Reforming the Data Protection Directive

Caspar Bowden

independent privacy advocate

(formerly: Chief Privacy Adviser, Microsoft Europe
Director, Foundation for Information Policy Research, www.fipr.org)

International Data Protection conference

Warsaw 21st September 2011

Review of Data Protection Directive

"official" agenda

- Principle of "Accountability"
- *faux* "Privacy by Design"
- "flexible" export rules
- DPA forum shopping ?
- Right to be Forgotten
- Privacy by Default (?)

the real privacy agenda

- "anonymisation"
- right to online access
- Privacy engineering for data minimization
- Art.29 does not enforce against major Internet companies
- future of Safe Harbor ?
- DPA computer science competence ?

Anonymization is a legal fiction*

- DPD exempts “anonymized” data from regulation
- ….but what do we mean by anonymous ?
  - external “auxiliary” data, “intersection attacks”
  - Shmatikov 2008 Netflix paper + algorithm
  - social network relationships cannot be anonymised
    - Dwork 2008 – “structural steganography”
- Paul Ohm’s paper “Broken Promises…”
- *Transparent Government, not Transparent citizens (Ch.4 2011 UK report, Dr.Kieron O’Hara)
  - *“the EU Directive forbids the release of almost everything, while the UK Act, which supposedly implements it, will allow rampant reidentification!”*

The problem of what is personal: "identifiable" by whom ?

- DPD split definition of personal data between an Article and Recital 26 ("or by any other person")
  - drafting of most important term of DP was defective in '95
- UK and IE are "leaky buckets"
  - export pseudonymous data with impunity
  - how did UK allow 1 month of all UK telephone records to be exported to US (in 2005)?
    - telephone numbers hashed but social structure intact!
    - biggest privacy breach from any EU country ?
- EU internal market **regulatory arbitrage**

Can Safe Harbor apply to data processors ?

- Notice
  - (x) processor doesn't know who are the data subjects
- Choice
  - (x) ..therefore processor can't manage individuals' choices
- Onward Transfer (notice and choice)
  - (x) *a fortiori* from above
- Access
  - (x) processor doesn't know who are data subjects
- Security
  - (x) processor doesn't know what controller is doing
- Data integrity
  - (x) processor cannot ensure accuracy, relevance, not excessive
- Enforcement
  - (x) *a fortiori* from above

Conclusions

1. Privacy engineers need a precise definition of personal data so they can minimize what they collect, protect what they process, and give the user access to their own data safely and privately. Unless regulators apply clear, consistent and forceful pressure to innovate in privacy technology, vested interests will suppress
2. Safe Harbor is not applicable to Cloud Computing (as processors)
3. Need new central authority for EU transnational enforcement
4. DPAs should create career structure for privacy engineers in computer science

   "technological neutrality" should not mean:

   "if it's technical, DPAs don't need to understand it" !
5. if the right of subject access is to be useful, people need a legal right to online subject access to personal data in online services
   - need Subject-Access-by-Design ?
   - core idea in privacy technology of "private credentials"
   - authenticating the data subject does not require identification
   - still not understood in legal/policy community after 20 years !

Recommended reading

"Online Privacy: Towards Informational Self-Determination on the Internet"

(Dagstuhl Perspectives Workshop 2011)

EU Data Protection Directive EC 95/46

- Article 2
  - *(a) 'personal data' shall mean any information relating to an identified or identifiable natural person ('data subject'); an **identifiable** person is one who can be identified, directly or indirectly, in particular by reference to an identification number or to one or more factors specific to his physical, physiological, mental, economic, cultural or social identity;*

The central question: Identifiable by whom?

- Recital 26
  - *Whereas the principles of protection must apply to any information concerning an identified or identifiable person; whereas, to determine whether a person is identifiable, account should be taken of all the **means likely reasonably to be used** either by the controller **or by any other person** to identify the said person; whereas the' principles of protection shall **not apply to data rendered anonymous in such a way that the data subject is no longer identifiable**.*

- Q. "means" – must refer to legal methods as well as statistical methods?
  - Otherwise passport/health-insurance/credit-card numbers would not be personal data (in themselves)
- Q. "means likely reasonably to be used" – must refer to the *method* not the *individual person*?
  - Because otherwise implies a person can be deprived of ECHR Art.8 privacy rights because in a minority!
- Q. "or by any other person" – seems to imply that the following rider "whereas…rendered anonymous" has to be understood as anonymous with respect to "any other person", not merely with respect to the putative data controller?
- Conclusion: if identification might likely reasonably occur by lawful process or statistical means (but even if only happens to a minority of individuals) it **is personal data**.

Data Retention Directive 15th March 2006 (EC 06/24)

Recital 3: Articles 5, 6 and 9 of **Directive 2002/58/EC** lay down the rules applicable to the processing by network and service providers of traffic and location data generated by using electronic communications services. Such data **must be erased or made anonymous when no longer needed for the purpose of the transmission** of a communication, except for the data necessary for billing or interconnection payments.

Recital 9: …. Because retention of data has proved to be such a necessary and effective investigative tool for law enforcement in several Member States, and in particular concerning serious matters such as organised crime and terrorism, it is necessary to ensure that retained data are made available to law enforcement authorities for a certain period, subject to the conditions provided for in this Directive. The adoption of an instrument on data retention that complies with the requirements of Article 8 of the ECHR is therefore a necessary measure.

Article 5 - Categories of data to be retained

- 1. Member States shall ensure that the following categories of data are retained under this Directive:
- (a) data necessary to trace and **identify** the source of a communication:…
- (2) concerning Internet access, Internet e-mail and Internet telephony:
  - (i) the user ID(s) allocated;
  - (ii) the user ID and telephone number allocated to any communication entering the public telephone network;
  - (iii) the name and address of the subscriber or registered user to whom an **Internet Protocol (IP) address**, user ID or telephone number was allocated at the time of the communication;
- (b) data necessary to **identify** the destination of a communication:…
- (2) concerning Internet e-mail and Internet telephony:
  - (i) the user ID or telephone number of the intended recipient(s) of an Internet telephony call;
  - (ii) the name(s) and address(es) of the subscriber(s) or registered user(s) and user ID of the intended recipient of the communication;
- (c) data necessary to identify the date, time and duration of a communication:…
- (2) concerning Internet access, Internet e-mail and Internet telephony:
  - (i) the date and time of the log-in and log-off of the Internet access service, based on a certain time zone, together with **the IP address, whether dynamic or static**, allocated by the Internet access service provider to a communication, and the user ID of the subscriber or registered user;

Since the function of Data Retention is to identify users from IP addresses, surely they are "personal" ?