



09/06/2015

Big Data: Wrongs and Rights

Andrew Cormack, Chief Regulatory Adviser, @Janet_LegReg

- » Volume + Variety + Velocity
- » Data-driven, not hypothesis-driven
 - › Don't know in advance what you might find
 - › Minimal time to think about it after you do

- » Big data for education
- » “huge potential...to contribute to our understanding of the different variables impacting on the effectiveness of learning, student success and retention” [Prinsloo & Slade]
- » Big Data “raises important social, legal and ethical questions, among which [are] concerns with regard to the privacy and data protection rights of these individuals” [Art29WP]
- » Both envisaging what **might** happen, but we don't know...
- » BTW, that's **our** data...

- » Course material hard to understand – needs improving...
- » Individual struggling – needs help...
- » Students like you...
- » Which modules will student do best on...
- » Which applicants will do best here...
- » Improve wifi, but only where students are studying...
- » Great lecture, noone using Facebook...
- » No, you weren't studying last night...

- » “Right”?
- » “Wrong”?
- » “OK in some conditions”?

Challenge: How to “do the right thing”?

Thinking about

- » What protection does it give the individual?
- » What guidance does it give the analyst?

“Research subjects must be informed fully about the purpose, methods and intended possible uses of the search, what their participation...entails and what risks, if any, are involved.”

“... must participate in a voluntary way, free from any coercion” [Nuremberg Code/Helsinki Declaration]

- » Individual should be “fully informed” in advance
 - › But whole point of BD is we don’t know in advance 😞
- » Tells analyst how to ask, not what (not) to ask for

“If you sign up for an Oyster online account [rather than leaving your card pseudonymous], you can see what's on your Oyster card, see your journey history and apply for refunds. Any Oyster cards that you add to your account will automatically be protected against loss or theft”
[Oystercard website]

- » Benefit clear to individual; no mention of disbenefit
- » No guidance to analyst on what's acceptable
 - › Except “what keeps me off the front page of Metro”

“Processing is necessary for the purposes of the legitimate interests pursued by the controller...except where such interests are overridden by the interests for fundamental rights and freedoms of the data subject” [95/46/EC Art.7f]

- » Individual’s rights can always “override”
 - › When risk to DS doesn’t justify benefit to DC [Art29WP]
- » Analyst required to ensure rights/interests protected
 - › Reduce (potential) harm; improve protection of rights

Big data: balancing interests

- » Analysis of Big Data (“finding patterns”)
 - › Legitimate Interests looks like a good model
 - Must have (+disclose) clear, legitimate, beneficial purpose
 - Must protect data/individuals’ rights while investigating it
 - Definitely not a carte blanche...
- » Intervention from Big Data (“applying patterns”)
 - › Consent generally preferable
 - Now you **can** fully inform the individual
 - Who is free(-ish) to choose personalised or generic treatment

- » Not just privacy/DP rights that may be affected
 - › Exclusion, discrimination, chilling effect on speech, etc.
 - › Must include these in the balance
 - And stop analysis when risk is too great
- » Is human or automated intervention better?
[law places additional constraints on automated]
 - › “based on your engagement, here are other seminars”
 - › “based on your engagement, here’s your mark”

- » Technologies?
 - › Reduce risk to individuals?
 - › Protect against misuse?
- » Processes?
 - › Right ways to use tools (technical, legal, ethical)
- » Norms?
 - › Things we always do
 - › Things we never do

- » Big data for mutual benefit, not exploitation
- » Paul Bernal: “Symbiotic web”

Andrew Cormack
Chief Regulatory Adviser, Jisc Technologies

Andrew.Cormack@jisc.ac.uk

<https://community.ja.net/blogs/regulatory-developments>



Except where otherwise noted, this work is licensed under CC-BY-NC-ND