information

MDPI

Article

Big Picture on Privacy Enhancing Technologies in e-Health: A Holistic Personal Privacy Workflow

Stefan Becher *, Armin Gerl *, Bianca Meier and Felix Bölz

Faculty of Computer Science and Mathematics, Chair for Distributed Information Systems, University of Passau, 94032 Passau, Germany; bianca.meier@uni-passau.de (B.M.); felix.boelz@uni-passau.de (F.B.)

* Correspondence: Stefan.Becher@uni-passau.de (S.B.); Armin.Gerl@uni-passau.de (A.G.)

Received: 6 June 2020; Accepted: 5 July 2020; Published: 8 July 2020

Abstract: The collection and processing of personal data offers great opportunities for technological advances, but the accumulation of vast amounts of personal data also increases the risk of misuse for malicious intentions, especially in health care. Therefore, personal data are legally protected, e.g., by the European General Data Protection Regulation (GDPR), which states that individuals must be transparently informed and have the right to take control over the processing of their personal data. In real applications privacy policies are used to fulfill these requirements which can be negotiated via user interfaces. The literature proposes privacy languages as an electronic format for privacy policies while the users privacy preferences are represented by preference languages. However, this is only the beginning of the personal data life-cycle, which also includes the processing of personal data and its transfer to various stakeholders. In this work we define a personal privacy workflow, considering the negotiation of privacy policies, privacy-preserving processing and secondary use of personal data, in context of health care data processing to survey applicable Privacy Enhancing Technologies (PETs) to ensure the individuals' privacy. Based on a broad literature review we identify open research questions for each step of the workflow.

Keywords: formal languages; GDPR; privacy enhancing technologies; privacy languages

1. Introduction

In the age of digitisation the importance of developments such as Big Data, artificial intelligence and Industry 4.0 increases, especially in the health care sector. These technologies are the base to various applications, like patients predictions for an improved staffing [1] or monitoring the health status of elderly people and helping them in emergency situations [2]. To gain information, which is an added value for health care applications, e.g., scientific research, the collection and processing of personal data is essential. However, the privacy of the users must also be respected in the process. This includes, among other things, that they know which data are collected, how they are processed and who receives their personal data. In May 2018, the EU-wide General Data Protection Regulation (GDPR) [3] was enforced to build a legal framework for the whole process of data collection and data processing. To protect the privacy of the users it is specified that for special categories of data, e.g., health data, the explicit consent of the patient to the processing is necessary. The consent of the patient has to be given freely and in an informed way, such that is essential that the patient is presented the processing of his personal data in a transparent and understandable way. Furthermore, the patient may consent only to a subset of the processing purposes he is presented, e.g., Alice allows the processing of her health data for research and her treatment, but not for advertisement purposes. Assuming Alice has given her consent, it has to be ensured that her personal data are only processed for these purposes, even when it is transferred to third parties (for the processing of the personal data in the context of the agreed purposes). The processing entities, e.g., research institutions and hospitals, are accountable

and responsible for the privacy-preserving processing of Alice's data. This includes the application of de-identification methods, e.g., anonymization and pseudonymization, to ensure the individual's privacy. Furthermore, GDPR also includes various rights for the individual, i.e., Data Subject Rights which can be claimed by every individual. If they fail to comply to these duties, significant fines are possible ([3], Art. 83). All these rights and duties, i.e., legal requirements for processing personal data, require in-depth knowledge and immense effort to be implemented. The use of Privacy Enhancing Technologies (PETs) can overcome this complexity. Especially, the establishment of standard definitions of policies and processing rules in an holistic approach can ease the necessary effort to manage privacy policies, e.g., consent decisions, and their enforcement.

To establish the importance of privacy, i.e., a globally common technical understanding of privacy, we will elaborate the global pandemic of COVID-19 in the following. A very recent example shows that these two aspects are not always easy to reconcile and that some mediation between them is necessary: The disease (COVID-19) determines the life of people all over the world [4]. There are over 6.28 million confirmed infected people and over 379,000 who died as a result with or without previous illness (as of 04.06.2020).Since there is no medication or vaccine for this novel virus as yet, measures must be taken worldwide to protect human life [5]. For example in Israel, infected people are tracked by using the location data of their smartphone and also their credit card data [6]. The USA intends to track and isolate the spread of the virus by using methods of artificial intelligence, like face recognition, thermal images and heat measurements to detect a increased body temperature and raising an alarm if someone is conspicuous [7]. Additionally, in Germany there is a discussion about the introduction of a COVID-19 tracking app. Next to the discussion about the medical effectiveness, there are concerns about the legal implications regarding privacy. An approval is only possible if the data collection and processing by this app complies with the rules of the GDPR (to be more specific the national data protection framework based on the GDPR). These examples show that the respective governments have taken measures, which can affect the personal rights of each individual. In addition, personal and sensitive data are already collected in many cases. However, it is not clear how or if they are protected. These measures should help to make the spreading of the virus more controllable and to determine chains of infection. It can also help to ensure that the health care system is not overloaded such that as many people as possible receive an appropriate treatment [8,9].

Moreover, the development of a vaccine against COVID-19 requires worldwide cooperation of data sharing and processing [10]. The specific use case, on which we want to focus in this paper, is the collection and processing of personal data by COVID-19 apps for a global tracking and measurement of the spread of COVID-19 and its prevention (see Figure 1).

Many countries are currently trying to develop their own COVID-19 app or are already using it, examples are the Stopp Corona App (Austrian Red Cross, Vienna, Austria), Hamagen (Ministry of Health, Jerusalem, Israel) and Stay Home Safe (Hongkong, China) [11]. This work focuses on GDPR as a legal framework, because it has worldwide impact. However, the legal frameworks from other countries must also be considered, like the Health Insurance Portability and Accountability Act of 1996 (HIPAA) in the USA [12].

These apps work in a very similar way. The location data of each user are recorded and with the help of Bluetooth sensors such that it is possible to identify other people in the immediate vicinity. If someone is now tested positive for COVID-19, all people who have been in contact with the infected person are notified and sent to go into self-quarantine.

The current status is as follows: Each country has its own app and therefore its own data on which analyses can be carried out. These serve as a basis for science and research. The much more important point, however, would be to combine the data from different countries and thus create a data warehouse, which puts these different sources in one centralised database. The benefit of this sharing would be, that researches have a larger data volume. This leads to extended possibilities, e.g., experiments become more representative for the pandemic or global cooperative strategies for the limitation of the spread of the virus can be implemented. Likewise, different measures in different

countries and their consequences can be observed and new insights can be drawn from them to develop a 'best practice'.

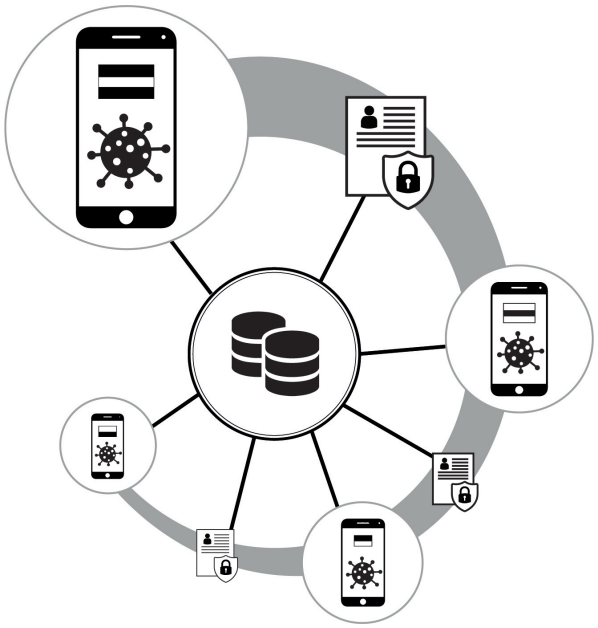

Figure 1. Data collection and processing by COVID-19 apps leading to a common and centralised database for tracking the spread of COVID-19.

Next to these technical possibilities, that arise with collecting, processing and sharing data, the side of protecting the users personal rights also has to be considered. In this use case, every user has to give his explicit consent to process his personal data, like location data (GPS) or phone numbers, for a specific purpose like medical research or tracking. With this small example, it comes to mind, that privacy has an important part for implementing and using such apps. The user must have the control over his personal data all the time. Especially, the user of the app should have the possibility to express his own preferences for consenting to purposes of processing or the data that is collected and shared. In addition, the data collecting and processing steps have to be presented in a transparent and clear way, while they have to be enforced for processing and secondary use. To standardize this process and make it even more controllable for the user and the controller, e.g., hospital or research institute, standard definition of privacy policies and the users' preferences, i.e., privacy and preference languages, are necessary. In addition policies for transfer of data between controllers should be formalised to streamline the compliance to the legal frameworks.

The main contributions in this work are:

1. The definition of a holistic 'big picture' of the personal privacy workflow in context of health care
2. A broad state-of-the-art analysis of technologies for each sub-sector of the personal privacy workflow which is relevant for health data
3. Highlighting open research gaps in existing approaches motivating future research opportunities

In this work, we want to focus on how the GDPR and the privacy constructs can be included in the data processing and collecting workflow. Before an overview of the privacy workflow, description of the meaning of single workflow elements and how they work together, is given in Section 3, related work and case studies are presented in Section 2. Furthermore, this workflow can be divided in three different parts, which are explained in more detail in the Sections 4–6. After this analysis, the current issues, research gaps and possible research opportunities for future work are be elaborated in Section 7. Lastly, Section 8 concludes this work.

2. Related Work

There are various relevant use cases demonstrating the importance of a more formal privacy processing supported by a well-defined privacy workflow. First, there is a necessity of privacy enhancing technologies in the context of health care, i.e., IoT-Health [13] where personal data must only be used to improve a patients treatment. Next, artificial intelligence systems are deployed at contexts which provide an access to Big Data. Unfortunately, there are some privacy threads which are naturally induced by this area of science. For example, the training data can be altered in such a way that the system provides the attackers desired results after launch of the system. As another example, the knowledge of a speech recognition system can be used to artificially produce altered phrases of a real-world voice [14]. Next, autonomous vehicles need access to a variety of different sensors. Those sensors capture information necessary to provide a secure operation of the vehicle. However, the high amount of collected data also includes privacy related data which must not be used for any other purpose [15]. The rise of information systems, i.e., smart homes, modern warehouses, as for autonomous vehicles, results in an interconnection of different sensors working together while collecting all kind of data [16]. In addition, the context of of cloud computing [17] can be named, where all uploaded data possibly might be privacy relevant. For instance, more and more mobile apps tend to utilise cloud services which leads to several privacy issues [18]. Another example would be the outsourcing of services towards a cloud [19].

As shown, there are many fields of application in the context of a well-defined privacy workflow. In the following, we will provide an area-specific overview of privacy-related work, beginning with related survey in the healthcare sector.

Hathaliya and Tanwar [20] show that the healthcare industry has been revolutionised to now using cloud computing, fog computing and Internet of Things (IoT) technologies whereas technologies are surveyed to provide security and privacy. The use of wearable devices by patients for collecting data and its transfer to various stakeholders is common to this work. Similarly, Iqridar Newaz et al. [21] survey security and privacy issues for healthcare and focus on possible attacks due to design flaws in medical systems and review potential solutions. Majdoubi and Bakkali [22] survey smart city technologies, e.g., IoT, Big Data and the cloud, showing various privacy concerns. They conclude that the different stakeholders, which add complexity, and their individual privacy concerns are not well presented in the literature.

However, in other sectors privacy is also an essential topic. Deep et al. [23] survey security and privacy issues in the context of IoT. They argue that these devices become more and more spread and thus the inherit security and privacy risks become more eminent. In their survey they examine the security and privacy issues for each layer of the IoT protocol stack and proposed solutions, including blockchain technologies. Ferrag et al. [24] present open challenges on security and privacy issues for an IoT-based agriculture use case. Within their survey they also focus on blockchain-based technologies as a solution. Weixiong et al. [25] survey security and privacy issues in the context of network communication, especially wireless networks. Their focus lays on the comparison of different encryption mechanisms and how they are perceived by users.

Another aspect of privacy, next to the technical realisation, is the individuals' view on the usage of, e.g., privacy policies. Linden et al. [26] analyzed pre- and post-GDPR privacy policies to unveil the impact of the GDPR in terms of the presentation, textual features, coverage, compliance and specificity of privacy languages. They showed that due to GDPR privacy policies have been revamped covering the privacy rights and required information well, but also became significantly longer. Ebert et al. [27] investigate which privacy concerns and privacy preferences individuals have for two different contexts, i.e., loyalty cards and fitness tracking. They showed that the context may be a significant moderator of the users' concerns and preferences. Additionally, they support the statement that some concerns are more prominent than others for users, e.g., information about Data Subject Rights is more relevant than information on contact persons. Esmaeilzadeh [28] investigates the effects of perceived transparency of privacy policies in an healthcare scenario on the individuals decision to disclose information.

The study shows that the perceived transparency is a significant factor to gain the trust of individuals. Johnson et al. [29] study the impact of the individuals' privacy choice or preference in an online advertising context deriving the cost of an individuals' opt-out choice for the advertisement industry, but also show that opt-out users tend to be more technologically adept.

Especially for privacy policies different attempts have been made to formalise them, each with a different focus. Therefore, surveys on privacy policy languages can be found, which classify them according to various taxonomies. Leicht and Heisel [30] survey privacy policy languages on how well they express the recent data protection regulation, i.e., GDPR. Therefore, they focus on the expressiveness of formalised privacy policies. Kumaraguru et al. [31] classify privacy languages according to their target use cases by differentiating sophisticated access control languages, web privacy policy languages, context sensitive languages, and enterprise privacy policy languages. Kasem-Madani and Meier [32] introduce a multidimensional categorisation for security and privacy languages, whereas their taxonomy differentiates several types of languages, namely Security, accountability, availability, privacy, data carriage, data usage control, and network and device management. Morel and Pardo [33] classify privacy languages according to their features, audience, conditions and content.

To conclude the literature review, it can be seen that privacy is a essential topic in various domains, but especially in the healthcare domain in which sensitive patient data are handled and have to be protected. Furthermore, it can be derived that different stakeholders, e.g., the user and the processing companies, have to be considered for a holistic view on privacy-preserving data processing including third party transfers of personal data. The privacy policy is often the center of attention for privacy research including how it is formalised via privacy policy languages, how it is composed to be transparent, and how it is perceived by individuals. We argue that all these aspects have to be considered for an holistic approach—the personal privacy workflow—to implement privacy in technical systems.

3. Personal Privacy Workflow

To understand the privacy related challenges that have to be overcome, we envision a future health care system (see Figure 2). This future of health care is a combination of various services that are integrated and inter-connected to effectively and efficiently process personal data [34].

The personal privacy workflow details the factors that have to be considered for processing an individual's personal data under consideration of the (personal) privacy-related context for the processing. The individuals preferences determine which processing purposes are agreed to, thus limiting or enabling certain processing of the individual's data. Furthermore, this personal decision-making also affects the transfer of data to third parties. Similarly, policies and contracts between companies have to be considered in addition to the individuals' preferences for privacy. Hereby, the proposed personal privacy workflow differentiates itself from models that considered only the data-flow by also considering the privacy-related context, e.g., the individuals' privacy preferences, the personalised privacy policy, or Technical and Organisational Measures. Thus, not only Privacy Enhancing Technologies (PETs) but also legal frameworks are considered for the personal privacy workflow.

To process personal data, it first has to be collected. Hereby, we envision that personal data and health data are not only gathered by medical examinations, treatments in hospitals or general practitioners, but also collected continuously from various devices and services of the daily life. This includes sensors and smart devices in the personal space—the user environment—of the individual, e.g., the household of the patient, e.g., to detect the activity of the patient or to detect emergency situations [35], or the car which senses stress or fatigue of the driver [36]. In addition, personal devices like fitness or health trackers continuously collect data about various aspects of the individual, from the pulse rate, amount of daily sleep, activity, location (tracking), or blood sugar [37]. These data points, or especially the long-term data can be valuable to optimise ones' personal health, get diagnosed by medical practitioners, treatment, emergency detection and alerting of

ambulances, or personalised offers for health care insurances (deductions or increased fees). Naturally, more sophisticated services and health care solutions will require more or higher quality personal data. The added value of such services can be clearly envisioned, but there are also risks. For example, an individual, Bob, with chronic deceases may not be accepted by health care insurances or only for horrendous fees, which may even lead to the situation that this individual can no longer afford to pay his mandatory medical treatments.

To avoid such a situation, the individual Bob should always be in control over the collection and processing of his personal data, i.e., which of his data are used for which purpose by which entities. These should be defined in the privacy policy of the devices and services that Bob uses. The challenge is that various services and devices exist in the future user environment of Bob. The proper personalisation of each of the privacy policies, the management or revision of the decisions due to updates to the services' privacy policies or the personal preference of Bob, is a major challenge and requires significant effort from the individual.

Therefore, we see it as a necessity that technical solutions support this process. The privacy preferences of Bob should be defined in an electronic format or standard, a preference language, that is then used as a basis to negotiate the privacy policies of all devices. The definition of the privacy preferences requires suitable user interfaces—Preference User Interfaces (P-UIs)—that enable users to define their preferences in a comprehensible way.

However, this scenario not only requires the formalisation of privacy preferences in an electronic format, but also the definition of privacy policies in a machine- and human-readable way. Therefore, lets assume Alice voluntarily decides to download a COVID-19 tracking app. This app needs to access the GPS data of Alice's mobile device but also might ask for optional data regarding the health status of Alice, possible chronic diseases, her age, etc., for further analyses of possible risks of infection. Requested data are used within the controller environment for various processing purposes, e.g., determination of contacts to track down chains of infection, which is documented in the controllers privacy policy (PP). This policy is presented to Alice in a transparent and easily understandable way by specific user interfaces—Policy Negotiation User Interfaces (PN-UIs). These interfaces summarise policy actions and allow for personalisation of the policy, e.g., Alice might refuse to share her age or possible chronic diseases with the service.

According to her preferences consent is given to stated processing purposes and finally Alice agrees to the policy. Therefore, at least her GPS data are shared with the service. At Data Collection (DC) the agreed privacy policy and the personal data of Alice (PP + D) are collected by the controller and stored together. This principle is called sticky policies. Then the policy and the data might be used for several Business Processes (BPs) according to stated policy actions, e.g., location tracking. Assuming that the privacy policy is based on a privacy language, misuse of the data can be technically prevented.

Finally, let us assume Charlie, who was diagnosed with COVID-19, was treated in a hospital and consented to share all of his data for research. This implies that not only the controller who collected the data, i.e., the hospital, has a processing right on them but also other chosen third party controllers. Therefore, Charlie's personal data might be queried by external entities or transferred to them, in order to conduct further analyses based on a large amount of collected patient data. In business scenarios personal data might be traded for money or other assets. These actions are characterised as Data Transfer (DT) within the C2C Environment. So-called Technical and Organisational Measures (TOMs) and Service Level Agreements (SLAs) are used by third party controllers to document what the data are used for and what technical measures are taken to protect them. Therefore, collaboration, security and privacy are defined and ensured among controllers. Furthermore, control over his personal data is granted to Charlie by Data Subject Rights according to GDPR ([3], Art. 12–23).

If personal data are made publicly available for research there are special requirements to protect the data and privacy of the user against potential misuse. For example, a malicious entity might have external knowledge about Charlie, which he could combine with a public data set to infer Charlie's record and therefore find out that he was treated for COVID-19. In addition, the adversary could

consecutively query the database and join the results to infer information which would be hidden otherwise. Therefore, data within these data sets must be protected by anonymization methods to fulfill certain privacy model criteria and in addition measures for Inference Detection and Prevention (ID/P) as well as access control need to be taken.

The proposed personal privacy workflow for health care data models the whole process of Data Collection based on a user's preferences, Business Processes performed on the collected data and Data Transfer to third parties. The difference to data workflows is that during each of the workflow steps various privacy-related requirements or conditions, e.g., denoted in a privacy policy or the user's preferences, have to be considered. These requirements not only define the processing of personal data but also each other, e.g., the preferences of a user influence the individual's agreed-to privacy policy and the individual's privacy policy can influence the requirements for transferring data or whether the individual's data are transferred at all to third parties. In the following, various approaches from the literature will be discussed for each of the environments, of the personal privacy workflow in the context of health care.

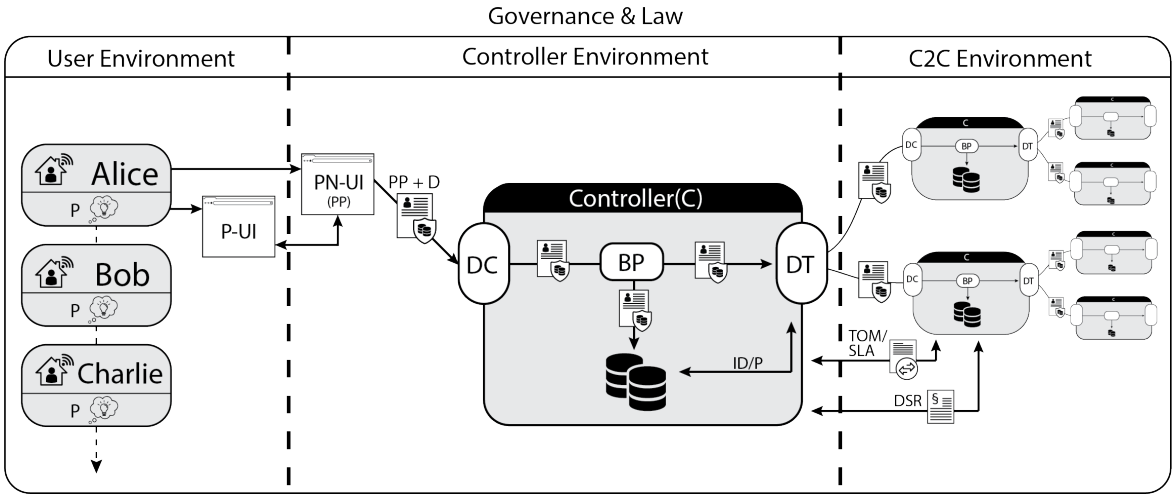

Figure 2. Personal privacy workflow. User preferences (P) are defined over Preference User Interfaces (P-UIs) and matched against privacy policies (PPs) which are presented and negotiated by Policy Negotiation User Interfaces (PN-UIs). Privacy policies and data (PP + D), so-called sticky policies, are used by controllers for Data Collection (DC), Business Processes (BPs) and Data Transfer (DT). At (DT) data are technically protected by Inference Detection and Prevention (ID/P) and legally by Technical and Organisational Measures (TOMs), Service Level Agreements (SLAs) and Data Subject Rights (DSRs).

4. Controller Environment

Within the controller environment, companies, i.e., controllers, offer various services, which might need certain data from the user in order to work properly. Which data are collected as well as for what actions they are used is stated in the controllers' privacy policy (PP). The policy is presented via a policy negotiation user interface (PN-UI) to the user who then can personalise and finally accept it. This creates a legally binding between the user and the controller and allows for processing of the data as stated in the privacy policy which is regulated by GDPR ([3], Arts. 12 and 13). The claimed data has several uses for the controller. On the one hand, it can be analysed to optimise business processes or to send personalised advertisements to users. On the other hand, the data might be shared or traded with other third party controllers, in order to outsource some business processes or simply to sell it. This means, personal data always has (monetary) value for the controller. At the same time, the privacy of the user has to be protected, otherwise, according to GDPR, huge fines might be imposed ([3], Art. 83). Therefore, technical measures have to be taken to protect privacy according

to the privacy policy the user agreed on and to ensure accountability of the controller. This can be achieved by using access control mechanisms and privacy languages.

4.1. Access Control

Access control mechanisms in the field of data privacy are approaches or language frameworks, which limit access to data to solely authorised entities. Therefore, cyber attacks, which aim on compromising personal user data, can be prevented whereby privacy of the user is protected. In the literature various approaches exist to enforce access control (see Table 1). Prominent categories of these approaches are Role-Based Access Control (RBAC), Attribute-Based Access Control (ABAC) and Context-Aware Access Control (CAAC) [38].

Role-Based Access Control [39] grants permissions to access data based on roles which are assigned to users. This model is widespread and often uses hierarchically structured roles to separate user groups. Therefore, it scales very well with huge numbers of users and enables for fast access control. Motta and Furuie [40] proposed a RBAC model for electronic patient records based on organisational roles. It also features contextual information by processing environmental information at access time, like patient relationships.

Attribute-Based Access Control [41] determines access to data or services based on characteristics of the user, i.e., attribute sets. As opposed to RBAC, where soley the role of the user determines if access is granted or not, ABAC takes more attributes of the user into consideration. For example, access could be granted if an emergency situation is happening. This model can utilise several attributes for diverse services, which makes it more flexible than the classic RBAC model. A privacy-aware s-health access control system (PASH) [42] is an example for a framework which implements ABAC. It is specifically designed for smart health scenarios and uses attribute-based encryption of ciphertext policies. This results in partially hidden access policies whereas attributes are encrypted.

Context-Aware Access Control [38] utilises context information, i.e., dynamic information, which can be centred around the user, data resources or the environment. Often, CAAC builds upon RBAC, whereas in addition access can be granted or limited based on constraints like time and location. In addition, the relationship between users can be used to determine access. An intelligent context-aware framework (ICAF) [43] is an access control framework which embeds context into access control policies. Therefore, it features a context model for access control, a situation model for purposes and a policy model to protect data in a dynamic environment. RelBOSS [44] implements Relationship-Aware Access Control (RAAC) as part of CAAC. The framework considers social relationships as dynamic context which is extracted from user profiles. Granularity and strengths of relationships can dynamically be captured and processed at runtime. Furthermore, approaches for Fuzzy Context-Aware Access Control (FCAAC) [45] are suited for imprecise fuzzy data. This concept suits the health sector well, whereas information are present, like 'critical' or 'highly critical' health status. In this case it is hard to draw a precise border between the two conditions. Finally, Kayes et al. [46] proposed a framework, which introduces CAAC within fog computing environments (FB-CAAC). This model addresses missing support of CAAC by enabling flexible access control for distributed cloud services.

Table 1. Overview of selected access control mechanisms from the literature, categorised into Role-Based Access Control, Attribute-Based Access Control and Context-Aware Access Control.

Category	Approach / Framework			
Role-Based Access Control	Motta and Furuie [40]			
Attribute-Based Access Control	PASH [42]			
Context-Aware Access Control	ICAF [43]	RelBOSS [44]	FCAAC [45]	FB-CAAC [46]

These approaches offer various mechanisms to limit access to data to only authorised entities. While RBAC might be less flexible than other approaches, like ABAC, it is still widespread and has shown to be a valid access control model for health care data. In addition, more advanced models, like CAAC, often use RBAC as a basis and therefore achieve dynamic up-to-date access control in cloud systems. Most of the existing approaches for access control access data from centralised sources. However, with the increasing need to being able to access data from multiple distributed sources, these approaches don't scale for this scenario due to latency and processing overheads. Therefore, Kayes et al. [47] proposed an approach which reduces processing overhead by unified policy specifications introduced in a CAAC framework. This approach enables efficient access to data from multiple sources and therefore closes the gap for this research field.

Originated form access control languages, privacy languages not only consider technical enforcement of access control, but also legal frameworks (i.e., GDPR [3] in EU countries). Therefore, personalisation and understandability of privacy policies, consent management and enforcement of privacy-preserving technologies, e.g., privacy models, are additionally implemented.

4.2. Privacy Language

Privacy policies are usually presented to users in form of a huge plain text document. Besides the fact that this representation is not user friendly at all, it also has further technical and legal downsides.

Plain text policies are designed to be read and understood by humans, but not machines. Although techniques exist that extract information from plain text documents, these techniques are not perfect and the information cannot be interpreted exactly by a machine. Therefore, no technical measures can be taken to ensure that data are handled exactly as stated in the policy. There are only limited possibilities to personalise plain text policies which we already defined as a necessary feature within the user environment. To overcome the stated problems, privacy languages can be utilised for the definition of processible privacy policies.

Various privacy languages have been proposed in the literature [30] which can be roughly categorised into access control privacy languages, which only grant usage rights to authorised entities, transparency privacy languages, which aim to inform about policy actions, or de-identification privacy languages, which technically enforce masking of data (see Table 2). In the following, some examples are given for each category.

Some representatives for access control privacy languages are Ponder [48], Rei [49], SecPAL [50], XACML [51] and PPL [52]. Access control privacy languages ensure only authorised access and usage of data by binary decisions, i.e., yes or no. Therefore, unauthorized misuse of data is technically prevented.

Ponder [48] features obligation based policies, which consist of event triggered rules. These hierarchically structured rules are used to gain role-based access control in networks. Based on the role an entity has, access to certain data can be granted to other entities. In health care access to data could be limited to only the responsible doctor or researcher by giving them the proper role.

Policies implemented by Rei [49] consist of tuples of action and condition which enforce domain specific restrictions. The policy is formally defined in RDF, by using parts of deontic logic. To enforce role-based access control, Data Subjects and policy objects get unique identifiers. Rei offers similar advantages for health related scenarios like Ponder by utilising the identifiers.

SecPAL [50] is designed to work within large scale distributed systems. Therefore, its logic-based policies, which are defined as predicate triples, feature complex access control mechanisms. Fine-grained access control is enabled by mapping requests to logical authorisation queries. This could allow specific access to patient data based on several criteria in large scale health databases.

XACML [51] is a standardised access control language, which allows for distributed policy definition. Rule elements, which consist of conditions and effects, are joined with policy elements to gain fine-grained access rules. Multiple access rules form a policy set. Its distributed policy definition could fit for large data warehouses shared between several nations, e.g., COVID-19 data.

PPL [52] features the same access control mechanisms as XACML as it builds up on it. On top PPL adds further control mechanisms for secondary data usage, which only allows data to be used for purposes as defined in the policy. Therefore, personalised policies are stored alongside the user data. This concept is called sticky policies. In health scenarios PPL offers the same advantages as XACML while also adding additional features to secure secondary data usage.

Examples for transparency privacy languages are P3P [53] and CPExchange [54]. Transparency privacy languages have a focus on informing users about policy actions in a transparent and understandable way. Therefore, policies are usually structured based on purposes.

The first real contender for a privacy language was P3P [53], which was standardised and also used in real applications. Its policies are defined in machine-readable XML format. Besides this XML policy a separate human-readable policy is needed. P3P mainly aims to inform users about policy actions. Another focus of the language is to ease policy negotiation by automatically matching APPEL [55] preferences. In health care standardised policies based on P3P could be utilised for automated consent in order to ease the process for the patient to release his data for research. Even though, researchers have shown several drawbacks of P3P, like a limited vocabulary, lack of formal semantics and inconsistencies within policies [56,57].

CPExchange [54] builds upon P3P and aims to exchange private data as meta-data attached to business data. The exchange of information is supported by a privacy model as well as a fine-grained authorisation mechanism. Information regarding the participating entities of a trade and jurisdictional information is stored alongside the elements inherited from P3P. Therefore, it is legally ensured that released health information is only used as the user consented to.

Finally, de-identification privacy languages are less common than the other categories. While several privacy languages have been proposed which technically enforce privacy actions, the definition and enforcement of de-identification methods is almost always absent. Therefore, EPAL [58], PPL [52], P2U [59] and SPECIAL [60] technically enforce privacy actions while to the best of our knowledge only LPL [61] also implements de-identification methods. SPECIAL considers de-identification conceptually but doesn't enforce it.

LPL [61] is a GDPR compliant privacy language, which also considers technical enforcement of privacy policies. Therefore, sticky privacy policies are layered to ensure that no agreements between the user and the controller are lost once data are transferred to third parties. In this cases new layers of agreements can't be less strict than the already existing ones. LPL also features fine-grained personalisation of policies, allowing the user to set minimum anonymization degrees for each attribute of the collected data. When the data are used by the controller or some third party they are at least anonymized to the degree the user demanded. Therefore, patients might be more often willing to release their health data for research, as accessed data are always anonymized.

Table 2. Selection of privacy languages from the literature categorised into access control privacy languages, transparency privacy languages and de-identification privacy languages.

Category	**Privacy languages**				
Access control privacy languages	Ponder	Rei	SecPAL	XACML	PPL
Transparency privacy languages	P3P	CPExchange			
De-identification privacy languages	LPL				

To sum up, privacy policies do not only have to be personalisable and transparent to the user, but also enforceable by the responsible company.

Privacy languages are a valid approach to ensure the above mentioned properties and can be utilised to technically document that controllers comply with the given legal framework of the GDPR. Therefore, accountability of the controllers is given.

4.3. Accountability

Controller (companies), e.g., COVID-19 app providers as well as data integrators, are responsible for the processing of personal data. GDPR ([3], Arts. 24 and 32) states that appropriate mechanisms have to be put in place to consider privacy aspects throughout the development of technical systems (applications, services, data warehouses) ([3], Privacy by Design Art. 25). Furthermore, the principle of privacy by default states that the default settings, e.g., when a user registers for a service, have to be privacy-friendly and can only be relaxed in a secondary step by the user, e.g., by explicit consent to the processing of their personal data for specific purposes. Controllers are responsible and accountable for complying to these rules and have to document their actions and processes to achieve those ([3], Art. 5). Furthermore, they can be audited by supervisory authorities ([3], Art. 51).

Thus, a privacy language, expressing the privacy policy as the core documentation for the agreed purposes and data that is allowed to be processed for each individual, can be used to enable accountability. Therefore, any privacy language, e.g., P3P [53], EPAL [58], P2U [59], PPL [52], SPECIALs' Usage Policy Language [60] or LPL [61] could be used while also privacy languages have been proposed that have a special focus on accountability, like AAL [62] and A-PPL [63] (see Table 3). Furthermore, if the decision is documented or logged in an immutable and transparent way, audits may be eased supporting the tasks of supervisory authorities. To achieve such an immutable and transparent documentation the blockchain technology seems suitable. Here, each decision or update to a privacy policy could be stored transparently and immutable if sensitive information is protected, e.g., by encryption. Vargas and Camilo [64] developed such an approach to conduct consent management via blockchain. Therefore, participants within the network are defined as stated in the GDPR, namely data subject, data controller and data processor. This approach already considers the integration of supervisory authorities as a fourth participant even though the concept is not tested for real scenarios. Dorri et al. [65] proposed an optimization of blockchain for IoT and smart home scenarios which showed promising results. This means the technology could also be used in resource limited scenarios and mobile devices and therefore cover almost every area where data protection is relevant. One open question that remains is if blockchain scales for large data sets, e.g., ones containing health data. Angraal et al. [66] have shown that even in smaller applications there are scalability issues of blockchain limiting its efficiency.

Table 3. Overview of selected privacy languages from the literature, categorised if accountability is considered by design or not.

Category	**Preference Languages**					
Accountability not directly considered	P3P	EPAL	P2U	PPL	SPECIAL	LPL
Accountability by design	AAL	A-PPL				

4.4. Privacy-Preserving Processing

Analysis of data is mandatory for many services to work properly and especially for research to achieve precise results based on a huge sample size. These data are collected by controllers, stored within data warehouse environments and might be shared with third parties. Usually, if enough data about a user are available, it is possible to identify the person related to the data record. This is especially critical if sensitive information, like health status is present. Let us stick with the example of COVID-19 apps, as introduced above, which might collect GPS data and contacts to track down chains of infection. These information might reveal the home address, workplace, close contacts of the user, etc. if joined with external knowledge. Therefore, adversaries can get access to critical information about a person and might use it against them. For example, information about prior medical conditions, i.e., infection with COVID-19, may lead to worse chances on the job market.

To protect the users identity as well as its personal data and sensitive information de-identification is necessary. One could argue, that if all user data are completely masked no misuse is possible.

While this might be true, the data would become useless for analysis which would heavily affect research and company interests. Therefore, a trade-off between privacy and information content, often called utility, has to be made. This is achieved by de-identifying personal data up to the point when it does not uniquely identify the related person but still carries parts of the information value for analysis. Common methods for de-identification of data records are anonymization, which usually goes hand in hand with privacy models, and pseudonymization techniques (see Table 4).

Note that the term 'privacy model' is contested by the term 'confidentiality criteria'. The later term 'confidentiality criteria' is based on the work of Sicherman et al. [67]. It is furthermore used in the works of Biskup and Bonatti [68–71]. The term 'privacy model' is associated with the definition of privacy via probabilistic indistinguishable conditions, while 'confidentiality criteria' defines privacy via precise indistinguishable conditions. Due to the more common use of the term 'privacy model' in the literature, it is used in the remaining of this work.

Anonymization masks the personal data of a user to prevent the leakage of his identity. Methods for anonymization are applied on attribute level, meaning that each attribute can be masked to a different degree. Common methods are suppression, which gradually replaces characters of an attribute with placeholders, generalization, which replaces attributes by more general representations based on a given hierarchy, and deletion, which simply deletes the given attribute.

To ensure anonymity within the whole data set privacy models are used in combination with anonymisation methods. Privacy models define criteria the data set must fulfill after the anonymisation and usually define a threshold for the likelihood of re-identification of single records. Common privacy models are k-anonymity [72] and differential privacy [73] which both have been shown to work within the context of health care [74]. For example, k-anonymity claims that every record within the anonymized data set must be indistinguishable from at least $k - 1$ other records. As a k-anonymous record is related to k entities the original owner is not uniquely re-identifiable. In addition, in the context of GPS tracking of COVID-19 apps privacy models can ensure a proper de-identification of GPS data. Besides its application in health care data sets differential privacy can provide location-based privacy [75]. Therefore, it might be a good fit for COVID-19 apps.

Pseudonymization usually replaces data with tokens. To preserve additional data utility, similar entries result in the same token. As a result, the identifiers are de-identificated but still interconnected to each other. There are various methods for the token generation, including pseudo-random seeds [76], cryptographic methods [77] or hashing [78].

For instance, assuming there is a medical data set including different cancer types. Anonymization would result into a loss of the data utility. For pseudonymization the most straight forward approach would be a keyed hashing of the plain entries. In this case the information of which subjects have the same type of cancer are preserved.

There are also more domain specific techniques, which preserve the data structure of the identifier, enabling the processing of different operations on it. For example, it is possible to pseudonymize IP addresses with artificial ones by simultaneously preserving the sub net hierarchies [79]. In addition, there might be the necessity of calculating the distance of two numbers after de-identification, which can be achieved by the distance-preserving pseudonymization [80].

However, there are also drawbacks, especially for weak pseudonymization approaches. Possible vulnerabilities are side channel attacks like pattern behavior and anomaly detection [81]. In addition, dictionary attacks [77], i.e., hashing identifiers and comparing the results with tokens and more algorithm related threads might occur.

Moreover, a combination of anonymisation and pseudonymization is possible.

Table 4. Overview of selected approaches for privacy-preserving processing from the literature separated into anonymization, privacy models and pseudonymization.

Method	Approaches		
Anonymisation	Suppression	Generalisation	Deletion
Privacy Model	k-Anonymity	Differential Privacy	
Pseudonymisation	Pseudo-random Seeds	Cryptographic Approaches	Hashing

Inference Detection and Prevention

When data are collected from various sources it has to be ensured that not more than intended information about individuals (personal data inference) can be derived. Therefore, various techniques have been proposed to prevent such inference (see Table 5).

Guarnieri et al. [82] developed ANGERONA, a DBIC mechanism to secure against probabilistic inference attacks. Especially medical data sets are vulnerable to probabilistic attacks, e.g., patients whose parents suffer from hereditary diseases are more likely to also suffer from the same disease. In addition, additional factors can influence the likelihood of a disease to break out, e.g., if the patient is a smoker. Therefore, if an attacker has access to such information the probability for a successful inference of this sensitive attribute rises. ANGERONA blocks certain queries, if the probability of a successful inference exceeds a given threshold.

Chen and Chu [83] proposed a query-time inference violation detection system which utilizes a semantic inference graph. This graph is based on a semantic inference model which represents possible inferences from all given attributes to any sensitive information. When data are requested, the query log of the user or attacker is used to calculate the probability of inferring sensitive data. Once this likelihood exceeds a given threshold the query is denied by the system. This also works for multi-user attacks based on query sequences to prevent collaborative inferences.

To prevent inferences before they can occur in queries Qian et al. [84] eliminate inferences within multilevel database systems. By their definition an inference within the database occurs if a user with a lower clearance is able to infer data only accessible to users with a higher clearance without external knowledge. The developed tool is called DISSECT. It scans the database for possible inferences within the data schema by identifying different sequences of foreign key relationships with the same entity. To prevent inferences the classification of affected foreign key relationships is raised.

Instead of focusing on functional dependencies within the database schema, Yip and Levitt [85] analyse the stored data itself to find inferences. They propose a rule-based approach for data level inference detection. A prototypical implementation has shown that for several thousand records it takes a few seconds to scan for inferences. Considering real data sets with millions of records this approach might not be practical for query-time inference detection and it is questionable if it scales for offline scans of the database.

Table 5. Overview of selected approaches for inference detection for privacy-preservation at query-time, database-level and data-level from the literature.

Author	Approaches
Guarnieri et al. [82]	Query-time inference detection if adversary has external knowledge
Chen and Chu [83]	Query-time detection of semantic inferences based on query log for sensitive data
Qian et al. [84]	Multi-level database inference detection based on foreign key relations in the schema
Yip and Levitt [85]	Data-level inference detection based on analysing the stored data with inference rules

These approaches have shown that inference detection on both database and query level are a valid approach to further protect the users privacy. The open question is which rules are required, i.e., which data attributes must be prevented to be combined, and if the proposed approaches scale

with real-world data sets consisting of millions or billions of records. Furthermore, the concept of purpose-based processing has to be considered such that the proposed rules may differ for purposes and requesting entities. Thus, a layer on top of the privacy languages and privacy preferences has to be situated to detect and prevent privacy inferences. The following section details technical possibilities for users to take control over their privacy.

5. User Environment

Within the user environment individuals, i.e., users, interact with edge devices, like computers or smartphones. These devices run software and apps which might collect personal information of the user or access sensor data. Depending on the type of service the software offers, data collection and processing can be essential in order to make the service work or optional whereas users might benefit from additional features and controllers might process additional data. For example, a COVID-19 app that tracks a users movement patterns to investigate chains of infection must access the device's GPS data. Without the processing of this data the app does not work. Besides, there might be optional purposes that are not necessary for the core function of the app but can provide benefits, e.g., for research. An optional purpose might prevent sharing of collected data with other countries or companies. These data could be used to improve countermeasures against COVID-19 worldwide. The user might also have the possibility to limit the quality of his personal data, e.g., by demanding certain degrees of anonymization. For example, the age of 52 could be generalised to 50–70. While this protects his privacy more, it also leads to a less precise result for researching how the age influences severity of the disease.

Every user as an individual is unique and has their own preferences in terms of sharing personal data with service providers. As privacy policies offer various personalisation options, the negotiation process can get quite time-consuming. To ease this process so-called preference languages have been developed.

5.1. Preference Language

Preference languages are machine-readable languages that allow users to express their preferences regarding privacy actions in formal rules. These rules can be matched against privacy policies at policy negotiation and ease the process for the user. The amount of help the user receives can differ for each preference language and for each supporting tool. A basic interaction could be to hint the user if the preferences match the policy, leaving it up to himself to check for mismatches. From this basic case the amount of help rises from summarising mismatching policy statements to consent recommendation or even automated consent. This process requires the privacy policy to be defined by a machine-readable privacy language. In the following, some outstanding preference languages of the last two decades will be briefly described (see Table 6).

The first preference language proposed and used in real applications was APPEL [55]. Its XML-styled rule sets are designed to match P3P policies and automate the negotiation process with such policies. This perfect compatibility is one of the main advantages of APPEL but is unfortunately also the reason why it was abandoned over the years. P3P's fixed vocabulary was too limiting for many controllers to properly express their privacy policies and they turned their back on this technology. Therefore, APPEL was also not used anymore. Furthermore, APPEL rules were known to be error-prone and semantic inconsistencies were observed. Before the decline of the P3P usage happened, attempts were made to fix at least the latter problems of APPEL. XPref [56] is an attempt to make APPEL rules more error-robust and to gain higher expressiveness within the rule sets. To achieve this goal the APPEL rule body is replaced by a subset of XPath. Furthermore, accepting rules are introduced as well as rule connections. To fix semantic inconsistencies SemPref [86] added semantics to both P3P and APPEL. This prevented the definition of two or more syntactically different APPEL rules with the same semantic meaning.

Besides P3P based preference languages further approaches have been researched that combine preferences and a privacy language into a single language. One of these attempts is SecPAL4P [87] which uses SecPAL assertions to define user preferences as claims and privacy policies as promises. In addition, PPL [52] follows the same approach and directly integrates preferences in the PrimeLife policy language while obligations provide further customisable notifications for users. With the rise of mobile technologies new requirements for preferences emerged. CPL [88] has its focus on context information which are mainly present on mobile devices. Therefore, the definition of preferences which feature location or time constraints is possible. Context-based preferences could be especially useful for employees that need to share information while being on the company grounds or at certain time periods. Finally, YaPPL [89] is specifically designed to consider GDPR requirements while its preferences are also feasible for resource limited edge devices or sensors in IoT scenarios.

Table 6. Selection of preference languages from the literature categorised into P3P based languages, symmetric approaches which combine both a privacy language as well as a preference language into a single language and languages meeting modern requirements, like supporting mobile devices by introducing preferences for context information and meeting GDPR requirements.

Category	**Preference Languages**		
P3P based	APPEL	XPref	SemPref
Symmetric Approach	SecPAL4P	PPL	
Context Information	CPL		
GDPR Requirements	YaPPL		

While preference languages can offer various advantages for users there is still a low chance for them to be used at all if no further help is provided. The sole application of a preference language would mean that users have to learn complex technical languages. As stated by the privacy paradox [90] on the one hand users are concerned about their personal data being misused but on the other hand do not value their privacy enough to invest time to protect it. Therefore, preference languages need to be extended by user interfaces that allow for a quick and easy setup process such that the user experience and user acceptance is provided. Users should not experience their privacy settings as 'work without benefits'.

5.2. Preference User Interfaces

Preference User Interfaces (P-UIs) are tools that add a graphical layer on top of a preference language. This drastically eases the setup process of privacy preferences, as users only have to interact with predefined checkboxes or drop-down menus instead of leaning complex languages. In addition, preferences defined by P-UIs are less error-prone than those coded by the user himself, because well designed interfaces prevent contradictions by default. In the literature several P-UIs have been proposed for the prominent preference languages.

To ease the definition of APPEL [55] preferences and to ease the interaction with P3P [53] policies, Privacy Bird [91] was developed. It offers several predefined preference rules, that users can select and also provides preference profiles, that select certain preference rules. The tool adds a small icon to the browser window showing a little bird that represents the matching status of the preferences and the privacy policy. This matching is done once a website based on a P3P policy is loaded as APPEL and P3P support automated policy negotiation. Therefore, the user instantly notices if any mismatch occurs. In this case Privacy Bird hints which parts of the policy mismatch the preferences and offers a summary of the P3P policy for the user to show up himself. Studies of Cranor [91] have shown that users tend to read policies more often if Privacy Bird is used.

Another tool that creates preferences for P3P policies is Preference Cockpit [92]. It provides a wizard-based setup which guides the user through the whole process. Therefore, it offers several levels

of detail, from beginner to expert mode. Preferences defined by this tool are related to service types, like web-mail, online shopping or gaming, whereas for each category a fine-grained selection of data is possible. All the choices made during the setup are summarised in the cockpit window afterwards, to make choices traceable.

The send data dialog [93] is a P-UI that was developed during the PrimeLife project [94] to ease negotiation with PPL [52] policies. This tool shows a dialog window every time data are requested by a service. It summarizes which data will be used for which purpose and shows mismatches between preferences and policy actions. Furthermore, it allows to change the preferences during policy negotiation, which adds a dynamic component to the typically static preference matching. This reflects real user behaviour as preferences might change over time and it is unrealistic to assume that users are considering every possible use case in a setup. By offering the possibility to dynamically generate new preference rules at a later points in time the setup process shrinks which might encourage more users to invest time into it, due to the privacy paradox. This principle is also utilized by almost all mobile apps and is known as the ask-on-first-use model.

6. C2C Environment

Within the C2C Environment controllers interact with other third party controllers. This might include trading data or selling it for money, sharing data with institutes for research or to provide collaborative services, outsourcing certain processing tasks, and more. As third parties might be trusted or untrustworthy, certain measures have to be taken in order to ensure legally correct handling of the data according to the agreement between the user and the controller. Therefore, so-called Technical and Organisational Measures (TOMs) and Service Level Agreements (SLAs) are defined. Furthermore the personal data of the user are protected by Data Subject Rights (DSRs) according to GDPR ([3], Arts. 12–23).

6.1. Technical and Organisational Measures

If two companies, i.e., controllers, exchange personal data, e.g., company A outsources the processing of some personal data for a marketing purpose to company B, then company B has to disclose the Technical and Organisational Measures that are put into place to protect the personal data. Therefore, Technical and Organisational Measures (TOM) are documented and negotiated. This can be especially challenging when the companies are located within different countries, as each country usually has different standards and laws for privacy and security. This is essential in the context of the legal framework of GDPR, which states that Technical and Organisational Measures (TOM) have to be applied and documented to ensure secure and private processing of personal data ([3], Arts. 25 and 32, Recital 78). In the context of data protection, TOMs define which guarantees are given to protect personal data. Therefore, various measures have to be documented and shared between companies. These measures cover various aspects of physical and digital security that are implemented to protect the personal data. For example, this includes physical access control to data centers, server rooms, or computers that are utilized to use (storage, processing, etc.) the personal data. In addition, access control mechanisms and strategies for data (read, write, update, delete) have to be detailed, as well as the secure transmission of data. The topics reliability, data integrity and recoverability are seen as novel aspects addressed for data protection. These topics, as well as various others, have to be documented and shared between companies when personal data are shared to show minimal compliance for its protection. Many of the aspects of TOMs are also represented in ISO 27001 control objectives, e.g., human resources security, physical and environmental security or access control [95].

Considering our use case, TOMs would set the conditions for the transfer of collected personal data from different COVID-19 apps to a centralized data warehouse. The controller of the COVID-19 app is responsible for the protection of the personal data, thus requiring the controller of the data warehouse to protect the personal data with similar or better standards. For example, data have to

be transferred only by encrypted network connections or the data have to be encrypted for storing in the data warehouse. In addition, physical access to the server rooms is restricted to only authorised personnel. The implemented measures for the protection of the personal data will be documented within the TOMs and shared to the COVID-19 app controller for verification. If the implemented measures meet the requirements, then the data will be transferred. Furthermore, the TOMs can be audited and updated over time to ensure up-to-date measures. Note that the protection of the usage of the personal data for only the consented data is not covered by TOMs in our scenario, but by sticky privacy policies that are transferred to the data warehouse with the personal data. Although various checklists, guidelines and questionnaires support this documentation of TOMs, no standardized electronic format for its exchange and negotiation has been put into place.

However, for Service Level Agreements (SLA) various proposals for electronic formats in different domains have been made. SLA describe contracts and agreements for, e.g., B2B processes, in order to ensure a reliable and secure environment, thus aspects of TOMs are covered. Therefore, an electronic SLA format may be the baseline for a future electronic TOMs representation. SLAs can express various requirements both functional and non-functional. Their main intention is to formulate agreements or contracts for B2B processes. Various approaches for SLAs can be found in the literature:

WSLA defines SLAs of web services [96]. For this description the XML schema is used, which makes it very general and therefore allows to define various metrics for services. A service (ServiceDefinition) is based upon various parameter (SLAParameter). A parameter is defined using metrics (Metric) which is defined with functions using other metrics. The hereby defined SLA can be associated with obligations detailing the target of the service level that has to be guaranteed.

With WS-Agreement a protocol for defining agreements between services is defined. It is a specification developed by the Grid Resource Allocation Agreement Protocol Working Group (GRAAP-WG), which is composed of a schema for the agreement description and management. Additional ServiceTerms and GuaranteeTerms can be contextualised. Furthermore, conditions may specify the target value for the service level. Thus this specification does not define the syntax, Oldham et al. [97] proposes a suitable ontology using QoSOnt [98] to define Quality of Service (QoS) aspects. On top of the ontology, a service is proposed for reasoning and inference management.

SLAng describes SLAs which accommodate end-to-end Quality of Service (QoS), typically for B2B cloud use cases. QoS has many facets and requires complex agreements between network, storage and middle-ware services. This includes the definition of the retention of data, which is part of the requirements for legal compliance [99,100].

The Unified Service Description Language (USDL) describes business, operational and technical parameters of services while context specific legal requirements, e.g., terms of use or copyright, are taken into account. Its legal module addresses the need for legal compliance in service networks and in trading services on marketplaces. USDL is designed to incorporate business processes that can be easily comprehended by its users [101].

SLA policies mainly focus on formulating inter business agreements instead of policies between users and companies, e.g., privacy policies. In general, the standardisation of SLAs requires several components. Interfaces for negotiation and management of SLAs are necessary to communicate between companies efficiently. Data models for SLAs, as focused on in this work, have to be defined, as well as SLA vocabularies. Lastly, the specific legal frameworks have to be considered [102].

It can be observed that various languages for the definition of SLAs have been proposed in the literature, whereas web services and services in general are focused (see Table 7). Although SLAs define minimum required qualities for a service, they do not consider all the documentation required for GDPR, i.e., TOMs. The various aspects of TOMs can be derived from various sources and are focused on the protection of personal data with physical and electronic measures. The ISO 27001 covers such aspects, thus may provide a baseline for the standardisation of TOMs in an electronic format like SLAs.

Table 7. Overview of languages and their description to define SLAs and TOMs.

Type	Solution	Description
SLA	WSLA	SLAs for Web Services
	WS-Agreement	Protocol and Scheme to define SLAs between Services
	SLAng	SLAs for end-to-end QoS in Cloud Services
	USDL	SLAs with Legal Context for Business Services
TOM	-	No electronic representation

6.2. Data Subject Rights

GDPR ([3], Arts. 12–23) strengthens the rights of Data Subjects within the European Union. These articles demand transparency, confirmation about processing and access to personal data, deletion of personal data, data portability etc. (see Table 8). To ensure the enforcement of Data Subject Rights (DSRs) technical measures are needed. Privacy languages already cover [3] (Arts. 12–14). Even P3P as the first well known privacy language had a strong focus on transparency of policy actions. More recent approaches also integrate privacy icons in their Policy Negotiation UIs to make information even more accessible to users, e.g., LPL with the LPL policy viewer [103]. Therefore, Ref [3] (Art. 12) is fully covered by privacy languages. Ref [3] (Arts. 13 and 14) are also fulfilled by several privacy languages. PPL combined with its Preference UI, the Send Data Dialog, informs the user every time when data are collected about its usage and lets the user decide himself if he agrees on it. Furthermore, an attempt was made to technically meet the requirements of [3] (Art. 20) for data portability. The Data Transfer Project [104] is a collaboration of several service providers which aims to ease the process of switching from one service to another. Therefore, users can download their personal data from service providers but also share their data with other services. This might be helpful for backups of their own personal data as well as encouraging users to test new services.

To the best of our knowledge, [3] (Arts. 15–19) and [3] (Arts. 21–23) are not realised yet by any technical solution, leading to a huge research gap for Data Subject Rights. There are several open topics which might be especially interesting to tackle with already existing privacy languages in mind. For example, Ref [3] (Art. 15) is conceptually supported by most of the modern privacy languages, e.g., PPL, LPL or SPECIAL, as their policies are already structured by purposes, recipients, data categories, retention etc., even though none of the privacy languages can technically export this information to the user, if requested. In addition, Ref [3] (Art. 17), the right to be forgotten, has some interesting implications on already existing technologies. In context of backups, which are omnipresent in data warehouse structures or company-intern databases it has to be ensured that each entry of the data record is properly deleted. Therefore, if the record is stored in any backup file, it also needs to be deleted there. In the context of the blockchain technology the same challenge remains to be an open topic. While there already exist promising technical approaches to realise data subject rights there is still a huge amount of open research questions, that might be addressed by future efforts.

Table 8. Overview of Data Subject Rights and proposed technical solutions to solve them.

Data Subject Right	Technical Solutions
Art. 12 Transparent information, communication and modalities for the exercise of the rights of the data subject	Privacy language, privacy icon
Art. 13 Information to be provided where personal data are collected from the data subject	Privacy language
Art. 14 Information to be provided where personal data have not been obtained from the data subject	Privacy language
Art. 15 Right of access by the data subject	

Table 8. *Cont.*

Data Subject Right	Technical Solutions
Art. 16 Right to rectification	
Art. 17 Right to erasure ('right to be forgotten')	
Art. 18 Right to restriction of processing	
Art. 19 Notification obligation regarding rectification or erasure of personal data or restriction of processing	
Art. 20 Right to data portability	Data Transfer Project [104]
Art. 21 Right to object	
Art. 22 Automated individual decision-making, including profiling	
Art. 23 Restrictions	

6.3. Trading

Besides the necessity of personal data for certain services to function, e.g., COVID-19 apps to track locations, or the possibility for users to personalise services to fit their preferences, trading of data might be a controllers' strategy to increase their revenue. As an example, the income of Facebook in the first quarter of 2020 solely generated from advertisements was $17.44 billion according to Facebook Investor Relations [105]. These advertisements are personalised based on analysed user data, which is shared with third parties. Facebook's huge revenue generated by trading data shows that this is a valid business concept, and in fact, sharing and trading data became a gigantic worldwide market.

This development motivated diverse technical solutions to either protect the shared data or to ease the trading process. In this context, several privacy languages have been proposed dealing with sharing of data. PPL [52], A-PPL [63] and SPECIAL [60] consider the processing of personal data for secondary use. Therefore, measures are provided to protect personal data by limiting the access to data for authorised purposes only. P2U [59] targets secondary sharing of data in a more business oriented way by enabling the negotiation of a price for personal data trading. Hereby, the concept of sticky policies is introduced to link personal data with the corresponding data even after they have been transferred.

There are also other approaches than those based on privacy languages. Truthfulness and Privacy preservation in Data Markets (TPDM) [106] uses homomorphic encryption and identity-based signatures to integrate privacy-preservation into data markets. Evaluations of TPDM showed low computation overheads for large scale data markets, making it potentially useful in real world applications. Bataineh et al. [107] proposed a technical solution for a marketplace for data trading. Instead of data being shared or traded by controllers it allows the users to sell their personal data themselves whereas higher quality data also lead to higher prices. Therefore, users are directly compensated for sharing their data by monetary rewards.

6.4. Multi-Party Privacy

Enforcing privacy is highly difficult when dealing with settings where the data are distributed among multiple parties. Traditional approaches targeting personal privacy require direct data access, and thus may not work if there is a range of various involved parties. For instance, point-of-interest related data usually is owned by different entities. An arising question would be how a map service could provide information without infringing privacy. This could be solved by introducing certain privacy enhancing protocols [108].

7. Discussion and Future Research Directions

In this work we detailed the personal privacy workflow, which covers the whole process of data being collected, processed, shared or traded by controllers after the user gave his permission. Therefore,

we separate the workflow into controller environment, user environment and C2C environment. The controller environment capsules data handling practices of controllers, summarised within privacy policies. Data collected by the controller, according to GDPR, must only be handled as stated within the policies in order to protect the user's privacy. Privacy policies can technically be implemented by a privacy language which makes them machine-readable and processible. In addition, access control can either be integrated in a privacy language or is established by external access control mechanisms. Therefore, data protection can be enforced while further personalisation options are provided to the user. Users negotiate, i.e., personalize, privacy policies via Policy Negotiation UIs which aim to ease this process.

Overall, the controller environment is a well researched area in terms of privacy protection. Over the past two decades several privacy languages have been proposed with diverse features and most of them are suited well for health care scenarios with large scale data sets. For example, EPAL has a focus on business intern representations of privacy policies, A-PPL aims for accountability and LPL ensures de-identification of data before usage. P3P as the first well known privacy language also was widely used when it came up and also became standard. In addition, efforts were made to propose privacy languages which meet the legal framework of the GDPR, e.g., LPL and SPECIAL. In terms of privacy-preserving processing of data various approaches for pseudonymization, anonymization and privacy models exist, whereas several of them are used in real applications. For example, k-anonymity and differential privacy are both privacy models that are used to ensure anonymity within real health care data sets. Furthermore, various approaches for access control exist whereas RBAC, ABAC and CAAC are the most common categories. Promising new approaches for CAAC are suited for distributed cloud or fog computing environments, fuzzy data and relationship awareness whereby a huge area of applications could be covered when combined with privacy languages.

The user environment capsules the preferences of the users. As each user has unique desires for privacy, preference languages model the privacy preferences of the user as rules. Preference UIs assist at the definition of preferences and aim to make the process less error-prone. At policy negotiation preferences are matched against the privacy policy and can offer various help. This can be hints, if the preferences match the policy statements, but also sophisticated consent recommendation systems or even automated consent to purposes.

Similar to the controller environment various approaches have been proposed for the user environment. On the one hand, preference languages allow for the definition of preference rules in an electronic format, similar to privacy languages. Various preference languages have been proposed with diverse focuses, e.g., CPL introduces location and time based preferences and YaPPL covers GDPR requirements while limiting resource requirements to enable preferences for IoT scenarios. Unfortunately, proposed preference languages tend to be less complete than their privacy language counterparts. For example, CPL doesn't consider purposes in its model which are present in almost all modern privacy languages and also reflects GDPR ([3], Art. 15). While YaPPL is one of the most complete and modern preference languages it misses measures for negotiation with privacy policies. Similar shortcomings can be found in several preference languages, which results in a research gap for a complete preference language covering all of the necessary functionality. On the other hand, Preference UIs, which ease the definition process and therefore are necessary for preference languages to be usable at all, according to the privacy paradox [90], are mostly outdated and do not consider newly introduced GDPR requirements. The existing P-UIs feature some promising functionalities, like Privacy Bird directly showing the matching status as an icon in the browser window or the send data dialog allowing for dynamic preference definition on-the-fly, which could be adopted in a modern P-UI. That is a second gap within the user environment that future research could focus on.

Finally, the one problem that is present in all of the existing approaches is missing standardisation. Usually, preference languages are designed to negotiate with exactly one counterpart privacy language, e.g., APPEL with P3P or SecPAL4P and PPL featuring a symmetric approach and cover both preferences and policy in a single language. The problem that arises from missing standardisation is, if multiple

preference and privacy languages are used in real applications. Even if a user defines his preferences in one preference language, it will only work on services that are based on the counterpart privacy language. Therefore, measures have to be taken to enable standardised negotiation between preference languages and privacy languages to maximise the coverage of services with a single preference setup. To the best of our knowledge, this has not been considered in the literature yet.

The C2C environment capsules data transfer interactions of a controller with third party controllers. In this context, data might be transferred to outsource some processing, to make use of certain third party services, to make data available for research or as part of some trading transactions. The shared user data also need to be protected, if shared with a third party controller in order to maintain privacy. Therefore, additional agreements between the controllers are specified. The Technical and Organisational Measures (TOMs) define measures taken by the third party to protect personal user data while Service Level Agreements (SLAs) define agreements to ensure a reliable and secure environment. Furthermore, the privacy of the user is protected by Data Subject Rights according to GDPR.

Even though there are strong legal requirements regarding data protection of shared data with third party controllers there is a lack of supporting technologies in this area. For TOMs there is no electronic representation at all, meaning that these are only plain text documents. Therefore, stated measures can't be technically enforced or controlled which leaves room for deviations of claimed data handling practices. For SLAs various electronic formats, e.g., WSLA or WS-Agreement, exist with a main focus on (web) services. While this is a desirable development, it is not enough to fully cover all of the necessary documentation on side of third parties according to GDPR as SLAs only cover minimum requirements for qualities of a service. To achieve full coverage of GDPR requirements all of the content stated in the TOMs also would need to be in an electronic format. This missing support is a gap that future research might address. The technologies used for the definition of SLAs could be used as a basis to also create electronic formats of TOMs. Regarding Data Subject Rights a problem that still remains is the right to erasure according to GDPR ([3], Art. 17). With various backups of personal data existing, it is difficult to enforce deletion in each backup layer as well as controlling, if it was done appropriately.

To sum up, research gaps within the user environment remain for a preference language which covers the full bandwidth of functionalities, a modern and GDPR-conforming preference UI solution and standardisation for the negotiation between preferences and privacy policies. In the C2C environment there is a gap of electronic formats for the definition of TOMs to cover all needed documentation according to GDPR and how to ensure deletion of personal data in backup files. These are topics future research might focus on.

8. Conclusions

Personal data became a valuable asset which affects many parts of our daily lives. Massive data warehouses are omnipresent and carry a vast amount of sensitive user data. These huge databases have many applications and can offer advantages for both users and service providers. Users can benefit from sharing their data as this might unlock certain additional features of a service, e.g., location recommendations if GPS data are shared, or personalise the service to better fit the users preferences. In addition, service providers, i.e., controllers, can greatly get use out of personal data. On the one hand, they can be used for analysis and targeted advertisement, to improve business processes, to find out which products are demanded the most to increase production and to raise their sales numbers. In addition, users might invest in some premium features of a personalized service. On the other hand, personal data, as a valuable asset, can be shared or traded with other third parties. Several companies, e.g., Facebook, have proven that this is a valuable business model by earning billions of dollars a year with trading data for targeted advertisements [105]. Finally, research can greatly benefit from databases filled with personal data. This is especially important in the health care sector, as a huge sample size leads to more precise results. The current pandemic of COVID-19 has

shown several usages of personal data in form of COVID-19 apps which use GPS data to track down chains of infection. The apps analyse contacts via Bluetooth and warn the user if they were in contact with an infected person. Therefore, the user might also be infected and should be tested for the disease. Analysis of medical records of treated COVID-19 patients might help to improve treatment of future cases and could potentially speed up the process of producing a vaccine. Therefore, every data record is important and the best possible results could be obtained if all nations would work together on a shared database.

In addition to these advantages that the use of personal data brings with it, there are also several disadvantages from a data protection perspective. Thus, a balance between the protection of personal data and its use has to be found. We argue that privacy languages and preference languages are necessary to standardise the whole process. They can be a tool to help negotiate in this conflict of interest. Privacy languages can structure privacy policies to make them machine-readable. This means that the process of collecting and passing on data can be transparently automated and thus becomes more controllable. Thus, the controller is supported in fulfilling his duties. In contrast, preference languages support the user in controlling their personal sensitive data. By using these formalisation of privacy protection rules, it is possible to exploit the advantages of data use without denying the individual rights to their data and thus to fully comply with the GDPR even in extreme situations. Through the use of privacy languages, it can be ensured that the personal data are made anonymous in an adequate form and thus no identification of individual people is possible. The anonymized data can form an important basis for science and research, e.g., statistics and forecasts on the further course of the pandemic can be produced with the help of these statistics. An example showing that collecting and combining data can lead to success is Taiwan. Despite their proximity to China, they managed to keep the number of infected people very low by linking travel and disease data that were related to each other [109].

In this work we introduced the personal privacy workflow process, which we separated into controller environment, user environment and C2C environment. For each of the categories we discussed several technologies, proposed in the literature, which realise the digitisation of the workflow. Especially for the controller environment, which capsules data handling practices of controllers, several sophisticated approaches exist to support the necessary processes. Well researched de-identification methods, including pseudonymization, anonymization and privacy models, combined with methods for access control and inference prevention effectively protect sensitive user data. Several privacy languages already deal with the legal framework of the GDPR while offering a broad range of functionalities to create a transparent environment for users who negotiate privacy policies. This is supported by Policy Negotiation User Interfaces which include personalisation options and privacy icons to further ease the policy negotiation process.

For the user environment, which capsules the preferences of the user, also some promising approaches exist for preference languages. Most of the preference languages focus on specific functions for preference definition, e.g., location constraints, while other aspects are left out. Even though there are preference languages which also consider GDPR, we found that none of them offer such a broad and complete functionality as their privacy language counterparts. While there are some promising approaches for Preference User Interfaces, like dynamic preference definition on-the-fly or icons in the browser window to directly show the matching status, the majority of them is outdated and does not consider GDPR requirements. Furthermore, one important point that is often not considered is standardisation. Current approaches of preference languages always focus on negotiation with a single privacy language and are therefore specifically designed for it. However, considering that several privacy languages exist and also might be used in real applications, measures have to be taken to enable standardised negotiation between preferences and (potentially incompatible) privacy policies. Therefore, we identify research gaps within the user environment regarding functionally incomplete preference languages, missing standardisation and outdated Preference User Interfaces. These are important topics which future research might focus on.

Finally, regarding the C2C environment, where data are transferred between controllers, we found several gaps in research. For trading of personal data several approaches have been proposed to technically implement the process. In this context privacy languages consider secondary use of data and might enable negotiation of prices while protecting personal data. In addition, approaches independent of privacy languages exist which focus on data market technologies and allow users to share their data for money. As sharing or trading personal user data is a critical process in terms of privacy, there are strict legal regulations. On the one hand, GDPR requires third party controllers to document how they handle user data. This is documented within Service Level Agreements and the Technical and Organisational Measures. While electronic formats exist for Service Level Agreements, which define agreements to ensure a reliable and secure environment, none exist for Technical and Organisational Measures, which document which measures are taken to protect data. Therefore, no technical control is possible for the latter which infers a gap in research. A possible approach might be to adapt electronic representations of Service Level Agreements or use them as a basis. On the other hand, GDPR grants the user more control over his personal data by Data Subject Rights. Some of the Data Subject Rights, i.e., Ref [3] (Arts. 12–14, 20), are covered by privacy languages or other technical approaches. For the remaining articles no technical solution has been proposed yet, to the best of our knowledge. This results in another research gap, which contains several interesting implications on already existing technologies, e.g., how [3] (Art. 17) The right to be forgotten can be applied on backup files or blockchains. To sum up, within the C2C environment open research topics remain for technical solutions for Technical and Organisational Measures and Data Subject Rights.

By our work we hope to motivate further research in the remaining open topics of privacy-preserving data handling. GDPR was an important step towards a better privacy which in our data-centred world is more important than ever before. Now new technologies have to emerge that fulfill all of the stated requirements.

Author Contributions: S.B. and A.G. did the conceptualisation and supervision of the work. All authors performed literature research, contributed to the initial draft, discussed the results and contributed to the final draft of the work. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Raghupathi, W.; Raghupathi, V. Big data analytics in healthcare: Promise and potential. *Health Inf. Sci. Syst.* **2014**, *2*, 3. [CrossRef] [PubMed]
2. Arcelus, A.; Jones, M.H.; Goubran, R.; Knoefel, F. Integration of Smart Home Technologies in a Health Monitoring System for the Elderly. In Proceedings of the 21st International Conference on Advanced Information Networking and Applications Workshops (AINAW'07), Niagara Fa-lls, ON, Canada, 21–23 May 2007; Volume 2, pp. 820–825.
3. Parliament, E.; The Council of the European Union. General Data Protection Regulation, 2016. Regulation (EU) 2016 of the European Parliament and of the Council of on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC. *OJ L* **2016**, *119*.
4. Epidemiology Working Group for NCIP Epidemic Response, Chinese Center for Disease Control and Prevention. The Epidemiological Characteristics of an Outbreak of 2019 Novel Coronavirus Diseases (COVID-19) in China. *Zhonghua Liu Xing Bing Xue Za Zhi* **2020**, *41*, 145.
5. World Health Organization. *Coronavirus Disease (COVID-19) Situation Report—135*; Technical Report; World Health Organization: Geneva, Switzerland, 2020.
6. Berke, A.; Bakker, M.; Vepakomma, P.; Raskar, R.; Larson, K.; Pentland, A. Assessing disease exposure risk with location histories and protecting privacy: A cryptographic approach in response to a global pandemic. *arXiv* **2020**, arXiv:2003.14412.

7. Wu, Y.C.; Chen, C.S.; Chan, Y.J. The outbreak of COVID-19: An overview. *J. Chin. Med Assoc.* **2020**, *83*, 217. [CrossRef] [PubMed]
8. World Health Organization. *Protocol for Assessment of Potential Risk Factors for Coronavirus Disease 2019 (COVID-19) among Health Workers in a Health Care Setting, 23 March 2020*; Technical Report; World Health Organization: Geneva, Switzerland, 2020.
9. Wu, Z.; McGoogan, J.M. Characteristics of and important lessons from the coronavirus disease 2019 (COVID-19) outbreak in China: Summary of a report of 72 314 cases from the Chinese Center for Disease Control and Prevention. *JAMA* **2020**, *323*, 1239–1242. [CrossRef] [PubMed]
10. Allam, Z.; Jones, D.S. On the coronavirus (COVID-19) outbreak and the smart city network: universal data sharing standards coupled with artificial intelligence (AI) to benefit urban health monitoring and management. In *Healthcare*; Multidisciplinary Digital Publishing Institute: Basel, Switzerland, 2020; Volume 8, p. 46.
11. Li, J.; Guo, X. COVID-19 Contact-tracing Apps: A Survey on the Global Deployment and Challenges. *arXiv* **2020**, arXiv:2005.03599.
12. Annas, G.J. HIPAA regulations-a new era of medical-record privacy? *N. Engl. J. Med.* **2003**, *348*, 1486–1490. [CrossRef]
13. Jaigirdar, F.T.; Rudolph, C.; Bain, C. Can I Trust the Data I See? A Physician's Concern on Medical Data in IoT Health Architectures. In Proceedings of the Australasian Computer Science Week Multiconference (ACSW 2019), Sydney, NSW, Australia, 29–31 January 2019; Association for Computing Machinery: New York, NY, USA. [CrossRef]
14. Dilmaghani, S.; Brust, M.R.; Danoy, G.; Cassagnes, N.; Pecero, J.; Bouvry, P. Privacy and Security of Big Data in AI Systems: A Research and Standards Perspective. In Proceedings of the IEEE International Conference on Big Data, Los Angeles, CA, USA, 9–12 December 2019; pp. 5737–5743.
15. Taeihagh, A.; Lim, H.S.M. Governing autonomous vehicles: Emerging responses for safety, liability, privacy, cybersecurity, and industry risks. *Transp. Rev.* **2019**, *39*, 103–128. [CrossRef]
16. Lowry, P.B.; Dinev, T.; Willison, R. Why security and privacy research lies at the centre of the information systems (IS) artefact: Proposing a bold research agenda. *Eur. J. Inf. Syst.* **2017**, *26*, 546–563. [CrossRef]
17. Xiao, Z.; Xiao, Y. Security and Privacy in Cloud Computing. *IEEE Commun. Surv. Tutor.* **2013**, *15*, 843–859. [CrossRef]
18. Henze, M.; Inaba, R.; Fink, I.B.; Ziegeldorf, J.H. Privacy-Preserving Comparison of Cloud Exposure Induced by Mobile Apps. In Proceedings of the 14th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Services (MobiQuitous 2017), Melbourne, Australia, 7–10 November 2017; Association for Computing Machinery: New York, NY, USA, 2017; pp. 543–544. [CrossRef]
19. Dahlmanns, M.; Dax, C.; Matzutt, R.; Pennekamp, J.; Hiller, J.; Wehrle, K. Privacy-Preserving Remote Knowledge System. In Proceedings of the 2019 IEEE 27th International Conference on Network Protocols (ICNP), Chicago, IL, USA, 8–10 October 2019; pp. 1–2.
20. Hathaliya, J.J.; Tanwar, S. An exhaustive survey on security and privacy issues in Healthcare 4.0. *Comput. Commun.* **2020**, *153*, 311–335. [CrossRef]
21. Iqridar Newaz, A.; Sikder, A.K.; Ashiqur Rahman, M.; Selcuk Uluagac, A. A Survey on Security and Privacy Issues in Modern Healthcare Systems: Attacks and Defenses. *arXiv* **2020**, arXiv:2005.07359,
22. El Majdoubi, D.; El Bakkali, H. Towards a Holistic Privacy Preserving Approach in a Smart City Environment. In *Innovations in Smart Cities Applications*, 3rd ed.; Ben Ahmed, M., Boudhir, A.A., Santos, D., El Aroussi, M., Karas, İ.R., Eds.; Springer International Publishing: Cham, Switzerland, 2020; pp. 947–960.
23. Deep, S.; Zheng, X.; Jolfaei, A.; Yu, D.; Ostovari, P.; Kashif Bashir, A. A survey of security and privacy issues in the Internet of Things from the layered context. *Trans. Emerg. Telecommun. Technol.* **2020**, e3935. [CrossRef]
24. Ferrag, M.A.; Shu, L.; Yang, X.; Derhab, A.; Maglaras, L. Security and Privacy for Green IoT-Based Agriculture: Review, Blockchain Solutions, and Challenges. *IEEE Access* **2020**, *8*, 32031–32053. [CrossRef]
25. Weixiong, Y.; Lee, R.; Seng, A.K.S.; tuz Zahra, F. Security and Privacy Concerns in Wireless Networks—A Survey. *TechRxiv* **2020**. [CrossRef]
26. Linden, T.; Khandelwal, R.; Harkous, H.; Fawaz, K. The Privacy Policy Landscape After the GDPR. *Proc. Priv. Enhancing Technol.* **2020**, *2020*, 47–64. [CrossRef]

27. Ebert, N.; Ackermann, K.A.; Heinrich, P. Does Context in Privacy Communication Really Matter? A Survey on Consumer Concerns and Preferences. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (CHI '20), Island of Oahu, HI, USA, 25–30 April 2020; Association for Computing Machinery: New York, NY, USA, 2020; pp. 1–11. [CrossRef]
28. Esmaeilzadeh, P. The effect of the privacy policy of Health Information Exchange (HIE) on patients' information disclosure intention. *Comput. Secur.* **2020**, *95*, 101819. [CrossRef]
29. Johnson, G.A.; Shriver, S.K.; Du, S. Consumer Privacy Choice in Online Advertising: Who Opts Out and at What Cost to Industry? *Mark. Sci.* **2020**, *39*, 33–51. [CrossRef]
30. Leicht, J.; Heisel, M. A Survey on Privacy Policy Languages: Expressiveness Concerning Data Protection Regulations. In Proceedings of the 12th CMI Conference on Cybersecurity and Privacy (CMI), Copenhagen, Denmark, 28–29 November 2019; pp. 1–6.
31. Kumaraguru, P.; Cranor, L.; Lobo, J.; Calo, S. A Survey of Privacy Policy Languages. In *Workshop on Usable IT Security Management (USM '07) at Symposium On Usable Privacy and Security '07*; ACM: New York, NY, USA, 2007.
32. Kasem-Madani, S.; Meier, M. Security and Privacy Policy Languages: A Survey, Categorization and Gap Identification. *arXiv* **2015**, arXiv:1512.00201.
33. Morel, V.; Pardo, R. Three Dimensions of Privacy Policies. *arXiv* **2019**, arXiv:1908.06814,
34. Gerl, A.; Meier, B. Privacy in the Future of Integrated Health Care Services—Are Privacy Languages the Key? In Proceedings of the 2019 International Conference on Wireless and Mobile Computing, Networking and Communications (WiMob), Barcelona, Spain, 21–23 October 2019; pp. 312–317.
35. Ding, D.; Cooper, R.A.; Pasquina, P.F.; Fici-Pasquina, L. Sensor technology for smart homes. *Maturitas* **2011**, *69*, 131–136. [CrossRef] [PubMed]
36. Walter, M.; Eilebrecht, B.; Wartzek, T.; Leonhardt, S. The Smart Car Seat: Personalized Monitoring of Vital Signs in Automotive Applications. *Pers. Ubiquitous Comput.* **2011**, *15*, 707–715. [CrossRef]
37. Wu, Q.; Sum, K.; Nathan-Roberts, D. How Fitness Trackers Facilitate Health Behavior Change. *Proc. Hum. Factors Ergon. Soc. Annu. Meet.* **2016**, *60*, 1068–1072. [CrossRef]
38. Kayes, A.S.M.; Kalaria, R.; Sarker, I.; Islam, M.; Watters, P.; Ng, A.; Hammoudeh, M.; Badsha, S.; Kumara, I. A Survey of Context-Aware Access Control Mechanisms for Cloud and Fog Networks: Taxonomy and Open Research Issues. *Sensors* **2020**, *20*, 2464. [CrossRef] [PubMed]
39. Sandhu, R.S.; Coyne, E.J.; Feinstein, H.L.; Youman, C.E. Role-Based Access Control Models. *Computer* **1996**, *29*, 38–47. [CrossRef]
40. Motta, G.H.M.B.; Furuie, S.S. A contextual role-based access control authorization model for electronic patient record. *IEEE Trans. Inf. Technol. Biomed.* **2003**, *7*, 202–207. [CrossRef]
41. Wang, L.; Wijesekera, D.; Jajodia, S. A Logic-Based Framework for Attribute Based Access Control. In Proceedings of the 2004 ACM Workshop on Formal Methods in Security Engineering (FMSE '04), Washington, DC, USA, 25–29 October 2004; Association for Computing Machinery: New York, NY, USA, 2004; pp. 45–55. [CrossRef]
42. Zhang, Y.; Zheng, D.; Deng, R.H. Security and Privacy in Smart Health: Efficient Policy-Hiding Attribute-Based Access Control. *IEEE Internet Things J.* **2018**, *5*, 2130–2145. [CrossRef]
43. Kayes, A.S.M.; Han, J.; Colman, A. ICAF: A Context-Aware Framework for Access Control. In *Australasian Conference on Information Security and Privacy*; Springer: Berlin/Heidelberg, Germany, 2012. [CrossRef]
44. Kayes, A.S.M.; Han, J.; Colman, A.; Islam, M. *RelBOSS: A Relationship-Aware Access Control Framework for Software Services*; Springer: Berlin/Heidelberg, Germany, 2014. [CrossRef]
45. Kayes, A.S.M.; Rahayu, W.; Dillon, T.; Chang, E.; Han, J. Context-Aware Access Control with Imprecise Context Characterization Through a Combined Fuzzy Logic and Ontology-Based Approach. In *OTM Confederated International Conferences "On the Move to Meaningful Internet Systems"*; Springer: Berlin/Heidelberg, Germany, 2017; pp. 132–153. [CrossRef]
46. Kayes, A.S.M.; Rahayu, W.; Watters, P.; Alazab, M.; Dillon, T.; Chang, E. Achieving security scalability and flexibility using Fog-Based Context-Aware Access Control. *Future Gener. Comput. Syst.* **2020**, *107*. [CrossRef]
47. Kayes, A.S.M.; Rahayu, W.; Dillon, T.; Chang, E. Accessing Data from Multiple Sources Through Context-Aware Access Control. In Proceedings of the 17th IEEE International Conference on Trust, Security and Privacy in Computing and Communications/12th IEEE International Conference on Big Data Science and Engineering (TrustCom/BigDataSE), New York, NY, USA, 1–3 August 2018. [CrossRef]

48. Damianou, N.; Dulay, N.; Lupu, E.; Sloman, M. The Ponder Policy Specification Language. In Proceedings of the International Workshop on Policies for Distributed Systems and Networks (POLICY '01), Bristol, UK, 29–31 January 2001; Springer: London, UK, 2001; pp. 18–38.
49. Kagal, L. *Rei: APolicy Language for the Me-Centric Project*; Technical Report; HP Labs: Palo Alto, CA, USA, 2002.
50. Becker, M.; Fournet, C.; Gordon, A. SecPAL: Design and semantics of a decentralized authorization language. *J. Comput. Secur.* **2010**, *18*, 619–665. [CrossRef]
51. Rissanen, E.; Bill Parducci, H.L. *eXtensible Access Control Markup Language (XACML) Version 3.0*; Technical Report; OASIS: Burlington, MA, USA, 2013.
52. Ardagna, C.; Bussard, L.; De Capitani Di Vimercati, S.; Neven, G.; Pedrini, E.; Paraboschi, S.; Preiss, F.; Samarati, P.; Trabelsi, S.; Verdicchio, M.; et al. Primelife policy language. In *W3C Workshop on Access Control Application Scenarios*; W3C: Cambridge, MA, USA, 2009.
53. Wenning, R.; Schunter, M.; Cranor, L.; Dobbs, B.; Egelman, S.; Hogben, G.; Humphrey, J.; Langheinrich, M; Marchiori, M.; Presler-Marshall, M.; et al. *The Platform for Privacy Preferences 1.1 (P3P1.1) Specification*; Technical Report; W3C: Cambridge, MA, USA, 2006.
54. Bohrer, K.; Holland, B. *Customer Profile Exchange (CPExchange) Specification*; Version 1.0.; OASIS, Burlington, MA, USA, 2000.
55. Cranor, L.; Langheinrich, M.; Marchiori, M. *A P3P Preference Exchange Language 1.0 (APPEL1.0)*; Technical Report; W3C: Cambridge, MA, USA, 2002.
56. Agrawal, R.; Kiernan, J.; Srikant, R.; Xu, Y. XPref: A preference language for P3P. *Comput. Netw.* **2005**, *48*, 809–827. [CrossRef]
57. Yu, T.; Li, N.; Antón, A.I. A Formal Semantics for P3P. In Proceedings of the 2004 Workshop on Secure Web Service (SWS '04), Fairfax, VA, USA, 29 October 2004; ACM: New York, NY, USA, 2004; pp. 1–8. [CrossRef]
58. Ashley, P.; Hada, S.; Karjoth, G.; Powers, C.; Schunter, M. *Enterprice Privacy Authorization Language (EPAL 1.2)*; Technical Report; IBM: Armonk, NY, USA, 2003. Available online: https://www.w3.org/Submission/2003/SUBM-EPAL-20031110/ (accessed on 2 June 2020).
59. Iyilade, J.; Vassileva, J. P2U: A Privacy Policy Specification Language for Secondary Data Sharing and Usage. In *IEEE Security and Privacy Workshops*; Technical Report; IEEE: Piscataway, NJ, USA, 2014.
60. Bonatti, P.A.; Kirrane, S.; Petrova, I.; Schlehahn, E.; Sauro, L. *Deliverable D2.1-Policy Language V1*; Technical Report; Scalable Policy-Aware Linked Data Architecture for Privacy, Transparency and Compliance-SPECIAL; Zenodo: Geneva, Switzerland, 2017.
61. Gerl, A.; Bennani, N.; Kosch, H.; Brunie, L. LPL, Towards a GDPR-Compliant Privacy Language: Formal Definition and Usage. In *Transactions on Large-Scale Databases and Knowledge-Centered Systems (TLDKS)*; Lecture Notes in Computer Science (LNCS) 10940; Springer: Berlin/Heidelberg, Germany, 2018; Chapter 2, pp. 1–40. [CrossRef]
62. Benghabrit, W.; Grall, H.; Royer, J.C.; Sellami, M.; Azraoui, M.; Elkhiyaoui, K.; Önen, M.; De Oliveira, A.S.; Bernsmed, K. A Cloud Accountability Policy Representation Framework. In Proceedings of the 4th International Conference on Cloud Computing and Services Science (CLOSER 2014), Barcelona, Spain, 3–5 April 2014; SCITEPRESS-Science and Technology Publications, Lda: Setubal, Portugal, 2014; pp. 489–498. [CrossRef]
63. Azraoui, M.; Elkhiyaoui, K.; Önen, M.; Bernsmed, K.; De Oliveira, A.S.; Sendor, J. A-PPL: An Accountability Policy Language. In *Data Privacy Management, Autonomous Spontaneous Security, and Security Assurance: 9th International Workshop, DPM 2014, 7th International Workshop, SETOP 2014, and 3rd International Workshop, QASA 2014, Wroclaw, Poland, 10–11 September 2014. Revised Selected Papers*; Springer International Publishing: Cham, Switzerland, 2015; pp. 319–326. [CrossRef]
64. Camilo, J. Blockchain-based consent manager for GDPR compliance. In *Open Identity Summit 2019*; Roßnagel, H., Wagner, S., Hühnlein, D., Eds.; Gesellschaft für Informatik: Bonn, Germany, 2019; pp. 165–170.
65. Dorri, A.; Kanhere, S.S.; Jurdak, R.; Gauravaram, P. Blockchain for IoT security and privacy: The case study of a smart home. In Proceedings of the 2017 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops), Kona, HI, USA, 13–17 March 2017; pp. 618–623.
66. Angraal, S.; Krumholz, H.M.; Schulz, W.L. Blockchain Technology. *Circ. Cardiovasc. Qual. Outcomes* **2017**, *10*, e003800. [CrossRef]

67. Sicherman, G.L.; De Jonge, W.; Van de Riet, R.P. Answering Queries Without Revealing Secrets. *ACM Trans. Database Syst.* **1983**, *8*, 41–59. [CrossRef]
68. Biskup, J.; Bonatti, P.A. Lying versus refusal for known potential secrets. *Data Knowl. Eng.* **2001**, *38*, 199–222. [CrossRef]
69. Biskup, J.; Bonatti, P. Controlled query evaluation for enforcing confidentiality in complete information systems. *Int. J. Inf. Secur.* **2004**, *3*, 14–27. [CrossRef]
70. Biskup, J.; Bonatti, P.A. Controlled Query Evaluation for Known Policies by Combining Lying and Refusal. *Ann. Math. Artif. Intell.* **2004**, *40*, 37–62. [CrossRef]
71. Biskup, J.; Bonatti, P. Controlled query evaluation with open queries for a decidable relational submodel. *Ann. Math. Artif. Intell.* **2007**, *50*, 39–77. [CrossRef]
72. Sweeney, L. k-Anonymity: A model for protecting privacy. *Int. J. Fuzziness Knowl.-Based Syst.* **2002**, *10*, 557–570. [CrossRef]
73. Dwork, C. Differential Privacy. In *Automata, Languages and Programming*; Bugliesi, M., Preneel, B., Sassone, V., Wegener, I., Eds.; Springer: Berlin/Heidelberg, Germany, 2006; pp. 1–12.
74. Dankar, F.K.; El Emam, K. The Application of Differential Privacy to Health Data. In *EDBT-ICDT '12, Proceedings of the 2012 Joint EDBT/ICDT Workshops*; Association for Computing Machinery: New York, NY, USA, 2012; pp. 158–166. [CrossRef]
75. Andrés, M.E.; Bordenabe, N.E.; Chatzikokolakis, K.; Palamidessi, C. Geo-Indistinguishability: Differential Privacy for Location-Based Systems. In Proceedings of the 2013 ACM SIGSAC Conference on Computer & Communications Security (CCS '13), Berlin, Germany, 4–8 November 2013; Association for Computing Machinery: New York, NY, USA, 2013; p. 901–914. [CrossRef]
76. Lablans, M.; Borg, A.; Ückert, F. A RESTful interface to pseudonymization services in modern web applications. *BMC Med. Inform. Decis. Mak.* **2015**, *15*, 2. [CrossRef] [PubMed]
77. Noumeir, R.; Lemay, A.; Lina, J.M. Pseudonymization of radiology data for research purposes. *J. Digit. Imaging* **2007**, *20*, 284–295. [CrossRef] [PubMed]
78. Brekne, T.; Årnes, A.; Øslebø, A. Anonymization of ip traffic monitoring data: Attacks on two prefix-preserving anonymization schemes and some proposed remedies. In *International Workshop on Privacy Enhancing Technologies*; Springer: Berlin/Heidelberg, Germany, 2005; pp. 179–196.
79. Fan, J.; Xu, J.; Ammar, M.H.; Moon, S.B. Prefix-preserving IP address anonymization: Measurement-based security evaluation and a new cryptography-based scheme. *Comput. Netw.* **2004**, *46*, 253–272. [CrossRef]
80. Kerschbaum, F. Distance-preserving pseudonymization for timestamps and spatial data. In Proceedings of the 2007 ACM workshop on Privacy in electronic society, Alexandria, VA, USA, 29 October 2007; pp. 68–71.
81. Jawurek, M.; Johns, M.; Rieck, K. Smart metering de-pseudonymization. In Proceedings of the 27th Annual Computer Security Applications Conference, Orlando, FL, USA, 5–9 December 2011; pp. 227–236.
82. Guarnieri, M.; Marinovic, S.; Basin, D. Securing Databases from Probabilistic Inference. In Proceedings of the 2017 IEEE 30th Computer Security Foundations Symposium (CSF), Santa Barbara, CA, USA, 21–25 August 2017, pp. 343–359.
83. Chen, Y.; Chu, W.W. Protection of Database Security via Collaborative Inference Detection. *IEEE Trans. Knowl. Data Eng.* **2008**, *20*, 1013–1027. [CrossRef]
84. Qian, X.; Stickel, M.E.; Karp, P.D.; Lunt, T.F.; Garvey, T.D. Detection and elimination of inference channels in multilevel relational database systems. In Proceedings of the 1993 IEEE Computer Society Symposium on Research in Security and Privacy, Oakland, CA, USA, 24–26 May 1993; pp. 196–205.
85. Yip, R.W.; Levitt, E.N. Data level inference detection in database systems. In Proceedings of the 11th IEEE Computer Security Foundations Workshop (Cat. No.98TB100238), Rockport, MA, USA, 11 June 1998; pp. 179–189.
86. Li, N.; Yu, T.; Antón, A. A semantics-base approach to privacy languages. *Comput. Syst. Sci. Eng. CSSE* **2006**, *21*, 339.
87. Becker, M.Y.; Malkis, A.; Bussard, L. A Framework for Privacy Preferences and Data-Handling Policies. In *Microsoft Research Cambridge Technical Report, MSR-TR-2009-128*; Microsoft Research Cambridge: Cambridge, UK, 2009.
88. Kapitsaki, G.M. Reflecting User Privacy Preferences in Context-Aware Web Services. In Proceedings of the 2013 IEEE 20th International Conference on Web Services, Santa Clara, CA, USA, 28 June–3 July 2013; pp. 123–130. [CrossRef]

89. Ulbricht, M.R.; Pallas, F. YaPPL-A Lightweight Privacy Preference Language for Legally Sufficient and Automated Consent Provision in IoT Scenarios. In *Data Privacy Management, Cryptocurrencies and Blockchain Technology*; Springer: Cham, Switzerland, 2018; pp. 329–344. [CrossRef]
90. Kokolakis, S. Privacy attitudes and privacy behaviour: A review of current research on the privacy paradox phenomenon. *Comput. Secur.* **2015**. [CrossRef]
91. Cranor, L.F.; Guduru, P.; Arjula, M. User Interfaces for Privacy Agents. *ACM Trans. Comput.-Hum. Interact.* **2006**, *13*, 135–178. [CrossRef]
92. Kolter, J.; Netter, M.; Pernul, G. Visualizing past personal data disclosures. In Proceedings of the ARES'10 International Conference on Availability, Reliability, and Security, Krakow, Poland, 15–18 February 2010; pp. 131–139.
93. Angulo, J.; Fischer-Hübner, S.; Pulls, T.; Wästlund, E. *Towards Usable Privacy Policy Display & Management—The PrimeLife Approach*; HAISA: Coron, Philippines, 2011.
94. PrimeLife. PrimeLife-Bringing Sustainable Privacy and Identity Management to Future Networks and Services, 2008–2011. Available online: http://primelife.ercim.eu/ (accessed on 13 January 2020).
95. Disterer, G. ISO/IEC 27000, 27001 and 27002 for Information Security Management. *J. Inf. Secur.* **2013**, *2013*, 92–100. [CrossRef]
96. Ludwig, H.; Keller, A.; Dan, A.; King, R.P.; Franck, R. *Web Service Level Agreement (WSLA) Language Specification*; Ibm Corporation: Endicott, NY, USA, 2003; pp. 815–824.
97. Oldham, N.; Verma, K.; Sheth, A.; Hakimpour, F. Semantic WS-agreement partner selection. In Proceedings of the 15th international conference on World Wide Web, Montreal, QC, Canada, 2006; pp. 697–706.
98. Dobson, G.; Lock, R.; Sommerville, I. QoSOnt: A QoS ontology for service-centric systems. In Proceedings of the 31st EUROMICRO Conference on Software Engineering and Advanced Applications, Porto, Portugal, 30 August–3 September 2005; pp. 80–87.
99. Lamanna, D.D.; Skene, J.; Emmerich, W. SLAng: A language for defining service level agreements. In Proceedings of the Ninth IEEE Workshop on Future Trends of Distributed Computing Systems, San Juan, Philippines, 28–30 May 2003; pp. 100–106. [CrossRef]
100. Meland, P.H.; Bernsmed, K.; Jaatun, M.G.; Castejón, H.N.; Undheim, A. Expressing cloud security requirements for SLAs in deontic contract languages for cloud brokers. *Int. J. Cloud Comput.* **2014**, *3*, 69–93. [CrossRef]
101. Oberle, D.; Barros, A.; Kylau, U.; Heinzl, S. A Unified Description Language for Human to Automated Services. *Inf. Syst.* **2013**, *38*, 155–181. [CrossRef]
102. Kennedy, J. Towards Standardised SLAs. In *Euro-Par 2013: Parallel Processing Workshops*; an Mey, D., Alexander, M., Bientinesi, P., Cannataro, M., Clauss, C., Costan, A., Kecskemeti, G., Morin, C., Ricci, L., Sahuquillo, J., et al., Eds.; Springer: Berlin/Heidelberg, Germany, 2014; pp. 105–113.
103. Gerl, A.; Meier, B. The Layered Privacy Language Art. 12–14 GDPR Extension–Privacy Enhancing User Interfaces. *Datenschutz Und Datensicherheit-DuD* **2019**, *43*, 747–752. [CrossRef]
104. Data Transfer Project. Data Transfer Project Overview and Fundamentals, 2018. Available online: https://www.zurich.ibm.com/security/enterprise-privacy/epal/Specification/ (accessed on 2 June 2020).
105. Facebook. Facebook Reports First Quarter 2020 Results, 2020. Available online: https://investor.fb.com/investor-news/default.aspx (accessed on 2 June 2020).
106. Niu, C.; Zheng, Z.; Wu, F.; Gao, X.; Chen, G. Trading Data in Good Faith: Integrating Truthfulness and Privacy Preservation in Data Markets. In Proceedings of the 2017 IEEE 33rd International Conference on Data Engineering (ICDE), San Diego, CA, USA, 19–22 April 2017; pp. 223–226.
107. Bataineh, A.S.; Mizouni, R.; Barachi, M.E.; Bentahar, J. Monetizing Personal Data: A Two-Sided Market Approach. *Procedia Comput. Sci.* **2016**, *83*, 472–479. [CrossRef]
108. Wang, W.; Liu, A.; Li, Z.; Zhang, X.; Li, Q.; Zhou, X. Protecting multi-party privacy in location-aware social point-of-interest recommendation. *World Wide Web* **2019**, *22*, 863–883. [CrossRef]
109. Wang, C.J.; Ng, C.Y.; Brook, R.H. Response to COVID-19 in Taiwan: Big data analytics, new technology, and proactive testing. *JAMA* **2020**, *323*, 1341–1342. [CrossRef] [PubMed]