Session 8

- Datafication, dataveillance and privacy
- The General Data Protection Regulative (EU's GDPR)
- Guest presentation:
  - Ghazaal Sheiki on Automated Fact Checking

Datafication, dataveillance and privacy

Dataveillance

- Surveillance “enacted through sorting and sifting datasets in order to identify, monitor, track, regulate, predict, and prescribe”
  - makes sense out of big data by restructuring them around entities such as
    - *people*, *transactions*, animals, other lifeforms, *products*, other *artefacts* and things, *places* and *organisations*
  - was always possible
  - is made easier today by digital technologies
  - the scale is different and
  - the data may never be erased

Data footprints and shadows

- *Data footprints* are data directly left behind
- *Data shadows* are secondary data
  - repurposed, recombined, derived by others
  - central for data-driven governments and other organisations
- From data *oligopticons*:
    - many single perspectives, limited view
  - to data *panopticons*:
    - fewer all-encompassing, flexible views
- *Data brokers* seek to pull as much data together as they can to provide detailed intelligence and data products about people, institutions and places


Presidio Modelo, Isla de la Juventud, Cuba

Profiling

- Profiling in the past:
  - divided consumers into (socio-economical, age, gender) groups
- Modern profiling:
  - down to household and individual level
- Better-tailored offers, products and services for some
  - but excluding unprofitable consumers, areas, groups
  - less favourable, dynamic and individual pricing
- Data become ontologically primary:
  - the data about you constrain your available options
  - what the data say (about you) is more important than what people say
  - the data / their processing – may not be accessible or comprehensible

Secondary data uses

- Control creep is when data collected for one purpose is used for another, e.g.,
  - transport data for national security
  - rental car navigation systems
- Anticipatory governance is using predictive analytics to assess likely future behaviours or events
  - where to prioritise police patrols
  - danger of creating self-reinforcing loops
  - the data shadows precede the people
- Dangerous together with ontological primacy of data:
  - (contributing to) creating outcasts, criminals, terrorists
  - by analytically predicting them to be

Critique of data-driven societies

- Data-driven societies (cities, institutions...) have potential benefits, but...
  - *technocratic viewpoint:*
    - society is a mechanical system that can be optimised
    - no critique/reflection on the social system itself
    - levelling bumps rather than fixing their cause
    - “instrumental rationality”, “solutionism”
  - *commercial lock-in:*
    - dependence on (commercial) technology providers
    - reusable assets: “one-system fits all” solutions
    - hard to control the process, change provider...
    - loss of (sole) data control and ownership

New societal challenges

- A self-reinforcing cycle:
  - *bigger datasets + larger computing clusters + stronger computing skills + mergers/acquisitions* ⇒
  - *new AI / ML capabilities and services* ⇒
  - *stronger market position* ⇒
  - *more data and resources* ⇒ ...and so on
- A world dominated by a few big companies?
  - *Amazon, Apple, Facebook/Meta, Google/Alphabet, Microsoft...*
  - *Alibaba, Baidu, JD.com (JingDong), Tencent…*
- A threat to liberal welfare states?
  - in particular to smaller nations

New security challenges

- Information privacy relies on information security
- Common PCs, laptops, servers, and networks are (becoming (rather)) secure
- Other types of digital devices are not:
  - tablets, smartphones, faxes, printers, scanners, photocopiers
  - all kinds of networked devices (pacemakers!)
- Mirroring trend in safety:
  - aviation, automotive, nuclear, military
  - trend towards internetworked standard software and hardware

Privacy

A cultural shift

- Is privacy dead?
  - “better UX”, “good for business”, “people don't mind”, “nothing to hide”, “blend into the crowd”…
  - young people may care less about privacy
  - social media, smartphones...
  - generation shift or life phase?
- Privacy *by rule/regulation* is challenged, but there is also:
  - privacy *by citizen / user* behaviour
  - privacy *by self-regulation*
  - privacy (and data protection) *by design*
  - privacy *by better law* (and surveillance forms)

Privacy

- UN's Human Rights, article 12:
  "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."
- The European Human Rights Convention (article 8-1)
  "1. Everyone has the right to respect for his private and family life, his home and his correspondence.
  2. There shall be no interference by a public authority with the exercise of this right except ... national security, public safety ... economic well-being ... prevention of disorder or crime ... protection of health or morals, or for the protection of the rights and freedoms of others."
- The Norwegian Constitution ("Grunnlov") §12 (from 2014!)

The privacy concept

- Early concepts:
  - *right to privacy* (US, 1890): to be left alone, in seclusion
    - related to the growth of print media and photographs
  - *personal protection*: upholding of personal integrity ... of peoples' private lives, autonomy and self-expression
  - cultural, Anglo-Saxon
  - related to, but not the same as, *secrecy*

The privacy concept

- Forms of privacy:
  - identity privacy (to protect personal and confidential data)
  - bodily privacy (to protect the integrity of the physical person)
  - territorial privacy (to protect personal space, objects, and property)
  - locational and movement privacy (to protect against the tracking of spatial behaviour)
  - communications privacy (to protect against the surveillance of conversations and correspondence)
  - transactions privacy (to protect against monitoring of queries/searches, purchases, and other exchanges)

Privacy and ICT

- Focus in privacy law has shifted:
  - from personal space, honour and reputation
  - to personal information and ICT
- Honour still plays a role
  - but privacy is more often justified in societal terms
  - privacy gives freedom to participate in society
  - as a voter, citizen, consumer, entrepreneur...
- ICT exacerbates existing privacy concerns
  - processing, storing, sharing, searching
  - undeteced(-able) theft, copying, contamination
  - repurposing, recombination, derivation

Table 13.2 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	Use of information collected for one purpose 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 him/her and participate in its handling and use, including being barred from being able to access and correct errors in those data

Table 13.2 A taxonomy of privacy

Domain	Privacy breach	Description
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 his/her 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 tranquility or solitude
	Decisional interference	Incursion into the data subject's decisions regarding his/her private affairs

Fair Informaiton Privacy Principles (FIPPs, OECD 1980)

- *Notice*: about collection / generation, sharing and purpose
- *Choice*: opting in or out of use and disclosure
- *Consent*: about generation and disclosure
- *Security*: from loss, misuse, unauthorised access, disclosure, alteration, destruction
- *Integrity*: reliability, accuracy, completeness, timeliness
- *Access* ("innsyn"): check and verify data about you
- *Use*: use only for intended purpose
- *Accountability:* responsibilities and mechanisms of data holders (data owners and processors)
  - ...following up is a central challenge, *in Norway: "Datatilsynet"*

Table 13.3 Fair Information Practice Principles

General principle	General description	Original OECD principle and description
Notice	Individuals are informed that data are being generated and the purpose to which the data will be put	*Purpose Specification Principle.* The purposes for which personal data are collected should be specified not later than at the time of data collection and the subsequent use limited to the fulfilment of those purposes or such others as are not incompatible with those purposes and as are specified on each occasion of change of purpose
Choice	Individuals have the choice to opt in or opt out as to whether and how their data will be used or disclosed	*Openness Principle.* There should be a general policy of openness about developments, practices and policies with respect to personal data. Means should be readily available of establishing the existence and nature of personal data, and the main purposes of their use, as well as the identity and usual residence of the data controller

Table 13.3 Fair Information Practice Principles

General principle	General description	Original OECD principle and description
Consent	Data are only generated and disclosed with the consent of individuals	*Collection Limitation Principle.* There should be limits to the collection of personal data and any such data should be obtained by lawful and fair means and, where appropriate, with the knowledge or consent of the data subject
Security	Data are protected from loss, misuse, unauthorised access, disclosure, alteration and destruction	*Security Safeguards Principle.* Personal data should be protected by reasonable security safeguards against such risks as loss or unauthorised access, destruction, use, modification or disclosure of data
Integrity	Data are reliable, accurate, complete and current	*Data Quality Principle.* Personal data should be relevant to the purposes for which they are to be used, and, to the extent necessary for those purposes, should be accurate, complete and kept up-to-date

Table 13.3 Fair Information Practice Principles

General principle	General description	Original OECD principle and description
Access	Individuals can access, check and verify data about themselves	*Individual Participation Principle.* An individual should have the right: (a) to obtain from a data controller, or otherwise, confirmation of whether or not the data controller has data relating to her/him; (b) to have communicated to her/him, data relating to her/him within a reasonable time; at a charge, if any, that is not excessive; in a reasonable manner; and in a form that is readily intelligible to her/him; (c) to be given reasons if a request made under subparagraphs (a) and (b) is denied, and to be able to challenge such denial; and (d) to challenge data relating to her/him and, if the challenge is successful, to have the data erased, rectified, completed or amended

Table 13.3 Fair Information Practice Principles

General principle	General description	Original OECD principle and description
Use	Data are only used for the purpose for which they are generated and individuals are informed of each change of purpose	*Use Limitation Principle*. Personal data should not be disclosed, made available or otherwise used for purposes other than those specified within the notice, except: with the consent of the data subject, or by the authority of law
Accountability	The data holder is accountable for ensuring the above principles and has mechanisms in place to assure compliance	*Accountability Principle*. A data controller should be accountable for complying with measures which give effect to the principles stated above

Sources: Compiled from OECD (1980) and Minelli et al. (2013)

Other principles

- *Data minimisation*:
  - only generating data necessary for the task
  - only retain data while they are needed for that task
  - only use the data for this task
  - *...less strictly adhered to in practice?!*
- *Sensitive personal data:*
  - have stricter rules...

EU's General Data Protection Regulative (the GDPR)

The General Data Protection Regulation (GDPR)

- The toughest privacy and security law in the world
- Drafted and passed by the European Union (EU)
- «Re-codes the FIPPs for the data age»
- Imposes obligations onto organizations anywhere
  - so long as they target or collect data related to people in the EU
- Harsh fines against those who violate it
  - up tens of millions of euros or 4% of revenue

EU's Global Data Protection Regulative (GDPR)

- *Regulation (EU) 2016/679* on the protection of natural persons with regard to the processing [and free movement] of personal data ... repealing Directive 95/46/EC
- *Directive (EU) 2016/680* on the protection of natural persons with regard to the processing [and free movement] of personal data … for ... the prevention, investigation, detection or prosecution of criminal offences or the execution of criminal penalties... repealing Council Framework Decision 2008/977/JHA
- *Corresponding Norwegian regulations from 2018*
    - https://www.datatilsynet.no/

History of the GDPR

- The 1950 European Convention on Human Rights
  - right to privacy
  - "Everyone has the right to respect for his private and family life, his home and his correspondence."
  - protection of this right through legislation
- The European Data Protection Directive (1995)
  - minimum data privacy and security standards
  - each member state based its own implementing law
- Data protection as a fundamental right (Article 8 of the Lisbon Treaty, 2009)
- The GDPR
  - entered into force in 2016 after passing European Parliament
  - all organizations were required to be compliant by May 25, 2018

Privacy law in Norway

- *Protection of personality* (“personlighetsvern”, Norway, 1902):
  - right to a secluded private life (“privatlivets fred”)
- The *person registry protection law* (“Personregisterloven”, 1978)
- Law of *protection of personal information* (*“Personopplysningsloven”*, Norway, 2000)
  - implemented *Directive 95/46/EC* ... on the protection of individuals ... processing [and movement] of personal data (EU, 1995)
- *Constitutional paragraph* (§102) in 2014
- Revised law of *protection of personal information* (*“Personopplysningsloven”*, Norway, 2018)

New themes in EU Regulation 2016/679

- Larger geographical area
- Provisions for new technologies:
  - including big data, social media...
- Stronger rights for citizens
  - easier access to your own data
  - right to data portability
  - right to object to data processing (e.g., profiling, automated decisions)
  - explicit right to be forgotten
  - right to be informed when data security is breached
- Transparent and easily accessible data protection policies
- Data protection by design
- Less bureaucratic

Scope

- The GDPR
  - does *not* apply to "purely personal or household activity"
  - only applies to organizations engaged in "professional or commercial activity"
    - if you process the personal data of EU citizens or residents
    - if you offer goods or services to such people
    - even if you're not in the EU
- "Extra-territorial effect"...
- Penalties...

Scope

- "Extra-territorial effect":
  - the GDPR applies to
    - organisations that are based in the EU
      - even if the data are being stored or used outside of the EU
    - organisations that are not in the EU if:
      - the organisation offers goods or services to people in the EU, *or*
      - the organisation monitors the online behavior of people in the EU
      - EU embassies and other special territories
- Penalties
  - max €20 million or 4% of global revenue (whichever is higher)
  - data subjects have the right to seek compensation for damages

Terminology

- Personal data
  - any information that relates to an individual who can be directly or indirectly identified
  - *examples?*
    - *directly?*
    - *indirectly?*
    - *derived?*

Terminology

- Personal data
  - any information that relates to an individual who can be directly or indirectly identified
  - e.g., id numbers, bank accounts, names and email addresses, location information, biometric data, web cookies, IP addresses
    - some of them suffice alone, others must be combined
  - pseudonymous data
- Data processing
  - any action performed on data
  - whether automated or manual
  - e.g., collecting, recording, organizing, structuring, storing, using, erasing...

Terminology

- Data subject
  - the person whose data is processed
  - customers, site visitors, etc.
- Data controller
  - the person who decides why and how personal data will be processed
  - owners or employees who handle data
- Data processor
  - a third party that processes personal data on behalf of a data controller
  - special rules for these individuals and organizations

Personal and sensitive data

- *Personal* data:
  - data that directly or indirectly identifies or can be attributed to a person
  - name, addresses
  - personal identifiers of various kinds
  - identifiers of personal equipment of various kinds
- *Sensitive* personal data:
  - race/ethnicity
  - political opinions
  - religious/personal beliefs
  - trade union membership
  - health/medical information
  - marital status/sexual life
  - age
  - gender
  - criminal record

Table 13.1 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	
Credit/debit card number	Gender	
Biometric information (e.g., fingerprints, DNA, irises, face)	Criminal record	

GDPR themes

- Data protection principles
- Accountability
- Data security
- Data protection by design and by default
- JustificationConsent
- Data Protection Officers (DPOs)
- People’s privacy rights

Data protection principles

- Lawfulness, fairness and transparency: processing must be lawful, fair, and transparent to the data subject
- Purpose limitation: only process data for the legitimate purposes specified explicitly to the data subject when collected
- Data minimization: only collect and process as much data as absolutely necessary for the purposes specified
- Accuracy: personal data must be kept accurate and up to date
- Storage limitation: store personally identifying data for as long as necessary and for the specified purpose
- Integrity and confidentiality: processing must be done in such a way as to ensure appropriate security, integrity, and confidentiality (e.g., by using encryption)
- Accountability: the data controller must demonstrate GDPR compliance with all of these principles

Accountability

- Data controllers must be able to demonstrate they are GDPR compliant
  - "If you think you are compliant with the GDPR but can't show how, then you're not"
- Compliance demonstration:
  - designate data protection responsibilities to your team
  - maintain detailed documentation of the data you're collecting:
    - how it is used, where it is stored, which employee is responsible for it, etc.
  - train your staff and implement technical and organizational security measures
  - have Data Processing Agreement contracts in place with third parties you contract to process data for you
  - appoint a Data Protection Officer (not all organizations need one)

Data security

- "Appropriate technical and organizational measures" are required
- Technical measures, e.g.:
  - requiring your employees to use two-factor authentication
  - contracting only with cloud providers that use end-to-end encryption
- Organizational measures, e.g.:
  - staff trainings
  - adding a data privacy policy to your employee handbook
  - limiting access to personal data for employees
- Data breach
  - must inform data subjects within 72 hours
  - exception if safeguards used (such as encryption)

Data protection by design and by default

- Everything must consider data protection "by design and by default"
  - must consider the data protection principles in the design of any new product or activity
  - mandatory use of technical safeguards like encryption
  - higher legal thresholds to justify data collection

Privacy by design

- Designing privacy into IT artefacts from the start, instead of adding it as an afterthought
- Seven principles:
  - design-embedded privacy
  - privacy as the default setting
  - proactive not reactive & preventative not remedial
  - full functionality – positive sum, not zero-sum
  - end-to-end security – full life-cycle protection
  - visibility and transparency
  - respect for user privacy – keep it user-centric
- A related principle embedded in new EU regulations

Table 17.1 The 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 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

Table 17.1 The principles of privacy by design

Principle	Description
End-to-end security – full lifecycle 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 and empower individuals

Source: Cavoukian (2009)

Justification

- Legal justifications for processing data:
  - specific, unambiguous consent to process the data from the data subject
  - processing is necessary for the contract with the data subject
  - needed to comply with a legal obligation
  - needed to save somebody's life
  - necessary for the public interest or some official function
  - legitimate interest to process someone's personal data
    - the most flexible lawful basis
    - *the press has legitimate interests*
    - overridden by "fundamental rights and freedoms of the data subject"
      - especially for children

Justification

- The lawful justification for your data processing must be
  - documented
  - the data subject notified (transparency!)
  - good reason needed to change (+ documentation and notification again)

Consent

- Consent from a data subject to process their information must be
  - “freely given, specific, informed and unambiguous.”
  - children under 13 can only give consent with permission from their parent
  - documentary evidence of consent must be kept
- Requests for consent must be
  - “clearly distinguishable from the other matters”
  - presented in “clear and plain language”
- Withdrawal of consent
  - data subjects can withdraw their consent whenever they want
  - data controllers/processors must honor their decision
    - they cannot change their legal justification

Data Protection Officers (DPO)

- Not every data controller or processor needs to appoint a DPO
- Mandatory when:
  - public authorities other than a court acting in a judicial capacity
  - core activities require you to monitor people systematically and regularly on a large scale
  - core activities are large-scale processing of special categories of data:
    - provisions relating to specific processing situations
    - data relating to criminal convictions and offenses
- There are benefits of appointing a DPO in any case...

Provisions relating to specific processing situations

- Chapter 9 in the GDPR
- Special provisions for processing relating to:
  - *freedom of expression and information*
  - public access to official documents
  - national identification numbers
  - employment context
  - archiving purposes in the public interest, scientific or historical research purposes or statistical purposes
  - churches and religious associations

Freedom of expression and information

- Member states must by law reconcile:
  - the right to the protection of personal data (the GDPR)
  - the freedom of expression and information
- Freedom of expression and information includes
  - *journalistic purposes* & academic, artistic or literary expression.
- Member states shall provide exemptions or derogations of:
  - the general principles
  - the rights of the data subject, controller and processor
  - transfer of personal data to third countries or international organisations
  - independent supervisory authorities
  - cooperation and consistency
  - specific data processing situations

People's privacy rights

- Privacy rights for data subjects
  - aim to give individuals more control over the data they loan to organizations
  - the right to be informed
  - the right of access
  - the right to rectification
  - the right to erasure (to be forgotten)
  - the right to restrict processing
  - the right to data portability
  - the right to object
  - rights in relation to automated decision making and profiling

Limitations of the legislation

- Predictive profiling:
  - need a «'right to reasonable inferences', that specifically concerns predictive profiling and 'high risk inferences' that cause harms or make important decisions based on poor-quality data»
    - the right to determine if an inference is reasonable
    - being able to question why the inference is necessary for a purpose
    - being able to question the veracity of the data and calculations
    - the right to inspect the data and decision after the fact

Limitations of the legislation

- Group profiling
  - individuals as dividuals in different contexts:
    - politics, education, commerce, interests, values, preferences...
  - algorithmic identities comprising digital data
    - and treated on the basis of these data
- Particularly problematic wrt marginalised groups
- From personally identifiable information (PII)
  - to *demographically identifiable information (DII)*

Limitations of the legislation

- Predictive profiling:
  - need a «'right to reasonable inferences', that specifically concerns predictive profiling and 'high risk inferences' that cause harms or make important decisions based on poor-quality data»
    - the right to determine if an inference is reasonable
    - being able to question why the inference is necessary for a purpose
    - being able to question the veracity of the data and calculations
    - the right to inspect the data and decision after the fact
- Group profiling
  - individuals as dividuals in different contexts:
    - politics, education, commerce, interests, values, preferences...
  - algorithmic identities comprising digital data
    - and treated on the basis of these data

Getting around the legislation

- End-user license agreements (EULAs), cookie preferences...
  - an example of privacy policies
  - purposefully long and unreadable
  - reserves future modification rights, old versions can disappear...
  - liability disclaimers for companies, not assurance for customers
- Anonymising data
  - deidentification/anonymisation, pseudonymisation (hard)
  - danger of recombination with other datasets
  - creating aggregated / derived (tertiary) data
  - error infusion that preserves overall statistics

Privacy and big data

- The paradox:
  - big data can “make societies more secure, safe competitive, productive, efficient, transparent, and accountable”
  - yet “do so through processes that monitor, discipline, rerpess, presuade, coerce, and exploit people”
    - and thus weaken / threaten privacy
- The data are not good or bad
  - nor are their uses only good or bad
  - often simultaneous liberation/benefits and coercion/costs