Beyond De-Identification
Record Falsification
to Disarm Expropriated Data-Sets

Roger Clarke

Xamax Consultancy Pty Ltd, Canberra
ANU Research School of Computer Science, UNSW Law
Australian Privacy Foundation, Internet Society of Australia

http://rogerclarke.com/DV/RFED {.html, .pdf}

Bled eConference – 18 June 2019

32nd Bled eConference

Humanizing Technology for a Sustainable Society

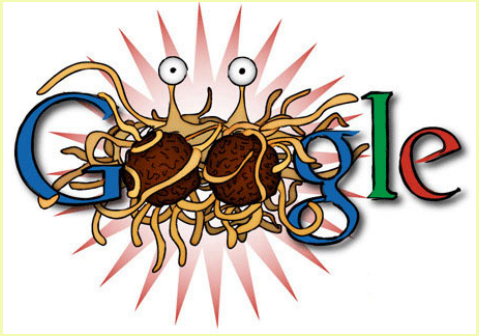
Problem Statement

- ‘Big Data’ / Data Science:
  - Expropriates Personal Data
  - Exploits Loop-Holes in Data Protection Laws
  - Uses the pretext that the data is De-Identified

Problem Statement

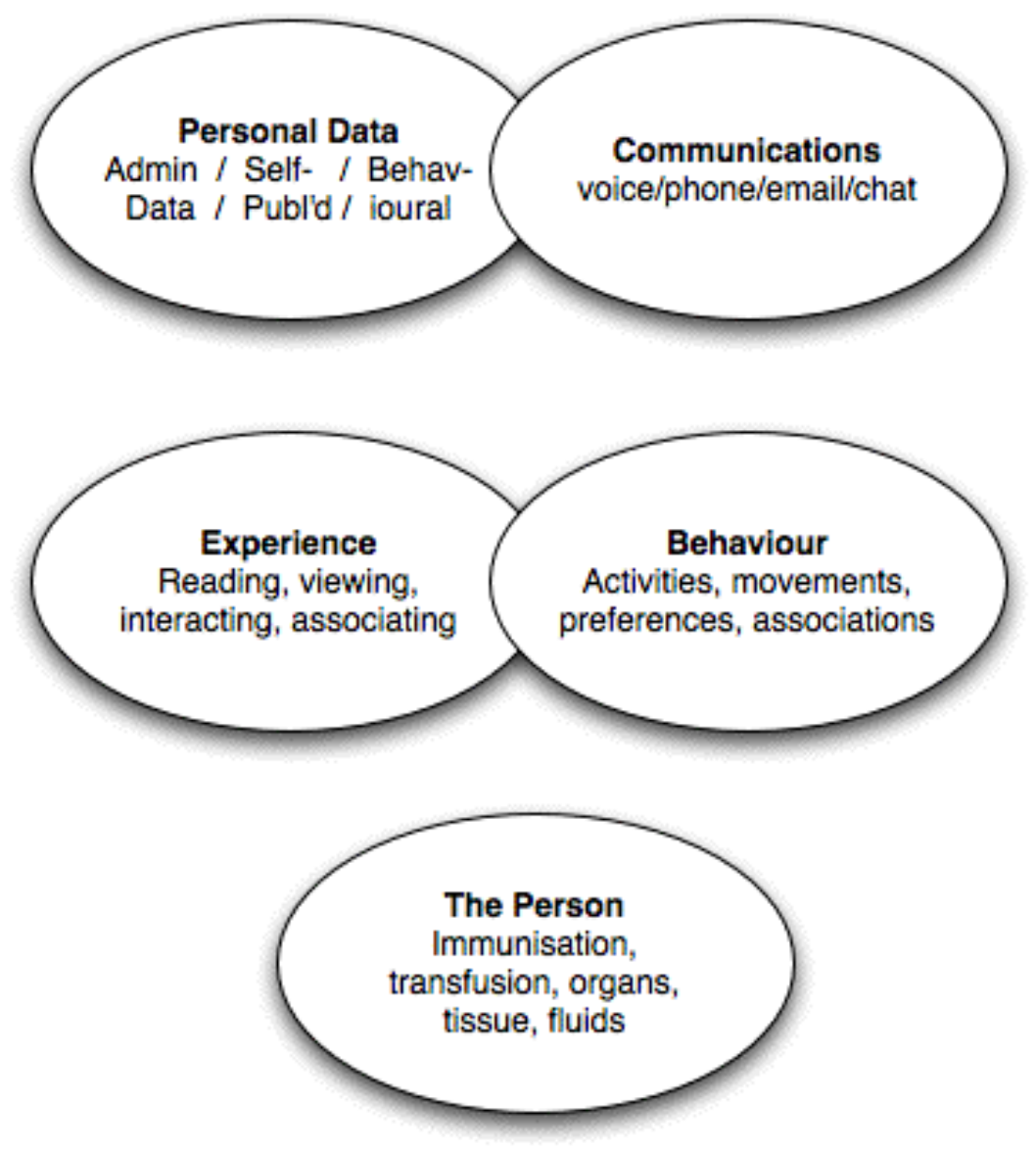
- ‘Big Data’ / Data Science Expropriates Personal Data, and exploits Loop-Holes in Data Protection Laws, under the pretext that the data is de-identified
- “After more than a decade of research, there is comparatively little known about the underlying science of de-identification” (Garfinkel 2015, p.39)
- **De-Identification Techniques don't work**
- **Re-identification Techniques <u>do</u> work**
- Privacy is a fundamental human right
- The assumption that Data Utility is the primary value needs to be replaced by ‘Privacy-First’

Privacy



- **Privacy** is the interest that individuals have in sustaining 'personal space', free from interference by other people and organisations
- **Data Privacy** is the interest that individuals have in controlling, or at least significantly influencing, the handling of data about themselves
- Information Privacy underpins the protections of **other privacy dimensions**:
  - Privacy of Personal Behaviour
  - Privacy of Personal Experience
  - Privacy of the Physical Person

Privacy Dimensions



http://www.rogerclarke.com/DV/Intro.html#Priv

Harms arising from Privacy Breaches

- **Physical**
  Discovery of identity or location leads to assault and worse
- **Psychological**
  Closed doors, drawn curtains, 'jumping for joy'; loss of control over one's life, image, and respect, undermining social cohesion
- **Economic**
  Stifling of non-conformist, risk-taking, inventive and innovative behaviour, undermining cultural, scientific and economic change
- **Political**
  Actual repression, and self-repression (the 'chilling effect'); Embarrassments, stigmas, reduced pool of political contributors
- **Philosophical**
  Human dignity, integrity, autonomy, self-determination

Low Quality Data 'Science' Heightens the Risk of Harm

Data is lifted **out of context** and 're-purposed'

Data is **merged or linked** with other data-sets

Faulty inferences arise because:

(1) **Data quality** is generally not high

(2) **Comparisons** of data-content are often unreliable

(3) **Data meaning** is often unclear or ambiguous

(4) **Data meanings** in multiple data-sets are commonly inconsistent or incompatible

(5) **Data scrubbing** cleans up some problems, moves the dirt somewhere else, and creates new problems

(Clarke 2016, 2018)
rogerclarke.com/EC/BDQAS.html
rogerclarke.com/EC/GDA.html

Categories of 'Persons-at-Risk'

Social Contexts

- Celebrities and notorieties at risk of extortion, kidnap, burglary
- Short-term celebrities such as lottery-winners, victims of crime
- **Victims of domestic violence**
- Victims of harassment, stalking
- Individuals subject to significant discriminatory behaviour
- People seeking to leave a former association, e.g. ex-gang-members

Political Contexts

- **Whistleblowers, Media Sources**
- Dissidents
- Human Rights Activists
- **Candidates for Political Office**

Organisational Contexts

- **Corporate executives, esp. M&A**
- Government executives
- **Undercover operatives**
- Law enforcement and prison staff
- Mental health care prof'ls, counsellors

Legal Contexts

- Judges, lawyers and jurors, particularly in highly-charged cases
- **Police Informants**
- Witnesses, especially **people in Protected Witness Programs**
- Ex-prisoners re-integrating with society

Clarke (2014)
http://www.rogerclarke.com/DV/UPETs-1405.html#Tab2

The Research Questions

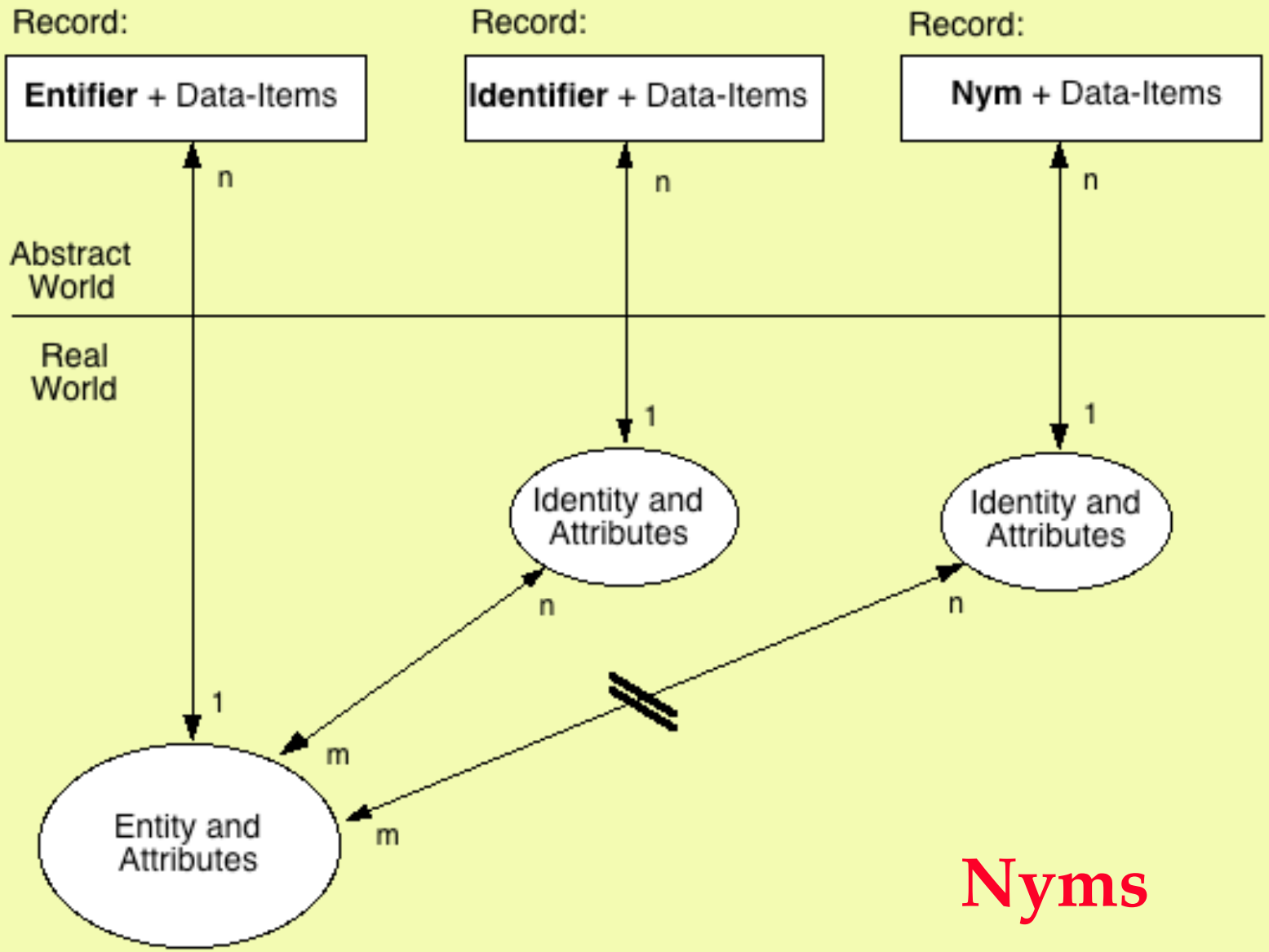
(1) Does De-Identification satisfy the requirements of current data protection laws?

(2) Whether or not it does so, **does De-Identification protect the interests of individuals?**

(3) **If answer** (1) or **(2) is 'No', what approach needs to be adopted in order to satisfy those needs**, while also addressing the interests of data-exploiters in industry, government and academe?

Identity

- **Anonymity**
  A characteristic of an Identity, whereby it cannot be associated with any particular Entity, from the data, or in combination with other data
- **Entity**
  A real-world thing
- **Entifier**
  A set of Data-items that distinguish an entity from similar entities
- **Identity**
  A real-world thing, but of virtual rather than physical form
- **Identifier**
  A set of Data-items that distinguish an identity from similar identities
- **<u>Anonym</u>**
  An Identifier that <u>cannot</u> be associated with any particular Entity, whether from the data itself, or by combining it with other data



Nyms

Clarke (2001, 2010)
rogerclarke.com/ID/IdModel-1002.html

'De-Identification'
Alternative Interpretations

(1) **Remove 'Identifiers'**
(Common, necessary, far from sufficient)

(2) (1) + **'Perturbate' the data-set**
(Common, necessary, but lacks a criterion)

(3) (2) + Process the data-set to **address the risks of merger, linkage or comparison of data-sets**
(Very uncommon, necessary, lacks a criterion)

(4) (3) + **Demonstrate the process's reliability**
(Hardly seen in literature or practice to date)

Conventional De-Identification Techniques

- **'Privacy-Preserving Data Mining' (PPDM)**

  Denning 1980, Sweeney 1996, Agrawal & Srikant 2000

- **Processing of the Data-Set before Release**

  Replacement, suppression, generalisation, perturbation

  UKICO (2012), DHHS (2012) Slee (2011)
  See also Garfinkel (2015), Polonetsky et al. (2016)

Re-Identification

The re-discovery or inference of an association between a record and a real-world (id)entity, despite any prior attempts to de-identify the record

Some techniques target specific individuals; whereas others are conducted on a statistical basis

Sweeney (2000), Narayanan & Shmatikov (2008), Acquisti & Gross (2009), Ohm (2010), Slee (2011)

Lots of Examples of Re-Identification

- "human mobility traces are highly identifiable with only a few spatio-temporal points" (Song et al. 2014, p.19)
- ”[credit card records with] four spatiotemporal points are enough to uniquely reidentify 90% of individuals ... [and] knowing the price of a transaction increases the risk of reidentification by 22%" (de De Montjoye et al. 2015, p. 536)
- successful re-identification of patients in a ‘de-identified’ open health dataset (Culnane et al. 2017, Teague et al. 2017)

Conventional De-Identification FAILS because it does not deliver Anonymity

Re-identification is easy where:

(1) The data-set contains **large numbers of data-items**

(2) **Unique values** exist within individual data-items

(3) **Unique combinations of values** exist across multiple data-items; and/or

(4) **Comparison data-sets are available**, e.g. electoral rolls, subscription lists, profiles on social networking sites, data broker offerings

‘Advanced’ De-Identification Techniques

Two families (D'Acquisto et al. 2015, p.30):

- **k-anonymity** and extensions
  p-sensitive k-anonymity, l-diversity,
  t-closeness, (n,t)-closeness
- **differential privacy** and variants
  crowd-blending privacy, BlowFish

k-Anonymity

- A framework for quantifying the amount of manipulation required of quasi-identifiers in order to achieve a given level of 'privacy' (Sweeney 2002)
- A data-set satisfies k-anonymity **iff each sequence of values in any quasi-identifier appears with at least k occurrences.** So 'privacy' merely means 'crowd-hiding'
- Bigger k is better (i.e. hide in a bigger crowd)
- **BUT** the technique addresses only some of the threats; attempts at repair have failed; in practice the value of 'k' is always set very low

Differential Privacy

- Mathematical techniques that reduce privacy risk by **adding non-deterministic noise to the results before release** (Dwork 2006, 2008)
- An algorithm is differentially private **if the probability of a given output is only marginally affected if one record is removed from the dataset**
  So again only a weak proxy for 'privacy'

- **BUT** dependent on assumptions re data, attacker, other data, attack-type, motivations; some claims debunked (Narayanan & Shmatikov 2010, Zang & Bolot 2011, Narayanan & Felten 2016, Zook et al. 2017, Ashgar & Kaafar 2019); statistical attacks are feasible (O'Keefe & Chipperfield 2013, pp. 441-451)

Conclusions about De-Identification

- **At best**, the result of the process is data that is **'mostly de-identified' or 'moderately perturbed'**
- The processes are complex and onerous
- More advanced forms are seldom implemented
- De-identification is a failure
- **Rich data-sets cannot be reliably de-identified**
- Organisations are routinely breaching public expectations and maybe also data protection law

Data-Utility has been the Objective with Privacy as a Mere Constraint

- "The goal is to keep the data 'truthful' and thus **provide good utility for data-mining applications, while achieving less than perfect privacy**" (Brickell & Shmatikov 2009, p.8)
- "The effort that is necessary to identify a single unit in the data set is higher than the actual benefit the potential intruder would gain by the identification" (Bleninger et al., 2010)
- "**Most data releasers today ... adopt the utility-first approach**" (D'Acquisto et al. 2015) pp.27-37)
- 'Re-identification risk' is defined as merely "the percentage of de-identified records that can be re-identified" (Garfinkel 2015, p. 38)
- O'Keefe et al. (2017) applies the threshold test of "when data is sufficiently de-identified given [the organisation's] data situation"

‘Humanising Technology’ requires:
Privacy as the Objective
Data-Utility as a Constraint

(1) Human rights law requires that the interests of people be a primary consideration

(2) Breach causes harm to individuals that may be far greater than the benefit to the breacher

(3) The many categories of 'persons-at-risk’ may suffer particularly serious harm

The Privacy-First Criterion

It is impossible to use an expropriated data-set:

- to discover any person's identity or location; or
- to usefully associate any data with an individual

Privacy-First Approaches

1. **Risk Avoidance**, by not using empirical data (Instead, Generate Synthetic Data)
2. **Risk Prevention**, by making the data unusable (Instead, Falsify the Empirical Data)

(1) Synthetic Data

- **Synthetic Data** does not relate to any individual, but "**has characteristics that are similar to real-world data** [with] frequency and error distributions of values [that] follow real-world distributions, and dependencies between attributes [that are] modelled accurately" (Christen & Pudjijono 2009. p.507)
- "It is possible ... to **construct an artificial database, for which sanitization provides** both complete utility and **complete privacy, even for the strongest definition of privacy** ..." (Brickell & Shmatikov 2009, p.7)

(2) Known Irreversible Record Falsification (KIRF)

- **Convert record-level data to synthetic data that represents a plausible phenomenon, not a real one**
- **Ensure widespread knowledge** of the fact of that processing, and of the standard achieved:
  (1) **by organisations** – so that they know it is unusable in relation to individuals
  (2) **by affected individuals** and their advocacy organisations – to ensure confidence and avoid motivating people to obfuscate or falsify

Test-Cases for Known Irreversible Record Falsification

- The combination of psychological and social data with stigmatised medical conditions
- Data about undercover operatives in national security and law enforcement contexts
- …
- …
- Every category of ‘Persons-at-Risk’ (Slide 8)

Can Data Utility be Rescued?

- Context-dependent, so **there's no general solution**
- For any given use, it may be feasible to apply **use-specific falsification processes** to produce a data-set that preserves the statistical features that are critical for that particular analysis
- It is likely that **circumstances exist in which it is infeasible to anonymise**, and hence the data-set cannot be released
- **Data-holders can provide services for 3rd parties**, conducting analyses and releasing non-sensitive data; or generating synthetic data

Duncan et al. (2001), Brickell & Shmatikov (2009), Friedman & Schuster (2010), current research?

Next Steps

- Keep searching for relevant existing literature
- Search for exemplars and testbeds
- Use k-anonymity with a very high value for k
- **Apply data perturbation and KIRF to existing data-sets, focussing on the Test-Cases**
- Begin with data-sets of convenience
- Move on to rich data-sets, e.g. those from Census, social data and health care fields that are commonly subjected to expropriation

32nd Bled eConference

Humanizing Technology for a Sustainable Society

Beyond De-Identification
Record Falsification to Disarm Expropriated Data-Sets

- Abandon the utility-first approach
- Adopt privacy as the objective, and relegate data-utility to a constraint
- Ban the release of all personal data-sets that are rich enough to support re-identification
- Apply Known Irreversible Record Falsification (KIRF) as the operational criterion
- Invest in Synthetic Data Techniques

Beyond De-Identification
Record Falsification
to Disarm Expropriated Data-Sets

Roger Clarke

Xamax Consultancy Pty Ltd, Canberra
ANU Research School of Computer Science, UNSW Law
Australian Privacy Foundation, Internet Society of Australia

http://rogerclarke.com/DV/RFED {.html, .pdf}

Bled eConference – 18 June 2019

Threat 'Models'

Victims of Domestic Violence

Discovery by a specific organisation and any informants of:

- individual identity
- the source documents / content / items of information
- the individuals to whom the d / c / i have been passed
- the individual's current location
- the individual's future locations

Whistleblowers

Discovery by a specific individual and any informants of:

- current location
- future locations

~~Protest Organisers~~

~~Discovery by 'the government' of:~~

- ~~individual identity~~
- ~~the movement's social network~~
- ~~the movement's plans and logistical arrangements~~
- ~~denial of service by 'the government'~~

Indicative Risk Assessment for a Whistleblower

Asset – Freedom

Harm – Denial of Freedom

Threats – Discovery of:

- Disclosure of suppressed information / documents
- Identities of persons involved in the disclosure
- Their Location
- Sufficient grounds to act

Vulnerabilities – Exposure of:

- Disclosure
- Identities
- Human entities underlying the relevant Identities
- Location of those persons

Security Safeguards re:

- Disclosures
- Actions, dates and times, physical and net locations,
- Identities
- Entities
- Locations

http://www.rogerclarke.com/DV/UPETs-1405.html#Tab3
https://freedom.press/encryption-works (Lee 2013)

Data Protection
A Weak Proxy for Protection of People's Privacy

- Data protection laws:
  - protect data not people
  - don't address behaviour, experience, safety
  - are riddled with loopholes
- **Non-EU countries' outdated data protection laws** are highly permissive of expropriation of personal data
- The **GDPR's** Art. 6 (Purpose Limitation Principle) is ripped apart by the **Art. 89 exemptions**
- These **Loop-Holes are mercilessly exploited**
- There is a risk of open warfare with the public, through encouragement of obfuscation and falsification of data

Corollaries of Known Irreversible Record Falsification

- If falsification of a record to the point of unusability cannot be achieved, then the record is unsuitable for expropriation, and no empirical derivative of it may be disclosed
- If undisclosable records constitute a sufficient proportion of the data-set as a whole, then the data-set as a whole cannot be disclosed