacy protection notion
- it strengthened the notion of
  - privacy rights for individuals
    - began a legacy of discussion on those rights

This right has been interpreted as:

- The right of a person to choose
  - seclusion from the attention of others
    - if they wish to do so, and
- The right to be immune from scrutiny
  - or being observed in private settings,
    - such as one's own home

2. Privacy – Concrete Rights: 2. Limited access

Limited access

- Refers to a person's ability
  - to participate in society
    - without having other individuals and organizations
    - collect information about them
- Various theorists have imagined privacy as
  - a system for limiting access to one's personal information
- "Nothing is better worthy of legal protection than private life, or, in other words,
  - the right of every man to keep his affairs to himself,
    - and to decide for himself
    - to what extent they shall be the
    - subject of public observation and discussion"
- "the condition of being protected from unwanted access by others – either physical access, personal information, or attention"

2. Privacy – Concrete Rights: 3. Control over information

Control over one's personal information;

- "Privacy is the claim of individuals, groups, or institutions to
  - determine for themselves
    - when, how, and to what extent
    - information about them is communicated to others"
- "Privacy is
  - not simply an absence of information
    - about us in the minds of others
- rather it is the control we have over information
  - about ourselves"
- Control over personal information
  - is one of the more popular theories of the meaning of privacy
    - In most modern views of IT-privacy
    - the concept of control is in the core

2. Privacy – Concrete Rights: 4. Secrecy

- Privacy is sometimes defined as an
  - option to have secrecy

The right to "conceal information about themselves that others might use to their disadvantage"

- In some cases it implies the right to lie
- If privacy is secrecy then
  - rights to privacy do not apply for any information which is already publicly disclosed

2. Privacy – Concrete Rights: 4. Secrecy

- Be careful:
  - Divulging some information
    - which is known in some context or environment
    - to other contexts where it is not known
  - is problematic
- When privacy-as-secrecy is discussed,
- it is usually imagined to be a
  - selective kind of secrecy in which individuals
    - keep some information secret and private, while they choose to make
    - other information public and not private

2. Privacy – Concrete Rights: 5. Personhood

- The theory of privacy as personhood
  - is constructed around a normative end of privacy
    - namely protection of the integrity of the personality
- Privacy is sometimes imagined to be a
  - fundamental aspect of personhood
    - which is to say that there is something natural in being a human
    - which requires humans to conceal some information

2. Privacy – Concrete Rights: 5. Personhood

Privacy protects individuals from threats which are

- "Demeaning to individuality"
- "An affront to personal dignity"
- "An assault on human personality"
- Privacy seeks promoting personhood
  - Rather than making general rules for
    - regulating information without that particular outcome

2. Privacy – Concrete Rights: 6. Intimacy

- Analogous to how personhood:
- The intimacy theory imagines privacy to be
  - an essential part of the way that
    - humans have strengthened or intimate relationships with other humans
- Privacy matters because
  - "there is a close connection between our ability to control who has access to us and to information about us
    - and our ability to create and maintain different sorts of social relationships with different people"
- Because part of human relationships includes
  - individuals volunteering to
    - self-disclose some information
  - but withholding other information
  - there is a concept of
  - privacy as a part of the process by means of which humans establish relationships with each other

3. Basic principles of the Fair Information Practices

1. Existence of personal data collections should be public knowledge
   - Transparency
2. Individuals have a right to review and correct their information
   - Control
3. The minimum information necessary should be collected, and where appropriate, consent of the included individuals should be obtained
   - Data Minimization and Consent
4. Personal data should be accurate and complete and retained only for a given time period
   - Accuracy
5. Data should only be used for the purpose originally intended
   - Purpose
6. Data should be protected by security safeguards against unauthorized access, modification or use
   - Security

3. Privacy as a set of Principles

- Personal data shall be
  - processed fairly and lawfully
  - obtained only for a specified and lawful purpose
    - not be further processed in any manner incompatible with that purpose
  - not be kept longer than is necessary for that purpose (or those purposes)
  - accurate and, where necessary, kept up to date

3. Privacy as a set of Principles

- GDPR gives rights to data subjects
  - access
  - rectification
  - erasure
  - notification
  - data portability
  - etc
- Personal data shall be processed in accordance with those rights
  - Appropriate security and organizational measures shall be taken
    - against unauthorized or unlawful processing of personal data and
    - against accidental loss or destruction of, or damage to, personal data
- Personal data shall not be transferred to a third country
  - unless that country ensures an adequate level of protection
    - of the rights of data subjects regarding processing of personal data

3. Privacy Principles

- Personal data
  - should not be processed at all,
    - except when certain conditions are met
- These conditions fall into three categories:
  - legitimate purpose
  - proportionality
  - transparency

3. Privacy Principles: Purpose

Purpose

Personal data shall be

- collected and processed
- only for a legitimate and specific purpose

Legitimate purpose =

- the purpose is to provide a service or, based
  - on a contract with the data subject, or
  - on a functionality that is legitimate (say, by law)

3. Privacy Principles: Proportionality

Personal data may be processed only insofar as it is

- adequate, relevant and not excessive
  - for the service or functionality being provided

Data Minimization

For the given specific purpose,

- personal data collection and processing must be minimal

3. Privacy Principles: Transparency

Consent

Data subject and the processing party

- must agree how data is processed

Data Subject Control

The data subject should have the right to

- access all data processed about him, and
- demand rectification, deletion or blocking of data that is
  - incomplete
  - inaccurate
  - isn't being processed in compliance with the DP rules
  - not required anymore for the provision of the service

3. Privacy Principles: Transparency

- Transparency provides to users an understanding of the system
  - Enabling them to intervene and assert control
- In general the data subject has no means to know:
  - which data are collected
  - for which purposes his data are used
  - how it is aggregated into profiles (individual or collective)
  - which decisions are made based on these profiles
- Transparency means that the public understands:
  - how the system works,
  - how it affects users,
  - how they can intervene in the systems,
  - how to develop practices for using the system

3. Privacy Principles: Transparency

- Data collection should be transparent
  - both individual and collective data
    - how, why, which data are collected
    - users are aware of the collection and use of data
- Decision-making is made transparent
  - Clash with IP (e.g, algorithms as trade secrets)
    - Users/communities may discover undesirable practices of organizations

How can be this enforced? What are the problems?

3. Privacy Principles: Transparency, cont

Individual participation, Notice and Access

Data Subject must be aware about

- personal data sources, processing,

Knowledge of the logic involved in automatic processing

Data subjects have the right to review their data and to withdraw it

Accountability

For instance: The controller must

- notify the supervisory authority
  - before he starts to process data

4. Protection against certain harmful activities

Notice the different Nature of the Examples:

- A newspaper reports the name of a rape victim
- Reporters deceitfully gain entry to a person's home and
  - secretly photograph and record the person
- New X-ray devices can see through people's clothing,
  - amounting to what some call a virtual strip-search
- Despite promising not to sell its members' personal information to others,
  - a company does so anyway
- A company markets a list of five million elderly incontinent women
- Commercial harm to individuals because of incorrect data
  - insurance coverage refusal, loss of credit worthiness, and denial of employment

4. Protection against certain harmful activities

Notice the different Nature of the Examples:

- Dept of Interior used a thermal imaging device outside a home
  - The device could not "penetrate walls or windows to reveal conversations or human activities
    - The device recorded only heat being emitted from the home
- The device showed an unusual amount of heat radiating from the garage
- . . . and to grow marijuana indoors, one needs to provide a large amount of light in order for the plants to photosynthesize

4. Privacy is Protection against harmful activities 1. Tort

1. Intrusion Upon Seclusion

Unreasonable and offensive intrusions into the solitude of another

- not necessarily physical intrusions
- Important is not the type of information obtained
  - but the offensive manner in which the information is obtained
  - intrusion does not require a publication
    - Such may constitute a distinct tort
    - or enhance the damages suffered

Definition

"One who intentionally intrudes, physically or otherwise, upon the solitude or seclusion of another, or his private affairs or concerns, is subject to liability to the other for invasion of his privacy, if the intrusion would be highly offensive to a reasonable person"

4. Privacy is Protection against harmful activities 2. Tort

2. Publication of private embarrassing facts:

- 1. A publication to the general public
  - or to a large number of persons, not to a few
  - without permission or privilege
  - excludes publicity given to matters already publicized or in the public record
- 2. Private contents
  - as opposed to the public life of the individual
- 3. No legitimate public interest or concern
  - Public has no legitimate concern
- 4. Embarrassing facts
  - The material published must bring
    - humiliation or shame to a person of
    - ordinary sensibilities

4. Privacy is Protection against harmful activities 2. Tort

- The Second Restatement of Torts §652D (1977) has the note:
  - "It has not been established with certainty that liability for publication of private embarrassing, but truthful facts is consistent with the free-speech and free-press provisions of the First Amendment to the Constitution"

4. Privacy is Protection against harmful activities 3. Tort

3. Appropriation of a person's name or likeness for the defendant's economic benefit

- This branch of invasion of privacy law recognizes an
  - individual's right to privacy from commercial exploitation
- It also recognizes a person's
  - property right to exploit his or her own name or image
    - for his or her own economic benefit
- Therefore, courts and commentators have often
  - described the tort not as an "invasion of privacy"
  - but rather as an interference with the "right of publicity"
    - Related to Intellectual Property

4. Privacy is Protection against harmful activities 4. Tort

4. Publicity Placing One in a False Light

The Restatement (Second) of Torts §652E (1977) provides: . . . placing somebody before the public in a false light, if

- (a) the false light in which the other was placed
  - would be highly offensive to a reasonable person
- (b) the actor had knowledge of the falsity
  - of the publicized matter or
    - acted in reckless disregard of it

4. Protection against certain harmful activities

- Warren and Brandeis spoke of privacy as an
  - incorporeal injury rather than a physical one
- They noted that the law started to recognize
  - nonphysical harms and that "modern enterprise and invention have,
    - through invasions upon privacy,
  - subjected him to mental pain and distress,
    - far greater than could be inflicted by mere bodily injury"
  - Privacy, they say, involves "injury to the feelings"

4. Protection against certain harmful activities

But . . . Privacy violations are more than harm to individuals

- Some activities may cause no direct harm to an individual
  - But may increase the risk (mostly: the probability)
    - that a certain harm could happen
  - For instance, if data is stored unencrypted,
    - the probability that an attacker discloses the data is higher
  - If personnel is not trained, risks are more probable

4. Privacy is Protection against harmful activities: Solove

Privacy is a set of "family resemblances"

- Solove: . . .
  - "not a single concept but a set of 'family resemblances'"
- "I suggest abandoning the traditional way of
  - conceptualizing privacy and instead
  - understanding it with Ludwig Wittgenstein's notion of 'family resemblances'"
- "Certain concepts might not have a single common characteristic
  - rather, they draw from a common pool of similar elements"
- "Privacy consists of many different yet related things"
- "In terms of generality,
  - I argue that privacy should be
    - conceptualized from the bottom up rather than the top down,
    - from particular contexts rather than in the abstract"

4. Protection against certain harmful activities

Taxonomy of legal Dimensions of Privacy

D.J. Solove: Specific activities that pose privacy problems

- A. Information Collection
  - Surveillance
  - Interrogation
- B. Information Processing
  - Aggregation
  - Identification
  - Insecurity
  - Secondary Use
  - Exclusion

4. Protection against certain harmful activities

Taxonomy of legal Dimensions of Privacy

D.J. Solove: Specific activities that pose privacy problems

- C. Information Dissemination
  - Breach of Confidentiality
  - Disclosure
  - Exposure
  - Increased Accessibility
  - Blackmail
  - Appropriation
  - Distortion
- D. Invasion
  - Intrusion
  - Decisional Interference

6. Privacy as a Contract btw org and user

Negotiated Privacy Techniques

Including commercial "infomediaries"

- Entities that could store consumers' data
  - help facilitate the drafting of contracts
  - set the terms of the exchange and use of data
- Around yr 2000 there were great expectations that privacy problems could be solved through a mix of decentralized storage and private contracts
  - These hopes have not turned to be true

6. Privacy is: A Contract btw org and user

A privacy policy (in the US) is a document

- disclosing what information a business gathers
  - and what they do with it
    - not a guarantee that your data will be treated
    - with any particular sensitivity or regard
- Most privacy policies tell three things:
  - What personal data is this business collecting?
  - How and why (for what uses) does the company collect that data?
  - With whom, and under what circumstances, does the company share that data?
    - e.g with third-party data brokers and marketers for advertisements

6. Privacy is: A Contract btw org and user

- Example: Bank of America

	Do we share?	Can you limit the sharing?
For everyday business purposes, respond to court and legal investigations, report to credit bureaus	Yes	No
For our marketing purposes	Yes	No
For joint marketing with financial companies	Yes	No
For our affiliates' everyday business purposes - Info on your transactions and experiences	Yes	No
For our affiliates' everyday business purposes - Info on your credit worthiness	Yes	Yes

6. Privacy is: A Contract btw org and user

- Example: Bank of America

	Do we share?	Can you limit the sharing?
For non-affiliates to market to you and services endorsed by another organization - for all credit card and Sponsored Accounts	Yes	Yes
For non-affiliates to market to you - for accounts other than above	No	N/A

- See `http://www.consumerreports.org/privacy/3-common-misconceptions-and-1-important-truth-about-pr`

6. Privacy is: A Contract btw org and user

Most policies strive to be accurate but vague, covering all possible cases

- Example, a company with a sizable online presence
  - Under "types of information we collect," they say:
    - (Next slide)

6. Privacy is: A Contract btw org and user

We collect your personal information

including name, address, phone number, email address, location information, and more, from:

- Information collected when you interact with us
  - transactions, completion of forms, registration or surveys, your participation in our marketing programs
- Information automatically collected when you
  - visit or use our Site or view our online ads, such as
    - via cookies and device information, and in Stores, such as through your use of our Wi-Fi Services
- Information from other sources,
  - such as companies that help us to update our records
- See `http://www.consumerreports.org/privacy/3-common-misconceptions-and-1-important-truth-about-pr`

6. Privacy is: A Contract btw org and user

We will collect

- your data <this and this exactly>
- For <this and this purpose> only

We will use your data

- In <this and this way>
- And associate / link it with <this and this information>

We will store your data

- In <this and this way>
- And will be deleted <when>

6. Privacy is: A Contract btw org and user

We will disclose your data

- In <this and this form> form
- With <these anonymity guarantees>

You will have

- the following ownership rights to your data (and related data), and/or the right to view, verify, and challenge that information
- the following privacy guarantees

6. Privacy is: A Contract btw org and user

- See slides on "Differential Privacy"

Is there a "Privacy Development Life Cycle" ?

- The European Commission and
  - the Federal Trade Commission (FTC)
    - have recently embraced privacy by design (PbD)
- However, what PbD means in practice is still unclear
  - Indeed, despite strong expressions of support for PbD, its meaning remains elusive
- The regulatory faith in PbD reflects a belief that
  - Privacy improves if firms considered privacy
    - at the beginning of any development process
    - rather than "adding it on" at the end
  - Probably true, but there is
    - insufficient relevant data to support of this view
    - No studies to determine if "PbD" achieves better privacy

"Privacy by Design" or "Privacy Engineering": Methodology

- Adopted as resolution by the 32. Annual International
- Conference of Data Protection and Privacy Commissioners meeting
- Seven "foundational" principles:
  - 1. Proactive not reactive, Preventative, not Remedial
  - 2. Privacy as the default
  - 3. Privacy Embedded into Design
  - 4. Full functionality - Positive Sum not Zero Sum
  - 5. End-to-end security - Lifecycle Protection
  - 6. Visibility and Transparency
  - 7. Respect for User Privacy

PbD approach

- Cavoukian's seven principles are more aspirational than practical or operational
  - Moreover, Cavoukian associates PbD with many other concepts including accountability, risk management, FIPs,and privacy impact assessments ("PIAs")
  - This breadth tends to dilute, rather than clarify, Cavoukian's definition of PbD

PbD approach

- Principles 1-3 express the
  - importance of considering privacy issues early in the design process and
  - setting defaults accordingly, but
    - they stop far short of offering concrete design guidance
- Cavoukian offers
  - practical advice in several technology-specific papers
  - fails to systematize or summarize required design principles

PbD approach

- Principle 4 (Full functionality - Positive Sum not Zero Sum) seems unrealistic:
  - personal data is a highly valuable asset
  - privacy controls only try to limit the exploitation of this valuable commodity
- Principle 5 emphasizes lifecycle management
  - a key aspect of privacy engineering
- Principle 6 resembles the familiar transparency principle found in all FIPs
- Principle 7 summarizes the earlier principles

PbD Methodology

- Despite its comprehensiveness, it is not clear from Cavoukian's document,
  - what "PbD" actually is and how it should be translated into the engineering practice
- Most of the principles include the term "PbD" in the explanation of the principle itself
- Example: Principle (3), Privacy Embedded into Design:
  - "Privacy by design is embedded into the design and architecture of IT systems [. . .] It is not bolted as an addon, after the fact
  - The result is that privacy becomes an essential
- component of the core functionality being delivered

Privacy is a Processes

- . . . intended to safeguard all personal information under the care and control of organizations
- . . . based upon openness and accountability

Goals and Methods are:

- limit collection, use and disclosure of personal data
- involve individuals in the data lifecycle
- apply appropriate safeguards in a continuous manner
- Keep all private data confidential
  - Encrypt at rest and in motion
- 4 eyes principle
- Security and audit processes
- Personnel training

Privacy is a Processes

- Lederer et al. (`http://repository.cmu.edu/hcii`) offer design guidelines for
  - improving privacy practices in technical systems
    - giving "five pitfalls" for designers to avoid
- tools should combine
  - feedback (understanding) and control (action) mechanisms to
    - "make their consequences known and do not require great effort to use"
    - resulting in socially meaningful privacy practices

Privacy Processes: 5 Pitfalls

Understanding Pitfalls

- 1. Obscuring potential information flow
  - Designs should not obscure the nature and extent of a system's potential for disclosure
    - Users can make informed use of a system only
    - when they understand the scope of its privacy implications
- 2. Obscuring actual information flow
  - Designs should not conceal the actual disclosure of information through a system
  - Users should understand what information is being disclosed to whom

Privacy Processes: 5 Pitfalls

Action Pitfalls

- 3. Emphasizing configuration over action
  - Designs should not require excessive configuration to manage privacy
  - They should enable users to practice privacy as a
  - natural consequence of their normal engagement with the system
- 4. Lacking coarse-grained control
  - Designs should not forgo an obvious, top-level mechanism for halting and resuming disclosure
- 5. Inhibiting established practice
  - Users should be supported in transferring established social practices to emerging technologies

Personal Data in IT Systems

- Collection and processing of personal information by organizations is useful and necessary
  - Hospitals: health records
  - Businesses: purchase records
  - Tax authorities: tax records
- Organizations have an interest in providing reliable services and protecting user privacy
  - The organization is "trusted" to, e.g, respect the preferences expressed in privacy settings
- Privacy problems arise when personal information is misused

Questions when considering privacy technologies

- What, explicitly, is the set of "privacy threats" that
  - you want to address by technology?
    - surveillance by a government
    - profiling by an organization
    - disclosure to the broader public
  - What aspects are left out of the scope?
- What are the trust assumptions in regard of
  - behaviour of the entities in the system?
  - available technical infrastructures?
    - what level of assurance is provided by them?

Goals of PETs (Examples)

Provide individuals with means to

- 1. Obtain informed consent
- 2. Determine (control) the collection and use of their information
- 3. Articulate those privacy expectations through policies
- 4. Means to review the privacy settings (policies) and change them
- 5. Means to enforce those polices by organizations
- 6. Means to understand and verify the processing of personal data and the applied DP

Goals of PETs (Examples)

Providing organizations with means to

- 1. Define and enforce data security policies
- 2. Prevent / detect the (ab)use of personal information
  - for unauthorized purposes
- 3. "Accountability" Tools
  - in order to prove compliance with Regulation

"Data Protection" (or classic) Approach

- Data subject provides her data
  - to a T3P ("Trusted Third Party")
    - A trusted Data Controller, who is responsible
    - for the privacy protection
- One or several Data Processors
  - operate securely
- Threat model
  - External parties, errors, malicious insider
- Controller/processors: main "users" of these technologies
  - Policies, access control, audits (accountability)

"PET-based" Approach

- In the Classic Approach
  - the user must trust the collector and processors
- The Emerging Situation is to use PETs
  - which greatly reduce the trust that
    - must be given to the T3Ps
- Most of those PETs
  - and most of the ones we will review in this course
    - try simply to limit the amount of data sent to the Data Controller/Processor
    - and in particular, not to disclose identities or to avoid identifiability

Examples

- Instead of giving your full name, etc
  - you only present an "anonymous credential"
    - which is good enough to prove that you are authorized
- Instead of paying with your credit card
  - you use some type of anonymous electronic money or vouchers

Main characteristics of DP technologies

- PETs should provide the ability to enforce
  - acceptable data collection and usage, as articulated
  - policy settings defined by the user
  - policies defined by the organization
  - Limitation: no emphasis on technically minimizing data collection
    - does not preempt the creation of large databases
- The organization is trusted to enforce the policies through access control mechanisms
  - Limitation: no protection guarantees toward organizations that want to violate user privacy by abusing the data that they hold
- Mimicry of off-line bureaucracy in digital systems
  - Mostly ignoring the properties of digital systems

Privacy settings

- Means for the user to specify her privacy preferences
- Many proposed technologies aim at making it easier for users to configure these settings
  - default "suites" of privacy settings
  - privacy wizards that automate the configuration of settings
- Note: the enforcement of the policy (settings) is done by the organization

Purpose-based access control

- Policy that establishes who can access which information
  - specifies which are the allowed uses of collected information
- Policy enforcement
- Ensure that the purpose of a data access is compliant with the policy

Auditing

- Verify that the policies are being respected
- Mechanism:
  - log the data access and processing operations
  - examine those logs to detect policy violations
- Proposals to generate auditing specifications that produce logs that are both minimal and sufficient for the audit
- Note: this may even enable additional privacy violations (that involve exploiting the information in the logs)
- Data protection technologies

Weakness

- Data subject has already lost control of her data
  - In practice, very difficult for data subject to verify how her data is collected and processed: organizations are opaque
  - Need to trust data controllers (honesty, competence) and hope for the best

Problems of data protection technologies

- Data minimization (proportionality) often ignored
- Informed consent?
  - Privacy policies long and difficult to understand
  - Complex systems, difficult to understand potential abuses
  - Lack of choice
- Trust assumptions may not be realistic
  - Incompetence
  - Incentives?
  - Cost of securing the data
  - Purpose (function creep)
  - Malicious insiders

Problems with DP alone

- Lack of transparency and DP enforcement: Once the data is under the control of an organization it is very difficult for individuals to verify how their data is actually being used
- Misalignment of incentives: Organizations that collect and process user data are not necessarily competent and honest, security is expensive
  - Incentives to amass and use personal data for financial gain (without regard for users' privacy concerns
  - Large number of reported privacy breaches (due to a lack of appropriate data security practices)
- Placing high levels of trust in organizations should be avoided whenever possible, as it leaves individuals vulnerable
  - Note: the lack of trust may refer to the difficulty of securing the technical infrastructure rather than generalized paranoia ("everybody is evil")

Problems of data protection technologies

- Technologically enforced?
  - Like security, privacy must be technologically supported
  - Technology must provide assurances where possible
  - Laws are necessary but not sufficient to protect privacy/ security
  - Example: legal interception interface abuse
  - Privacy/security needs cannot just be satisfied with good intentions
- How can you check that your data is not being abused?
- Weak enforcement, low penalties
- No protection for "anonymous" data

Privacy Enhancing Technologies

- System model
  - Subject provides as little data as possible
- Reduce as much as possible the need to "trust" other entities
- Threat model
  - Strategic adversary with certain resources motivated to breach privacy (similar to security systems)
  - Adversarial environment: communication provider, data holder

Goals of PETs

. . . could be:

- Create an individual autonomous sphere
  - free from intrusions from both a state
    - possibly in collusion with corporations
  - and from the pressure of social norms
- Inspired by the definition of privacy as
  - "the right to be let alone" [Warren-Brandeis1890]
- Disclosure of information is by default prevented
  - or information is minimally disclosed in a way that
    - cannot be linked back to the individual
  - Individuals may still disclose information voluntarily to others

Some strategies to protect Privacy

Anonymous Communication

Service provider can observe

- who accesses the service (if required))

Anonymity

Service provider can observe

- that someone accesses the service
- But cannot observe the identity of the user

Obfuscation of Data

"Noise" is added to the data

- Without damaging too much the usability of the data

Main characteristics of PETs

- Privacy notions are predefined as properties that are hard-coded in the technology itself
  - The goal of the technology is then to ensure that these formally defined privacy properties hold
  - Limitation: narrow privacy definitions
  - But the settings can be often/sometimes changed (parameters)
- Focus on preventing data disclosure
  - minimizing data disclosure beyond what seems intuitively possible
  - data cannot be abused for privacy invasive purposes if it has not been made available
  - Limitation
  - disclosure is sometimes desirable and/or necessary

Main characteristics of PETs

- Minimize the need to trust others with appropriately handling identifiable and linkable data
  - while still guaranteeing other security properties such as service integrity
  - Limitation:
  - often in direct conflict with business models
  - based on extracting value from data
  - transfers the trust to the technology

Deployment

- Data protection technologies
  - Compliance is a strong driver
  - Hidden costs of securing large databases
- Privacy Enhancing Technologies
  - Can reconcile aggressive data minimization and service integrity guarantees
  - Active research, lots of proposed solutions
  - Poor deployment
  - not evident what sort of protection is offered by a specific technology

PETs: Anonymization of Credentials

Tools for Anonymity / Pseudonymity / Attribute-based Credentials

. . . at application layer

- should offer protection against linking of activities

Anonymous (say, group) authentication or authorization Identity Management Systems (IDMS) with:

- Credential Management Systems
- Pseudonym Management Systems

PETs: Communication Anonymization

Communications Anonymity and Anonymizers

- Should offer protection against communication censorship and traffic analysis
  - Anonymizers
  - Onion Routing, Tor
  - Censorship resistant communications
  - Privacy-Enhanced Routing in wireless networks

PETs: Data Sharing and Databases

Privacy Preserving Data Sharing, Publishing, and Mining

Should offer protection against de-anonymization algorithms

- Database privacy: k-anonymization, k-anonymity, L-diversity, t-closeness, etc
- Privacy Preserving Data Publishing (PPDP)
- Privacy Preserving Data Mining (PPDM)
- Differential Privacy

PETs: Data Analytics

Privacy-Enhanced Analytics of Big Data

This is not a single technique

- but a coherent / comprehensive approach toward privacy in big data

User awareness and promotion of PETs

Decentralized (instead of centralized) data analytics

- The policy makers need to encourage and promote decentralized privacy preserving analytics models, both at research and at implementation levels

PETs: Data Analytics (cont)

Selectiveness

- securely accessing only the information that is actually needed for a particular analysis
- instead of collecting all possible data to feed the analysis

PETs: Data Analytics (cont)

Automated policy definition and enforcement

in a way that one party cannot refuse to honour the policy of another party in the chain of big data analytics

- Often, certain privacy requirements of one controller are not be respected by another
- privacy preferences of the data subjects may also be neglected or not adequately considered
- Semantics and relevant standards, as well as cryptographically enforced rules

For Big Data, Consent needs

- (a) new models
- (b) standards and regulation
- (c) automated enforcement mechanisms

PETs: Data Analytics (cont)

Consent, Transparency and control

For big data, "traditional" notice, consent, enforcement mechanisms

- privacy icons (next slides)
- sticky policies
  - Machine-enforceable policies that stick to data
    - specify allowed usage and obligations as it moves between servers/services
- personal data stores (PDS), as well as
  - personal data servers,
  - personal data lockers/vaults
  - personal clouds
    - platforms and protocols to support unified repositories of user data
    - that could be managed locally by the user or outsourced

fail to provide proper transparency and control

PETs: Private Data Exchange

Zero-Knowledge proofs (ZKP)

- A proof that only delivers the fact that a claim is true
  - But nothing else

Interactive ZKP

The proof only convinces the person who actively interacts with the prover

- But the proof is not transferable

ZKP of Knowledge

A ZKP that an entity knows a particular secret with a certain property

- But no information about the secret itself

PETs: Private Data Exchange

Oblivious transfer (OT)

A protocol between

- a sender
  - who has a set of items (pieces of information)
- a receiver
  - who queries for one of those items

The sender transfers back the item to the receiver

- but remains oblivious (unaware) as to what piece (if any)
  - has been queried or transferred
- Private information retrieval

PETs: Private Data Exchange

Private Search / Private Information Retrieval (PIR)

Service provider may identify user but

- Cannot observe details of the access to the service
  - Which records were accessed
  - Which search keywords were used
  - Which content was downloaded

Secure multiparty computation

- Parties calculate a function of their inputs
  - But do not gain information about the inputs of the others
- Parties have assurance that the other participants in the protocol are not cheating

PETs: Notice / Transparency: Privacy Icons

The problem

Users need to know how companies intend to use their data

- but privacy policies, terms of service are long-winded, complex documents
  - that encapsulate a lot of situation-specific detail

Privacy Icons may be seen as a very simple

- "Privacy Specification Language"

A simple solution: Privacy Icons

- Adding a Privacy Icon to a privacy policy
  - means "the icons meaning is true and preempts anything else in the privacy policy document."
- People can easily understand how their personal data will be transacted
- A company still has the flexibility of using the icons and creating comprehensive, detailed, and meaningful policies
- Privacy Icons are legal declarations
- They will be machine readable
  - enabling users to communicate their preferences through e.g his web browser

Proposed Icons

- Result of working groups convened by Aza Raskin
  - Seeking to develop boilerplate legal text to back up each of the icons
- Initial designs provided here
  - by Michael Nieling & Ocupop
  - designers of the official HTML5 logos
- `https://wiki.mozilla.org/Privacy_Icons`
  - Not yet final !

Retention period

Figure: Retention: (a) 3 months vs (b) indefinitely

- (a) Your data is deleted before 1, 3, 6, or 18 months
- (b) Your data is stored indefinitely unless you opt-out

Third-party use

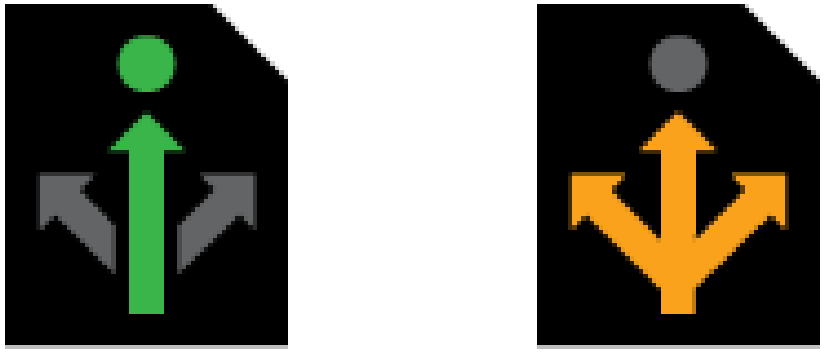

Figure: (a) Intended Use Only vs (b) Limited re-use

- (a) The site is not trading or selling your data
  - It will only share your data with other organizations
    - in order to carry out the intended transaction
- (b) The website is selling or trading it with
  - another organization, government, or person
    - Example: a shopping website collects your
    - shopping preferences, frugality, and IP address and
    - sells this info to data aggregators or e-commerce sites directly

Ad networks

Figure: (a) No ad share vs (b) Ad share with opt-out

- (a) Besides the information exposed via on-page advertisement, the site does not share the data it collects about you with advertisers
- (b) The site either shares the data it has about you with marketing or advertising companies or allows those companies to collect info about you while on its site

Law enforcement

Figure: (a) Statutory process vs (b) Transparent process

- (a) When the organization gets a request for your data, but in a legally insufficient form, like a phone call or letter
  - they don't comply because they require the government to comply
    - with the legal process provided by the law before getting users' data
- (b) The organizations might provide your data to a government that asks for it
  - without following the legally required process,
    - but always follows a publicly-documented and consistent process

PETs: Policy Languages

Privacy Policy Languages and Policy Negotiation

- User's Control via Policies
- Privacy Policy Languages and Policy Negotiation
  - P3P
  - EPAL
  - geopriv
- Privacy-preserving data usage control
  - Privacy enhanced "DRM-like privacy"

PETs: Rough Classification

Social Network Data Protection

- protection of shared content
- protection of user interactions
- tools for assistance with privacy relevant decision-making

Location Privacy

- space-temporal cloaking
- mix zones
- privacy-preserving access to location-based services

PETs:

Tools for Data Minimization,

- to avoid sending identifiers or quasi-identifiers
  - Pseudonym Management
  - Attribute-based authorization
- to avoid linking and tracking
  - Location Mixing
  - Obfuscation (Noise)

Tools for Policies

- Specification of Purpose
- Specification of Deletion, etc

PETs:

Tools for Privacy-enhanced access by Data Subject

- Consent Management
- "Privacy Dashboard"

Tools for management of Cryptology options/functions

- Security Negotiation
- Use of TPMs

PETs: Main Applications

- Auctions
- Anonymous e-cash
- eVoting
  - see next slide
- Private search (e.g., DuckDuckGo)
- Privacy Preserving Recommender systems

What PETs should provide, Example: Auctions

Open ascending price auction

- People place increasing bids
  - until no one wants to bid more

Sealed first-price auction or blind auction

- Bidders simultaneously submit sealed bids
  - no bidder knows the bid of any other participant

Vickrey auction

- Can be easily automated
- Works like sealed first-price auction but:
  - winning bidder pays second-highest bid
  - rather than his or her own

What PETs should provide, Example: Auctions

Bidders want to pay as less as possible

- the winner would like to pay less than his maximal amount
  - This happens if other bidders stop participating before the bid reaches his maximum

The auction, to be fair,

- must keep the maximum amount you are willing to pay private
- If the auctioneer knows your maximum
  - he can collude with a bidder (or create one) and
  - force the price to be always just below your maximum

What PETs should provide, Example: eCash Requirements

A user can withdraw coins from the bank, and spend each coin

- wherever he likes
  - on whatever he likes
  - without the bank or an external partner being informed
- The payment should be anonymous and unlinkable
  - Although rendering the service may require address, etc
  - If the amount exceeds a certain limit, regulations may require disclosing the identity and transaction details to the government

Double-spending of a coin must be refused

What PETs should provide, Example: eVoting Requirements

- The outcome of the election matches voter intent
  - Not necessarily the exact number of votes intended for each candidate
- Ballot secrecy (coercion protection), Plausible deniability
  - the system should ensure that nobody can figure out how you voted
  - even if you want to or try to reveal to these other parties how you voted
- Voter authentication
  - only authorized voters should be able to vote
  - voters should only be able to vote once, or whatever legally permitted
- Enfranchisement
  - all authorized voters should have the opportunity to vote

PETs: Environments

- Web privacy
  - Understand/control what web sites collect, maintain regarding personal data
- Mobile data privacy
  - Location Privacy
- Social Networks
- Health Care

Social Networks

- People share personal thoughts, political views, and pictures
  - They all can convey substantial information
- Example: Peter was attending an event
  - has met Joan there
  - perhaps someone takes a picture of both
  - Photographs are digital
    - devices taking them are increasingly connected to the Internet
    - photographs frequently are posted on social networks

Social Networks

- It is difficult to control who has access to which content
  - There is a large amount of pictures shared on a social network
  - Facebook receives more than 350 million new photo uploads every day
    - Is this a problem?
- There is already SW to deduce relationships from photos
  - posted on social networks
    - Is Joan in love with Peter?

Big Data

- Data is an asset with an enormous economic value
  - data industry has steadily grown exponentially
  - as well as its impact on many sectors
    - healthcare
    - transportation
    - e-commerce
- Big Data may make a big difference:
  - better, more accurate services

Big Data vs Privacy

- A common (mis-)understanding in Big Data is that
  - data is an asset that
    - belongs to the parties generating/processing the data
  - They are free to decide
    - what they want to do with that asset to achieve their business goals
- This contradicts fundamental privacy principles:
  - it should be the "Data Subject"
    - the person that determines how the data can be used and by whom
    - and not the "data generator/controller"

Big Data vs Privacy

- In Big Data, after the data is collected
  - an analysis may unveil several purposes for which it can be used
- This again contradicts fundamental privacy principles:
  - "Specification of purpose is an essential first step
    - in applying data protection laws . . .
    - and a pre-requisite for applying other data quality requirements"

Big Data: . . . Will this help?

- Is public analysis of private data a meaningful/achievable Goal?
  - That is, to get utility of "statistical analysis"
    - while protecting privacy of every individual participant
- Is this a reasonable hope or is this wishful thinking?
  - "Privacy-Preserving" Sanitization
    - will allow to create reasonably accurate answers
    - to meaningful queries
  - without disclosing private information of individuals?
    - (and what does this exactly mean?)

Location and Mobility databases

Small amounts of data can provide lots of information

1930 Edmond Locard showed:

- 12 points are needed to uniquely identify a fingerprint

Unicity on mobility DBs

De Montjoye:

- 4 spatio-temporal points even with low resolution are
  - enough to uniquely identify
    - 95% of the people

Anonymized, coarse or blurred mobility datasets

- provide little anonymity

Questions?