CO-CREATING AUTONOMY

Group data protection and individual self-determination within a data commons

Janis Wong and Tristan Henderson

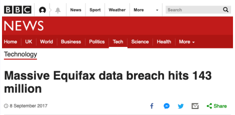
School of Computer Science
University of St Andrews
janiswong.org
jccw@st-andrews.ac.uk

University of
St Andrews

FOUNDED
1413

THE DATA PROTECTION DIVIDE

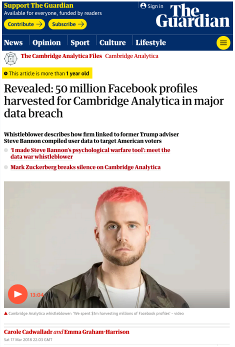
Data subjects are powerless against the increasingly sizeable and international data controllers in our data-driven society



Technology

Massive Equifax data breach hits 143 million

8 September 2017



The Cambridge Analytica Files Cambridge Analytica

This article is more than **1 year old**

Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach

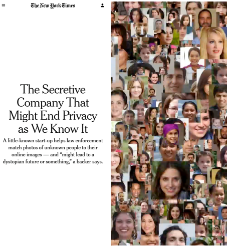
Whistleblower describes how firm linked to former Trump adviser Steve Bannon compiled user data to target American voters

- 'I made Steve Bannon's psychological warfare tool': meet the data war whistleblower
- Mark Zuckerberg breaks silence on Cambridge Analytica

Cambridge Analytica whistleblower: 'We spent $1m harvesting millions of Facebook profiles' – video

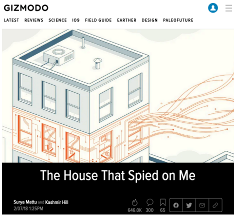
Carole Cadwalladr *and* Emma Graham-Harrison

Sat 17 Mar 2018 22.03 GMT



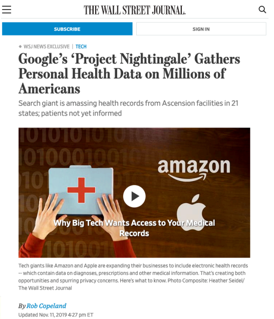
The Secretive Company That Might End Privacy as We Know It

A little-known start-up helps law enforcement match photos of unknown people to their online images — and "might lead to a dystopian future or something," a backer says.



The House That Spied on Me

Surya Mattu and Kashmir Hill
2/07/18 1:25PM



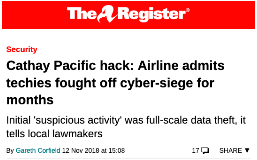
WSJ NEWS EXCLUSIVE | TECH

Google's 'Project Nightingale' Gathers Personal Health Data on Millions of Americans

Search giant is amassing health records from Ascension facilities in 21 states; patients not yet informed

Tech giants like Amazon and Apple are expanding their businesses to include electronic health records -- which contain data on diagnoses, prescriptions and other medical information. That's creating both opportunities and spurring privacy concerns. Here's what to know. Photo Composite: Heather Seidel/ The Wall Street Journal

By *Rob Copeland*
Updated Nov. 11, 2019 4:27 pm ET



Security

Cathay Pacific hack: Airline admits techies fought off cyber-siege for months

Initial 'suspicious activity' was full-scale data theft, it tells local lawmakers

By Gareth Corfield 12 Nov 2018 at 15:08

LEGAL & TECH SOLUTIONS

- **Law:** provide data subjects with rights
  - General Data Protection Regulation
  - California Consumer Privacy Act
- **Legal Frameworks:** facilitate data stewardship
  - Data trusts
  - Data Collaboratives
- **Privacy tools:** granular control for data subjects
  - Databox
  - Solid
- **Data reuse with privacy-by-design:** support data controllers
  - Data Transfer Project
  - OpenGDPR
  - Jumbo Privacy

DATA SUBJECT DATA PROTECTION DRAWBACKS

- Existing solutions rely on data subjects having a high level of understanding of both the law and technological resources available for individual redress, often *after* data collection
  - Cambridge Analytica: academic tensions, broad reaching data collection, and wider implications on democracy
- Responsibilisation of the data protection process from data controllers to data subjects
  - Data subjects cannot access Apple's Siri recordings of them (Veale et al., 2018)
- Heavy reliance on under-resourced data protection authorities to enforce data protection laws and act on breaches
  - Irish Data Protection Commission was given 27% of its requested increase by the Government despite increased responsibilities post-GDPR (The Irish Times, 2019)

DATA SUBJECT DATA PROTECTION DRAWBACKS

- Existing solutions rely on data subjects having a high level of understanding of both the law and technological resources available for individual redress, often *after* data collection
  - Cambridge Analytica: academic tensions, broad reaching data collection, and wider implications on democracy
- Responsibili ess from data controllers t
  - Data sub dings of them (Veale e

Data subjects lack a meaningful voice in creating solutions that involve protecting their own personal data

- Heavy reliance on under-resourced data protection authorities to enforce data protection laws and act on breaches
  - Irish Data Protection Commission was given 27% of its requested increase by the Government despite increased responsibilities post-GDPR (The Irish Times, 2019)

FINDING A CO-CREATED, COLLABORATIVE SOLUTION

- Based on Elinor Ostrom's work on the commons (1990):
  - **Boundary setting:** Communities can set and enforce rules of the use of a common-pool resource (CPR)
    - Transparency, accountability, citizen participation, and management effectiveness
  - **Bottom-up norms:** Autonomous human decisions and activities following community rules
  - **Polycentricity:** A complex form of governance with multiple centres of decision-making, each of which operates semi-autonomously and collectively
  - **Design principles:** Govern the limitations of the commons and allow for iterative changes to regulate the CPR

DATA COMMONS

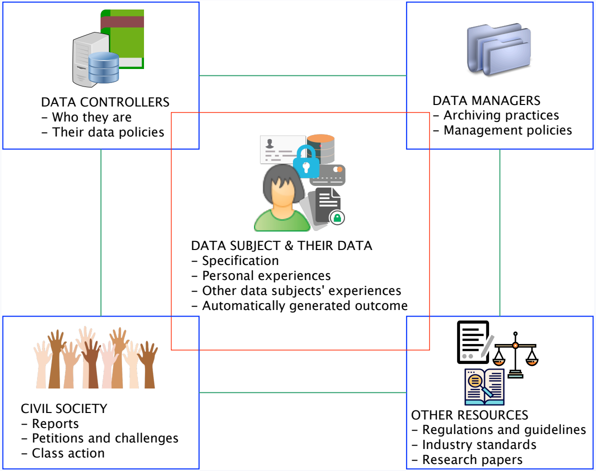
- Collaboration and co-creation with data subjects can limit the harms from mass data collection, processing, and sharing
- A data protection-focused commons allows data subjects to collectively curate, inform, and protect each other
    - Common property is personal data from data subjects
    - Contextualise privacy beyond control
    - Collective exercise of data protection rights
    - Create new norms using shared knowledge
    - Identify accountability to stakeholders
        - Data controllers, data managers, researchers, civil society

DATA PROTECTION-FOCUSED DATA COMMONS

- The data subject specifies to what extent they would like their data to be protected
- No prior knowledge of law, policies, or tech required
- The data subject's data protection outcome is automatically generated from the system based on their specification and stakeholder information
  - E.g. An airline is decided for the data subject based on their privacy policies and any recent scandals
- Data subject decisions may override existing preferences, policies, or standards set by stakeholders
  - E.g. Data subjects can request their data be kept in research papers even though it may be deleted online
- Data subjects can review their outcome, add their experiences, and participate in the co-creation process

DATA COMMONS SET UP

The data subject and their personal data are the most important. Other stakeholders are considered in context of the data subject's data protection.



DATA CONTROLLERS
- Who they are
- Their data policies

DATA MANAGERS
- Archiving practices
- Management policies

DATA SUBJECT & THEIR DATA
- Specification
- Personal experiences
- Other data subjects' experiences
- Automatically generated outcome

CIVIL SOCIETY
- Reports
- Petitions and challenges
- Class action

OTHER RESOURCES
- Regulations and guidelines
- Industry standards
- Research papers

DATA COMMONS FOR DATA CURATION - USE CASE

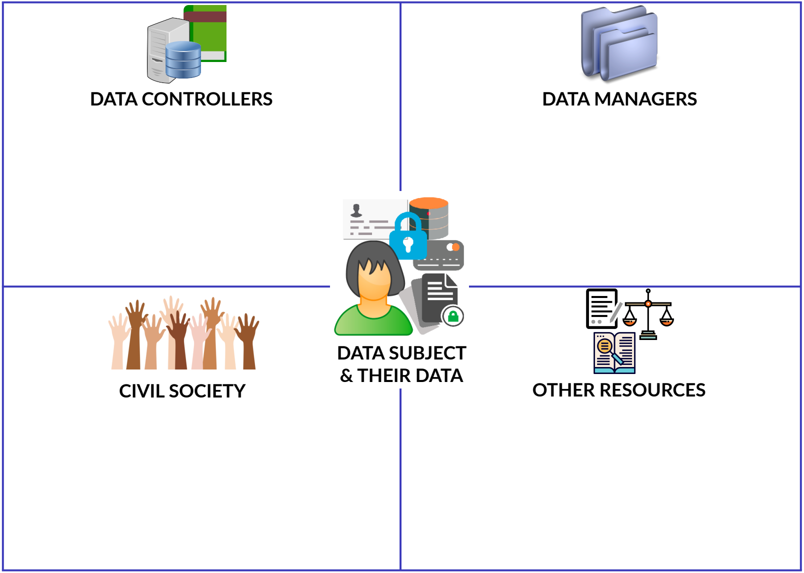
The 2014 Hong Kong Umbrella Movement, where curated data in public domains could have wider ramifications due to socio-political sensitivities (Tromble and Stockmann, 2017).

Photo by Pasu Au Yeung [CC BY (https://creativecommons.org/licenses/by/2.0)]

DATA COMMONS FOR DATA CURATION - ISSUES

- Organising on public platforms could result in crackdown by authorities and government due to its political nature
- Researchers can scrape online platforms to curate data for research without the consent of data subjects
- Data subjects may not be aware of what research has been done based on their public domain data

DATA COMMONS FOR DATA CURATION - MAPPING



DATA CONTROLLERS

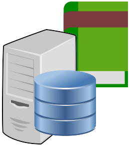
DATA MANAGERS

DATA SUBJECT & THEIR DATA

CIVIL SOCIETY

OTHER RESOURCES

DATA COMMONS FOR DATA CURATION - MAPPING



DATA CONTROLLERS

- Data controllers available
- Data controller policies on data, research, and archiving for data curation
- If, for example, Twitter suffered from a data breach, it can be addressed by automatically supporting data subjects in exercising the rights to erasure and sending notifications to data controllers to request their data be removed

DATA SUBJECT & THEIR DATA

DATA COMMONS FOR DATA CURATION - MAPPING

DATA MANAGERS

- Allow data subjects to identify the most applicable data commons by keywords used by data managers' data and metadata
- Provide further advice on how data subject personal data can be best protected through existing data management mechanisms

DATA SUBJECT & THEIR DATA

DATA COMMONS FOR DATA CURATION - MAPPING

CIVIL SOCIETY

- Expert reports from civil society on the rights of individuals and past problems in data curation
- Class actions in place for collective action on any issues that may arise as a result of data protection issues
- Information on relevant cases pertaining to other data subjects' or organisations' actions and contact details for support and redress

DATA SUBJECT
& THEIR DATA

DATA COMMONS FOR DATA CURATION - MAPPING

OTHER RESOURCES

- Other data subject’s experiences and outcomes from exercising their data subject and information rights
- Recent news and scandals on data controllers
- Update the data commons with any new regulatory changes, policy amendment and calls for consultations on data curation
- Relevant researcher findings from their work

DATA SUBJECT & THEIR DATA

DATA COMMONS FOR DATA CURATION - MAPPING

DATA CONTROLLERS

- Data controllers available
- Data controller policies on data, research, and archiving for data curation
- If, for example, Twitter suffered from a data breach, it can be addressed by automatically supporting data subjects in exercising the rights to erasure and sending notifications to data controllers to request their data be removed

DATA MANAGERS

- Allow data subjects to identify the most applicable data commons by keywords used by data managers' data and metadata
- Provide further advice on how data subject personal data can be best protected through existing data management mechanisms

DATA SUBJECT & THEIR DATA

CIVIL SOCIETY

- Expert reports from civil society on the rights of individuals and past problems in data curation
- Class actions in place for collective action on any issues that may arise as a result of data protection issues
- Information on relevant cases pertaining to other data subjects' or organisations' actions and contact details for support and redress

OTHER RESOURCES

- Other data subject's experiences and outcomes from exercising their data subject and information rights
- Recent news and scandals on data controllers
- Update the data commons with any new regulatory changes, policy amendment and calls for consultations on data curation
- Relevant researcher findings from their work

WHAT NEXT?

- Interviewing experts and participants of existing commons frameworks on identifying and understanding challenges in the commons development process
    - If this is you, please get in touch!
- Creating a prototype of the data protection-focused data commons based on expert and participant experience
- Testing the prototype with data subjects to measure how data subjects find the data commons in terms of is use and whether they find a participatory and co-creation data protection process useful

REFERENCES

- Ostrom, E. (1990). Governing the Commons: The Evolution of Institutions for Collective Action. Cambridge, UK: Cambridge University Press.
- The Irish Times (2019). Is Ireland breaching EU rules by underfunding data regulator? Accessed from https://www.irishtimes.com/business/technology/is-ireland-breaching-eu-rules-by-underfunding-data-regulator-1.4047897, 15 February 2020.
- The Observer (2018). Revealed: 50 million Facebook profiles harvested for Cambridge Analytica in major data breach https://www.theguardian.com/news/2018/mar/17/cambridge-analytica-facebook-influence-us-election, accessed 15 February 2020.
- Tromble, R. and Stockmann, D. (2017). Lost Umbrellas: Bias and the Right to Be Forgotten in Social Media Research. In *Internet Research Ethics for the Social Age: New Challenges, Cases, and Contexts*. 75-94. Oxford, UK: Peter Lang. doi:10.3726/b11077
- Veale, M., Binns, R., and Ausloos, J. (2018). When Data Protection by Design and Data Subject Rights Clash. *International Data Privacy Law* 8 (2). 105-123. doi:10.1093/idpl/ipy002

CONCLUSIONS

- Data sharing must be carefully managed and regulated to prevent data misuse
- Data protection law, legal frameworks, and technological solutions tend to focus on controller responsibilities as opposed to protecting data subjects from the beginning of the data collection process
- Using data curation as a use case, we propose that a co-created data commons can protect individual autonomy over personal data through collective curation and rebalance power between data subjects and controllers

🏠	janiswong.org	tnhh.org
✉	jccw@st-andrews.ac.uk	tnhh@st-andrews.ac.uk
🐦	@janiswong_	@tnhh