MONASH University

MONASH
INFORMATION
TECHNOLOGY

Ethics and security

Lei Yang

GROUP OF EIGHT AUSTRALIA

Ethics

Introduction

The field of big data ethics itself is defined as outlining, defending and recommending concepts of **right and wrong practice** when it comes to the use of data, with particular emphasis on personal data. **Big data ethics** aims to create an **ethical and moral code** of conduct for data use.

Reference:
https://blog.hurree.co/the-ethics-of-big-data#:~:text=The%20field%20of%20big%20data%20ethics%20itself%20is%20defined%20as,of%20conduct%20for%20data%20use

Five main areas of ethics

Informed Consent

Privacy

Ownership

Algorithm bias and objectivity

Big data divide

Informed Consent

To consent means that you give uncoerced permission for something to happen to you. Informed consent is the most careful, respectful and ethical form of consent. It requires the data collector to make a significant effort to give participants a reasonable and accurate understanding of how their data will be used.

Privacy

The ethics of privacy involve many different concepts such as liberty, autonomy, security, and in a more modern sense, data protection and data exposure.

You can understand the concept of big data privacy by breaking it down into three categories:

The condition of privacy
The right to privacy
The loss of privacy and invasion

- Names
- Phone numbers
- Email addresses
- Profile descriptions
- Follower and engagement data
- Locations
- LinkedIn profile links
- Connected social media account login names

Ownership

When we talk about ownership in big data terms, we steer away from the traditional or legal understanding of the word as the exclusive right to use, possess, and dispose of property. Rather, in this context, ownership refers to the **redistribution of data, the modification of data, and the ability to benefit from data innovations**.

Algorithm bias and objectivity

Algorithms are designed by humans, the data sets they study are **selected and prepared** by humans, and humans have bias.

Big data divide

The big data divide seeks to define the current state of data access; the understanding and mining capabilities of big data is isolated within the hands of **a few major corporations**. These divides create ‘haves’ and ‘have nots’ in big data and exclude those who lack the necessary financial, educational and technological resources to access and analyse big datasets.

security

Introduction

Big data security is the process of monitoring and protecting a company's important business data with the goal of ensuing safe and compliant ongoing operation.

Ref: https://www.datamation.com/big-data/big-data-security/

How does big data security work

Stage 1: Data Sources

Big data sources come from a variety of sources and data types. You need to secure this data in transit, from sources to the platform.

Stage 2: Stored Data.

Protecting stored data takes mature security toolsets including encryption at rest, strong user authentication, and intrusion protection and planning.

Stage 3: Output Data.

These analytics output results to applications, reports, and dashboards. This extremely valuable intelligence makes for a rich target for intrusion, and it is critical to encrypt output as well as ingress

Big Data Security Technologies

Encryption

Centralized Key Management

User Access Control

Intrusion Detection and Prevention

Physical Security