applied sciences

MDPI

Review

An Analytical Review of Industrial Privacy Frameworks and Regulations for Organisational Data Sharing

Seyed Ramin Ghorashi [1,2,*], Tanveer Zia [1,2,3], Michael Bewong [1,2] and Yinhao Jiang [1,2]

[1] School of Computing, Mathematics, and Engineering, Charles Sturt University, NSW 2678, Australia; tzia@csu.edu.au (T.Z.); mbewong@csu.edu.au (M.B.); yjiang@csu.edu.au (Y.J.)
[2] Cyber Security Cooperative Research Centre, WA 6027, Australia
[3] School of Arts and Sciences, University of Notre Dame, Australia
* Correspondence: sghorashi@csu.edu.au

Abstract: This study examines the privacy protection challenges in data sharing between organisations and third-party entities, focusing on changing collaborations in the digital age. Utilising a mixed-method approach, we categorise data-sharing practices into three business models, each with unique privacy concerns. The research reviews legal regulations like the General Data Protection Regulation (GDPR), highlighting their emphasis on user privacy protection but noting a lack of specific technical guidance. In contrast, industrial privacy frameworks such as NIST and Five Safes are explored for their comprehensive procedural and technical guidance, bridging the gap between legal mandates and practical applications. A key component of this study is the analysis of the Facebook–Cambridge Analytica data breach, which illustrates the significant privacy violations and their wider implications. This case study demonstrates how the principles of the NIST and Five Safes frameworks can effectively mitigate privacy risks, enhancing transparency and accountability in data sharing. Our findings highlight the dynamic nature of data sharing and the vital role of both privacy regulations and industry-specific frameworks in protecting individual privacy rights. This study contributes insights into the development of robust privacy strategies, highlighting the necessity of integrating comprehensive privacy frameworks into organisational practices for improved decision making, operational efficiency, and privacy protection in collaborative data environments.

Keywords: privacy regulation; privacy frameworks; data sharing; organisations; third-party entities; Facebook; Cambridge Analytica

Citation: Ghorashi, S.R.; Zia, T.; Bewong, M.; Jiang, Y. An Analytical Review of Industrial Privacy Frameworks and Regulations for Organisational Data Sharing. *Appl. Sci.* **2023**, *13*, 12727. https://doi.org/10.3390/app132312727

Academic Editor: Gianluca Lax

Received: 30 September 2023
Revised: 19 November 2023
Accepted: 20 November 2023
Published: 27 November 2023

1. Introduction

Organisations are increasingly engaging in collaboration with third-party entities such as vendors, suppliers, and business partners to leverage data collection and analysis for improved decision making and operational efficiency [1]. These partnerships allow access to specialised knowledge and resources, facilitating efficient data analytics and cost-effectiveness compared to developing in-house capabilities. Data sharing, a practice vital for knowledge growth and informed decision making, is prevalent across organisations regardless of their resources or technology. It is particularly significant in Internet of Things (IoT) contexts, where data exchange between devices and systems is key to enabling smart, connected environments. While organisations already benefit from using consumer data, further advantages can be gained by sharing data with external entities. This is evident in the growing importance of data sharing for businesses, especially in online social networking platforms like LinkedIn, Facebook, and X (previously known as Twitter), where user data are shared across various communities [2].

However, despite the benefits associated with sharing data with third-party entities, these organisations pose privacy risks to individuals who use their services. This is because these organisations collect, analyse, and share individuals' data, which often includes

personally identifiable information (PII) and sensitive data. Consequently, these data-sharing activities raise privacy concerns, as many organisations exploit data for marketing purposes by delivering targeted advertisements and other content [3]. For example, Google collects extensive user data across its platforms, such as maps and YouTube, for targeted advertising. While it is part of Google's business model to collect data about its users, it has raised concerns about privacy, as individuals may not be fully aware of the extent of the data collected or how it is used.

The personal data exchange between primary organisations and their third-party partners significantly raises the risk of privacy breaches. This issue is supposedly addressed by privacy regulations and industrial privacy frameworks. However, a critical problem lies in the existing privacy regulations, which, while mandating organisations to comply with legal standards and enforce privacy policies, often lack specificity in their guidelines. These regulations typically focus more on legal compliance and general privacy protection rather than providing detailed, technical directives essential for robust data protection. For example, organisations are required to disclose how they collect, process, and share personal data, but the regulations fall short of offering explicit, technical instructions on implementing these practices securely. Terms like "Safeguarding" used in these regulations are often too broad and ambiguous, leaving organisations without clear, actionable steps for effectively protecting personal data, thus creating a gap in practical data protection measures.

Consequently, both primary and their third-party collaborators are obligated to lean towards the legal aspects and thus often encounter a lack of agreement between the emphasis on legal compliance outlined in regulations and the lack of technical elements essential for establishing robust data security. This difference can lead to the development of privacy policies that, while legally compliant with the law, may not fully address the complexities of data protection in today's digital world. Another issue can also be the level of data importance in protecting personal data. While all organisations have the responsibility of implementing privacy policy measures to protect and maintain the integrity of personal data, third-party entities collaborating with these organisations may not bear the same level of responsibility. For example, the Facebook–Cambridge Analytica scandal highlights how third-party entities might misuse data due to inadequate privacy protections. Cambridge Analytica accessed millions of Facebook users' data without proper consent, using it for political profiling and targeting, which Facebook failed to prevent [4]. In similar cases, in the Target data breach, the intrusion into Target's network was through a third-party HVAC vendor, showcasing how third-party entities might lack the security protocols necessary to protect sensitive data.

Industry approaches, such as privacy frameworks, provide essential guidelines and best practices for privacy preservation and risk management, focusing on information, standardisation, and privacy risk identification. These frameworks not only align with privacy regulations but also offer technical advice to help organisations protect personal data effectively. Despite previous studies [5–9] examining major privacy regulations and the need for compliance tools, research on these industrial privacy frameworks, particularly in the context of organisational data sharing, remains limited. This study aims to bridge this gap by comparing how privacy frameworks stand against regulations in promoting legal compliance and a robust approach to data sharing and protection, beneficial for both organisations and individuals. Our analysis focuses on three operational models—business to business (B2B), business to consumer (B2C), and consumer to consumer (C2C), to provide insights into the application of these frameworks in varied organisational contexts.

This study highlights the critical need for organisations to not only adhere to privacy regulations but also develop comprehensive privacy policies that address prevalent challenges. Our research reveals that a common issue is the creation of privacy policies that aim for regulatory compliance but often lack sufficient technical detail, potentially leading to privacy breaches. In response, we propose that adopting industrial privacy frameworks can be an effective solution. These frameworks provide organisations with essential tech-

nical guidance and customisable best practices, addressing shortcomings identified in current privacy strategies. Our findings, also discussed in the literature [10,11], highlight the inadequacies in existing privacy regulations and the general awareness among users, advocating for more complex and technically sound approaches to privacy management. Through a case study, we demonstrate a practical balance between regulatory compliance and effective data privacy management, showing how adaptable solutions based on these frameworks can significantly enhance privacy protections within organisations.

This paper is organised as follows: Section 2 presents the existing challenges and privacy risks associated with organisational data sharing. Section 3 provides an overview of organisational data-sharing models, while Section 4 presents insights into privacy regulations and privacy frameworks. Section 5 conducts an in-depth analysis of the case study, followed by a discussion of the findings in Section 6. Finally, in Section 7, we draw our study to a conclusion.

2. Overview of Privacy Challenges and Disclosure Risks in Organisational Data Sharing

In this section, we will introduce and define the aspects of the data-sharing models, the current problems with privacy policies in organisational data sharing, and an overview of their privacy risks and privacy disclosure.

2.1. Challenges in Organisational Data Sharing

Data sharing is a widely acknowledged practice that enhances organisational efficiency and performance by providing insights into processes and technologies. However, it can present challenges, particularly in relation to privacy concerns [12,13]. The nature of data sharing varies among organisations, depending on their business models, such as B2B, B2C, or C2C, which dictate distinct relationship models and privacy policies [14]. Privacy policies are crucial for organisations as they ensure compliance with privacy regulations and safeguard customer privacy. These policies serve as legal documents, informing individuals about how their personal data are collected, processed, and shared [15,16]. Nevertheless, the implementation of privacy policies can differ between organisations, leading to discrepancies between policy statements and actual practices.

The shift towards cross-organisational data sharing, especially in models like B2B, B2C, and C2C, can enhance organisational performance by exchanging personal data for services or goods [17]. However, there is currently no standardised guidance on disclosing personal data-collection details in privacy policies, resulting in variations between organisations' policies [14]. This lack of standardisation can confuse consumers, who may not fully grasp how their data is managed. Intra-organisational data sharing, as practiced by Meta (the parent company of Facebook, Instagram, and WhatsApp), is valuable when internal systems and customer demographics align [12,18]. However, compliance with privacy regulations can vary based on the sensitivity of the data collected. For example, healthcare organisations require stricter compliance due to the sensitivity of medical data compared to e-commerce organisations [14].

Privacy policies play a crucial role in protecting customer privacy, but comprehending these documents can be a challenge due to their often lengthy and complex nature [19,20]. Research conducted in the United States revealed that citizens would need to spend an average of 40 min per day just to read all the privacy policies they encounter [21]. Moreover, privacy policies may be influenced by other privacy regulations from various regions worldwide.

One of the significant difficulties organisations encounter when striving for transparency in their privacy policies is in how they use and disclose personal data. Some organisations adopt a permissive approach, sharing personal data with business partners or affiliates, while others primarily use this data internally, as seen with Meta disclosing user information across its subsidiary platforms like Facebook and Instagram [18]. Although these organisations attempt to explain in their privacy policies how personal data are protected, they often fail to provide information about how it is protected or where the

data are utilised. While their privacy policies may meet legal requirements from industry to government levels, the vague language employed can increase the privacy risk associated with personal data.

Ensuring the accuracy of information within privacy policies is vital for maintaining customer trust regarding how organisations handle personal data. For instance, Meta, the parent company of Facebook and Instagram, previously provided false information in 2014 regarding automated profile matching between Facebook and WhatsApp. Subsequently, Meta updated its terms of service, requiring users to agree to the new terms. In 2021, Meta modified its privacy policy to allow the sharing of personal information with its affiliates [22,23]. This highlights the evolving nature of privacy policies and their potential impact on users' data privacy.

Privacy policies serve a critical role in defining an organisation's procedures for processing and disclosing personal data [16]. These procedures, known as privacy practices, are essential commitments made by organisations. However, they are not always consistently followed, often due to factors like data breaches or unauthorised access by malicious actors. For example, in October 2016, a significant data breach occurred at the well-known ridesharing company Uber, resulting in the exposure of 57 million customer and driver personal records. Shockingly, Uber chose not to inform the affected individuals or regulatory authorities about the breach. Instead, they paid a ransom to the hackers to cover up the incident. This breach only came to light a year later, in November 2017, raising substantial concerns about Uber's lack of transparency and failure to comply with data breach notification regulations [24].

Privacy policies offer numerous advantages beyond merely collecting personal data. They also provide crucial information about partners and affiliates with whom data is shared. One of the key benefits of privacy policies is their ability to inform end users about the specific personal data collected. However, these policies often fall short of providing detailed information about the types of personal information shared with affiliates or third-party organisations. An illustrative example is the acquisition of the messaging app WhatsApp by Facebook in 2016. At that time, the app's privacy policies did not explicitly mention that Facebook would collect certain user data, such as phone numbers, for targeted advertising purposes [25]. This situation raised regulatory concerns, suggesting that the information provided to users was not sufficiently clear.

Even when organisations comply with privacy regulations by implementing privacy policies to protect personal data or share data with affiliates or third-party organisations, it still presents privacy risks. These risks can potentially empower malicious actors to identify and exploit the data, thereby increasing the risk of privacy breaches [26]. The interaction between privacy policies, data sharing, and the potential for privacy breaches highlights the current challenges and the need to implement alternative data-protection methods within organisations.

2.2. Organisational Privacy Disclosure Risks

In today's data-driven world, where organisations rely on exchanging high quantities of data with each other to gain valuable understandings, there are challenges in sharing that data, such as privacy risks of disclosure. Privacy risk refers to the potential threats posed to personal information, which can eventually disclose the information of individuals. In the context of organisational data sharing with third-party entities, the potential privacy disclosure risks could take two forms, such identity disclosure and attribute disclosure.

Identity disclosure [27] refers to the risk of re-identification of individual's identities from shared data. The privacy risk arises when personal data are shared with third-party organisations, revealing combinations of different attributes (QID), allowing the adversaries to recognise the data records by profile-mapping the QID together. The attribute disclosure [28] is achieved by exposing specific QIDs of individuals that were supposed to remain confidential. When data are shared, it might contain sensitive attributes such as medical conditions or financial status. These attributes, although may not directly identify

the individual's identities, can still lead to privacy breaches if linked to other publicly available information or combined with other attributes to create a detailed profile.

Two of the main privacy disclosures, identity and attributes, have helped researchers to study many approaches to reduce privacy risks; however, it is important to understand membership disclosure, which could help researchers and regulators to identify privacy risks in organisational data sharing. Membership disclosure [29] occurs when an individual's affiliation to a particular group is revealed without their consent. Usually, this disclosure occurs from data modelling or analysis when an adversary comes to an understanding of an individual's presence in a dataset.

3. Organisational Data-Sharing Models

The main value of any organisation is the ability to analyse data based on the information they have. As each organisation has a different type of operation, they could deal with different types of data, and it would suggest what analysis of data they would perform. This analysis is usually conducted on the business side, and the information collected is from their consumers.

A business entity is usually referred to as an organisation that is engaged in commercial, industrial, or professional activities that produce and sell goods or services. The activities of these organisations can define their missions as either for-profit or not-for-profit (non-profit) [30]. There are also organisations that rely on external entities for a specific function or operation. These external entities are called third-party organisations, which are not part of the primary organisation but collaborate together [31]. The consumers, on the other hand, are usually individuals or businesses that purchase and consume the market goods or services of another business [32].

The entities within an organisation are motivated by day-to-day data. These data make up information that is collected from a specific entity, or it is driven data, meaning the data have been populated based on a task or process. There are two types of data that create this different information: first is consumer process data, and the second is business process data. Consumer process data contain personal information about an individual, which describes the identity or the attitude and behaviour of an individual towards a business product or service. Business process data is data about the business. These could include basic information about a business, employee details, and information about the operations, business products/services, and marketing/sales.

As discussed above, most organisations fall into a business model category that defines the operation of their business, their business relationships, and their consumers. The main business models are three kinds: B2B, B2C, and C2C.

Business-to-Business, or B2B, is a form of intercompany transaction between two businesses, such that one is a manufacturer or provider and the other becomes a retailer. The transactions between the businesses include trades, purchases, services, resources, and technologies [33,34]. B2B data sharing is normally data about other businesses or consumers that can be used for marketing activities or to make decisions. The data that businesses usually share can be divided into two categories: business process data and consumer process data.

Business-to-Consumer, or B2C, is a relationship between a business and a consumer or business acting as a consumer who are engaged in transactions, such as data, products, and services. The business is usually the entity providing the products and services, and the consumer entity is usually the end user of the goods. Depending on the type of consumer entity, they also provide data to the businesses. These data can be about themselves, or they could be data about how they interact with the products of the businesses [35].

The data that the businesses collect are based on the consumer (business). The businesses take advantage of this collected data to understand their consumers better and determine what their personalities are based on the products and services they provide. There are generally three types of data that businesses are interested in [36]. The personal data of their consumers, the engagement activities of their consumers associated with their

business, and the behavioural data of their consumers when it comes to purchasing and using their products and services [36,37]. These data are personal data, engagement, and behavioural data.

Consumer-to-consumer, or C2C, is a business model that allows the transaction of goods and services between two consumers. This model is also known as peer-to-peer (P2P) in the e-commerce business model. This is very similar to the B2C model; however, the consumers are interacting with one another. The C2C platforms are typically electronic market platforms created by businesses to reach more consumers online.

In a C2C platform, the consumers require basic data about one another to communicate. However, the businesses that are facilitating these platforms will collect more data about the consumers' activities. In regard to the process of interaction between the consumers, they may only require the personal information of the consumer. However, the businesses may also collect engagement and behavioural data as part of the process.

Each business model delivers a unique strength based on who the targeted audiences are. In the context of privacy, understanding these business models can help authorities regarding the privacy risks involved. By recognising the type of audiences and the type of data that businesses are dealing with, organisations can eliminate the possibilities of privacy risks by introducing privacy regulations. Nonetheless, privacy disclosure still may occur even if legal steps have been taken to be compliant with the privacy regulations [38]. These unforeseen consequences can happen because of a number of reasons, such as unintentional accidents caused by human error, a lack of understanding, or misuse of personal information. These accidents usually occur when incorporating third-party entities that have different policies or lack transparency when it comes to personal data handling.

Organisations have a legal obligation to comply with the privacy regulations as they collect, process, and share personal information [39,40]. In this regard, organisations are trying to minimise any potential privacy risk that threatens their business or the privacy of their information by implementing correct privacy practices. Despite the fulfilment of legal obligations, such as complying with privacy regulations and implementing privacy measures, privacy disclosure still occurs. Privacy disclosure is a major disruption for organisations financially and reputation-wise. Luckily, privacy-protection measures such as privacy frameworks exist to minimise the impact of privacy risk.

A misconception in privacy protections regarding the privacy framework and other measures, such as privacy by design, is the implementation costs and the disruption to the organisation. From an execution perspective, privacy frameworks require planning and a fair amount of consideration in the implementation of the standards. However, compared to privacy by design, the requirement is the implementation of the standards from the foundation level. Simply put, organisations need to re-design their operations for the handling of personal data from a privacy perspective. This means major changes to the collection, processing, and disclosure of the information operations. For simplicity, we will review privacy-protection methods that are efficient and quick in implementation.

4. Overview of the Privacy-Protection Regulations and Frameworks

This section presents an overview of privacy- and data-protection regulations and privacy frameworks in the context of organisational data sharing. It is important to acknowledge there are many other privacy regulations and privacy frameworks; however, after our analysis, we concluded that privacy regulations such as the General Data Protection Regulation (GDPR) are mandatory, and organisations have a legal obligation to comply with the laws. Nevertheless, privacy frameworks are voluntary tools available for organisations. Despite the differences between the privacy framework and regulations, we will analyse the guidelines of NIST and Five Safes with a case study to discuss privacy violations and how these privacy frameworks can provide appropriate guidelines.

4.1. The GDPR in Organisational Data Sharing

In the context of organisational data sharing, the enforcement of privacy regulations in recent years has brought a significant amount of transformation in the data-sharing practices [41]. These legal regulations have been implemented by the governments to protect the privacy of individuals by allowing them to have privacy rights when sharing their information. The most comprehensive privacy regulations, such as the GDPR, have been effective in governance, awareness, and monitoring of the usage of individual data, forcing organisations to protect privacy more proactively [42].

Privacy regulations have introduced a major change in how organisations share data. The requirements for receiving consent prior to data collection, as well as the organisational transparency in data sharing, have encouraged organisations to re-evaluate their data-acquisition and privacy practices, leading to provisioning clear, concise, and easily accessible privacy policies that explain data processing and data sharing [43]. This means organisations must clearly communicate the purposes for which data will be used, which entities will have access to these data, and how these data are protected.

These privacy regulations have also impacted business partners and third-party entities, which also need to show their compliance with the privacy regulations by including appropriate data-protection measures [44]. Any third-party entity that wishes to collaborate with first-party organisations must have due-diligence processes where organisations assess the data-handling processes before sharing any data. For example, if a hospital wishes to share a certain patient's health data with a research institute, the hospital must ensure the research institute complies with the necessary data-protection standards.

In recent years, many governments have developed privacy regulations to govern and maintain the privacy protection of their citizens around the world. The European Union released their privacy regulation in 2018, setting out strict rules and principles for the processing of personal data to ensure the privacy- and data-protection rights of individuals. The GDPR consists of significant protection of personal data throughout its lifecycle, most notably when sharing data with other organisations. For third-party organisations, they must establish GDPR-compliant data-processing agreements with their clients. Being GDPR-compliant entails the responsibility and the obligations of both parties to guarantee that the data processing is protected and processed in accordance with the GDPR.

The key principles of the GDPR are (1) lawful processing, (2) purpose limitation, (3) data minimisation, (4) accuracy, (5) storage limitation, (6) integrity and confidentiality, and (7) accuracy [45]. These principles define a summary of the GDPR; failure to comply with these rules can result in significant fines and legal consequences.

4.2. The NIST and Five Safes Privacy Framework

The industrial privacy frameworks have provided guidance for organisations to help them with data privacy and ensure the protection of individuals' personal data. This guidance involves standards, practices, and policies to follow to protect data, but mostly, the guideline's objectives are enhancing the protection level. These scopes are broad, including various approaches such as data collection, processing, storage, and sharing. The guidelines are designed to be applied to a wide range of organisations, such as non-profit organisations, businesses, or governmental organisations.

This study focuses on the NIST and Five Safes privacy frameworks due to the limited existing literature analysing their principles, particularly in privacy risk-mitigation scenarios. The NIST framework is a recent addition to its security counterpart. The widely used but older Five Safes framework, notably implemented by entities like the Australian Bureau of Statistics (ABS), presents a unique opportunity for exploration. While other frameworks, such as International Organisation for Standardisation/International Electrotechnical Commission (ISO/IEC) 27701 and privacy by design [46], have been extensively studied, our analysis seeks to fill the research gap between these two frameworks, providing new insights into their effectiveness and application.

Privacy frameworks have brought significant change within organisations and other entities, shaping their approach to data privacy and protection techniques, which includes enhancing transparency, increasing individuals' privacy rights, and mitigating privacy risks. One of the significant changes that influenced organisations was the transparency of their data processing with their users, which ultimately forced them to be compliant with privacy regulations. These include guidelines and practices to inform individuals properly and clearly about how their data will be used and shared. Additionally, these privacy frameworks have helped organisations comply with privacy regulations by implementing mechanisms and practices to be cautious when sharing information with third-party organisations or securing personal data in a manner that minimises privacy disclosure.

In the last couple of decades, many privacy frameworks have been developed to aid organisations with guidance and practices. One of these privacy frameworks is Five Safes, which was developed to help researchers, particularly in academics. However, with its wide recognition, it is also used in businesses and governmental organisations such as ABS [47]. The main objective of the Five Safes is to create a structured approach for these organisations for better data accessibility and useability whilst minimising privacy risks. The key five dimensions of the framework are (1) people, (2) projects, (3) settings, (4) data, and (5) outputs. These independent questions are used to conduct a risk assessment scheme for data accessibility and data.

In recent years, another privacy framework was developed by the National Institute of Standards and Technology (NIST), called the NIST privacy framework, designed to manage and protect individuals' privacy. Although this framework was developed very recently, it is believed to be widely used across many organisations and businesses of various sizes. This framework provides guidelines and standards for organisations to adopt to their privacy practices or procedures [48]. The scope of this framework covers a wide range of privacy concerns, including identification, governance and controlling privacy risks. The key main functions of the privacy framework are (1) identify, (2) govern, (3) control, (4) communicate, and (5) protect.

Both the NIST and Five Safes privacy frameworks are designed to provide voluntary guidelines and risk assessments to mitigate privacy risks. However, there are slight differences when considering the use cases of data collection, processing, and sharing. While the NIST provides comprehensive and detailed guidance applicable to various organisational contexts [49], the Five Safes framework is specific and mainly used in research settings. It focuses on ensuring responsible data usage for research, balancing data utility and privacy [50].

5. A Case Study of Organisational Data-Sharing Business Model Challenges

In this section, we will discuss privacy violations and any wrongdoing of the organisations that may have put the privacy of individuals at risk. We will also analyse privacy regulation violations as well as the privacy framework principles, such as NIST and Five Safes, for best practices.

In the event of privacy disclosure, a common mistake is made by the primary organisations. The negligence of the employers allowing the adversaries to have unauthorised access to confidential data is a type of human error that leads to a data breach. However, privacy disclosure is not always to be blamed on the primary organisation. Third-party organisations can also lead to privacy disclosure by lack of maintenance or ignoring privacy measures such as privacy regulations and privacy policies. For example, in July 2020, Morgan Stanley, a financial institute, notified its customers about a data breach that involved the disclosure of the customer's details, including names, addresses, date of birth, and financial status information. This incident occurred while Morgan Stanley hired a moving company to be in charge of the destruction of its information technology (IT) equipment that contained unencrypted data. The moving company was tasked to decommission the hard drives and the servers by partnering with an e-waste management company. Instead, the moving company sold the IT equipment, which had unwiped, unencrypted data, and it

ended up on an online auction [51]. Morgan Stanley managed to recover some of its hard drives; however, some are still missing.

In this section, we will study the Facebook–Cambridge Analytica scandal that was caused by a third-party entity. We will focus on the case from different perspectives of business models and the type of data collected, processed, and disclosed, highlighting the privacy violations as well as the privacy disclosure of each model.

The Facebook–Cambridge Analytica data breach incident is a notorious scandal that has impacted the lives of millions of individuals, from the disclosure of Facebook user privacy to involvement in election campaigns in recent years. The scandal occurred around the same time that the EU's privacy regulation (GDPR) was about to come into effect. Naturally, other privacy regulations, such as the UK's Information Commissioner's Office (ICO) and the US's Federal Trade Commission (FTC), were imposed on this matter for violations of regulations. However, it is in our best interest to explore the implications of this scenario if the privacy regulations had to be enforced by comprehensive privacy regulations such as the GDPR on privacy principle violations.

Previous studies [4,52–55] that examined the Facebook and Cambridge Analytica case study primarily focused their research on two aspects: the implications of privacy regulations and the misconduct of the organisations involved. This research explored the human perception of the Cambridge Analytica impact, reaching into public attitudes and reactions to the scandal's fallout. Additionally, these studies closely examined the unethical exploitation of personal data by organisations like Cambridge Analytica, uncovering the ethical violations and potential harm caused to individuals through data misuse. Furthermore, the weaknesses in privacy policies received considerable attention, with researchers highlighting how ambiguous or incomplete disclosures contributed to the exposure of individual privacy.

This research distinguishes itself from previous studies on the Facebook and Cambridge Analytica case by emphasising the essential role of industrial privacy frameworks in bridging the gap between legal compliance and technical implementation. By exploring the technical intricacies of data security and providing customisable solutions, this research serves as a valuable resource in data privacy, offering organisations insights to enhance their data-protection strategies.

5.1. Case Study: Facebook and Cambridge Analytica Scandal

In 2014, a researcher created a personality quiz application on Facebook that collected personal data about the users and user's friends. This created a database of millions of users' personal information [56,57]. In early 2018, it was reported that a consulting firm, Cambridge Analytica, was involved in obtaining the personal information of millions of Facebook users without their consent. Moreover, they processed the collected data allegedly to target the users with specific marketing campaigns for elections, such as the 2016 U.S. presidential election. An overview of the Facebook and Cambridge Analytica scandal is shown in Figure 1.

5.1.1. C2C: Facebook and User Data Policies

In 2013, Facebook allowed developers to design apps that could interact with the Facebook platform and access user data. This was only achievable if permission was granted by the user. A U.K.-based academic researcher named Aleksandr Kogan and his company, Global Science Research (GSR), developed an application that interacted with Facebook's platform. More specifically, it was designed to retrieve Facebook user data and friends' data. After installation of the app, users were presented with the privacy policy of the app, which stated that the app would collect data on the users and the users' friends, but it did not specify what the app would collect friends' data beyond what was necessary for the intended purpose. GSR claimed the app was used for academic research. However, it was reported that the data collected were later sold to a third-party organisation [4].

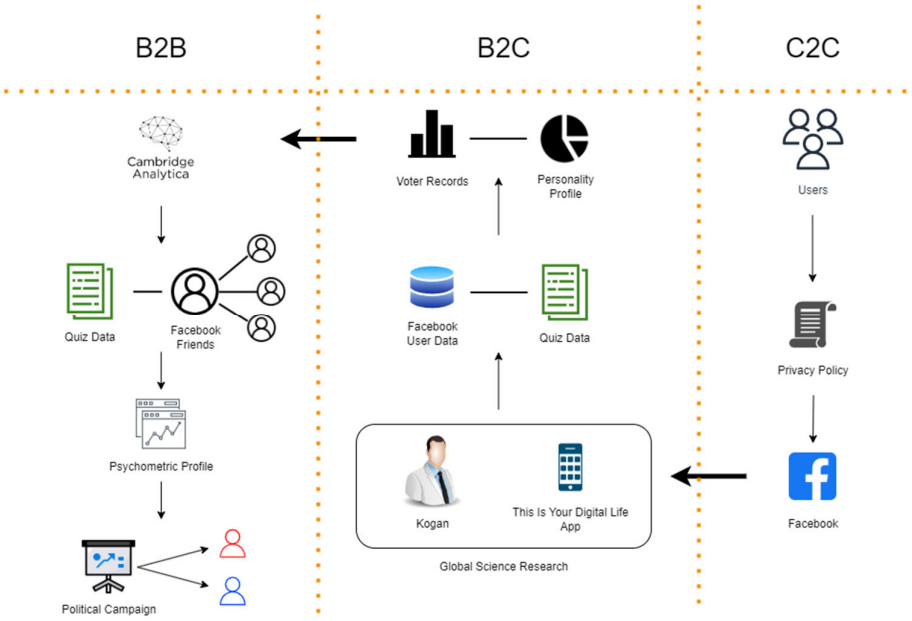

Figure 1. An overview of Facebook–Cambridge Analytica scandal in organisational data-sharing models (B2B, B2C, and C2C).

Facebook's policy back then allowed third-party developers to collect user data and their friends' data, up to 500,000 friends per user. The policy also had some restrictions in place; for example, Facebook prohibited the usage of user data for political manipulation or spamming. However, these restrictions were not effectively enforced.

In 2014, Facebook updated its privacy policies to limit the amount of data that third-party apps could collect. The updated policy reduced the number of friends' data that an app could collect from 500,000 to 100. However, Kogan's app had already collected data on millions of users before this change took effect.

In 2013, when third-party developers were allowed to gain access to the personal data of Facebook users, Kogan and his team, GSR, took this opportunity to collect significant personal data. Their aim was to collect personal data from Facebook users to conduct research experiments. At the time, Facebook's policies clarified that users must consent prior to any collection of data by third-party developers. Although the data-collection process was legally conducted, there were few misconducts during the data-processing stages.

In the event that took place between Facebook and GSR, both organisations have been deemed to violate privacy regulations of the GDPR based on multiple principles. The initial policies of Facebook explained that third-party developers cannot use the collected Facebook user data for political manipulation or spamming. However, Facebook was not diligent in ensuring the effectiveness of its own policies. Facebook failed to conduct rigorous audits or provide sufficient management, thus allowing developers like Kogan to exploit loopholes and collect data beyond what was necessary. In this regard, Facebook's accountability and acknowledgement of responsibility allowed developers, such as Kogan, to take advantage of this feature.

Simultaneously, Kogan and his team introduced an application to Facebook users, through which they presented the policies of the data collection and usage of personal data. However, there was an oversight, as they neglected to communicate important information. GSR failed to provide a clear explanation regarding the way collected personal data would be used and the information on the entities who would use the data later. As users opted to install the app to participate in the personality trait quiz experiment, they remained unaware of the usage of their data in the profile-matching. Both Facebook and GSR neglected to clearly communicate to users about the intention of their personal data. The

lack of transparency by both entities resulted in unlawfulness with regard to the privacy regulations and fairness with regard to the privacy of Facebook users.

Kogan failed to properly inform the users about the extent of data collection and the full intention of their personal data utilisation. This led to the disclosure of friends' user information, which they did not know. As Facebook's policy stated, only shared friends' User IDs were shared with third-party developers. However, more information was revealed about the friends, such as basic profile information, as well as their Facebook activities, including likes and shares.

Privacy frameworks such as NIST and Five Safes provide guidelines and principles that can help organisations assess and mitigate privacy risks. As such, the biggest privacy risk that concerns Facebook users is the ambiguity of their personal data used for something they did not consent to initially.

NIST Privacy Framework

1. Lack of transparency in organisations such as GSR and Facebook can cause serious privacy issues for their users. In this case, NIST's framework function, communication, offers developing guidelines and policies that clearly communicate the purpose, practice, roles, and responsibilities of individuals that are involved regarding the data processing. However, the data processing awareness category provides two specific guidelines (CM.AW-P1 and CM.AW-P2) for organisations to create a mechanism to send reports either internally or externally and create mechanisms to receive feedback. These features provide opportunities for organisations to establish communication methods, such as for data-processing purposes, practices, and associated privacy risks, and allow individuals to set their preferences for how their data are processed.
2. Regarding lawfulness and fairness, the NIST framework suggests two categories under governance. The first category (GV.PO-P1 and GV.PO-P5) indicates that organisations must establish privacy values and policies, including conditions on data usage and data-retention periods, and respect individuals' privacy rights towards data processing. The other guideline suggests that the organisation must understand and comply with any legal and regulatory obligations. Under governance, it also suggests (GV.MT-P1) monitoring and reviewing the privacy risks associated with data processing that may have been impacted by the business environment, new technologies, or any other legal obligations.

Five Safes Framework

The Five Safes privacy framework provides principles to manage and overcome privacy risks by proposing strategic questions about the usage of the data in different aspects. For instance, in the first scenario, Facebook, as well as the quiz policies, did not explicitly and clearly explain the importance of data collection, processing, and data sharing. However, in this case, the Five Safe's privacy framework provides two principles that could provide insights into how these risks could be mitigated. The principles are safe people and safe data.

1. The first principle, "Safe People", can be Kogan's team (GSR) implementing the quiz app on Facebook's platform. The primary organisation (Facebook) is required to control access to sensitive information by ensuring that only authorised individuals have permission to use and access such information. This principle allows organisations to have some level of responsibility to protect data from being misused. Typically, when more personal data is collected, responsibility also increases. This involves implementing robust measures such as authentication, authorisation, access control, and audit logs.
2. The principle "Safe Data" ensures that enough protection has been applied to the data. The data custodians (GSR) are responsible for implementing security and privacy measures to protect the data in storage and during transmission. The protective measures involve techniques such as encrypting the data during data transfer when

required and usage of anonymisation techniques to remove any identifiers or sensitive information before publishing the data.

5.1.2. B2C: Kogan and Facebook Users' Profile-Matching

By 2014, GSR developed a personality quiz application on Facebook called "This Is Your Digital Life". The app asked innocuous questions about personality traits, and this quiz was sent to around 300,000 users in the U.S. The quiz data collected information about the user's interests, preferences, values, and opinions based on five personality traits: openness, conscientiousness, extroversion, agreeableness, and neuroticism (OCEAN) [58].

At the completion of the quiz, the participants would be paid less than 5 USD; however, the condition of the quiz mandated the users to log in to their Facebook account. By participating in the quiz, the app would request access to the user's Facebook account data. A consent policy that was presented stated that the app would receive access to certain information, such as birth date, location; and Facebook's activities, including likes and shares. Although the consent policy stated that the app would collect data about users' friends lists, due to the vagueness of the language used in the policy, it did not explicitly state what information the app would collect from friends. The information that was collected from friends' data included their names, email addresses, and their Facebook likes and shares.

Upon the completion of the data collection about the Facebook users and users' friends' data, Kogan and his team combined the quiz results with the Facebook data to create a personality trait profile of each user. It was reported that they harvested 50–87 million users' data [59,60]. His team then used this data to match the Facebook profiles with voter registration records. This was achievable by linking the Facebook data with publicly available voter records using data such as name, age, and other identifiable information. Kogan then shared these data with a U.K.-based political consultancy firm called Cambridge Analytica (CA).

Following the massive aggregation of millions of users' personal data in 2014, Facebook changed their policies in response to the realisation of the extensive data collection by Kogan. Although Kogan stated that the collection of the data was used for academic research, the collection of data was highly criticised for being unclear and misleading. Firstly, the users were not aware of the type of data that was being collected, how it would be used, or who would access it. According to a privacy policy set in 2013, Facebook would share basic information or users' "public profile" as well as friends' User IDs with applications. However, it was discovered that friends' information (public profile) was also collected from the users. The app requested permission to access friends' information from the users, which is a violation of receiving acknowledgement or individual consent on behalf of the individual friends.

When Kogan combined the quiz data with the acquired Facebook data, such as likes and shares, they were able to ascertain characteristics and behaviours about the users by creating a personality profile. Based on this profile, some correlations were made that helped Kogan and his team to better understand the users' personalities. For example, from thousands of data points combined, they discerned the sexual orientation of certain men based on their "Liked" cosmetic brand, were able to distinguish some users' form of sociality (introvert and extrovert), and whether they liked a singer such as Lady Gaga or certain subjects such as Philosophy. It was also proved by the team that other preferences and affiliations could be predicted, such as political beliefs, religion, or even someone's skin colour. This was the second criticism: the app personalised each individual's political view based on their Facebook profile activities as well as their voter registration list. At this point, the approach undertaken was beyond what the app was intended to achieve. The collection of millions of US Facebook users to seek out psychological patterns is a violation of data minimisation as well as purpose limitation. As privacy regulations state, the data controllers must collect and process personal data that is specific, explicit, and necessary to the purpose that has been communicated to the users prior to the collection of the data.

From a privacy risk perspective, although Kogan integrated the personality profile data with voter registration information, there was a huge potential risk for identity disclosure. However, the exploitation of users' behaviours, as well as characteristics, is considered attributed disclosure. Analysis and understanding of user attributes are privacy concerns, even much riskier when users are not aware of it. In the context of the attribute disclosure, the User IDs, as well as basic information of the Facebook users, were available. While these data may not directly reveal an individual's identity, the linkage of these data with secondary data that contain non-identifiable information—in this case, quiz results—can lead to the identification of personal characteristics and behaviour. In this scenario, certain characteristics and behaviours of individuals were revealed to Kogan. Typically, skin colour and sexual orientation are considered sensitive information that individuals are entitled to have as private information, along with religious beliefs, ethnicity; and personal preferences, such as political beliefs. Although these data may not directly identify the individuals, the combination of such data with an external source may eventually reveal their identity.

Later, Kogan combined the personality profile with the voter's registration records to attain a better understanding of the users' political beliefs. Disclosure of such sensitive attributes not only violates the privacy rights of individuals, but it can also disclose their beliefs in politics. As mentioned before, beliefs such as politics are sensitive information because they can be exposed as a vulnerability in an individual by creating targeted advertising or propaganda, which can manipulate the individual's personal beliefs.

NIST Privacy Framework

The main issue in the scenario that could be mitigated using the NIST privacy framework is understanding the characteristics and behaviours of the users. This issue revolves around the process of combining different sources to disclose sensitive attributes. For this, in NIST's privacy framework, the protect and communicate function can be used for the primary organisation (Facebook) and the third-party organisation (Kogan and GSR), respectively.

1. For the primary organisations, the category identity management, authentication, and access control (PR.AC-P1-P6) provides practical practices and exercises to control and authenticate the usage of the data by third-party organisations. The aim of this category is to reduce the limit of the data to only those which are authorised and are authenticated with proof of credentials. For example, the credentials, authority, access control, and privileges of everyone are audited to uphold security and privacy.
2. Data processing awareness (CM.AW-P1-P5) and (CM.AW-P7-P8), when achieved by third-party organisations, can be a good privacy practice for effective awareness of data processing. Organisations such as GSR will have to adopt practices to promptly inform users about their data usage and analysis. This can be achieved by featuring mechanisms to send and receive reports between the primary organisation as well as the users. Other ways to achieve this are to ensure the records of the data are available and accessible to be reviewed, corrected, or deleted. Finally, there are mechanisms in place for the users to freely oppose the data processing by withdrawing or altering their data.

Five Safes Framework

The further usage of the data processing without any consent was a breach of privacy rights. As the intended purpose of the quiz app extended far beyond the scope, it led to the discovery of various personal attributes and behaviours. The Five Safes principles include the "Safe Project" and "Safe Settings". Both principles are to be implemented and determined by the primary organisation (Facebook).

1. The Safe Project principle aims to focus on the purpose of the project. Whether the obtained data are appropriate to be used in the project would be the question for the primary organisations. By limiting the amount of data analysis to what is required, it

ensures that data are used only for legitimate and well-defined purposes. For example, Kogan and his team had to explain to Facebook and its users the purpose of their research project, even if it meant explaining that they would need to analyse the behaviours and characteristics of the users and combine them with voter registration to create a personality profile. Clearly defining the objectives and scope of the research project can minimise the risk of privacy disclosure.

2. The other principle, Safe Settings, questions the safety of the environment in which data is processed. GSR should use measures like anonymisation to protect the data. Although for public information such as User ID and some other publicly available profile information, Safe Settings is not required, for data that are collected and processed with other external data, there must be privacy measures to protect the data. GSR was able to analyse the characteristics of Facebook users by identifying their sensitive attributes, such as their religion and skin colour. Additionally, by combining the data analysed with voter registration, they were able to identify the users with specific political beliefs.

5.1.3. C2C: Cambridge Analytica and Facebook User Friends' Data Analysis

Just before the 2016 U.S. presidential election, Kogan's company (GSR) collaborated with CA. This company had created a computer algorithm that could predict the outcomes of elections based on social media data. As part of the collaboration, GSR shared the profiles with the CA. The main purpose of CA was to use the data provided by the GSR company to help political campaigns and organisations better understand and engage with their target audiences. Instead, they used the data to create targeted political advertisements and propaganda during the 2016 U.S. presidential election.

CA was able to create a psychometric model of individual voters. These profiles contained information about their interest in politics, their behaviours, and the demographics [60]. These data were useful to create personalised and persuasive political messaging and target specific groups of voters. By leveraging these insights, CA was able to sway voters towards specific candidates.

The collected data from Facebook users were used as a significant approach to persuade individuals to overturn the US presidential election in 2016. It is reported that CA was hired by the Trump campaign to run data algorithms in which CA had created a psychometric profile model based on each user. The aim was to identify the voters and make specific ads that can help the campaign's strategic communications. One of the strategies of CA was to disregard persuading millions of voters and instead, using their data analysis, focus only on vulnerable voters who they believed to be hesitant, neurotic, or worried. Characteristics of such individuals were valuable to CA as they knew, with proper micro-targeting ads, they could swing the results in Trump's favour. An example of a personalised ad on social media would be focused on individuals with firearms. The targeted ad would carry messages such as "Did you know Hillary Clinton wants to take your gun away?", which only featured in the social media posts for a few hours to be visible to the targeted individual and eventually disappear off the social media.

The collection and processing of these data for such utilisation is a demonstration of privacy violations of the GDPR and APP. CA's approach to targeting and creating personalised ads is a violation of privacy principles. Although the usage of the data for such an approach mainly violates principles, including consent and transparency, significant issues in this case would be accountability, source, and accuracy. CA's data-driven approaches and strategies should have been compliant with the privacy regulations. Ensuring data disclosure protection and acknowledgement should have been practised before the usage of the data. Moreover, the source of the data is important in the data-sharing process. Organisations such as CA should have obtained the data legally from GSR by informing the users about the details of the data's processing. The accuracy of the data is also important in the analysis of information. If CA had obtained the Facebook data in

collaboration with the GSR, it would have ensured the data were validated and accurate. Any data that is not correct can produce misleading information.

Privacy regulation violations can pose privacy risks, too. From data analysis to psychometric profiling, targeted advertising based on personal characteristics has privacy consequences, which reveal information about certain individuals or groups. In this scenario, CA raised privacy risks when identifying individuals as well as micro-targeting for ads. Such risks can be prone to membership disclosure. As CA refers to specific individuals that are affiliated to a specific group or category for their work, it can be said that the individual's information regarding their affiliation towards certain beliefs, in this case political, was revealed and misused by the organisation.

In the US presidential election context, CA used their data to analyse users' psychometric profiles to identify the most vulnerable individuals. Their aim was to influence the individuals' political beliefs, so rather than focusing on the majority of the voters in the election, they focused on minorities who shared common attributes or characteristics. For example, CA identifies users who are described as "hesitant or worried". Another way CA micro-targeted political ads was to identify the groups of users with specific conditions—for example, running political ads appealing to groups who own firearms. Identified users with similar attributes were effectively categorised as members of a specific psychological or behavioural group. Identification and revelation grouping of users based on shared characteristics, conditions, or interests indicates a form of membership disclosure.

In the context of privacy disclosure from third-party organisations, the main issue is related to the primary organisation's management of the data. As CA and Kogan collected and processed data in such ways that led to privacy disclosure, Facebook's management of the data was poor. It was only after the privacy breach that Facebook updated their policies and tightened the security of its user's data. Privacy frameworks such as NIST and Five Safes introduce practices that can overcome privacy issues, including membership disclosure.

NIST Privacy Framework

To ensure the security and privacy of individuals' data, the NIST privacy framework provides the control function. This function contains three categories that provide data procedures, data management, and data authority. All the sub-categories are inclusive and applicable to resolve the privacy disclosure risk in B2B.

1. The initial sub-categories (CT.PO-P1-P4) discuss the importance of developing procedures and policies to keep an authority on data processing. The significance of these procedures and policies is to empower the primary organisations to exercise their control of the data by maintaining and establishing practices that ensure the respect of the individual's privacy rights and stay aligned with privacy regulations.
2. The sub-categories of data processing management (CT.DM-P1-P10) provide valuable measures for organisations to effectively manage data while upholding privacy regulations. These measures provide management and control to the primary organisation for accessibility, data transmission, data alteration, data deletion, data destruction, standardisation, and audit practices that respect the privacy preferences of the individual's data.
3. The last category, disassociated processing, contains five sub-categories (CT.DP-P1-P5) that focus on creating data-processing solutions to increase the disassociability of data with the individual's identity. The goal of this category is to develop processes that can limit the processing of data or events to associate with an individual—for example, by using privacy-preserving models such as differential privacy [61] to de-identify individuals or limiting the likability and attributes referencing an individual's behaviour or activities.

Five Safes Framework

One privacy risk is membership disclosure, where an individual's affiliation with certain groups is revealed through data processing. Specifically, this issue resolves around the targeted manipulation of vulnerable users who have particular characteristics or conditions and are associated with specific groups. The Five Safes privacy framework provides a guideline to mitigate the risks associated with membership disclosure. The "Safe Outputs" is a principle that provides an approach to ensure that data processing and analysis are conducted in a manner that respects the individual's privacy rights and minimises the privacy risks.

1. Before the data analysis is published, the final stages, according to the Five Safes privacy frameworks, ensure there are strict privacy controls implemented in the outputs of data analysis, so the statistical results are non-disclosive. For example, instead of directly targeting users with certain affiliations with specific characteristics or behaviours, the principle could control the third-party organisation by targeting a broader audience. By reducing the control of the targeted audiences, this principle can minimise the effect of political ads based on only a significant portion of the users interested in obtaining and keeping firearms. This principle can also help third parties to create generalised ads that are specifically targeted at users affiliated with certain groups that are prone to be vulnerable.

6. Discussion

Data-sharing models typically involve three primary categories: B2B, B2C, and C2C. Organisations tend to align with one or more of these models based on their objectives and operational strategies. Consequently, the purpose behind collecting and sharing data varies among organisations and users. In each of these models, data exchange occurs between two organisations. Typically, one organisation serves as the data supplier, sharing data, while the other organisation acts as the data user, responsible for collecting and processing the shared data.

In each of these models, the data-sharing methods and the types of data exchanged between the two entities are distinct. Moreover, there are inherent risks of privacy disclosures during the data-sharing processes. Given that the shared data in these models in the case study were often poorly managed, it is reasonable to anticipate various types of privacy risks, as each entity has unique requirements and intentions for using the data.

6.1. Ethical and Legal Implications of The Data Breach

For the Facebook and Cambridge Analytica scenario, we divided the cases into business data-sharing models that demonstrate the dependency of users, organisations, and third-party organisations for data sharing. We discussed the reasons for the data sharing and analysed the privacy violations and implications. As discussed before, privacy regulations are legal obligations for each organisation to comply with when they request permission to access personal data. This legal requirement mandates that organisations must be transparent, trustworthy, and responsible for the security and privacy of the users. However, third-party organisations may not always feel the same way as primary organisations.

The Facebook–Cambridge Analytica data breach incident stands out as a significant and far-reaching scandal that left a profound impact on the lives of millions of individuals. This incident brought to light a series of grave consequences, ranging from the compromised privacy of Facebook users, the unsettling involvement in election campaigns in recent years, and increased public awareness and concern over personal data security. Users are now more cautious about the information they share online. This scandal influenced the tech industry, and many organisations have revised their privacy policies and practices, emphasising user privacy and consent more than ever before [62,63].

These privacy violations raised several ethical concerns and violated many core principles of the GDPR. A survey shows that 44% of participants think organisations have become

more concerned about their privacy following the implementation of the GDPR [64], indicating that organisations have taken their customers' privacy more seriously, as they could be accountable for any wrongdoings. The Facebook–Cambridge Analytica incident raised serious concerns related to transparency, as users were left in the dark about how their data were being used and shared. Additionally, the breach highlighted weaknesses in communication, as users were not adequately informed or given a choice regarding their data usage. The incident also raised issues concerning data minimisation, as excessive amounts of personal data were collected without a genuine need [62]. Furthermore, fairness in data processing was undermined as user data was exploited for purposes that they had not consented to or were unaware of. This breach demonstrated a significant disconnect between user expectations and corporate practices, highlighting the moral imperative for companies to handle user data with greater care and honesty.

The scandal had substantial legal consequences. Initially, it led to intense scrutiny of Facebook's data practices, resulting in various legal challenges and fines. For example, Facebook faced a record-breaking 5 billion USD fine by the U.S. Federal Trade Commission for privacy violations [65]. Secondly, this incident forced governments to create stronger data protection laws worldwide. It reinforced the importance of the GDPR in the EU and inspired similar legislative efforts in other countries, aiming to empower users and ensure that organisations are held accountable for data privacy breaches [4].

6.2. Role of Industrial Privacy Frameworks in Mitigating Risks

In essence, this high-profile data breach highlights the critical importance of clear and transparent communication with users about how their data are handled. It also serves as a stark reminder of the need for robust governance within organisations and strict adherence to legal requirements, particularly in the realm of data privacy. The Facebook–Cambridge Analytica incident has left an indelible mark on the discourse surrounding data privacy, emphasising the vital role that user awareness, consent, and adherence to privacy regulations play in safeguarding individuals' personal information in the digital age.

Given the NIST and Five Safes privacy frameworks are used in different types of organisations, both deliver valuable guidance for addressing the privacy risks in the context of the Facebook–Cambridge Analytica scandal. The NIST privacy framework analysis [49] offers a wide range of procedures and technical control mechanisms, allowing organisations to make a robust privacy practice. The strongest guidelines (governance and communicate) allow organisations to make clear communications, governance, and legal compliance. Governance ensures that organisations establish clear policies and accountability structures for data privacy, which is critical in preventing incidents like the Facebook–Cambridge Analytica case, where unclear data-handling policies led to significant breaches. Communication, on the other hand, emphasises the importance of transparency with users regarding data collection, usage, and sharing practices. This aligns with the ethical need for consent and understanding, as users must be fully aware of how their data are being utilised. When organisations adhere to these guidelines, they can enhance their privacy practices, ensuring not only legal compliance but also fostering trust with their users. Table 1 illustrates a summary of the principles of the NIST privacy framework aligned with the principles of GDPR.

Meanwhile, the Five Safes framework provides guidance addressing privacy risks associated with targeted advertising and membership disclosure. The "Safe Outputs" principle provides a practical method to control data outputs and protect individual privacy rights [50]. This principle is vital in organisations to carefully evaluate the nature and extent of data being shared. By focusing on project-specific applications, the Five Safes framework assists organisations in clarifying the purpose and scope of data usage, which is essential in minimising unnecessary data collection and processing, which is a key issue in the Facebook–Cambridge Analytica incident. A summary of Five Safes principles aligned with the principle of GDPR is demonstrated in Table 2.

Table 1. Summary of NIST privacy framework principle of protection against privacy violations.

Scope	Breach Notification	Consent	Fairness	Lawfulness	Transparency	Data Minimisation	Purpose Limitation	Accountability	Confidentiality
Communicate CM.AW-P1-2, CM.AW-P1-P5, CM.AW-P7-P8	✓	✓	✓	✓	✓	-	✓	-	-
Governance GV.PO-P1, GV.PO-P5, GV.MT-P1	✓	✓	✓	✓	✓	✓	✓	✓	✓
Protect PR.AC-P1-P6	-	-	-	✓	-	✓	✓	✓	✓
Control CT.PO-P1-4, CT.DM-P1-10, CT.DP-P1-P5	-	-	✓	✓	✓	✓	✓	-	✓

Table 2. Summary of Five Safes framework principle protection against privacy violations.

Scope	Breach Notification	Consent	Fairness	Lawfulness	Transparency	Data Minimisation	Purpose Limitation	Accountability	Confidentiality
Safe People	-	-	-	✓	-	✓	✓	✓	✓
Safe Data	-	-	✓	✓	-	✓	✓	-	✓
Safe Project	-	✓	✓	✓	✓	✓	✓	✓	-
Safe Settings	-	-	✓	✓	✓	-	-	-	✓
Safe Outputs	✓	✓	✓	✓	✓	-	-	✓	-

6.3. Implications of the Study

In the broader context, this research highlights the critical importance of organisations adhering to privacy regulations and developing comprehensive privacy policies. This study shows the prevalent challenges within privacy policies, where the absence of technical guidance can inadvertently lead to privacy breaches. To address this pressing issue, this research studied the concept of industrial privacy frameworks, which offer organisations invaluable technical guidelines and best practices that can be customised to their specific needs. These frameworks not only ensure compliance with legal mandates but also effectively mitigate the risk of unintentional privacy breaches. However, a notable shortcoming of these frameworks is their lack of specificity in recommending particular privacy-protection methods. For example, methods like differential privacy, a standard privacy preservation method used by tech giants such as Google [66] and Apple [67] to anonymise data before it reaches their servers, are not explicitly discussed. Differential privacy is a crucial technology in the current data-centric environment, providing a mathematical guarantee of individual privacy while allowing for the collection of useful aggregation of data. Its absence in the guidance provided by frameworks like NIST points to a gap in addressing the practical implementation of advanced privacy-preserving techniques.

This study highlights the need for organisations to go beyond legal compliance with privacy regulations and to develop comprehensive privacy policies addressing prevalent challenges. It reveals that privacy policies often focus on regulatory compliance but lack technical depth, leading to potential breaches. Adopting industrial privacy frameworks is proposed in this study because it offers crucial technical guidance and adaptable best practices to rectify existing strategy limitations. Moreover, this study emphasises the inadequacies of current privacy regulations and user awareness, advocating for more sophisticated, technically grounded privacy management. A case study demonstrates achieving a balance between regulatory compliance and effective data privacy management, showing how these frameworks' adaptable solutions can enhance organisational privacy protection and serve as a model for effective data privacy governance.

This research serves as a pivotal link between regulatory compliance and pragmatic data privacy management, providing a tangible and adaptable solution to enhance privacy protection within organisations. In the context of the Facebook and Cambridge Analytica case, these findings align with the NIST privacy framework's emphasis on transparency, clear communication with users, and accountability, highlighting the framework's potential to address the privacy violations that occurred. An evaluation of the NIST and Five Safes privacy framework is demonstrated in Table 3. In our study, we conducted an analysis of the NIST and Five Safes privacy frameworks against the GDPR's key privacy regulation principles based on a thorough examination of official documents, articles and industry reports. The tabular formats in our paper are a synthesis of these findings, where the presence of a tick mark signifies that the framework addresses a specific principle. This approach was chosen to provide a clear, comparative overview and is reflective of the extensive review and discussion presented in our study.

Table 3. Evaluation of NIST and Five Safes privacy framework.

Category	NIST Privacy Framework	Five Safes Privacy Framework
Breach notification	Governance, Control	Safe Output
Consent	Communicate, Governance	Safe Project, Safe Outputs
Fairness	Communicate, Control	Safe Data, Safe Project, Safe Outputs
Lawfulness	Communicate, Governance, Protect, Control	Safe People, Safe Data, Safe Project, Safe Settings
Transparency	Communicate, Control	Safe Project
Data minimisation	Communicate, Protect, Control	Safe People, Safe Data
Purpose limitation	Communicate, Governance, Protect, Control	Safe People, Safe Data, Safe Project
Accountability	Governance, Protect	Safe People, Safe Project, Safe Outputs
Confidentiality	Governance, Protect, Control	Safe People, Safe Data, Safe Settings

7. Conclusions

In conclusion, the growing trend of organisational collaboration with third-party entities for data acquisition, while enhancing decision-making and operational efficiency, also introduces significant privacy concerns. These external partners often engage in the collection and utilisation of personal and sensitive data, necessitating compliance with government privacy regulations to protect user privacy through robust privacy measures. However, third-party organisations may have varying policies, emphasising the need for comprehensive privacy frameworks that, in conjunction with legal regulations, emphasise standardisation, privacy management, and risk assessment. Such frameworks offer holistic approaches to address the challenges arising in data-sharing partnerships.

In this study, we conducted an analysis of privacy breaches in the Facebook and Cambridge Analytica case with a specific focus on issues related to data collection, transparency, and accountability. We compared the NIST and Five Safes privacy frameworks to assess their potential to mitigate these privacy violations. Our analysis reveals that these frameworks serve distinct but complementary purposes, both offering valuable guidance in addressing the complexities of this scenario.

Therefore, this study urges organisations, policymakers, and individuals to actively engage in dialogues and initiatives that promote a culture of privacy that is deeply rooted in ethical practices and informed by robust frameworks like NIST and Five Safes.

In future work, there are several research directions to be explored in the practical implementation and adoption of privacy frameworks like NIST and Five Safes in organisations, which focuses on the real-world challenges and solutions in integrating these frameworks into data-sharing practices. Furthermore, the exploration of emerging technologies such as federated learning, implementation of differential privacy, and homomorphic encryption, and their impact on data privacy could present rich knowledge for future studies. These areas of research are crucial for advancing our understanding of how to use these cutting-edge technologies with robust privacy measures, thus developing a more secure and privacy-aware digital environment.

Author Contributions: Conceptualisation, S.R.G., T.Z., M.B. and Y.J.; formal analysis, S.R.G.; investigation, S.R.G.; writing—original draft preparation, S.R.G.; writing—review and editing, S.R.G., T.Z., M.B. and Y.J.; supervision, T.Z. All authors have read and agreed to the published version of the manuscript.

Funding: The work has been supported by the Cyber Security Research Centre Limited, whose activities are partially funded by the Australian Government's Cooperative Research Centres Programme P23_000_19_0027.

Conflicts of Interest: The authors Michael Bewong and Yinhao Jiang were employed by Charles Sturt University. The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Mariani, M.; Baggio, R.; Fuchs, M.; Höepken, W. Business intelligence and big data in hospitality and tourism: A systematic literature review. *Int. J. Contemp. Hosp. Manag.* **2018**, *30*, 3514–3554. [CrossRef]
2. Stieglitz, S.; Mirbabaie, M.; Ross, B.; Neuberger, C. Social media analytics—Challenges in topic discovery, data collection, and data preparation. *Int. J. Inf. Manag.* **2018**, *39*, 156–168. [CrossRef]
3. Nussbaum, E.; Segal, M. Privacy vulnerabilities of dataset anonymization techniques. *arXiv* **2019**, arXiv:1905.11694.
4. Isaak, J.; Hanna, M.J. User data privacy: Facebook, Cambridge Analytica, and privacy protection. *Computer* **2018**, *51*, 56–59. [CrossRef]
5. Aljeraisy, A.; Barati, M.; Rana, O.; Perera, C. Privacy laws and privacy by design schemes for the internet of things: A developer's perspective. *ACM Comput. Surv. Csur* **2021**, *54*, 1–38. [CrossRef]
6. Arellano, A.M.; Dai, W.; Wang, S.; Jiang, X.; Ohno-Machado, L. Privacy policy and technology in biomedical data science. *Annu. Rev. Biomed. Data Sci.* **2018**, *1*, 115–129. [CrossRef] [PubMed]
7. Koops, B.-J.; Leenes, R. Privacy regulation cannot be hardcoded. A critical comment on the 'privacy by design'provision in data-protection law. *Int. Rev. Law Comput. Technol.* **2014**, *28*, 159–171. [CrossRef]

8. Okoyomon, E.; Samarin, N.; Wijesekera, P.; Elazari Bar On, A.; Vallina-Rodriguez, N.; Reyes, I.; Feal, Á.; Egelman, S. On the ridiculousness of notice and consent: Contradictions in app privacy policies. In Proceedings of the Workshop on Technology and Consumer Protection (ConPro 2019), in Conjunction with the 39th IEEE Symposium on Security and Privacy, San Francisco, CA, USA, 20–22 May 2019.
9. Zhao, K.; Zhan, X.; Yu, L.; Zhou, S.; Zhou, H.; Luo, X.; Wang, H.; Liu, Y. Demystifying Privacy Policy of Third-Party Libraries in Mobile Apps. In Proceedings of the 2023 IEEE/ACM 45th International Conference on Software Engineering (ICSE), Melbourne, Australia, 14–20 May 2023; pp. 1583–1595.
10. Afriat, H.; Dvir-Gvirsman, S.; Tsuriel, K.; Ivan, L. NThis is capitalism. It is not illegal: Users' attitudes toward institutional privacy following the Cambridge Analytica scandal. *Inf. Soc.* **2020**, *37*, 115–127. [CrossRef]
11. Brown, A.J. Should I stay or should I leave?: Exploring (dis) continued Facebook use after the Cambridge Analytica scandal. *Soc. Media+ Soc.* **2020**, *6*, 2056305120913884. [CrossRef]
12. Yang, T.-M.; Maxwell, T.A. Information-sharing in public organizations: A literature review of interpersonal, intra-organizational and inter-organizational success factors. *Gov. Inf. Q.* **2011**, *28*, 164–175. [CrossRef]
13. Pearson, S. Privacy Management in Global Organisations. In *Communications and Multimedia Security, Proceedings of the IFIP International Conference on Communications and Multimedia Security, Canterbury, UK, 3–5 September 2012*; De Decker, B., Chadwick, D.W., Eds.; Springer: Berlin/Heidelberg, Germany, 2012; pp. 217–237.
14. Chua, H.N.; Herbland, A.; Wong, S.F.; Chang, Y. Compliance to personal data protection principles: A study of how organizations frame privacy policy notices. *Telemat. Inform.* **2017**, *34*, 157–170. [CrossRef]
15. Morton, A.; Sasse, M.A. Privacy is a process, not a PET: A theory for effective privacy practice. In Proceedings of the 2012 New Security Paradigms Workshop, Bertinoro, Italy, 18–21 September 2012; pp. 87–104.
16. Bargh, M.S.; van de Mosselaar, M.; Rutten, P.; Choenni, S. On Using Privacy Labels for Visualizing the Privacy Practice of SMEs: Challenges and Research Directions. In Proceedings of the DGO 2022: The 23rd Annual International Conference on Digital Government Research, Virtual Event, 15–17 June 2022; pp. 166–175.
17. Feasey, R.; de Streel, A. *Data Sharing for Digital Markets Contestability: Towards a Governance Framework*; Centre on Regulation in Europe asbl (CERRE): Brussels, Belgium, 2020.
18. Mohan, J.; Wasserman, M.; Chidambaram, V. Analyzing GDPR Compliance through the Lens of Privacy Policy. In *Heterogeneous Data Management, Polystores, and Analytics for Healthcare, Proceedings of the VLDB 2019 Workshops, Poly and DMAH, Los Angeles, CA, USA, 30 August 2019*; Springer: Cham, Switzerland, 2019; pp. 82–95.
19. Schwaig, K.S.; Kane, G.C.; Storey, V.C. Compliance to the fair information practices: How are the Fortune 500 handling online privacy disclosures? *Inf. Manag.* **2006**, *43*, 805–820. [CrossRef]
20. Zaeem, R.N.; German, R.L.; Barber, K.S. Privacycheck: Automatic summarization of privacy policies using data mining. *ACM Trans. Internet Technol. TOIT* **2018**, *18*, 1–18. [CrossRef]
21. McDonald, A.M.; Cranor, L.F. The cost of reading privacy policies. *Isjlp* **2008**, *4*, 543.
22. Griggio, C.F.; Nouwens, M.; Klokmose, C.N. Caught in the Network: The Impact of WhatsApp's 2021 Privacy Policy Update on Users' Messaging App Ecosystems. In Proceedings of the CHI '22: Conference on Human Factors in Computing Systems, New Orleans, LA, USA, 29 April–5 May 2022; pp. 1–23.
23. Reisinger, T.; Wagner, I.; Boiten, E.A. Security and privacy in unified communication. *ACM Comput. Surv. CSUR* **2022**, *55*, 1–36. [CrossRef]
24. Wong, J.C. Uber Concealed Massive Hack That Exposed Data of 57m Users and Drivers. Available online: https://www.theguardian.com/technology/2017/nov/21/uber-data-hack-cyber-attack (accessed on 23 September 2023).
25. Anthonysamy, P.; Rashid, A.; Chitchyan, R. Privacy requirements: Present & future. In Proceedings of the 2017 IEEE/ACM 39th International Conference On Software Engineering: Software Engineering in Society Track (ICSE-SEIS), Bueons Aires, Argentina, 20–28 May 2017; pp. 13–22.
26. Schwee, J.H.; Sangogboye, F.C.; Salim, F.D.; Kjærgaard, M.B. Tool-chain for supporting Privacy Risk Assessments. In Proceedings of the 7th ACM International Conference on Systems for Energy-Efficient Buildings, Cities, and Transportation, Virtual Event, 18–20 November 2020; pp. 140–149.
27. Andreou, A.; Goga, O.; Loiseau, P. Identity vs. attribute disclosure risks for users with multiple social profiles. In Proceedings of the 2017 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining, Sydney, Australia, 31 July–3 August 2017; pp. 163–170.
28. Hittmeir, M.; Mayer, R.; Ekelhart, A. A baseline for attribute disclosure risk in synthetic data. In Proceedings of the Tenth ACM Conference on Data and Application Security and Privacy, New Orleans, LA, USA, 16–18 March 2020; pp. 133–143.
29. Li, N.; Qardaji, W.; Su, D.; Wu, Y.; Yang, W. Membership privacy: A unifying framework for privacy definitions. In Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security, Berlin, Germany, 4–8 November 2013; pp. 889–900.
30. Hayes, A. Business. Available online: https://www.investopedia.com/terms/b/business.asp (accessed on 22 February 2022).
31. Kenton, W. Third Party. Available online: https://www.investopedia.com/terms/t/third-party.asp (accessed on 22 February 2022).
32. Kenton, W. Customer. Available online: https://www.investopedia.com/terms/c/customer.asp (accessed on 22 February 2022).
33. Chen, J. Business-to-Business. Available online: https://www.investopedia.com/terms/b/btob.asp (accessed on 22 February 2022).
34. Lucking-Reiley, D.; Spulber, D.F. Business-to-business electronic commerce. *J. Econ. Perspect.* **2001**, *15*, 55–68. [CrossRef]

35. Tamplin, T. Business to Consumer (B2C) Meaning. Available online: https://learn.financestrategists.com/finance-terms/b2c/ (accessed on 11 March 2022).
36. Norris, J. Types of Customer Data: Definitions, Value, Examples. Available online: https://www.the-future-of-commerce.com/2021/04/23/types-of-customer-data-definition-examples/ (accessed on 11 March 2022).
37. Freedman, M. How Businesses Are Collecting Data (And What They're Doing with It). Available online: https://www.businessnewsdaily.com/10625-businesses-collecting-data.html (accessed on 11 March 2022).
38. Li, Z.S.; Werner, C.; Ernst, N.; Damian, D. Towards privacy compliance: A design science study in a small organization. *Inf. Softw. Technol.* **2022**, *146*, 106868. [CrossRef]
39. Pearson, S.; Benameur, A. Privacy, security and trust issues arising from cloud computing. In Proceedings of the 2010 IEEE Second International Conference on Cloud Computing Technology and Science, Indianapolis, IN, USA, 30 November–3 December 2010; pp. 693–702.
40. Greenaway, K.E.; Chan, Y.E.; Crossler, R.E. Company information privacy orientation: A conceptual framework. *Inf. Syst. J.* **2015**, *25*, 579–606. [CrossRef]
41. Chinchih, C.; Frey, C.B.; Presidente, G. *Privacy Regulation and Firm Performance: Estimating the GDPR Effect Globally*; Oxford Martin School: Oxford, UK, 2022.
42. Sun, Y.; Lo, F.P.-W.; Lo, B. Security and privacy for the internet of medical things enabled healthcare systems: A survey. *IEEE Access* **2019**, *7*, 183339–183355. [CrossRef]
43. Davari, M.; Bertino, E. Access control model extensions to support data privacy protection based on GDPR. In Proceedings of the 2019 IEEE International Conference on Big Data (Big Data), Los Angeles, CA, USA, 9–12 December 2019; pp. 4017–4024.
44. Wolford, B. GDPR Compliance Checklist for US Companies. Available online: https://gdpr.eu/compliance-checklist-us-companies/ (accessed on 29 September 2023).
45. Naik, N.; Jenkins, P. Your identity is yours: Take back control of your identity using GDPR compatible self-sovereign identity. In Proceedings of the 2020 7th International Conference on Behavioural and Social Computing (BESC), Bournemouth, UK, 5–7 November 2020; pp. 1–6.
46. Perera, C.; McCormick, C.; Bandara, A.K.; Price, B.A.; Nuseibeh, B. Privacy-by-design framework for assessing internet of things applications and platforms. In Proceedings of the 6th International Conference on the Internet of Things, Stuttgart, Germany, 7–9 November 2016; pp. 83–92.
47. ABS Five Safes Framework. Available online: https://www.abs.gov.au/about/data-services/data-confidentiality-guide/five-safes-framework (accessed on 23 September 2023).
48. Nist Nist Privacy Framework: A Tool For Improving Privacy Through Enterprise Risk Management, Version 1.0. Available online: https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.01162020.pdf (accessed on 2 March 2021).
49. Carter, T.; Kroll, J.A.; Michael, J.B. Lessons learned from applying the NIST privacy framework. *IT Prof.* **2021**, *23*, 9–13. [CrossRef]
50. Desai, T.; Ritchie, F.; Welpton, R. Five Safes: Designing Data Access for Research. 2016. Available online: https://www2.uwe.ac.uk/faculties/bbs/Documents/1601.pdf (accessed on 25 September 2023).
51. Schwartz, M. Morgan Stanley's Hard Drive Destruction Investment Failure. Available online: https://www.bankinfosecurity.com/blogs/morgan-stanleys-hard-drive-destruction-investment-failure-p-3286#:~:text=Another%20surprise%20is%20the%20five,occurred%20at%20the%20banking%20giant (accessed on 25 August 2023).
52. Ahmed, J.; Yildirim, S.; Nowostaki, M.; Ramachandra, R.; Elezaj, O.; Abomohara, M. GDPR compliant consent driven data protection in online social networks: A blockchain-based approach. In Proceedings of the 2020 3rd International Conference on Information and Computer Technologies (ICICT), San Jose, CA, USA, 9–12 March 2020; pp. 307–312.
53. Hinds, J.; Williams, E.J.; Joinson, A.N. It wouldn't happen to me: Privacy concerns and perspectives following the Cambridge Analytica scandal. *Int. J. Hum. Comput. Stud.* **2020**, *143*, 102498. [CrossRef]
54. Perera, H.; Hussain, W.; Mougouei, D.; Shams, R.A.; Nurwidyantoro, A.; Whittle, J. Towards integrating human values into software: Mapping principles and rights of GDPR to values. In Proceedings of the 2019 IEEE 27th International Requirements Engineering Conference (RE), Jeju, Republic of Korea, 23–27 September 2019; pp. 404–409.
55. Sandoval-Almazan, R.; Valle-Cruz, D. Sentiment analysis of facebook users reacting to political campaign posts. *Digit. Gov. Res. Pract.* **2020**, *1*, 1–13. [CrossRef]
56. Shipman, F.M.; Marshall, C.C. Ownership, privacy, and control in the wake of Cambridge analytica: The relationship between attitudes and awareness. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems, Honolulu, HI, USA, 25–30 April 2020; pp. 1–12.
57. Rathi, R. Effect of Cambridge Analytica's Facebook Ads on the 2016 US Presidential Election. Available online: https://towardsdatascience.com/effect-of-cambridge-analyticas-facebook-ads-on-the-2016-us-presidential-election-dacb5462155d (accessed on 23 September 2023).
58. Grassegger, H.; Krogerus, M. The Data that Turned the World Upside Down. Available online: https://www.vice.com/en/article/mg9vvn/how-our-likes-helped-trump-win (accessed on 23 September 2023).
59. Smith, G. Artificial Intelligence and the privacy paradox of opportunity, Big Data and the Digital universe. In Proceedings of the 2019 International Conference on Computational Intelligence and Knowledge Economy (ICCIKE), Dubai, United Arab Emirates, 11–12 December 2019; pp. 150–153.

60. Bay, M. Social media ethics: A Rawlsian approach to hypertargeting and psychometrics in political and commercial campaigns. *ACM Trans. Soc. Comput.* **2018**, *1*, 1–14. [CrossRef]
61. Dwork, C. Differential privacy. In *International Colloquium on Automata, Languages, and Programming*; Springer: Berlin/Heidelberg, Germany, 2006; pp. 1–12.
62. Wagner, P. *Data Privacy-The Ethical, Sociological, and Philosophical Effects of Cambridge Analytica*; University of Arizona-College of Applied Science and Technology: Sierra Vista, AZ, USA, 2021.
63. Arora, N.; Zinolabedini, D. *The Ethical Implications of the 2018 Facebook-Cambridge Analytica Data Scandal*; The University of Texas at Austin: Austin, TX, USA, 2019.
64. Deloitte, L. A New Era for Privacy: GDPR Six Months on. Available online: https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/risk/deloitte-uk-riskgdpr-six-months-on.pdf (accessed on 23 September 2023).
65. Hu, M. Cambridge Analytica's black box. *Big Data Soc.* **2020**, *7*, 2053951720938091. [CrossRef]
66. Erlingsson, Ú.; Pihur, V.; Korolova, A. Rappor: Randomized aggregatable privacy-preserving ordinal response. In Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, Scottsdale, AZ, USA, 3–7 November 2014; pp. 1054–1067.
67. Tang, J.; Korolova, A.; Bai, X.; Wang, X.; Wang, X. Privacy loss in apple's implementation of differential privacy on macos 10.12. *arXiv* **2017**, arXiv:1709.02753.