

# **Hypernudge: Big Data as a Mode of Regulation by Design**

**Karen Yeung**

Professor of Law, Director, Centre for Technology, Ethics, Law & Society (TELOS)  
The Dickson Poon School of Law, King's College London

Distinguished Visiting Fellow, Melbourne Law School

Friday 8<sup>th</sup> April,  
Centre for Corporate Law, Melbourne Law School

# + What is Big Data?



- A new Industrial Revolution is dawning, powered by the engine of ‘Big Data’
- Transformation in service delivery (finance, education, health, policing, energy distribution, dating etc) - how is industry harnessing Big Data to transform personal digital data into economic value?
- ‘Big Data’ - no universal definition, but essentially a combination of a technology + process:
  - Technology: configuration of information-processing hardware and software capable of sifting, sorting and interrogating vast quantities of data in very short times
  - Process: mining data for patterns, distilling the patterns into predictive analytics, and applying the analytics to new data
- **Methodological technique** that utilises analytical software to identify patterns and correlations through the use of **machine learning algorithms** applied to (often unstructured) data contained in multiple data sets, converting these data flows into a **highly data-intensive form of knowledge** (Cohen 2012)

# + Why the Big Fuss?



- Critically, enables the discovery of useful correlations within datasets *not capable of analysis by ordinary human assessment* (or even conventional computing techniques)
- Big Data's value lies in **finding patterns** that can be derived from making connections about pieces of data, about an individual, about individuals in relation to others, about groups of people, or simply about the structure of information itself.
- 'Big Data is important because it refers to an analytic phenomenon playing out in academia and industry' (boyd & Crawford 2012) – this understanding adopted here

# + Structure and argument



- Lawyers - Concerns about Big Data typically expressed in terms of **privacy** and **security**
- But **something more is** at stake, which ultimately concerns the quality of our individual agency and capacity for democratic participation
- I will argue that Big Data can be understood as a mode of ‘design-based’ regulation, which utilises a deceptively simple mechanism of influence - ‘**nudge**’ to channel attention and decision-making in directions preferred by the choice architect in a subtle, unobtrusive yet **extremely powerful** manner
- Aim: evaluate the legitimacy of these techniques, understood primarily in terms of conformity with liberal framework of values, rights-based perspective
- But also multi-disciplinary. Although my approach is animated around liberal political theory, I also draw from regulatory governance scholarship, behavioural economics, information law scholarship, STS and surveillance studies

# + Regulatory governance



- **‘Regulatory governance’** scholarship: multi-disciplinary field of scholarly inquiry concerned with critically examining the dynamics and legitimacy of ‘regulation’
- **Regulation** ( ‘regulatory governance’) = ‘the organised attempt to manage risks or behaviour in order to achieve a publicly stated objective or set of objectives’ (Black 2014)
- **Regulators** not confined to state agencies, also NGOs and firms
- **Regulatory techniques:** the instruments employed by regulators to attain the desired social outcome. Many different forms. Lawyers typically focus on ‘command and control’. My work on emerging technologies has focused on the use of ‘design’ (code) and other ‘technological’ forms of control (Lessig 1999: in cyberspace, code is ‘law’)

# + Rules vs design as regulatory techniques



# + Choice architecture as design-based regulation

- **Choice architecture** is a design-based regulatory instrument, which refers to intentionally designing the ‘choice environment’ in which individual decision-making takes place
- Some forms are intentionally made visible in order to prompt conscious behaviour change. Eg speed hump
- Others less visible and more subtle: Airline terminals designed so that passengers must walk through retail area in order to proceed to boarding gates – to maximise shopping opportunities.



## + ‘Nudge’ as a form of soft design-based control

- ‘**nudge**’ = a particular form or aspect of choice architecture that alters behaviour in a predictable way without forbidding any options or significantly changing economic incentives (Thaler and Sunstein, *Nudge* 2008)
- Intellectual heritage: findings from lab experiments by cognitive psychologists concerned with understanding human decision-making, demonstrating that people routinely and systematically make ‘irrational’ decisions due to reliance on **cognitive heuristics** (mental shortcuts): Kahneman and Tversky
- Nudges **rely on these heuristics** to influence decision-making in a subtle, unobtrusive fashion which **preserves individual choice** (doesn’t formally alter the available range of options).





# Ordinary (static) nudge



- Critically, individuals make decisions in a **passive and unreflective manner** rather than through active, conscious deliberation
- Arrange presentation and layout of food items in a cafeteria by placing the healthy items in front of the junk food to encourage healthier eating, owing to the 'availability' heuristic

# + Big Data as a form of design-based control

- To understand how Big Data analytic techniques entail the use of nudge, we can distinguish between two broad configurations of Big Data driven decision-making processes
  1. **Automated decision-making systems:** many common transactions now entail the operation of automated decision making processes. Eg. ticket dispensing machines through to instant loan offers (eg Wonga). These systems automatically issue a ‘decision’ without any human intervention
  2. **Digital decision guidance systems** (‘persuasive systems’): these systems seek to guide or ‘help’ a human agent make decisions in ways identified as ‘optimal’ by the underlying software algorithm by offering ‘suggestions’ intended to prompt the user to make decisions preferred by the choice architect. Eg general internet search engines

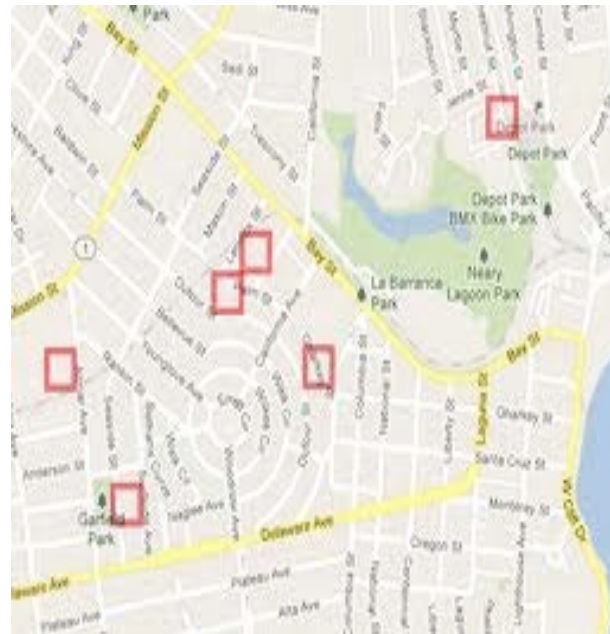
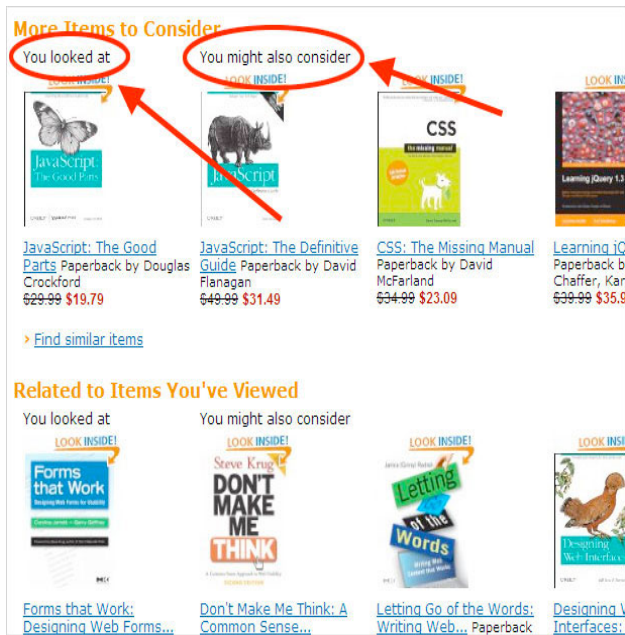
# + Big data techniques as Nudge

- For example: general internet search engines
- Big data decision guidance systems employ nudge by **highlighting correlations** between data items within a data set that would not otherwise be observable without continuous, real-time algorithmic processing
- Essentially, powerful tools for selection optimisation, conferring 'salience' on highlighted data





# Big Data as selection optimisation



- Other algorithmic selection optimisation techniques operate in a similar fashion, aimed at helping the user identify which data items to target from a very large population.
- Today: Big Data driven persuasive systems of commercial digital service providers, NOT
  - a research technique for data analysis and knowledge production
  - an instrument of state surveillance

# + Ordinary nudge vs hypernudge



Speed hump



Networked real-time navigation systems (eg Google Maps)

# + Big data as ‘hypernudge’

- Unlike conventional, static nudge, Big Data driven nudges entail automatic enforcement that is **dynamic and individualised**, with both the standard and its execution constantly and continuously updated, refined within a networked environment through real time data feeds, operating in three directions:
  - Continual, real-time revision of **individual’s choice environment**
  - Continual **feedback** provided to the choice architect
  - Continual monitoring and refinement at a **population-wide** level
- Enables the **personalisation of a user’s choice architecture** ( cf static nudges ) Nimble, unobtrusive and highly potent and powerful - hence ‘hypernudge’



# + Are Big Data driven ‘hypernudge’ techniques legitimate?

Despite enthusiastic embrace of nudge by policy-makers, considerable academic critiques of (ordinary) nudge

- ‘Libertarian Paternalism’

- Performance based critiques - Ineffective or unintended effects

## 2. Liberal manipulation critique

- Choice architects may pursue illegitimate motives (eg 2014 Facebook experiments – 700,000 user news feeds manipulated to test whether exposure to emotions led people to change their own Facebook posting behaviours). Critics: ‘mass experiment in emotional manipulation’ violating basic principles of research ethics

- Nudges as **deception**: even if legitimate purpose pursued, causal mechanism deliberately seeks to exploit cognitive weaknesses, hence manipulative and deceptive, fails to respect persons

- Nudges **are opaque** – nudges often lack transparency, akin to subliminal advertising. Algorithms as black boxes, lack of transparency & accountability, exacerbate risks of abuse

## + Does notice and consent overcome worries about manipulation?

- Can these objections to the opaque manipulative quality of hypernudge be overcome via individual consent to their use?
- Consider first from a liberal, rights-based perspective
- Right most directly implicated by Big Data driven hypernudging is the **right to informational privacy** given that they entail continuous monitoring of individuals, and the collection and algorithmic processing of personal digital data
- Contemporary data protection laws based on a model of **privacy self-management**: individuals decide how to weigh costs and benefits of personal data sharing





## + Problems with digital notice and consent

- Legal position: hypernudge ok if notice + user consent (via online ‘privacy notices’)
- But can these concerns *really* be overcome by notice + consent? .
  1. People don’t read or understand ‘privacy notices’ (Crannor 2008: would take 244 hours per year)
  2. People do not meaningfully assess the risks and benefits of data sharing due to problems of bounded rationality and aggregation, because many privacy harms are cumulative but users are required to assess isolated transaction
  3. Privacy preferences are highly malleable and easily manipulated by environmental cues (such as digital defaults)

Hence deep skepticism that click-through ‘acceptance’ of privacy notices = meaningful waiver of right to informational privacy





## Big Data and the notice and consent paradigm



Big Data exacerbates these problems because:

- Big Data mindset is all about collect as much data as you can first, work out what to do with it later: so can't specify intended purpose at time of collection
- 'Transparency Paradox' (Nissenbaum): the level of disclosure required to enable individuals to provide informed consent to the use of their data in highly dynamic, complex networked environments is too great and difficult for most people to comprehend, but simple disclosures fail to provide adequate information.
- 'Free' rather than 'fee' services model: users consent to share their personal data in return for services. Yet 'the power of free can get us to make many foolish decisions' (Ariely)



# Hypernudging, deception and consent

Notice and consent is unlikely to overcome liberal manipulation objections unless techniques expressly notified

- People have an independent moral right not to be deceived.
- So consent to data sharing ≠ consent to deception.
- Consider ‘digital gerrymandering’ – ie. social media platform utilising Big Data analytics actively to manipulate the voting behaviour of users during an election (Zittrain 2014).

“The content of your newsfeed is determined by an algorithm that has been constructed in ways intended to foster Facebook’s success as a commercial enterprise”

# + Authorising deception?



- Specific disclosure of the proposed deception is required:

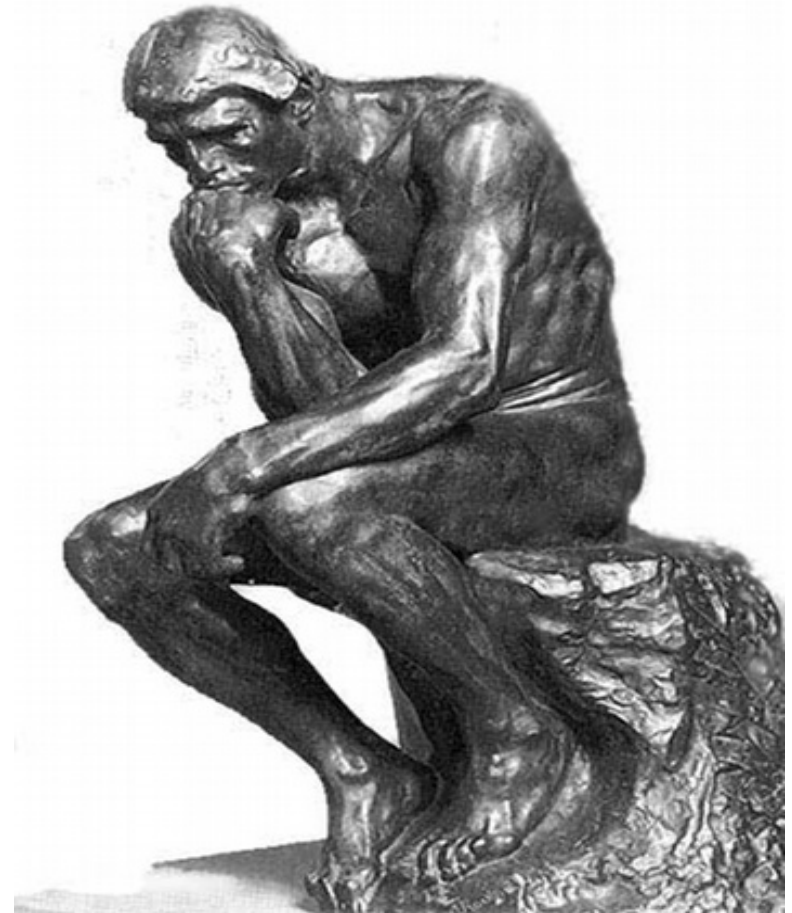
“The content of your newsfeed is determined by an algorithm that has been constructed in ways intended to encourage you to favour the political candidates favoured by Facebook.”

- Notice that material information is being withheld is unlikely to overcome concerns about deception (eg I inform you that I have omitted significant items from a dossier, but I don't tell you what has been omitted)
- Yet liberal-legal perspective, the threshold for notice and consent is low, generating no further concerns: individual autonomy respected while the market fosters innovation in digital services



# Post-liberal critiques: Insights from STS

- Why does law struggle to confront the deeper ethical, political and social concerns arising from data-driven persuasive technologies?
- Answer: rooted in the western liberal political tradition's understanding of the self and self-society relation
- Individuals as rational beings, separate and detached from their surrounding environment. Core idea of liberal personality as autonomy: freedom from interference from others, emphasis on individual choice.
- But choices cannot be made abstracted from context: our choices are always affected by the actions of others unless an 'interference' with our autonomy.
- When does the design of our surrounding choice architecture = 'interference'?



# + STS understandings of the self



- But critics point out that we are deeply enmeshed in identity-constituting relations, cultural and other connections, and we have little choice over many aspects of the self (eg embodiment)
- Cf STS approaches: adopt a much richer account of the self-society relation
- Individuals shape and configure their technologies but are also ‘configured’ by those technologies via a process of ‘co-production’
- Technological artefacts shape and mediate our relationship with the world around us and, over time, we come to perceive the world through the lenses that our artefacts create (Verbeek)

# + Hypernudge and surveillance capitalism

- On this reading, Big Data hypernudging operates as a ‘soft’ mechanism of surveillant control: via ‘modulation’: subtly molding the networked user’s understanding of the surrounding world by ‘tailoring their conditions of possibility’ (Cheney-Lippold 2011)
- But unlike traditional surveillance, which relies upon the coercive experience of living with the uncertainty of being seen (the Panopticon), it is subtle, unobtrusive yet highly potent
- It is central to an emerging ‘logic of accumulation’ which Zuboff calls ‘surveillance capitalism’ in which power is identified with the ownership of the means of behavioural modification cf industrial capitalism: power identified with ownership of the means of production (Zuboff 2015)

# + Why should we care?

## ■ Democracy impoverished

- The continuous, subtle molding of our individual choice environment by commercial Big Data barons produces individuals who are tractable, predictable neoliberal consumers rather than active, political citizens (Cohen 2012)
- Although we actively embrace them, our consent is akin to that of the gambling addict who willingly parts with his money for one more roll of the dice (Natasha Dow Schull: 'asymmetric collusion' between Las Vegas casinos and players to 'give players what they want')

- **Individual autonomy and agency diminished:** not flourishing, individual self-choosing agents envisaged (per liberal theory) but modulated individuals whose capacities and opportunities to undertake projects of authentic self-creation are blunted





# + Conclusion



- We are right to be concerned about our individual privacy and security in a data-driven age. But there is something much more fundamental at stake: the quality of our democracy, and the quality of our agency.
- “Google’s tools are not the objects of value exchange, but ‘hooks’ that ‘lure users into extractive operations that turn ordinary life into the daily renewal of a 21<sup>st</sup> century Faustian pact” (Shoshana Zuboff 2015)
- It is therefore essential that we find ways to secure meaningful and democratic accountability over the algorithms that exert ever more influence over our lives.

Karen Yeung ‘Hypermudge: Big Data as a Mode of Regulation by Design’ (2016)  
*Information, Communication & Society*, forthcoming, special issue on the Social Power of Algorithms.