Rob Kitchen, “Ethical, Political, Legal, Social Concerns,”

The Data Revolution (2014)

As the data revolution unfolds, the debates around these concerns are likely to intensify, especially as attempts are made to produce new legislation to address technological developments which enable new ways of generating, consolidating and analysing data, thus producing new issues and rendering old laws obsolete.

-Rules should be independent of technological developments.

Data Shadows and Dataveillance

Koops (2011), for example, reports that the Dutch Data Protection Authority estimates that the average **Dutch citizen is included in 250–500 databases**, with more socially **active people included in up to 1,000 databases**. These databases not only include individuals' digital footprints (data they themselves leave behind) but also individuals' data shadows (information about them generated by others), and increasingly provide data trails of location and interactions and transactions across space and time.

Ephermal footprints and shadows→ Persistant digital data shadows

- What are the methods that people can avoid their datashadows?

At best, they [governments] constitute oligopticons – limited views from partial vantage points from fixed positions with defined view sheds (Amin and Thrift 2002) – rather than slotting together to create a panopticon; an all-seeing, god's-eye view.

Privacy

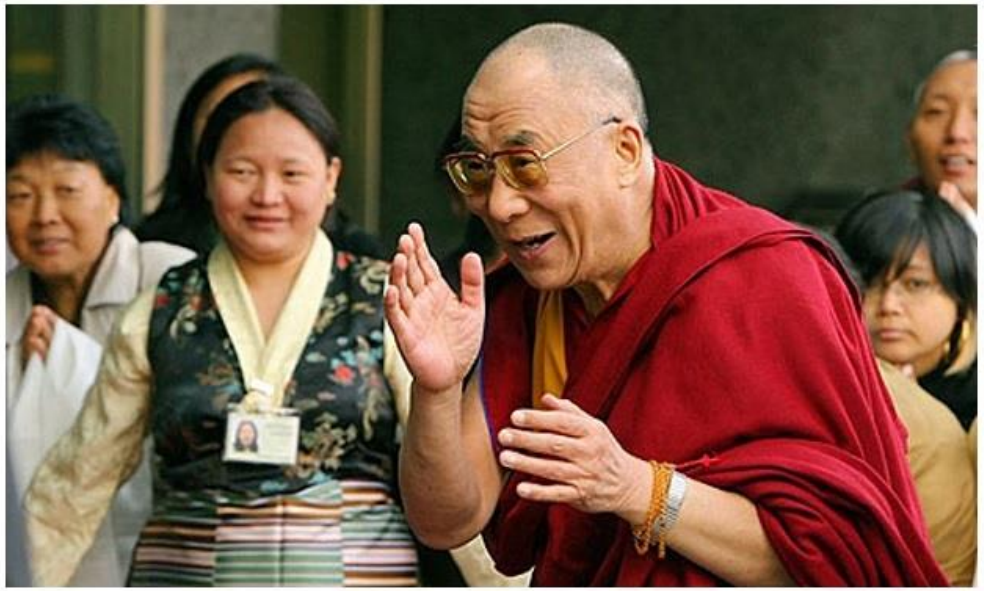
In the **United States** it is mostly covered under the rubric of **privacy laws**, whereas in the **European Union** it falls within the realm of **data protection** (Minelli et al. 2013). It is a term that is multidimensional in its meaning and is often used in **context-dependent** ways, but generally refers to acceptable practices with regards to accessing and disclosing personal and sensitive information (Elwood and Leszczynski 2011).

How privacy laws in the Unitet States are different from data protection in Europe?

There is little doubt that the concept of privacy is changing. As noted above, people are subject to much greater levels of scrutiny and modes of surveillance than ever before.

Proof! Just six degrees of separation between us

After checking 30 billion electronic messages, Microsoft researchers say the theory stands up



Just six degrees of separation or fewer between the Dalai Lama and everyone else. Photograph: Carl de Souza/AFP/Getty Images Carl de Souza/AFP

- Other than the direct surveillance the notion of privacy itself is changing. As an example photo sharing culture has totally changed.
- A 'degree of separation' is a measure of social distance between people. You are one degree away from everyone you know, two degrees away from everyone they know, and so on.
- Being connected = Being known = Less Privacy

Table 10.1 A taxonomy of privacy

Domain	Privacy breach	Description
Information collection	*Surveillance*	Watching, listening to, or recording of an individual's activities
	Interrogation	Various forms of questioning or probing for information
Information-processing	*Aggregation*	The combination of various pieces of data about a person
	Identification	Linking information to particular individuals
	Insecurity	Carelessness in protecting stored information from leaks and improper access
	Secondary use	Information collected for one purpose used for a different purpose without the data subject's consent
	Exclusion	Failure to allow the data subject to know about the data that others have about them and participate in its handling and use, including being barred from being able to access and correct errors in that data
Information dissemination	*Breach of confidentiality*	Breaking a promise to keep a person's information confidential
	Disclosure	Revelation of information about a person that impacts the way others judge their character
	Exposure	Revealing another's nudity, grief or bodily functions
	Increased accessibility	Amplifying the accessibility of information
	Blackmail	Threat to disclose personal information
	Appropriation	The use of the data subject's identity to serve the aims and interests of another
	Distortion	Dissemination of false or misleading information about individuals
Invasion	*Intrusion*	Invasive acts that disturb one's tranquillity or solitude
	Decisional interference	Incursion into the data subject's decisions regarding their private affairs

Source: Compiled from Solove (2006).

In a test of 101 smart phone apps, the *Wall Street Journal* found that 56 transmitted the phone's unique device identifier to other companies without users' awareness or consent, 47 sent the phone's location, 5 sent the users personal details, and 45 did not have any associated privacy policies that users could view (Efrati et al. 2011). TRUSTe found that only 19 per cent of 340 top apps link to a privacy policy and neither Apple nor Google stores require that apps have such policies (Coterill 2011).

- Applications' privacy policies are not accessible.

For some the notion of privacy is largely dead.

Table 10.2 Fair information practice principles

Principle	Description
Notice	Individuals are informed that data are being generated and the purpose to which the data will be put
Choice	Individuals have the choice to opt in or opt out as to whether and how their data will be used or disclosed
Consent	Data are only generated and disclosed with the consent of individuals
Security	Data are protected from loss, misuse, unauthorised access, disclosure, alteration and destruction
Integrity	Data are reliable, accurate, complete and current
Access	Individuals can access, check and verify data about themselves
Accountability	The data holder is accountable for ensuring the above principles and has mechanisms in place to assure compliance

Source: Minelli et al. (2013: 156).

Table 10.3 Types of protected information

Personally identifiable information (PII): any information that directly or indirectly identifies a person	Sensitive information: any information whose unauthorised disclosure could be embarrassing or detrimental to the individual	Other information that can be used to infer the identity of a person
Name	Race/ethnicity	Preferences
Postal address/zip code	Political opinions	Cookie ID
Email address	Religious/philosophical beliefs	Static IP address
Telephone/cell number	Trade union membership	
Social security number	Health/medical information	
Driver's licence number	Marital status/sexual life	
Financial account number	Age	
	Gender	
Credit/debit card number	Criminal record	

Source: Adapted from Minelli et al. (2013: 159).

Many companies do not feel compelled to present individuals with a privacy policy because they claim to generate anonymised data, thus falling outside of fair information practices. Or they present a policy at initial inception of engagement (e.g., when installing software or signing up for a service) **that is full of complex and ambiguous language and which often reserves the right to be modified at a later date without further consultation** (Rubinstein 2013).

(1) people do not read privacy policies;

(2) if people read them, they do not understand them;

(3) if people read and understand them, they often lack enough background knowledge to make an informed choice; and

(4) if people read them, understand them, and can make an informed choice, their choice might be skewed by various decision-making difficulties. (Solove 2013: 1888)

The EU:

- The right to data portability (i.e., the ability to transfer personal data from one service provider to another)
- a right to be forgotten, wherein individuals can seek to have their data deleted if there are no legitimate grounds for retaining them
- and that these rules apply to companies outside of the EU if they are active in the EU market

The US: Privacy by design wherein privacy is inherently built into every stage of product development

- Simplified choice for businesses and consumers that gives them the ability to make decisions about their data, including implementing a do not track
- mechanism and obtaining express consent for sensitive data, or before using data in a materially different manner than the purpose for which it was generated
- greater transparency about data, collection and usage, including reasonable access to data by those the data represent
- The ability to correct or suppress data (Federal Trade Commission 2012).

Table 10.4 The 7 foundational principles of *Privacy by Design*

Principle	Description
Proactive not reactive; preventative not remedial	IT systems should seek to anticipate privacy concerns rather than seeking to resolve privacy infractions once they have incurred
Privacy as the default setting	Privacy is automatically protected and does not require action on behalf of an individual
Privacy embedded into design	Privacy protections are core features of the design and architecture of IT systems and is not a bolt-on feature
Full functionality – positive-sum, not zero-sum	All legitimate interests and objectives are accommodated, rather than there being trade-offs between privacy and other considerations such as security
End-to-end security – full life cycle protection	Privacy is embedded into the system from ingestion to disposal
Visibility and transparency – keep it open	Component parts and operations are visible and transparent to users and providers alike and are subject to independent verification
Respect for user privacy – keep it user-centric	A system should be built around, protect the interests of, and empower individuals

Source: Cavoukian (2009).

- Is there any well designed example?

Data can be easily copied and distributed, digital rights management seeks to limit such practices and also make them easier to track. And yet, despite these threats, digital devices, services and data, and potential weak points in their configuration, are growing faster than the ability to secure them (Gantz and Reinsel 2011).

- Who should really solve this hardware/software problem?

Profiling, Social Sorting and Redlining

Discriminatory practices can also include dynamic and personalized pricing. It is already common for a supermarket chain to have the same goods differentially priced across stores dependent on the characteristics of the people who shop in them, or the prices of products to vary by volume (e.g., 1 for $1, or 3 for $2), or prices to vary across groups (e.g, students or senior citizens receiving discounts) (Varian 1996). The desire of many companies is the rolling out of such practices on an individualised basis, tailored to personal profiles, so that different people pay varying amounts for the same product (as with airline fares, but based on a personalised model). Prices will also vary dynamically and contextually, depending on circumstance.

- Isn’t that the same traditional trading system?
- What are the different contexts?

Secondary Uses, Control Creep

Control creep is where the data generated for one form of governance is appropriated for another (Innes 2001).

Control creep systematically undermines the rationale for data minimisation and its roll-out poses clear threats to civil liberties, with all citizens – both innocent and guilty - subject to its gaze and disciplinary action.

and Anticipatory Governance

Here, predictive analytics are used to assess likely future behaviours or events and to direct appropriate action.

Some companies such as Hewlett Packard are using predictive analytics to assess who might potentially leave the company and to proactively intervene to minimise employee churn (Siegel 2013). In such cases, a person's data shadow does more than follow them; it precedes them (Stalder 2002).

- What kinds of algorithms are they and how do they change employee's behavior?

Modes of Governance and Technological Lock-Ins

- People willingly confess their data (via social media, by joining loyalty card programmes, etc.)

More technocratic governance in nature:

- Instrumental Rationality
- Solutionism: wherein complex social situations can be disassembled into neatly defined problems that can be solved or optimised through computation. Here, there is a reification of big data; they can provide the answer to all problems (Mattern 2013). By capturing phenomena as realtime data it seemingly becomes possible to model, understand, manage and fix a situation as it unfolds.

Indeed, Mattern (2013) suggests that big data urbanism suffers from ‘datafication, the presumption that all meaningful flows and activity can be sensed and measured’.

Modes of Governance and Technological Lock-Ins

The concern around such a corporatisation of urban governance is three-fold (Kitchin 2014).

- First, that it actively promotes a neoliberal political economy and the marketisation of public services wherein city functions are administered for private profit (Hollands 2008).

- Second, that it creates a technological lock-in that makes cities beholden to particular technological platforms and vendors over a long period of time, creating monopoly positions (Hill 2013). The danger here is the creation of a corporate path dependency that cannot easily be undone or diverted (Bates 2012).

- Third, that it leads to ‘one size fits all smart city in a box’ solutions that take little account of the uniqueness of places, peoples and cultures and straitjackets city administrations into a narrowly visioned technocratic mode of governance (Townsend 2013). Indeed, IBM is now selling a product called ‘IBM Intelligent Operations Center’, which combines a number of the systems that were designed for Rio de Janeiro into a single product that can be applied to any city.