

James Perez. H

Writing for engineers: FA22

Topic: Big Data.

Target Audience: Average consumers.

Style: Technical report.

A brief yet thorough look into Big Data:

What is it? How does it affect the average consumer? And what can be done in order to protect yourself from these practices?

Introduction:

The main aim of this paper is to inform the reader, those being average citizens, about the term Big Data. This will be done by looking through a different piece of literature pertaining to the subject and involving a real-world case that had occurred regarding Big Data. Regarding the pacing of this report. The report will follow a structure where you will be first given a summary of all the information given and conclusive thoughts. This is for the sake of the time of the reader.

Summary/abstract:

In our current day and age, we are open to various online services that help us in our everyday life. These range from things like Google, Amazon, and social media services like

Facebook. The latter of which we will discuss more. The order will follow the introduction of Big Data, to the potential uses behind it. Lastly to give more information on the Facebook case. It will pertain to the Cambridge Analytica incident, where tens of millions of people's information was stolen and put into a Database that was designed to help a candidate win in the 2016 presidential election.

Body Paragraph

- What is Big data?
 - Big Data is the name given to processes in which, government agencies and private corporations gather extremely large groups of data.
 - These clumps of data are widely accepted to be “so intricate and vast that they cannot be comprehended by humans or analyzed with a matching relational database.”⁴
 - What types of Big Data are there?
 - Structured: Structured data is a much easier form of data to be understood. This is a result of the format used when acquiring this data. Leading to more easily established conclusions as to where they belong due to the unique information gathered.
 - Unstructured: Unlike the previously mentioned form, unstructured data is loose in its meaning. This is a result of not having a direct format to fill out but rather gathering many numerical values that are spread across the entire platform. “Pictures we post on Instagram or Facebook and videos we watch on other platforms are examples of unstructured.”⁴

- Semi-Structured: This form of data falls into more of a tricky position. Seeing as it pertains to both aspects of data being collected. It neither matches fully nor is too far off from similar points of other data.

These services, while incredibly helpful, also follow an agenda where they would collect private users' information for convenience.

- Potential uses of Big Data and how it affects the consumer.
 - This data is helpful for a variety of reasons: Google used this data in order to help complete other people's searches and prevent criminal activities through the use of analyzing people's searches that have keywords included in the search.
 - According to *New Frontiers In High Performance Computing and Big Data*, “The analysis of urban data is a valuable opportunity for development of Smart City application and to achieve improvements in urban policing and management issues {...} with a particular focus on security issues, one of the most challenging goals in big cities is to have accurate (short-term and long-term) crime forecasts in the city.”²
 - This point is brought up when discussing the use of gathering large quantities of data in order to create an algorithm that can determine when crime rates will increase at a given point of the year.

- Corporations such as Amazon used their data in an attempt to recommend products, and services, and even used other data to help give reason to the creation of new inventions.
 - This was further established through the use of amazon and their kindle services. Which would practice forms of price discrimination with their literature services.¹

- Real World implications of the potential misuse of Big Data and how to protect yourself from it.
 - Social media, in general, tends to have more harmful uses of this technology. The reasoning behind this is as follows: collecting data from not only their app or site but by scanning the user's device, selling data to other corporations, creating rather intrusive ads that target the user on a personal level, and recommending things that can be harmful to users.
 - The Cambridge Analytica Incident.
 - Leading up to the 2016 Presidential election, Facebook at had launched an app that was under the name of a “personality quiz”. This small app had the ability to gather information all kinds of information from the user and by extension their friends that also had a Facebook account. The creator of the said app would later sell this information to Cambridge Analytica. According to the *Office of the Attorney General for the District of Columbia*, “ OAG conducted a wide range of depositions with

Facebook's directors, former employees, and whistleblowers, and examined hours of Mr. Zuckerberg's public statements, including sworn testimony before the U.S. Senate and other law enforcement agencies. This evidence confirmed Mr. Zuckerberg's direct oversight of major decisions that led to Cambridge Analytica's {...} mass collection and manipulation {...} misrepresentation to users about the security of their personal information."³ This information would later lead to people being assigned to the likely hood of who they would vote for.

- The attorney general mentions that the best method in order to not allow your information to be exposed is through the following.
 - Use secure websites that have lock icons to verify.
 - Don't click on strange links or embedded links in emails.
 - Beware of random emails from users that you do not recognize.
 - And much more that can be found on their site and through this link provided.

<https://oag.dc.gov/consumer-protection/consumer-alert-online-privacy>

Conclusion:

In conclusion, through many breakthroughs in technology and data analysis methods, we can make sense of this vast amount of data. With that being said we then use this information to aid

us in practical uses from potentially stopping crime to potentially helping us search for something imparticular that someone else had attempted to find. All the sources used can be found in the work-referenced portion of this report.

Recommendations:

Format Style:

<https://templatelab.com/wp-content/uploads/2019/06/technical-report-template-09.jpg>

From: <https://templatelab.com/technical-reports/>

Work referenced:

Colin J. Bennett. *The Privacy Advocates : Resisting the Spread of Surveillance*. The MIT Press, 2008. EBSCOhost,

<https://search-ebSCOhost-com.ccny-proxy1.libr.ccny.cuny.edu/login.aspx?direct=true&db=e000xna&AN=237752&site=ehost-live>.

Fox, Geoffrey, et al. *New Frontiers in High Performance Computing and Big Data*. IOS Press, 2017. EBSCOhost,

<https://search-ebSCOhost-com.ccny-proxy1.libr.ccny.cuny.edu/login.aspx?direct=true&db=e000xna&AN=1632348&site=ehost-live>

AG Racine Sues Mark Zuckerberg for Failing to Protect Millions of Users' Data, Misleading Privacy Practices, Oagpress@Dc.gov, 23 May 2022,³

<https://oag.dc.gov/release/ag-racine-sues-mark-zuckerberg-failing-protect#:~:text=Racine%20today%20sued%20Facebook%20CEO,of%20privacy%20and%20data%20protection>.

Bandakkanavar, Ravi. "All You Need to Know about Big Data." *Krazytech*, 12 Dec. 2021,

<https://krazytech.com/technical-papers/big-data.>⁴

Other resources used:

<https://www.aresearchguide.com/writing-a-technical-report.html>

<https://www.ncvo.org.uk/help-and-guidance/strategy-and-impact/impact-evaluation/evaluation-and-impact-reporting/how-to-write-an-evaluation-report/#>
https://mcmassociates.io/textbook/report_design.html