SPSPA GenAI Student Declaration	
Module Number	**PHLM011**
Module Title	**Data Governance and Ethics**
Assignment Title	**Exploring Big Data: A Personal Reflection and Ethical Analysis of the Uber Data Breach**
Word Count	**3,279**

AI-supported assessment:

AI-supported use is permitted in this assessment. I acknowledge the following uses of GenAI tools in this assessment:

☒ *I have used GenAI tools for brainstorming ideas.*

☒ *I have used GenAI tools to assist with research or gathering information.*

□ *I have used GenAI tools to help me understand key theories and concepts.*

□ *I have used GenAI tools to identify trends and themes as part of my data analysis.*

□ *I have used GenAI tools to suggest a plan or structure for my assessment.*

☒ *I have used GenAI tools to give me feedback on a draft.*

□ *I have used GenAI tools to generate images, figures or diagrams.*

☒ *I have used GenAI tools to proofread and correct grammar or spelling errors.*

□ *I have used GenAI tools to generate citations or references.*

□ *Other [please specify]*

□ *I have not used any GenAI tools in preparing this assessment.*

I declare that I have referenced use of GenAI outputs within my assessment in line with the University referencing guidelines.

If I have used GenAI tools, I have ***kept a record*** *of the tools, prompts and outputs used and can produce these if necessary, at a viva, and demonstrate how I have built on this content to ensure that my work is original.*

Part A: Understanding Big Data and Its Ethical Dimensions

‘Big Data’ is a term used to describe vast, complex datasets that are beyond the capabilities of traditional data-processing techniques. The understanding and impact of Big Data has evolved over time, with many definitions and perspectives emerging across different fields, particularly within data science and business analytics.

A widely accepted model created to explain Big Data is Lacey’s (2001) 3 V’s, which define Big Data by three key characteristics: Volume, Velocity, and Variety. These attributes describe the sheer magnitude of data, the speed at which it is generated needs to be processed, and the diverse types of data it encompasses. Examples of Big Data that exhibit these characteristics include real-time data streams from stock markets or social media platforms, as well as transaction logs and sensor data in sectors like healthcare, finance, and retail. Lacey (2001) argued the importance of controlling these three dimensions for effective data management. Over time, the framework has been expanded, with additional V’s being added such as ‘Veracity’ referring to the quality and trustworthiness of the data and ‘Value’ representing the utility and economic worth of the data.

However, despite its widespread use, the 3V model has been criticised for oversimplifying the complexities of what constitutes ‘Big Data’. Kitchin and McArdle (2016) state that not all datasets that meet the characteristics mentioned are necessarily ‘Big Data’ in the conventional sense. Their research highlights that while many datasets may meet the V’s, only a select few fully encompass the breadth of characteristics typically associated with Big Data. They conclude that their key definitional boundary markets are traits of ‘velocity’ and ‘exhaustivity’. Velocity referring to the speed at which data is generated and processed, highlighting the need for near-instantaneous analysis. Exhaustivity, on the other hand, relating to datasets that strive to capture all relevant data points, aiming for comprehensive coverage without relying on sampling, such datasets that would record every transaction or monitor all sensor outputs in each setting.

One insight that I found particularly thought provoking comes from Clive Humby (2006), who famously likened data to crude oil at a senior marketing summit at Kellogg School. He argued that, like oil, data is a raw commodity that requires refinement and processing to unlock its true value. This analogy stood out to me as it is especially relevant to ‘Big Data’, highlighting the necessity of sorting and analysing vast datasets to inform business decisions and generate actionable insights in real-world contexts.

As the discussion around ‘Big Data’ becomes more nuanced, scholars have debated the mythology of Big Data and what it can achieve. One view that I found particularly interesting surrounding ‘Big Data’ was from Cukier and Mayer-Schoenberger (2013), who argue that the benefits of using vastly more data of variable quality outweigh the costs of using smaller amounts of very exact data. They suggest that, in the age of Big Data, we should embrace the inherent ‘messiness’ of data due to its sheer volume. I found this perspective eye-opening, as it challenges the traditional emphasis of high-quality, well-curated data. In fields such as medicine, healthcare, or urban planning, the idea that ‘more is better’ has become a guiding principle. The value is therefore gained by extracting insights from patterns that emerge from vast pools of data.

This approach resonates with Anderson's (2008) argument that with Big Data, "massive amounts of data and applied mathematics can replace every other tool that might be brought to bear". Anderson further asserts that theory is unnecessary when we have enough data to observe patterns and correlations. While I find this idea compelling as it offers exciting possibilities, it also raises significant concerns and debates.

Initially, I found this view appealing, as it seemed to suggest that we could gain all data-driven insights from purely numbers, bypassing traditional models or hypotheses. However, I believe that Big Data does not exist in a vacuum; it must be interpreted through the lens of theory, context, and human understanding. The idea that the data can 'speak for itself' only applies in certain contexts and overlooks the complex social, cultural, and ethical factors that influence the data we collect and how we interpret it. boyd and Crawford (2012) argue that models and theories of human relationships are needed to put behavioural data into context. If data is collected counting the number of interactions between people, we cannot assume that more contact with a person means that they have a 'strong tie' with one another. Furthermore, data from social media platforms or online behaviours might seem like an objective reflection of society, but these data sources are often skewed by factors such as fake accounts, bots, and the very algorithms that drive the platforms themselves. This can distort the insights we gain from them. Therefore, the context in which insights from Big Data are derived is crucial, and larger datasets are not always preferable in fields where potential errors or methodological concerns may arise.

Reflecting on the critiques of Big Data has significantly influenced my thinking about its ethical implications and societal impact. Cukier and Mayer-Schoenberger (2013) emphasise that the messiness of Big Data is an inherent feature that must be accepted. However, this poses important questions: at what cost do we embrace this messiness? In some specific contexts, such as in identifying consumer behaviour, the 'correlation over causation' approach may work effectively. However, when it comes to fields like healthcare or public policy, relying on correlations without understanding causal mechanisms can be dangerous. For example, identifying correlations in healthcare data, such as the link between certain genetic markers and diseases, might point to potential treatments. But without proper theoretical frameworks, we risk drawing improper conclusions, which could ultimately harm patients and lead to a failure to address the root causes of health issues.

Overall, this has made me more conscious not only of the risks associated with Big Data collection but also of the underlying power dynamics that come with its use. Those who control the data have an immense power in how they display and use it to influence decisions and manipulate behaviours. For instance, in industries like advertising and politics, data is often used to target vulnerable individuals with personalised messages designed to sway opinions or actions. I can now appreciate that Big Data is not just a tool for gaining insights or improving business practice, but also as a tool of control. This shift in perspective has led me to think more critically about privacy and consent, raising questions about how owns the data and how it should be used ethically. Therefore, Big Data presents an ethical dilemma: while it holds the potential to drive progress in areas like business and science, the mythology surrounding its infallibility must be challenged, and it must be governed responsibly to protect individual rights and prevent harm.

Part B: Ethical Analysis of the Uber Data Breach Case Study

The 2016 Uber data breach was a significant cybersecurity and privacy incident that exposed the personal data of around 57 million people worldwide (Chappell, 2018). Hackers gained unauthorised access to Uber's data in October 2016, obtaining sensitive information, including the names, email addresses, and mobile phone numbers of 50 million riders and 7 million drivers, as well as around 600,000 driver's license numbers (Lohrmann, 2017). This breach represented a severe compromise of user privacy which undermines individual's rights to control their personal information (Wirght, n.d.).

However, what made the Uber data breach truly significant was Uber's response to the incident. Instead of following legal protocol and disclosing the breach immediately, Uber chose to conceal it. The company's Chief Security Officer, Joe Sullivan, who was aware of the breach, chose to pay the hackers $100,000 to delete the stolen data in December 2016, without informing regulators or affected users (Chappell, 2018).

The concealment lasted over a year and it wasn't until November 2017 that new CEO Dara Khosrowshahi publicly revealed the breach, admitting that Uber had concealed it for more than a year (Chappell, 2018). As a result, the company faced severe public backlash, legal consequences, and reputational damage. Uber was fined over 148 million US dollars as part of a settlement with 50 U.S. states (Chappell, 2018). Further investigations were launched by regulatory authorities in the UK and Netherlands for Uber's failure to notify affected users and comply with data protection regulations (Somerville, 2018). This case raises crucial concerns regarding Uber's corporate responsibility, particularly in relation to data privacy, transparency, honesty and accountability. The company's choice to prioritise its own reputation over the safeguarding of personal data is a profound ethical issue, highlighting the broader implications of trust in digital platforms and the responsible use of data.

Consequentialism is an ethical framework that evaluates actions based on their outcomes. The most common form of consequentialism is utilitarianism, which seeks to maximise overall happiness and minimise harm. A utilitarian approach within consequentialism would evaluate Uber's actions by weighing the positive and negative consequences of their decision to conceal the breach. Classic utilitarianism, holds that an action is right if it "does the most good" for the greatest number, maximising overall well-being and minimising suffering for everyone affected (MacAskill et al., 2023).

From a consequentialist perspective, Uber's decision to conceal the breach and pay off the hackers was ethically flawed because the long-term consequences caused more harm than good. Initially, Uber may have hoped that by paying the hackers and keeping the breach secret, it could avoid short-term damage to its reputation. However, this strategy ultimately backfired. The most significant consequence was that millions of users were left vulnerable to identity theft, fraud, and other forms of personal data misuse because they were not informed that their data had been compromised (Chappell, 2018). The lack of transparency led to spiralling issues as users were unable to take steps to protect themselves, such as changing passwords or monitoring financial accounts. Therefore, the breach made individuals more susceptible to further harm from third parties.

Uber's concealment of the breach resulted in severe reputational damage for the company. When the breach became public in 2017, the case gravely undermined public trust in the company's commitment to data security. A consequentialist view would argue that only the outcome and consequences of an action matter, when evaluating the morality of an action. If Uber had chosen to admit to their breach immediately, the consequences of transparency could have been far less harmful. While there might have been an initial reputational hit, a timely admission would have allowed Uber to take immediate steps to mitigate the damage, such as notifying affected users, offering them identity protection services, and improving their security protocols (Sandel, 2020). This approach would have minimised long-term damage to customer trust and loyalty, and may have prevented further legal fines and public backlash that followed. Consequentialism argues that an action is good if it maximises people's welfare in the sense of making the majority of people who are affected by the action happier than they were before the action took place (Beaulieu & Leonelli, 2021). In this case, after the breach became public, the total amount of happiness experienced by Uber's customers would have dropped significantly. Therefore, blinded by the short-term repercussions of notifying the breach, Uber failed to prioritise the long-term well-being of its users and the company itself. As a result, Uber's brand was tarnished, and many users lost faith in the company's ability to safeguard their data, which damaged customer loyalty (Lohrmann, 2017).

Uber's cover-up of the breach led to massive fines and legal settlements. These financial consequences were a direct result of Uber's failure to notify affected users and regulatory bodies in a timely manner, which violated protection laws. Additionally, Uber faced regulatory scrutiny, being fined by the UK and Dutch authorities for non-compliance with privacy laws (Somerville, 2018). On top of these consequences, Uber may have emboldened criminal actors to continue exploiting similar vulnerabilities in other organisations. By paying the hackers, the company could have set a dangerous precedent, signalling that ransom payments could be effective in evading consequences for criminal activities. This outcome would cause more harm than good by undermining the deterrent effect that law enforcement could have had on such criminals.

The harmful consequences of Uber's actions including data misuse, loss of trust and legal penalties, clearly outweighed any short-term benefits that the company might have gained from concealing the breach. This analysis supports the argument that full disclosure and transparency are vital in data governance to minimise harm and protect public trust. However, consequentialism can be criticised for its sole focus on outcomes and potential uncertainty about future events. For example, if the breach had never been disclosed, Uber might have avoided the public backlash altogether. In theory, the hackers could have deleted the data, causing minimal long-term harm. Consequentialists would argue that this assumption is highly speculative and that Uber's failure to disclose the breach was reckless because it risked much greater harm in the long-run. Therefore, incorporating multiple ethical frameworks, such as virtue ethics, can help provide a more well-rounded analysis of Uber's actions. While consequentialism focuses primarily on the outcomes, virtue ethics offers insights into the moral character of decision-makers and the importance of virtues such as honesty, integrity, and courage. By combining both frameworks, we can consider not just the consequences but also the character flaws and leadership failures that contributed to the decision-making process. This pluralistic approach highlights the importance of not only achieving the best outcomes but also ensuring that ethical principles guide decision-making at every level of an organisation.

Virtue ethics is an ethical framework that focuses on the moral character of individuals and organisations. A virtue ethicist asks whether a decision or action reflects good character – "Is this what an honest, courageous, or generous person would do?" (MacDonald and Marcoux, 2023). Virtues are positive character traits such as honesty, integrity, courage, compassion, or generosity that society deems morally admirable (Koehn, 2016). Modern application of virtue ethics in organisations emphasise cultivating these virtues in corporate culture and leadership. For instance, ethical leadership research highlights that an effective moral leader must exemplify personal virtues such as honesty and integrity in their conduct (Mayer et al., 2012). In the case of Uber's 2016 data breach, Uber's response can be evaluated through the lens of virtue ethics by examining the character of the company's leadership and the moral virtues they displayed throughout the incident.

Uber's decision to conceal the data breach and pay off the hackers rather than report the incident and inform affected users was a clear reflection of the moral failings of its leadership. Virtue ethics would criticise the lack of honesty, responsibility, and courage displayed by acting Chief Security Officer, Joe Sullivan. Instead of demonstrating honesty and transparency, which are crucial in situations of crisis, Uber's leadership chose a path of deception and self-preservation. A virtuous company would have taken responsibility for the breach, informed the public, and worked to mitigate the harm caused. Instead, Uber's action demonstrated a lack of moral courage – the ability to face difficult truths and make the right decision despite potential personal or corporate costs (Lohrmann, 2017).

Furthermore, the company's handling of the breach underlines the absence of integrity within its corporate culture. Integrity, which is a core virtue in ethical leadership, involves being truthful and adhering to ethical principles, even when it is difficult or when no one is watching (Mayer et al., 2012). The decision to keep the breach quiet in the hopes of avoiding scrutiny reflects a corporate ethos where reputation and profits were valued over doing what was right by users and society. Uber's actions ultimately violated the trust that users had placed in the company, undermining the fundamental relationship between businesses and consumers.

Additionally, responsibility is another essential virtue that Uber's leadership failed to demonstrate. In a situation such as a data breach, the company has a moral duty to act in the best interest of its users, ensuring their data was protected and they were informed about potential risk (Lohrmann, 2017). By choosing to conceal the breach, Uber failed to uphold this duty and exposed its users to further harm. A company that embodies virtue ethics would have acted promptly, notifying their users and law enforcement to prevent the theft from causing excessive damage. However, Uber's leadership displayed a lack of empathy and concern for the welfare of users, and their actions were driven by self-interest rather than the moral duty to protect those affected.

Uber's culture, under former CEO Travis Kalanick, further demonstrates a broader lack of moral character in the organisation. The company's history, which includes a pattern of unethical behaviour, such as evading regulations and fostering an aggressive corporate culture, contributed to the breach's mishandling (Lohrmann, 2017). Ethical leaders within the organisation would have encouraged transparency, open communication and accountability. In a company with a strong emphasis on virtue ethics, leaders would cultivate a corporate culture focused on respecting the rights of individuals, upholding privacy, and acting with integrity in all manners.

However, virtue ethics also has limitations. One major critique of this ethical framework is that it can be subjective; different people may disagree on what constitutes virtuous behaviour, especially in ambiguous situations. If we consider the immediate short-term, Chief Security Officer Joe Sullivan may have thought he was acting virtuously by concealing the breach, potentially buying the company time to find a quick solution. Virtue ethics does not always provide clear guidelines for action, and it may be difficult to apply consistently across different scenarios. Furthermore, virtue ethics primarily focuses on character and internal motivations, but doesn't always give full analysis of the external outcomes like consequentialism does. Despite its limitations however, virtue ethics is valuable because it emphasises the importance of character and moral leadership in ethical decision-making, especially in a crisis. Regardless of having the company's best interests at heart, the decision to conceal the breach was ultimately ethically flawed and represented a failure in moral leadership that affected both the company and its users.

In conclusion, applying a pluralistic approach that combines consequentialism and virtue ethics provides a robust analysis of the ethical issues surrounding Uber's 2016 data breach. Consequentialism allows us to critically assess the outcomes of Uber's actions, revealing that the decision to conceal the breach resulted in significant harm to users, public trust, and the company itself. From a utilitarian perspective, the concealment of the breach led to negative consequences in the long-run such as identity theft risks, reputational damage, and legal penalties that far outweighed any potential short-term benefits (Lohrmann, 2017). This demonstrates that full disclosure and transparency would have minimised the long-term harm to Uber and its users, offering a clear path forward in addressing the breach.

In contrast to this, virtue ethics focuses on the moral character and leadership within Uber. The decision to hide the breach instead of confronting the issue transparently reflected a lack of responsibility and moral courage in leadership. Virtue ethics helps us understand that Uber's response was not only ethically wrong because of its harmful outcomes but also because it stemmed from moral failings at the top levels of the organisation (MacDonals and Marcoux, 2023).

Together, these frameworks provide a comprehensive ethical evaluation. While consequentialism highlights the detrimental effects of Uber's decisions, virtue ethics reveals the deeper issue of leadership failure and moral character. By integrating both perspectives, we are able to gain a deep understanding of why Uber's response to the breach was ethically flawed – an approach that considers the moral outcomes and ethical integrity of those involved. Therefore, a pluralistic approach strengthens our ability to analyse complex ethical situations and provides deeper insights into the moral responsibility of data governance, by considering both the outcomes of decisions and the ethical integrity of those making them.

Bibliography

Anderson, C. (2008). The End of Theory: The Data Deluge Makes the Scientific Method Obsolete. *Wired Magazine*. Available at: https://www.wired.com/2008/06/pb-theory/ [Accessed 10 Apr. 2025].

boyd, danah, & Crawford, K. (2012). CRITICAL QUESTIONS FOR BIG DATA: Provocations for a cultural, technological, and scholarly phenomenon. *Information, Communication & Society*, *15*(5), 662–679. https://doi.org/10.1080/1369118X.2012.678878

Chappell, B., 2018. *Uber Pays $148 Million Over Yearlong Cover-Up Of Data Breach*. NPR, 27 September. Available at: https://www.npr.org/2018/09/27/652119109/uber-pays-148-million-over-year-long-cover-up-of-data-breach [Accessed 4 April 2025].

Cukier, K. and Mayer-Schönberger, V. (2013). The Rise of Big Data: How It's Changing the Way We Think About the World. *Foreign Affairs*, 92(3). Available at: https://www.foreignaffairs.com/articles/2013-05-01/rise-big-data [Accessed 10 Apr. 2025]

Humby, C. (2006). Data is the New Oil. *Kellogg School of Management*. Available at: https://www.kellogg.northwestern.edu [Accessed 10 Apr. 2025].

Kitchin, R., & McArdle, G. (2016). What makes Big Data, Big Data? Exploring the ontological characteristics of 26 datasets. Big Data & Society, 3(1). https://doi.org/10.1177/2053951716631130

Koehn, D., 2016. *The Ground of Professional Ethics: A Philosophical Approach*. Chicago: University of Chicago Press.

Lacey, D. (2001). 3D Data Management: Controlling Data Volume, Velocity, and Variety. *META Group Research Note*, 6(70), pp. 1-12.

Lohrmann, D., 2017. *After Uber Data Breach: Lessons for All of Us*. Government Technology. Available at: https://www.govtech.com/blogs/lohrmann-on-cybersecurity/After-Uber-Data-Breach-Lessons-for-All-of-Us.html [Accessed 4 April 2025].

MacDonald, C. and Marcoux, A. (2023). *Ethical Theory: Virtue Theory*. [online] Torontomu.ca. Available at: https://pressbooks.library.torontomu.ca/cebe/chapter/ethical-theory-virtue-theory/ [Accessed 29 Apr. 2025].

Mayer, D. M., Aquino, K., Greenbaum, R. L., and Kuenzi, M., 2012. 'Who is the ethical leader? The role of employees' attributes in ethical leadership.' *Business Ethics Quarterly*, 22(3), pp. 1-16.

Somerville, H., 2018. *Uber to pay $148 million to settle data breach cover-up with U.S. states*. Reuters, 26 September. Available at: https://www.reuters.com/article/us-uber-cyber/uber-to-pay-148-million-to-settle-data-breach-cover-up-with-u-s-states-idUSKCN1M11C9 [Accessed 4 April 2025].

Wright, S.A., n.d. *Privacy Ethical Issues*. [Online] Available at: https://www.drstevenawright.com/privacy-ethical-issues/ [Accessed 4 April 2025].