9 easy squares: a framework for modern data protection issues

Eerke Boiten
April 2016

This slide pack merges the slides used at TRILcon16 with CO643 lecture slides.
Origin: RtbF (pre-Costeja)

• “I can't imagine how to have a meaningful conversation (much less write a law) about the Right to Oblivion without some framework to disentangle completely unrelated concepts, with completely unrelated implications.” Peter Fleischer, 2011
  http://peterfleischer.blogspot.co.uk/2011/03/foggy-thinking-about-right-to-oblivion.html

• “… data that is not really speech in any practical sense …”; “… stories …” Paul Bernal, CPDP 2013

[Hence still slight emphasis on web publication]
Objectives

- Knowledge on data protection legislation
- Awareness of contentious issues

Two dimensions, each with 3 possible values, give 9 contexts for personal data privacy. Some issues are in squares, others on boundaries.
First dimension: whose resource?

- **ME**: I (the data subject) am in full control of the resource (data controller).
  - No expectation of *linking* to other individuals.
  - No prescriptions on content or format
    
    *Example: personal web page, blog, email msg*

- **US**: The resource presents itself as a community or service, I am a member (?customer?product?) – there is an explicitly agreed relationship (T&Cs). The site shapes the contents and format. Links between members’ data. Published vs admin data.

  *Example: social network, comments section*

- **THEM**: The site has data about me but I have no direct control over that data. I may not even be aware that data/site exists.

  *Example: newspaper contents, ...*
Second dimension: type of content

• ** ATTributes** as traditionally stored in databases. Name, DoB, email address, .... Fixed number per person, format.

• **STOries**: explicitly generated personal information: texts, pictures, email contents, ...

  (What you get from facebook if you ask for “all info”)

• **BEHaviours**: implicitly generated personal info: mobile location, metadata, purchases

Note these may be roles: *location* can be any of 3.
<table>
<thead>
<tr>
<th>Attribute</th>
<th>ME Subject</th>
<th>US Participant</th>
<th>THEM Object</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>ATTRIBUTES</strong></td>
<td>control but identification effect to other contexts</td>
<td>traditional DPA context</td>
<td>traditional privacy: attributes in public domain, sensitive?</td>
</tr>
<tr>
<td><strong>“STORIES”</strong></td>
<td>full control</td>
<td>typical social network content</td>
<td>freedom of expression; cyberbullying</td>
</tr>
<tr>
<td><strong>BEHAVIOUR</strong></td>
<td>some leakage (e.g. dynamics)</td>
<td>metadata, browsing, hidden data and $$$</td>
<td>1984, data mining, deanonymisation</td>
</tr>
</tbody>
</table>
ME-ATTRIBUTES

• Full control, but should still worry re: visibility.
• Attributes disclosed here serve to identify / authenticate / disambiguate stories etc. found elsewhere. (Name, location, age, ...)
• Maintaining multiple identities on the internet: pseudonyms may sometimes be better and necessary. (No “real name” policy in this square.)
ME-STORIES

• Probably the reason to have the resource is publishing/storing “stories”
• Photos: identifying effect as in ME-ATT
• Like rest of this column: full control but erasure may leave a trace (e.g. Wayback Machine) which is part of ME-BEH
• Opportunities: SEO, painting a positive picture

ME-BEH not so interesting, until you sell/scrap your computer/smartphone/...
US-ATTRIBUTES

- Service: standard DP scenario. Service holds database with personal details of customers.
- Data minimisation, cancellation (if not gov't)
- Data published agreed in T&Cs
- For communities: privacy settings may apply to any publication
- Links to other members
US-STOries

• Essence of social networks and media collection sites
• Reluctant response to cancelling set off some alarm bells
• Privacy settings: other members, publication, search engines (complex!)
• Links to other members
US-BEHaviour

- Many services in US column are free. This square is where business model ("surveillance capitalism") lives.
- Aggregated? Anonymised? What does that mean in practice?
- Targeted advertising, selling to data brokers (into THEM-BEH)
- E.g. facebook search history
- Outside privacy controls
THEM-ATTRIBUTES

• Your attributes on a newspaper website: painful for security questions; identity theft
• Often unavoidable
• Search engines reach this and stories relating, this is where “delisting”/RtbF hits, freedom of expression versus privacy.
• Celebrities’ age/pregnancy/relationship/sexual orientation/...
• Where attributes not published, you may not know they are held (traditional DP problem)!
THEM-STOries

• Newspaper stories about you (etc) – no knowledge or contract
• Faces in pictures (UK data protection view now: not personal info if not likely processed/identified or linked).
• “Right to be Forgotten” (including Costeja ones) battles certainly here
• Social networks: cyber bullying, promoting tagging (pulls into US-STO), subtweeting
• Intelligence services: communication content
THEM-BEHaviour

• So they have data about your behaviour, but no agreement and likely no awareness
• Surveillance: intelligence services, data brokers obtained from “US” services, or harvested from web (“social media research”)
• “1984” scenarios
• Beacons in shopping malls, MAC sniffing bins
• De-anonymisation risks
• Big data
New paper due May 2016.
New European Data Protection Regulation: more later!
Aww there’s a youtube video too