



THE UNIVERSITY OF
MELBOURNE

Data Governance

Marc Cheong

—
School of Computing and Information Systems
Centre for AI & Digital Ethics
The University of Melbourne
marc.cheong [at] unimelb.edu.au





Learning Outcomes

1. Define the concept of data governance from an organisational perspective.
2. Understand how laws such as the famous GDPR (EU) help protect users.
3. Discover how your data is stored and managed by the organisations you entrust them with (and also organisations you never knew had access to), as well as your rights.
4. Critically evaluate research and storage practices on large amounts of data.



Related Reading

This module on Data Governance has two readings, with very interesting (and relevant) contemporary studies which encompass data science, policy, and ethics.

Data and its (dis)contents: A survey of dataset development and use in machine learning research.

Amandalynne Paullada, Inioluwa Deborah Raji, Emily M. Bender, Emily Denton, Alex Hanna. *arXiv [cs.LG]*, 9 Dec. 2020.

<https://arxiv.org/abs/2012.05345>

What if Facebook goes down? Ethical and legal considerations for the demise of big tech.

Carl Öhman, Nikita Aggarwal. *Internet Policy Review* 9(3), 1-21, 2020.

<https://policyreview.info/articles/analysis/what-if-facebook-goes-down-ethical-and-legal-considerations-demise-big-tech>



Outline

1. What is data governance? The ‘textbook definition’.
2. Data governance in practice. The ‘practitioner’s view’.
3. From organisations to users.
4. Laws that protect us (GDPR at a glance).
5. Organisations: *‘A Tale of Beacons and Pixels’*.
6. Users: Your rights, freedoms, and countermeasures.
7.  *Case Study & Reflection: Does your data outlive you/the tech giants?*
8.  *Case Study & Reflection: A ‘set’ of data versus a dataset.*



THE UNIVERSITY OF
MELBOURNE



What is data governance?

The ‘textbook definition’.

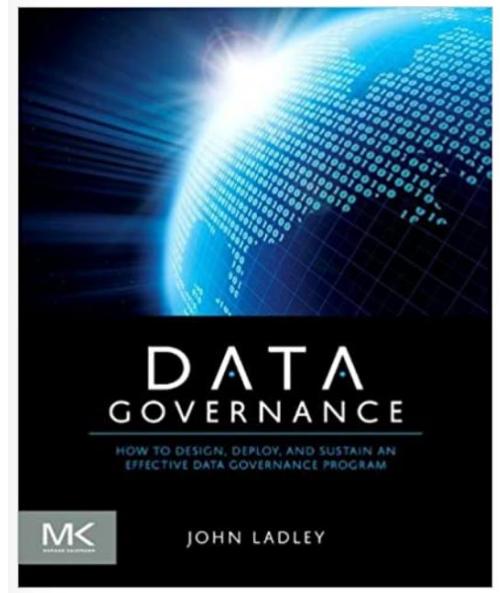
(Let’s pause for a moment to reflect before proceeding)



The ‘textbook’ definition #1

To appreciate the various ethical issues related to data governance, we first need to find out what it means!

"Data governance represents the program used by [an organisation] to manage the organi[s]ational bodies, policies, principles, and quality that will ensure access to accurate and risk-free data and information..."



- John Ladley (2012). *Data Governance: How to Design, Deploy and Sustain an Effective Data Governance Program (The Morgan Kaufmann Series on Business Intelligence) 1st Edition*. Morgan Kaufmann.



The ‘textbook’ (industry) definition #2

“The Data Governance Institute defines data governance as "a system of decision rights and accountabilities for information-related processes, executed according to agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods.” ...

“The Data Management Association (DAMA) International defines data governance as the "planning, oversight, and control over management of data and the use of data and data-related sources." ...
– Olavsrud (2020)

Image source: CIO Magazine

The screenshot shows a web page from CIO Australia. At the top, there's a navigation bar with the CIO logo, 'AUSTRALIA' (with a dropdown arrow), and links for 'LEADERSHIP', 'THE CIO SHOW', 'CIO SUMMIT', 'CIO50 AWARDS', 'INSIDER' (in green), a user icon, and a search icon. Below the navigation, the page title is 'What is data governance? A best practices framework for managing data assets'. It includes a sub-section 'FEATURE', a breadcrumb trail 'Home > Data Management', and a short description: 'Data governance defines roles, responsibilities, and processes for ensuring accountability for and ownership of data assets across the enterprise.' Below this is a author bio for 'Thor Olavsrud', a 'Senior Writer, CIO | 11 FEBRUARY 2020 22:00 AEDT', accompanied by a small circular profile picture.

Reflection.

Formal / system / process / model

Manage / management / oversight / control / planning

Rights / accountability

Who / what / where

Accurate / risk-free

DATA.



Image source: HowToGeek / Imaggentle/Shutterstock



THE UNIVERSITY OF
MELBOURNE



Data governance in practice. *The ‘practitioner’s view’.*



What tasks are there in data governance?

Let's see what a Data Governance Officer does at UniMelb

<https://secure.dc2.pageuppeople.com/apply/TransferRichTextFile.ashx?sData=Fwg6i4Eli-DgfNqV-PbYvXeEOtAs0yXwPaCK7AQYFAk1mpHE7B7I4z2iK7fPjnJtkHDeTmWnOXw~>

Key Responsibilities

- ▶ Define and document a data governance framework (people, process and tools) and implementation plans.
- ▶ Contribute to the on-going implementation of the data governance framework across the University.
- ▶ Provide consultation on governance and data management practices across the stakeholder community within the University.
- ▶ Develop data management processes in accordance with the data governance framework.



What tasks are there in data governance?

Let's see what a Data Governance Officer does at UniMelb

<https://secure.dc2.pageuppeople.com/apply/TransferRichTextFile.ashx?sData=Fwg6i4Eli-DgfNqV-PbYvXeEOtAs0yXwPaCK7AQYFAk1mpHE7B7I4z2iK7fPjnJtkHDeTmWnOXw~>

- ▶ Drive the implementation of data management processes across the University e.g. data stewardship.
- ▶ Monitor and review data assets and procedural controls in order to continually improve data management practices.
- ▶ Report on implementation of data governance framework and compliance across the University.

What tasks are there in data governance?

Let's see what a Data Governance Analyst does at UniMelb

<https://secure.dc2.pageuppeople.com/apply/TransferRichTextFile.ashx?sData=Fwg6i4Eli-DgfNqV-PbYvXeEOtAs0yXwPaCK7AQYFAnYbtT4zcnl02qpL-vZAUoPAQNdUPrhhDI~>

3. *Key Responsibilities*

- ▶ Maintain the enterprise business glossary. This will involve reviewing business terms submitted by project and BAU teams, cleaning up definitions to improve clarity and seeking agreement from key stakeholders across the University on enterprise terms.
- ▶ Work with project and BAU teams to promote the standardisation of business terms.
- ▶ Define and document guidelines for input of terms into the business glossary.
- ▶ Contribute to the on-going implementation of the data management strategy across the University.
- ▶ Provide consultation on data management practices across the stakeholder community within the University.



What tasks are there in data governance?

Let's see what a Data Governance Analyst does at UniMelb

<https://secure.dc2.pageuppeople.com/apply/TransferRichTextFile.ashx?sData=Fwg6i4Eli-DgfNqV-PbYvXeEOtAs0yXwPaCK7AQYFAnYbtT4zcnl02qpL-vZAUoPAQNdUPrhDI~>

- ▶ Contribute to the development of data management processes in accordance with the data governance framework.
- ▶ Contribute to the monitoring and review of data assets and procedural controls in order to continually improve data management practices.
- ▶ Develop supporting controls and processes to implement the data management strategy e.g. business glossary, metadata management, approval workflows etc.



Reflection.

Recall:

Formal / system / process / model

Manage / management / oversight / control / planning

Rights / accountability

Who / what / where

Accurate / risk-free

DATA.

By this stage we should have a pretty clear picture of what data governance is, and what tasks are required in this space, and how an organisation like UniMelb implements this (to protect our data etc).



THE UNIVERSITY OF
MELBOURNE



From
organisations to
users.

USers

You might have noticed the pun there – we are users of many services run by many different organisations.

Social media?

- Twitter/Facebook/TikTok...

Search giants?

- Google...

Device/platform manufacturers?

- Apple/Google...



There is no cloud
it's just someone else's computer

Also, we may be stakeholders even without consciously knowing!

- You're not an Amazon customer, but you use the Covidsafe app (which uses Amazon cloud storage!)



Our relationship with tech orgs. (us ❤️ tech?)

Point to ponder: is this relationship symmetrical?

- Consider, say, your relationship with a small business vendor selling coffee, and paying cash. The data you share with the vendor is just a coffee loyalty card on paper, your usual orders, etc. The vendor might use the data for e.g. how many total loyalty cards were used in a week to predict sales, your usual orders to personalise their service to you/thank you for being a loyal customer, etc.

Big tech knows more about users.

- Consider the Target case study →
- **Asymmetric?**



The image shows a snippet from a Forbes article. At the top right, the word "Forbes" is written in white. Below it, the date "Feb 16, 2012, 11:02am EST" is shown in a smaller font. The main title of the article, "How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did", is displayed in large, bold, dark text. Below the title, there is a small circular profile picture of a woman with short brown hair, identified as "Kashmir Hill Former Staff Tech". A subtitle at the bottom reads, "Welcome to The Not-So Private Parts where technology & privacy collide".

Image source: Forbes



Our relationship with tech orgs. (us ❤️ tech!)

Some scholars call for a ‘duty’ (a la Immanuel Kant’s philosophy) for us to thwart microtargeting.

JOURNAL OF MEDIA ETHICS
2018, VOL. 33, NO. 3, 133–148
<https://doi.org/10.1080/23736992.2018.1477047>



Social networks, the 2016 US presidential election, and Kantian ethics: applying the categorical imperative to Cambridge Analytica's behavioral microtargeting

Ken Ward

Department of Communication, Lamar University, Beaumont, Texas, USA

Reflection: Are we truly helpless as we constantly depend on big tech to stay connected, conduct business transactions, learn and study, especially during the 2019-2021 Covid pandemic?



THE UNIVERSITY OF
MELBOURNE



Laws that protect us. *(& GDPR at a glance).*

Disclaimer: I am not a lawyer



The information provided in this mini-lecture is summarized from various sources to explain how laws such as the GDPR and local laws protect users (at a glance).

This lecture won't make you an expert in the GDPR ☺



Remember these?

These GDPR banners seem to have popped up everywhere after mid-2018...

Cookies & Privacy

This website uses cookies to ensure you get the best experience on our website.

[More information](#) [Accept](#)

THIS WEBSITE USES COOKIES

We use cookies to personalise content and ads, to provide social media features and to analyse our traffic. We also share information about your use of our site with our social media, advertising and analytics partners who may combine it with other information that you've provided to them or that's collected from your use of their services

Necessary cookies Allow selection Allow all cookies

ary Preferences Statistics Marketing Show details ▾

Image source: Cookiebot / JQueryScript



The EU General Data Protection Regulation (GDPR)

Read more here:

<https://gdpr.eu/what-is-gdpr/>

The cookie popups?

“The data subject gave you specific, **unambiguous consent** to process the data.”

-- summary by Proton Technologies AG (2021)

Also: EU ePrivacy Directive
(<https://www.cookiebot.com/en/cookie-law/>)

The screenshot shows the GDPR.EU website with a navigation bar at the top featuring the logo, a European Union flag, and links for Home, Checklist, FAQ, GDPR, and News & Upd. The main content area is titled "Data protection principles" and contains text about the seven principles and a numbered list from 1 to 7.

If you process data, you have to do so according to seven protection and accountability principles outlined in [Article 5.1-2](#):

1. **Lawfulness, fairness and transparency** — Processing must be lawful, fair, and transparent to the data subject.
2. **Purpose limitation** — You must process data for the legitimate purposes specified explicitly to the data subject when you collected it.
3. **Data minimization** — You should collect and process only as much data as absolutely necessary for the purposes specified.
4. **Accuracy** — You must keep personal data accurate and up to date.
5. **Storage limitation** — You may only store personally identifying data for as long as necessary for the specified purpose.
6. **Integrity and confidentiality** — Processing must be done in such a way as to ensure appropriate security, integrity, and confidentiality (e.g. by using encryption).
7. **Accountability** — The data controller is responsible for being able to demonstrate GDPR compliance with all of these principles.



The EU General Data Protection Regulation (GDPR)

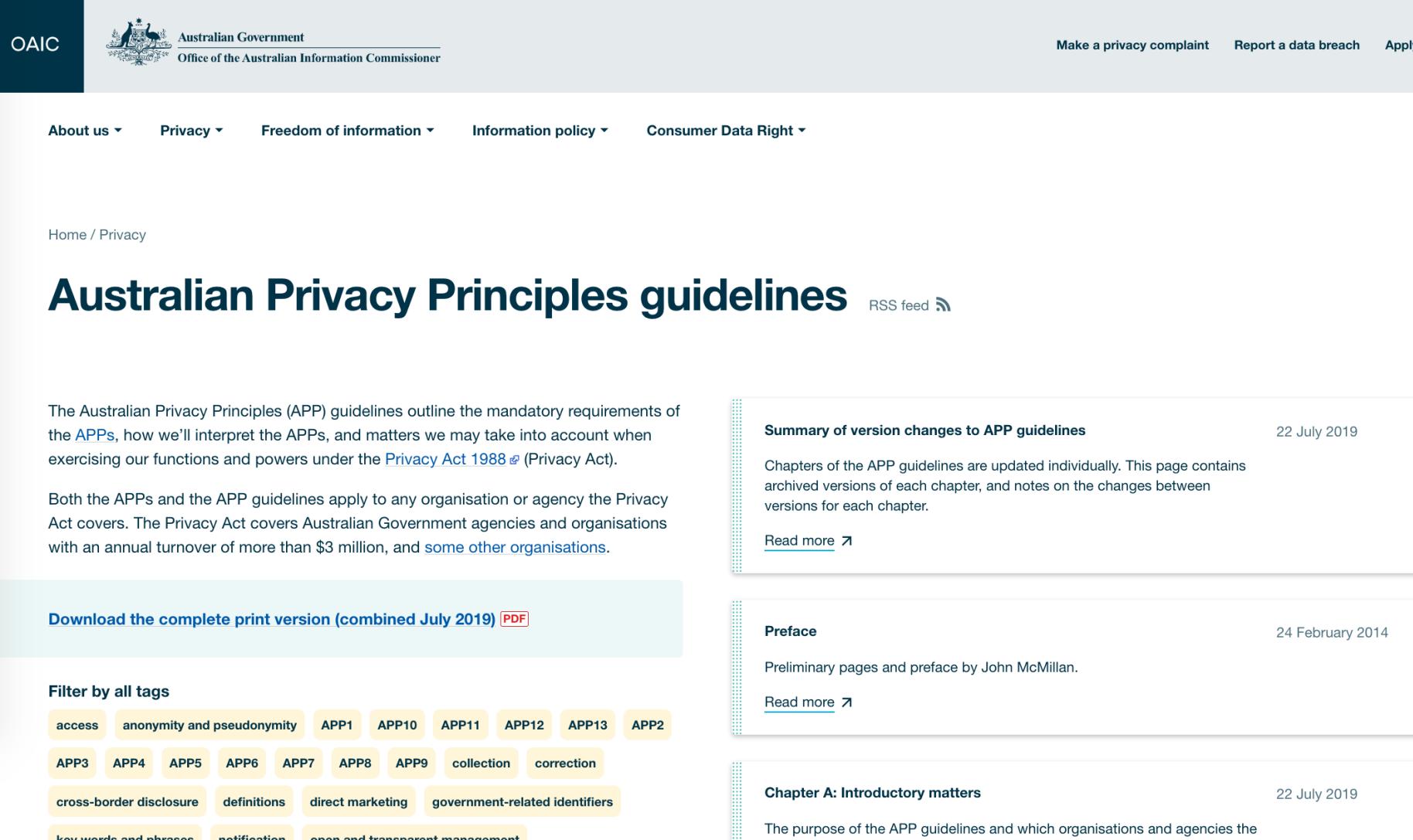
Amongst other things, any “organization that processes the personal data of people in the EU must comply with the GDPR...” (<https://gdpr.eu/faq/>)

(all quotes below are from <https://gdpr.eu/checklist/>)

- “data protection is something you now have to consider whenever you do anything with other people's personal data”, not just an afterthought
- “People have the right to see what personal data you have about them and how you're using it”
Though in practice, this is tricky...
- “You are also required to quickly communicate data breaches to your data subjects” and not just keep quiet about it!
- *NB: A ‘right’ to explainability – this is the subject of many scholarly works in legal studies and digital ethics. (Our very own Prof Tim Miller will be able to provide the technical details about what it means to have explainable AI)*

In Australia...

[https://www.oaic.gov.au/privacy/
australian-privacy-principles-guidelines/](https://www.oaic.gov.au/privacy/australian-privacy-principles-guidelines/)



The screenshot shows the homepage of the Office of the Australian Information Commissioner (OAIC). The top navigation bar includes links for 'About us', 'Privacy', 'Freedom of information', 'Information policy', and 'Consumer Data Right'. Below this, a secondary navigation bar includes 'Home / Privacy' and an 'RSS feed' icon. The main content area features a large heading 'Australian Privacy Principles guidelines' with a sub-section 'Summary of version changes to APP guidelines' dated 22 July 2019. Other sections visible include 'Preface' dated 24 February 2014 and 'Chapter A: Introductory matters' dated 22 July 2019. At the bottom left, there's a 'Download the complete print version (combined July 2019)' link and a 'Filter by all tags' section with various category buttons like 'access', 'anonymity and pseudonymity', 'APP1', etc.

OAIC Australian Government Office of the Australian Information Commissioner

About us ▾ Privacy ▾ Freedom of information ▾ Information policy ▾ Consumer Data Right ▾

Home / Privacy RSS feed

Australian Privacy Principles guidelines

The Australian Privacy Principles (APP) guidelines outline the mandatory requirements of the APPs, how we'll interpret the APPs, and matters we may take into account when exercising our functions and powers under the [Privacy Act 1988](#) (Privacy Act).

Both the APPs and the APP guidelines apply to any organisation or agency the Privacy Act covers. The Privacy Act covers Australian Government agencies and organisations with an annual turnover of more than \$3 million, and [some other organisations](#).

[Download the complete print version \(combined July 2019\)](#) [PDF](#)

Filter by all tags

access anonymity and pseudonymity APP1 APP10 APP11 APP12 APP13 APP2 APP3 APP4 APP5 APP6 APP7 APP8 APP9 collection correction cross-border disclosure definitions direct marketing government-related identifiers key words and phrases notification open and transparent management

Summary of version changes to APP guidelines 22 July 2019

Chapters of the APP guidelines are updated individually. This page contains archived versions of each chapter, and notes on the changes between versions for each chapter.

[Read more](#)

Preface 24 February 2014

Preliminary pages and preface by John McMillan.

[Read more](#)

Chapter A: Introductory matters 22 July 2019

The purpose of the APP guidelines and which organisations and agencies the

Reflection.

Laws protect us users.



Companies cannot abuse our trust and/or data.

We still need to be aware about how our data is used; gain an awareness of the issues involving data use/reuse; and find out how we can protect ourselves and others.

If a social media or search giant wants to profile you to monetize your data,
time for you to voice your “Objection”
(insert Phoenix Wright Ace Attorney meme here)

Image source: CAPCOM / Cinemablend.



THE UNIVERSITY OF
MELBOURNE



Organisations,
and

*'A tale of Beacons
and Pixels'*



Of Beacons and Pixels

Lane v. Facebook, Inc.

From Wikipedia, the free encyclopedia

Lane v. Facebook was a [class-action lawsuit](#) in the [United States District Court for the Northern District of California](#) regarding [internet privacy](#) and [social media](#).^[1] In December 2007, [Facebook](#) launched [Beacon](#), which resulted in users' private information being posted on Facebook without the users' consent. Facebook ended up terminating the Beacon program and created a \$9.5 million fund for privacy and security. There was no money awarded to Facebook users affected negatively by the Beacon program.



A screenshot of a dark-themed website page. At the top left is a button labeled "About Facebook Pixel". At the top right is a search bar with the placeholder text "Get help, tips, and more...". Below the search bar is a blurred background image of a person's face.

The Facebook pixel is an analytics tool that allows you to measure the effectiveness of your advertising by understanding the actions people take on your website.

You can use the pixel to:

- **Make sure your ads are shown to the right people.** Find new customers, or people who have visited a specific page or taken a desired action on your website.
- **Drive more sales.** Set up automatic bidding to reach people who are more likely to take an action you care about, like making a purchase.
- **Measure the results of your ads.** Better understand the impact of your ads by measuring what happens when people see them.

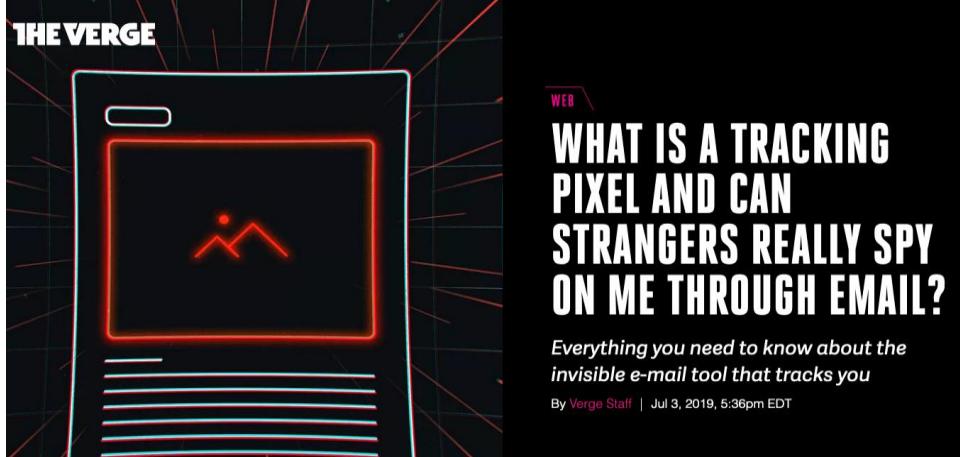
Topics



Image sources:
Wikipedia / Facebook

Implications

Your web-browsing life ‘off’ social media could potentially be picked up and used for ads, targeting, etc (and vice versa).



Facebook is not the only one.

Consider a shopping website and a news website using a tracking pixel system XYZ.

From Wikipedia (https://en.wikipedia.org/wiki/Web_beacon)

- “The image could be as small as a single pixel... (thus the name “tracking pixel”).
- “When a user opens the page or email where such an image was embedded, they might not see the image, but their web browser ... would...
- “...typically includes its IP address... existence of cookies previously sent by the host server ... [which can] associate it with a session identifier or tracking token that uniquely marks the interaction.”

Image source: The Verge



Privacy-consciousness-as-business-model

Support
mozilla

Get Help Volunteer

Find help...

Enhanced Tracking Protection in Firefox for desktop

Enhanced Tracking Protection in Firefox automatically protects your privacy while you browse. It blocks many of the trackers that follow you around online to collect information about your browsing habits and interests. It also includes protections against harmful scripts, such as malware that drains your battery.

Image source: Firefox (<https://support.mozilla.org/en-US/kb/enhanced-tracking-protection-firefox-desktop>)
Apple (<https://support.apple.com/en-au/guide/safari/ibrw35004465/14.0/mac/11.0>)

Safari User Guide

Communities Cont...

Select version:

macOS Big Sur 11.0

Search the user guide

Table of Contents

See who was blocked from tracking you in Safari on Mac

You can see who's been blocked from tracking you when you visit a website.

1. In the Safari app on your Mac, enter the page's name or URL in the [Smart Search field](#).
2. Click the Privacy Report button in the toolbar.





THE UNIVERSITY OF
MELBOURNE



Users: *Your rights, freedoms, and countermeasures.*

Our ethical duty?

We have a right to privacy!

- With events such as the Cambridge Analytica scandal (using Facebook data) reminding us...
- ... and initiatives by Firefox/Apple to encourage us.

From Ward (2018)'s claim that we have a duty to avoid microtargeting

- I might add, phenomena such as algorithmic personalisation (leading to Pariser's 2011 *filter bubble* phenomenon) has shown us that it's not just *us*, but those in our network are also at stake.

Some philosophy:

- We are *free* only if we recognise our responsibilities to others in promoting their *freedom to pursue their own projects*.
– Simone de Beauvoir (*Pyrrhus and Cinéas*, 1944).

JOURNAL OF MEDIA ETHICS
2018, VOL. 33, NO. 3, 133–148
<https://doi.org/10.1080/23736992.2018.1477047>

Routledge
Taylor & Francis Group

 Check for updates

Social networks, the 2016 US presidential election, and Kantian ethics: applying the categorical imperative to Cambridge Analytica's behavioral microtargeting

Ken Ward 

Department of Communication, Lamar University, Beaumont, Texas, USA



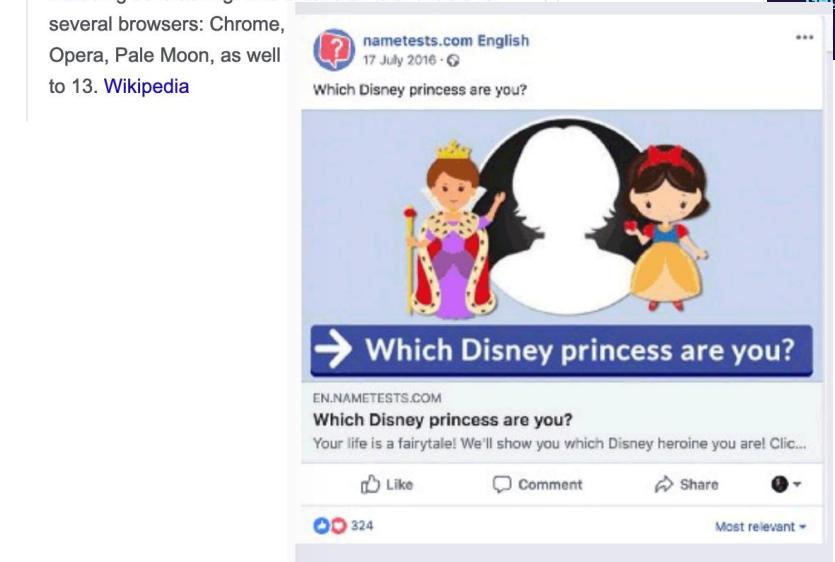


Find the odd one out.



uBlock Origin

uBlock Origin is a free and open-source, cross-platform browser extension for content-filtering, including ad-blocking. The extension is available for several browsers: Chrome, Opera, Pale Moon, as well to 13. [Wikipedia](#)



EN.NAMETESTS.COM Which Disney princess are you? Your life is a fairytale! We'll show you which Disney heroine you are! Clic...

Like Comment Share

324

Most relevant

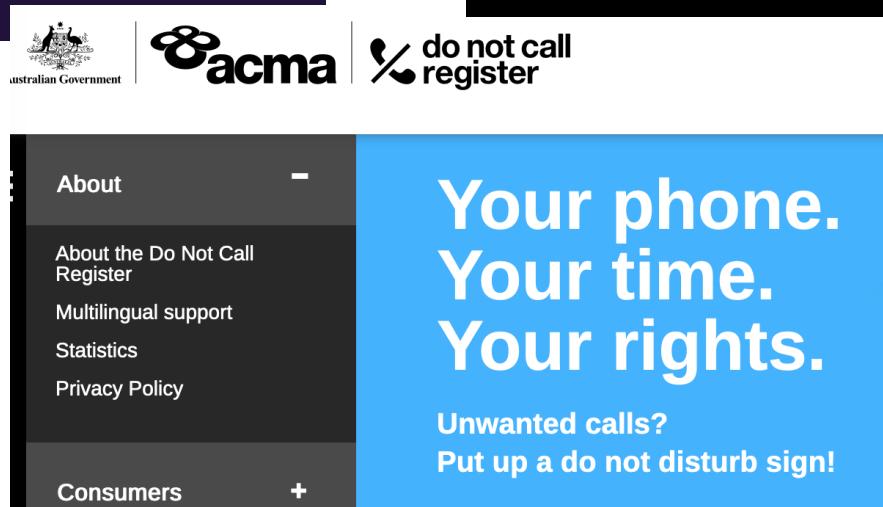


Image sources: Wikipedia via Google infobox / Mozilla / ACMA / TheHackerNews / Innersloth



THE UNIVERSITY OF
MELBOURNE



💡 Case Study: *Does your data outlive you/the tech giants?*

Kind note – a brief content warning:
One of the legal case studies within discusses
the topic of death.



Reading: Öhman & Aggarwal (2020)



INTERNET POLICY REVIEW
Journal on internet regulation

Volume 9 | Issue 3

What if Facebook goes down? Ethical and legal considerations for the demise of big tech

Carl Öhman

Oxford Internet Institute, University of Oxford, United Kingdom, carl.ohman@oii.ox.ac.uk

Nikita Aggarwal

Faculty of Law, University of Oxford, United Kingdom

Published on 11 Aug 2020 | DOI: 10.14763/2020.3.1488

Case Study: Dominicé & Haux (2020)

Santander Art and Culture Law Review 2/2020 (6): 251-260
DOI: 10.4467/2450050XSNR.20.018.13021

LEGAL COMMENTARIES

Antoinette Maget Dominicé*

antoinette.maget@lmu.de
orcid.org/0000-0001-9056-4544
Institute for Art History
LMU Munich
Zentnerstrasse 31
DE 80798 Munich, Germany

Dario Henri Haux*

dario.haux@unilu.ch
orcid.org/0000-0003-1889-6231
Faculty of Law
University of Lucerne
Frohburgstrasse 3
P.O. Box 4466
CH 6002 Lucerne, Switzerland

The Decision of the German Federal Court of Justice against Facebook:
Opportunity to Define Digital Heritage?

LEGAL COMMENTARIES

Antoinette Maget Dominicé and Dario Henri Haux

Abstract: The German Federal Court of Justice recently ruled (27.08.2020 – III ZB 30/20) that Facebook must grant parents direct access to the account of their deceased daughter. At the same time, the parents are prohibited from actively using the account. In this way, the judges established binding standards for the use of social network accounts of deceased users. Beyond inheritance and data protection law, the judgment provides an opportunity to prompt ongoing discussions about sustainable ways of safeguarding, as well as providing access to, digital content. Against the backdrop of a jurisprudence sensitized to the humanities, the two authors encourage a reflection on “spaces”, “containers”, and more generally on the significance of digital media for our everyday lives and future generations.

Keywords: digital inheritance, digital heritage, access, law & humanities, German Federal Court of Justice



Case Study: Dominicé & Haux (2020)

Facts in brief, quoted from the paper:

Role of tech companies?

- Dominice & Haux (2020):
“In response to this [first judgment in the paper]... Facebook sent a USB stick containing a PDF file of more than 14,000 pages to the parents. These pages, which were static content, included unstructured data copied from the account...”
- “Here the question arose whether the rights and obligations set out by the Berlin regional court in 2015 had been met. While Facebook claimed to have met the requirements, the family still felt that the “access” provided was insufficient. Hence, they brought the case back to the Berlin regional court...”
- “Whilst they should not be able to use the Facebook account actively, the parents should nevertheless be able to explore the content on the original platform. 15 In so ruling, the judges made clear that according to their legal evaluation, the obligations related to “access” under the judgment of the regional court of 17 December 2015 had not been fulfilled...”



Reflection.

Role of tech companies?

Who owns your digital life?

As consumers, what can we do with our data?

Open question: What happens when technologies / companies shut down?



THE UNIVERSITY OF
MELBOURNE



💡 Case Study: *A ‘set’ of data versus a dataset.*



Reading: Paullada et al (2020)

Data and its (dis)contents: A survey of dataset development and use in machine learning research

Amandalynne Paullada
Department of Linguistics
University of Washington

Inioluwa Deborah Raji
Mozilla Foundation

Emily M. Bender
Department of Linguistics
University of Washington

Emily Denton
Google Research

Alex Hanna
Google Research

Abstract

Datasets have played a foundational role in the advancement of machine learning research. They form the basis for the models we design and deploy, as well as our primary medium for benchmarking and evaluation. Furthermore, the ways in which we collect, construct and share these datasets inform the kinds of problems the field pursues and the methods explored in algorithm development. However, recent work from a breadth of perspectives has revealed the limitations of predominant practices in dataset collection and use. In this paper, we survey the many concerns raised about the way we collect and use data in machine learning and advocate that a more cautious and thorough understanding of data is necessary to address several of the practical and ethical issues of the field.



Reflection.

We have data of/about ourselves online.

To what extent should researchers use our ‘set of data’ as a dataset? And what are the implications?

- Harms? E.g. Representational?
- Autonomy?
- Consent?
- Use, reuse, and research ethics?
Tying it back to Data Governance.

This case study ties in with many other facets of digital ethics and responsible AI research.

We do not have all the answers – but hope this case study is a starting point for your own reflection.



THE UNIVERSITY OF
MELBOURNE

Thank you

