

Privacy, Deanononymisation and Transparency.

A Review for the Cabinet Office

Kieron O'Hara
30 March 2011

Will Transparency Pose a Threat to Privacy?

- Public data do not include personal data
- Departmental business plans seem relatively innocuous
- BUT
 - Where the state and the citizen interact, there will be issues
 - Cf. crime data
 - Demand-driven transparency
 - Health data
 - Education data
 - Court data ...
- Must ensure transparency of gov't \neq transparency of citizen

Balance



- Not a word I like here
 - How to ‘balance’ the citizen’s rights against the public interest?
 - Implies (falsely) that privacy and transparency are two opposite ends of a continuum
 - Is balance a process or a state?

3

Privacy isn’t (Solely) a Legal Matter

- Not just data protection
- ‘Mere’ legality is not enough
 - Law has many grey areas
 - Citizens’ perceptions do not follow the lines of DP
- Not sufficient to retain trust



4

Public Perceptions

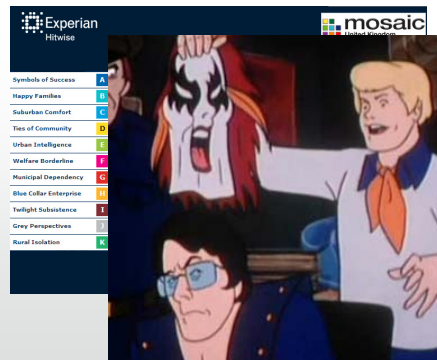
- Transparency depends on retaining confidence of citizens
- So perceptions are vital



5

Jigsaw Identification

- Specifically mentioned in remit
- Trade-off between data utility and privacy/anonymity



6

Squaring the Circle

- If two anonymised lines of a database differ somewhere they can be deanonymised with auxiliary information
- Main defences:
 - Disclosure control
 - Terms & conditions
 - Consent
- Inconsistent with transparency
- These technical issues are not well-understood
- Lack of empirical evidence

Theorem 1 Let $0 < \epsilon, \delta < 1$ and let D be the database. Let Aux be such that $aux = Aux(r)$ consists of at least $m \geq \frac{\log \delta - \log \epsilon}{-\log(1-\delta)}$ randomly selected attribute values of the target record r , where $\forall i \in \text{supp}(aux), \text{Sim}(aux, r_i) \geq 1 - \epsilon$. Then D can be $(1 - \epsilon - \delta, 1 - \epsilon)$ -deanonymized w.r.t. Aux .

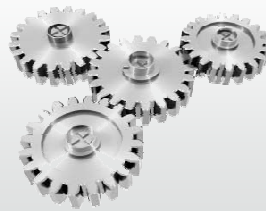
A Legal Fiction

- ‘Anony
 - Cf. ‘
- Yet the
 - Is th
- Alterna
 - Swe
 - Oh
 - Me



What's Missing

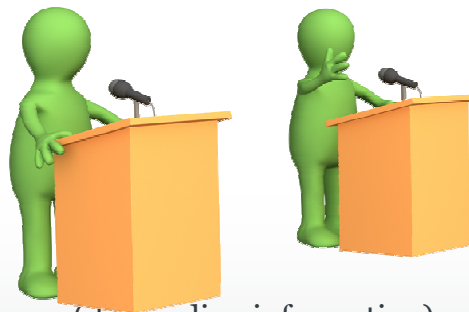
- There is no legal solution
 - Least of all a revised EU Data Protection Directive
- There is no technical solution
- There is no easy line to be drawn around private space
- What's missing are:
 - The processes
 - The institutions
 - The debates



9

Who Needs to be Involved?

- Transparency activists
- Privacy activists
- Technical experts
- Domain experts
- Information entrepreneurs (demanding information)



10

Conclusions

- This has so far been treated as a legal problem
 - It is as much technical as legal
- Transparency is a special context
 - Disclosure/query/access controls cannot work in general
- The role of an auditable debate trail
 - Quality of debate
 - Application of diverse sources of expertise
 - Media scrutiny
 - Public confidence
- Situation complicated but not hopeless
 - Embed privacy, don't bolt it on

11

Disclaimer

- Texts, marks, logos, names, graphics, images, photographs, illustrations, artwork, audio clips, video clips, and software copyrighted by their respective owners are used on these slides for non-commercial, educational and personal purposes only. Use of any copyrighted material is not authorized without the written consent of the copyright holder. Every effort has been made to respect the copyrights of other parties. If you believe that your copyright has been misused, please direct your correspondence to: kmo@ecs.soton.ac.uk stating your position and I shall endeavour to correct any misuse as early as possible.

12