

Ethical Implications of Data Management

Damian Gordon & Brendan Tierney



1

Workshop Agenda

- | | |
|--|------------|
| • Introduction to Ethics | 10 minutes |
| • What is Ethics | |
| • How is Ethics different to legal/laws. | |
| • Data Management & Ethics | |
| • Multifaceted Approach to Ethics | |
| • Captain America, Ironman and Thor of data ethics | |
| • Ethical Case Studies | 10 minutes |
| • Accessing Data | |
| • Anonymize data – does that really work | |
| • Cloud and data leaks | |
| • ML in the High Street | |
| • Workshop | 20 minutes |
| • Explore different issues on Ethics & Outputs | |
| • How to teach Ethics guidelines | |
| • Workshop Feedback & Key Takeaways | 10 minutes |

2

Ethics

Ethics are principles of behaviour based on ideas of right and wrong. Ethical principles often focus on ideas such as fairness, respect, responsibility, integrity, quality, reliability, transparency, and trust.

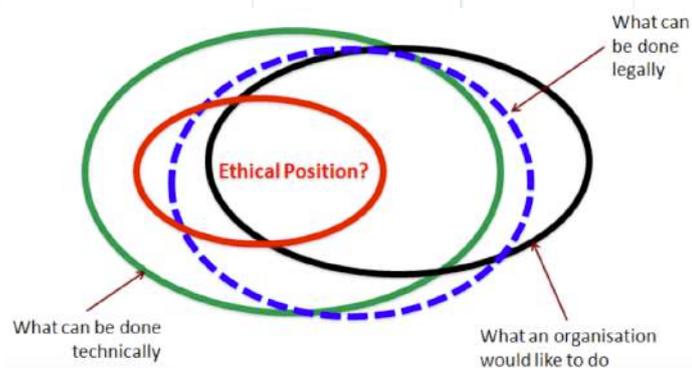
(are subjective rules that a person sets themselves about what is right or wrong)

- Data handling ethics are concerned with how to procure, store, manage, use and dispose of data in ways that are aligned with ethical principles.
- Handling data in an ethical manner is necessary to the long-term success of any organisation that wants to get value from its data.
- Unethical data handling can result in the loss of reputation and customers, because it puts at risk individuals whose data is exposed. In some cases, unethical practices are also illegal.
- Ultimately for data management professionals and the organisations for which they work, data ethics are a matter of social responsibility.

3

Ethical Position

- Many organisations fail to recognise and respond to the ethical obligations inherent in data management. Organisations are using data in ways not previously imagined and while laws codify some ethical principles, legislation cannot keep up to date with the risks associated with this.



4

Ethics vs Legal/Laws

**LEGAL COMPLIANCE
VERSUS
ETHICAL RESPONSIBILITY**

LEGAL COMPLIANCE	ETHICAL RESPONSIBILITY
Acting in accordance with the laws of a particular organization, company, etc.	Choice to comply with the code of ethics of the particular organization, company, etc.
Proactive	Reactive
Involves following the law	Involves doing what is right, sometimes even regardless of the law



Law	Ethics
Formal, written document	Unwritten principles
Interpreted by courts	Interpreted by each individual
Established by legislatures	Presented by philosophers, religious, professional groups
Applicable to everyone	Personal choice
Priority decided by court	Priority determined by individual
Court makes final decision	No external decision maker
Enforceable by police and courts	Limited enforcement

5

Ethics vs Legal/Laws

	Legal	Not Legal
Ethical	It is both legal and ethical to protect privacy when a customer makes online purchases from your website	It is not legal, but could be considered ethical to leak information that appeared on your employer's intranet to the media to stop an illegal activity that is occurring in your company
Not Ethical	It is legal but not considered ethical to call in sick to work when you are not really sick. Or to accept gifts from people in return for awarding a contract	It is neither legal nor ethical to sell e-mail address or your customer details without their permissions

6

Ethics vs Legal/Laws

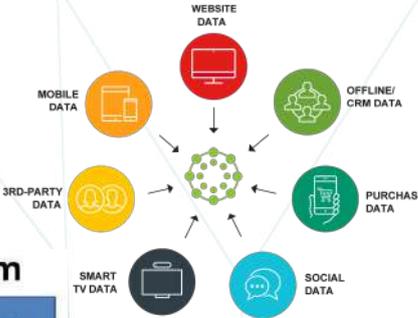




Country	Description
	Brazil's <i>Lei Geral de Proteção de Dados</i> (LGPD) was modeled directly after GDPR and is nearly identical in terms of scope, applicability, and financial penalties for non-compliance. Companies wishing to do business with Latin America's largest economy will have to comply with LGPD by February 2020
	Australia's Privacy Act came into effect in February 2018. Organizations with an annual turnover of over 3 million AUD will have to disclose data breaches that pose a "real threat of serious harm" within 30 days
	There is currently <u>no</u> data privacy law applicable to all industries on the federal level, every state in the Union has their own data privacy laws. California Consumer Privacy Act (CCPA), which has many provisions that overlap with GDPR. New York Privacy Act, 2019, other States to follow
	Japan's Act on Protection of Personal Information was amended in May 2017 and now applies to both foreign and domestic companies that process the data of Japanese citizens. Japan and the EU reached an agreement on "reciprocal adequacy" of their respective data protection laws.
	South Korea's Personal Information Protection Act has been in effect since September of 2011 and from the outset has included many GDPR-like provisions, including requirements for gaining consent, the scope of applicable data, appointment of a Chief Privacy Officer, and limitation and justification of data retention periods.
	Thailand Personal Data Protection Act (PDPA) was published May 2019 and will come into effect exactly a year later on 27 May 2020. The PDPA is similar to GDPR in a number of ways, including the broad definition of personal data, the requirement to establish a legal basis for collection and use of personal data, extraterritorial applicability, and potentially harsh penalties for non-compliance.

7

Data Management



Business Intelligence Ecosystem

Data acquisition

- Operational systems
 - ERP database
 - Operational database
 - CRM database
 - Specialised NoSQL
 - Flat files

Data Warehouse

- Structured data
 - Staging
 - EDW
- In-memory
- Analytics
 - Specialised NoSQL
 - Sandbox
- Semantic layer
 - HCatalog
- Structured access
 - Hive
- Unstructured data
 - Hadoop HDFS
 - MapReduce extracts

Data distribution

- Reports
- Ad-hoc queries
- Dashboards
- Analytical models
- Data exploration
- Downstream applications

8

Ethics in Data Analytics

How Target Figured Out A Teen Girl Was Pregnant Before Her Father Did

Every time you go shopping, you share intimate details about your consumption patterns with retailers. And many of those retailers are studying those details to figure out what you like, what you need, and which coupons are most likely to make you happy. Target, for example, has figured out how to data-mine its way into your womb, to figure out whether you have a baby on the way long before you need to start buying diapers.

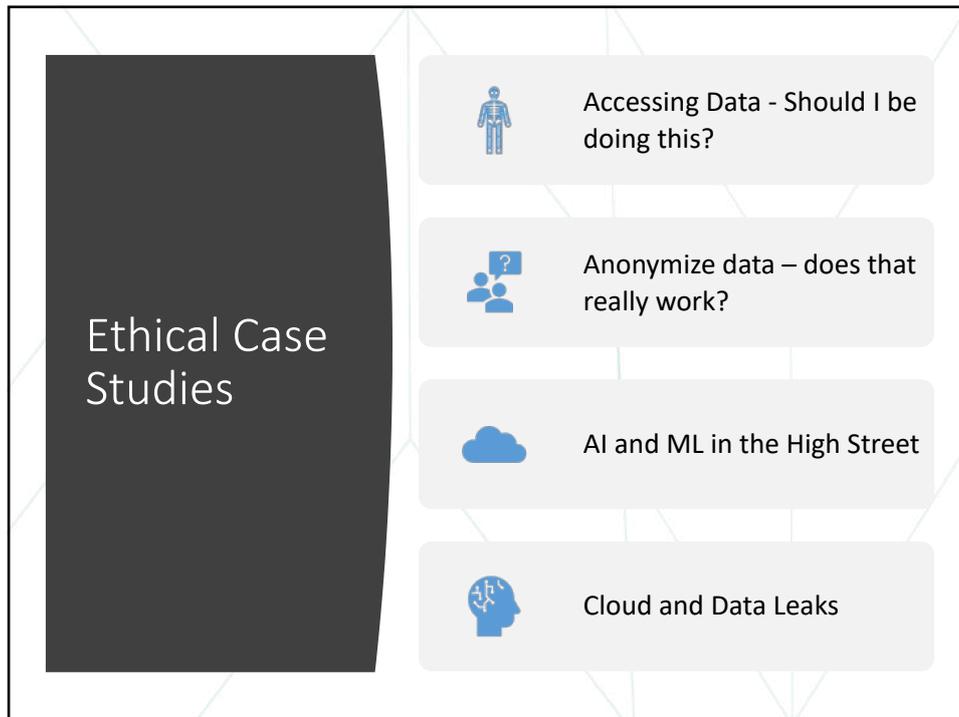
Forbes



9



10



11

Ethical Case Studies – Accessing Data

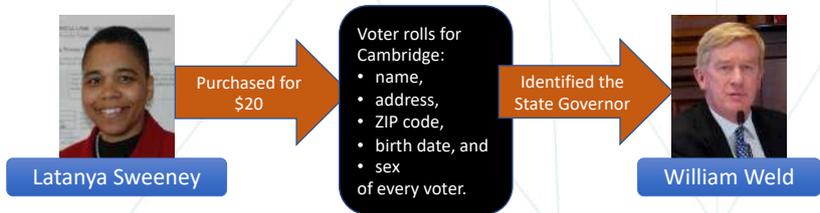
- Have you
 - Seen data you shouldn't have seen?
 - Used someone else login? (computer/network/application)
 - Gone looking for data or other things you know you shouldn't have?
- IT professionals are exposed to this every day
- Software developers are exposed to this every day
 - Development vs Test vs Production

12

Ethical Case Studies - Anonymizing data

- In Massachusetts, a government agency called the Group Insurance Commission (GIC) purchased health insurance for state employees. At some point in the mid-1990s, GIC decided to release records summarizing every state employee's hospital visits at no cost to any researcher who requested them.

They removed: <ul style="list-style-type: none"> • name, • address, • social security number, and • other "explicit identifiers" 	They did not remove: <ul style="list-style-type: none"> • ZIP code, • birth date, and • sex.
---	--



13

Ethical Case Studies - AI & ML in the High Street

TECH INSIDER
Amazon's warehouse-worker tracking system can automatically fire people without a human supervisor's involvement



Artificial Intelligence / Face Recognition

Why San Francisco's ban on face recognition is only the start of a long fight

The city government can't use the technology, but private companies still can, and regulating those uses is a thornier problem.

14

Ethical Case Studies - AI & ML in the High Street

Ads tailored to your age, sex and mood can tell if you're happy

Billboards with built-in cameras and facial detection software are targeting images at shoppers across the country



A billboard in Weymouth in London with built-in cameras and facial detection software.

People's online photos are being used without consent to train face recognition AI

Artificial Intelligence Mar 13



百度人脸识别技术在国际评测2015-2016 FDB 比赛中取得优异成绩



15

Ethical Case Studies - Cloud and Data Leaks

Who?	When?	What?	How?
Microsoft	2010	Non-authorized users were able to access employee info	Configuration issues
DropBox	2012	68 million user accounts	Password Issue
LinkedIn	2012	6 million user passwords	Password Issue
Yahoo	2013	One billion user accounts	"State sponsored attack" !?
Home Depot	2014	At least records of sales of hundred million dollars	Flaw in point-of-sale terminals
National Electoral Institute of Mexico	2016	93 million voter registration records	Poorly configured database
Apple iCloud	On-going?		Password Issue



16



17

Core Concepts

- The ethics of data handling are complex, but they centre on several core concepts:
- **Impact on People:** Because data represents characteristics of individuals and is used to make decisions that affect people's lives, there is an imperative to manage its quality and reliability.
- **Potential for Misuse:** Misusing data can negatively affect people and organisations, so there is an ethical imperative to prevent the misuse of data.
- **Economic Value of Data:** Data has economic value. Ethics of data ownership should determine how that value can be accessed and by whom.

18

Workshop – Groups of 3-4 people

- Think about the ethical issues associated with the Case Studies
- List the ethical issues for each case study, are there any common ethical issues?
- List what can be done to avoid these issues and what are the challenges. (max 5 for each)
 - By the Company
 - By the individual employees
 - By the customers

19

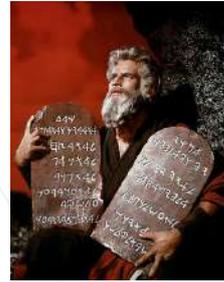
Workshop – Groups of 3-4 people

- List (up to ten) tenets of good ethical behavior, code for good ethical behavior in the workplace.
- List (between 5-8) recommendations for teaching Ethics to (under-graduate and post-graduate) students
- Based on your experiences, what recommendations would you give to new or early career Technology professionals
 - List between 5-8 recommendations
 - These should be different to those listed for teaching Ethics in previous point.

20

Here's one FROM ACM

- You should not use a computer to **harm other people**.
- You should not **interfere** with other people's computer work.
- You should not **snoop** around in other people's computer files.
- You should not use a computer to **steal**.
- You should not use a computer to **bear false witness**.
- You should not **copy or use proprietary software** for which you have not paid without permission.
- You should not use other people's computer resources **without authorization** or proper compensation.
- You should not appropriate other people's **intellectual output**.
- You should think about the **social consequences** of the program you are writing or the system you are designing.
- You should always use a computer in ways that ensure **consideration and respect** for your fellow humans.

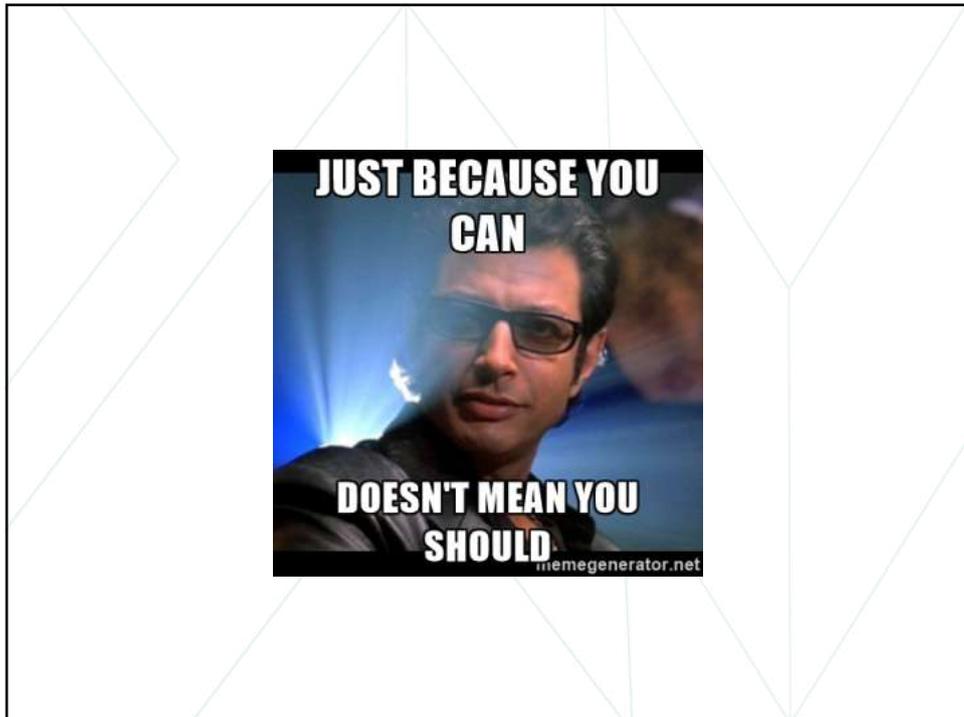


21

Workshop Feedback & Takeaways

- 1
- 2
- 3

22



23



24



25