Informational privacy

The digital person: technology and privacy in the information age by Daniel J Solove

Subhashis Banerjee, Subodh Sharma

Digital dossiers

Collection of personal information

- By both private and public sector
    - Companies build customers’ behavioural and financial profile for marketing
    - Governments (statedly) for efficiency, better regulation, tax compliance, crime control…
- Three types of information flows:
    - private DBs —> private DBs
    - public DBs —> private DBs
    - private DBs —> public DBs
- Miscellaneous legal and illegal ways to do this: mandatory; explicit solicitation, tracking cookies, spyware, DRM, etc.
- Also aggregation in both private and public DBs

Privacy threat: Orwell’s big brother metaphor

- ``Big Brother envisions a centralized authoritarian power that aims for absolute control”
- ``Totalitarian government. Rewrites the history, purges its critics, indoctrinates the population, burns books, and obliterates all disagreeable relics from the past”
- ``Big Brother’s goal is uniformity and complete discipline”
- ``Big Brother views solitude as dangerous. Employs various techniques of power to eliminate any sense of privacy. Recognises privacy as an essential dimension of the political structure of society and seeks to destroy it”
- ``In *1984,* citizens have no way of discovering if and when they are being watched. The *Panopticon* changes their behaviour. Threatens civil liberty and democracy”

Limitations of the Big Brother metaphor

- The metaphor is ubiquitous in privacy discourse
- Goals of data collection have often been rather benign—or at least far less malignant than the aims of Big Brother
- Rather, it is a story about a group of different actors with different purposes attempting to thrive in an increasingly information-based society
- Focuses only on a particular technique of power—surveillance
- However, marketers wish to observe behaviour so they can tailor goods and advertisements to individual differences. Control only to be able to sell
- Surveillance leads to conformity, inhibition, and self-censorship in situations where it is likely to involve human judgment
- Being observed by an insect on the wall is not invasive of privacy in the Big Brother sense. Computers are like insects. However, some future outcome may violate privacy
- Only some are subjected to Big Brother Surveillance by the state

The secrecy paradigm

- Privacy is invaded by uncovering one’s hidden world, by surveillance, and by the disclosure of concealed information.
- Harm occurs when one's hidden world is uncovered to the public
- The harm such invasions cause consists of inhibition, self-censorship, embarrassment, and damage to one’s reputation
- The law is heavily influenced by this paradigm. As a result, if the information isn’t secret, then courts often conclude that the information can’t be private
- A narrow view which holds that individuals should not expect privacy out of anything that is wilfully made available to third parties

The invasion conception

- The “invasion conception,” understands privacy to be a kind of invasion, in which somebody invades and somebody is invaded
- However, digital dossiers often do not result in any overt invasion
- People frequently don’t experience any direct injury when data about them is aggregated or transferred from one company to another
- Many of the problems of digital dossiers emerge from the collaboration of a multitude of different actors with different purposes
- Each step along the way is relatively small and innocuous, failing to cause harm that the invasion conception would recognize as substantial

Kafka’s *The Trial*

- Important decisions made about individuals by an opaque, insensitive, careless bureaucracy without involving the individuals.

- Arrested by who? For what? Who’s accusing? What authority is investigating?

- ``The power employed in *The Trial* has no apparent goal; any purpose remains shrouded in mystery. Nor is the power as direct and manipulative in design as that depicted by Orwell and Huxley. The Court system barely even cares about Joseph K. *The Trial* depicts a world that differs significantly from our traditional notions of a totalitarian state. Joseph K. was not arrested for his political views; nor did the Court manifest any plan to control people. Indeed, Joseph K. was searching for some reason why he was arrested, a reason that he never discovered. One frightening implication is that there was no reason, or if there were, it was absurd or arbitrary.”

The Kafkaesque metaphor is more appropriate

- Bureaucracy "dehumanises" people and turns them to a score, an algorithm, etc. Our digital biographies are often an inaccurate representation of us, not just because they are too "simplistic" but also because they are sometimes "factually incorrect" because of the carelessness of the bureaucrats.

- Individuals have no control: data is warehoused for undefined future use; it is used by unknown third parties for unknown purposes; power dynamics driven by huge information asymmetry making consent a "meaningless choice”

- Even though goals of the governments or companies are benign, and knowledge of individual data points is harmless, individuals remain vulnerable because linking (or aggregation) could lead to exponential loss of privacy

The Kafkaesque metaphor

- ``What *The Trial* illustrates is that power is not merely exercised in totalitarian forms, and that relationships to bureaucracies which are unbalanced in power can have debilitating effects upon individuals—regardless of the bureaucracies' purposes"

- ``Understanding the database privacy problem in terms of the Kafka metaphor illustrates that the problem with databases concerns the use of information, not merely keeping it secret"

- ``Privacy involves the ability to avoid the powerlessness of having others control information that can affect whether an individual gets a job, becomes licensed to practice in a profession, or obtains a critical loan" (Brazil!)

- ``Privacy involves the power to refuse to be treated with bureaucratic indifference when one complains about errors or when one wants certain data expunged"

Limitations of Information Privacy Laws

Follow Warren and Brandeis, *The Right to Privacy,*1890

Mainly concerned with *right to be left alone*

- Invasion of seclusion
- Public disclosure of private facts
- Projection in false light
- Appropriation

Limitations of Information Privacy Laws

Follow Warren and Brandeis, *The Right to Privacy,*1890

None are sufficient because

- Tort laws look at remedies for isolated actions which individually may be innocuous but collectively may be harmful
- Often these databases collect only public information
- Often disclosure isn't even public, or does not harm reputation
- It's hard to always assign commercial value to individuals' identity

Harm based discourse inadequate. US Constitutional laws provide some protection; also Puttaswamy I

Limitations of privacy self-management

- Consent is broken, as evidenced by the customary ``*I Agree*”
- Consent can be overridden
- Unfamiliarity with legal rights, technology
- Inability to envisage or judge potential harms of digitisation use cases, both to self and society
- Unfamiliarity with privacy management tools

Need an **accountability based framework***; it must be obligatory on the data controller to protect citizens’ rights*

Limitations of Market-based solutions

- Privacy as contract
    - personal information as property
    - limitations of consent
    - individuals cannot fine-tune
- Privacy as property
    - difficulty with assigning value
- Market self-regulation
    - difference in bargaining power
    - individuals need coordination to organise

Failure of privacy self-management

Asking for “consent” for data-sharing is often a meaningless or a false choice.

- Many **cognitive biases** operate on users making decisions about sharing their personal information (Solove, 2013; Acquisti & Grossklags, 2006).
- High degree of **information asymmetry** about how providers will use and share personal data.
- The **threat of denial of service** makes “taking consent” a false choice (Acquisti, 2004).

Architecture

- Privacy is not mere secrecy
- Privacy is not mere public disclosure
- Privacy is not mere identifiable invasion
- Privacy violation does not only happen when there is clearly identifiable harm with definite causal attribution

=> privacy protection must be **ex-ante rather than ex-post**

Architecture examples

- Panopticon is an *architecture for surveillance or control*
- Kafkaesque are *architectures of vulnerability*
- Some impacts of *architectures of vulnerability*
    - *identity loss*
    - *identity theft*
    - *new attack surfaces based on use cases*
    - *compelled speech*

Architectural solutions

- ``For problems that are architectural, the solutions should also be architectural”

- ``Privacy must be protected by reforming the architecture, which involves restructuring our relationships with businesses and the governments. Thus far the law does not do enough to redefine the underlying relationships that cause these symptoms.”

- ``**Regulate the relationships”**

- **``**Create structures to prevent harms from arising rather than merely providing remedies when harms occur”. ``Proactive rather than reactive”

- ``The protection of privacy does not mean an all-or-nothing tradeoff between the total restriction of information gathering versus the complete absence of regulation. Many privacy problems can be ameliorated if information uses are carefully and thoughtfully controlled”

.

Architecture for privacy

Private sector

- **Fiduciary relationship. Regulatory oversight**
- There must be no personal-data record-keeping systems whose very existence is secret
- There must be a way for an individual to find out what information about him is in a record and how it is used
- Data minimisation
- There must be a way for an individual to prevent information about him obtained for one purpose from being used or made available for other purposes without consent (**purpose limitation**)
- There must be a way for an individual to correct or amend a record of identifiable information about self
- Any organization creating, maintaining, using, or disseminating records of identifiable personal data must assure the reliability of the data for their intended use and must take reasonable precautions to prevent misuse of the data (**access control**)

Public records

- Records from birth to marriage to death: date of birth, place of birth, email address, home address, telephone number, SSN, Aadhaar, PAN, DL…
- Many professions require license from authorities: doctors, lawyers, engineers, insurance agents, nurses, police, accountants, and teachers
- Insurance records, medical records, …
- For public employees: many personal details are released to the public by way of personnel records, including home address, phone number, SSN, salary, sick leave, and sometimes even email messages; property details, medical reimbursements
- Property tax records; police records; court records
- RTI

Public records

Impact of technology

- Traditionally accessible only in very limited sense
- Now easy to search, copy, disseminate, join, mine…
- Beyond greater accessibility, technology may also lead to the retention of greater amounts of personal information in public records
- Government access. Laws.
- May be tremendously useful for efficiency of governance; economic analysis and econometrics; targeting welfare, immunisation, primary health care…; epidemiology; disaster management; tax compliance; policing; anti-terrorism…
- But both Orwellian and Kafkaesque risks increase
- Tension between transparency and privacy exacerbated (J Srikrishna committee recommended amendment of Section 8(1) in the RTI Act)

Access and aggregation

Rethinking privacy and transparency

- **Purposes of transparency**: accountability, watchdog for corruption control, determining ownership of land, verify identity, avoid fraud, check background of political candidates…
- **Access: The Public Is Private**:
    - in secrecy paradigm information is seen in this black-and-white manner; either it is wholly private or wholly public
    - invasion/secrecy paradigm for thinking about privacy in public records is insufficient due to the unknown harms that the linking effect might cause.
    - ``*possibility of constructing a sophisticated data center capable of generating a comprehensive womb-to-tomb dossier on every individual and transmitting it to a wide range of data users over a national network*" - Arthur Miller in *Assault on Privacy*, 1971
- **Privacy and freedom of speech**: cannot restrict press from publishing public record information
- Both transparency and privacy can be balanced through **limitations on the access and use of personal information** in public records.

Architecture for privacy

Goals

- **Minimisation**: only minimal information required should be collected and disseminated

- **Particularisation**: information should only be collected and used only for specific pre-approved purposes (**purpose limitation**)

- **Control**: there must be control, or (**regulatory**) oversight, of some body to see that these principles are being followed

- *Extent of control?*

Architecture for privacy

Scope

- The architecture cannot apply to everything
- Solove says the architecture should apply to wherever a "system of records" is being maintained
    - information in the records is more permanent and readily linked, leading to a clear possibility of *purpose violation*
    - individuals have little control over record taking systems as opposed to their neighbours (or even a stalking stranger)
    - the power dynamics with record systems are clearly unfavourable for the individuals

Architecture for privacy

Mechanisms of oversight

- Government should be allowed **access** to one's personal information only if
    - there is a legitimate state interest and a law
    - it can present facts and evidence before a neutral judge and obtain a warrant
- Fourth amendment only talks about restricting access and not about usage.
- Usage restriction and purpose limitation imperative for privacy protection

Architecture for privacy

Question(s) for us

Computer science principles and techniques for privacy architecture?