COVID-19 Apps and Individual Privacy

INFO30006 Report

Group TT

Darren Zhang (1086273)
Jason F. Suhartanto (1086250)
Vincent A. Hartono (1086226)

Semester 2, 2021.

Contents

1 Introduction

At the beginning of the COVID-19 pandemic and with the absence of a vaccine treatment, digital contact tracing has always been claimed to be the key to control and reduce the spread of the virus. This introduced the development of several automated tracing methods to be implemented across different countries as a public health strategy. Essentially, contact tracing apps are used to identify and notify people who have been in contact with a positive case. However, the centralised nature of the apps sparked concerns over privacy since user's control over data collection and protection is very limited. Additionally, increased surveillances are implemented in response to combat outbreaks, and new digital records are created from the increasing online activities.

Since then, some governments have been pushed to alter the configuration of the tracing apps to meet these concerns — to a decentralised system where private user information is stored on their own devices instead of a centralised server. While a decentralised model is widely regarded ethically preferable for its *privacy-preserving by design* architecture, a centralised system is proven to have a better efficiency on containing the spread. Therefore, a debate between these two systems needs to be reopened, on whether countries decide to prioritize privacy or a more efficient contact tracing for the benefit of their public health.

This report aims to discuss the key concepts of centralised and decentralised contact-tracing systems and explore their impacts on privacy and ethics. The earlier sections provide detailed technical explanations of the architecture and real-world implementations of both models, and later sections discuss their implications on privacy, security, and ethics as bases for our recommendations. Generally, our report recommends implementing a centralised system for its better tracing efficiency, as long as its benefit for the public health outweighs their higher privacy risks and no decentralised models can achieve better tracing performance.

2 Centralised vs Decentralised

Most countries have developed their COVID-19 tracing apps with each of those having its challenges in ensuring the privacy and security of millions of people (Vaudenay, 2020). Although the specific implementations might differ for each country, the underlying system architecture is mainly based on either a centralised, decentralised, or a hybrid model that combines both models (Ahmed et al., 2020). These three models have their own advantages and disadvantages in regards to privacy and security. The main coverage of this report, however, will be only on centralised and decentralised model.

2.1 Centralised Model

A centralised model is a model that relies on a central trusted system to store encrypted information, matching contacts between infected users, and deliver notification for identified users (Raman et al., 2021). This system allows the identity of the registered users to be known to the central system. Some examples of models that adopt a centralised architecture are Herald[1] and BlueTrace[2].

The registration process of a centralised model, as in step 1 and 2 of Figure 1, requires the

[1] Herald Protocol, as documented on https://heraldprox.io/, accessed on September 15, 2021.

[2] BlueTrace Protocol, as documented on https://bluetrace.io/, accessed on September 15, 2021.

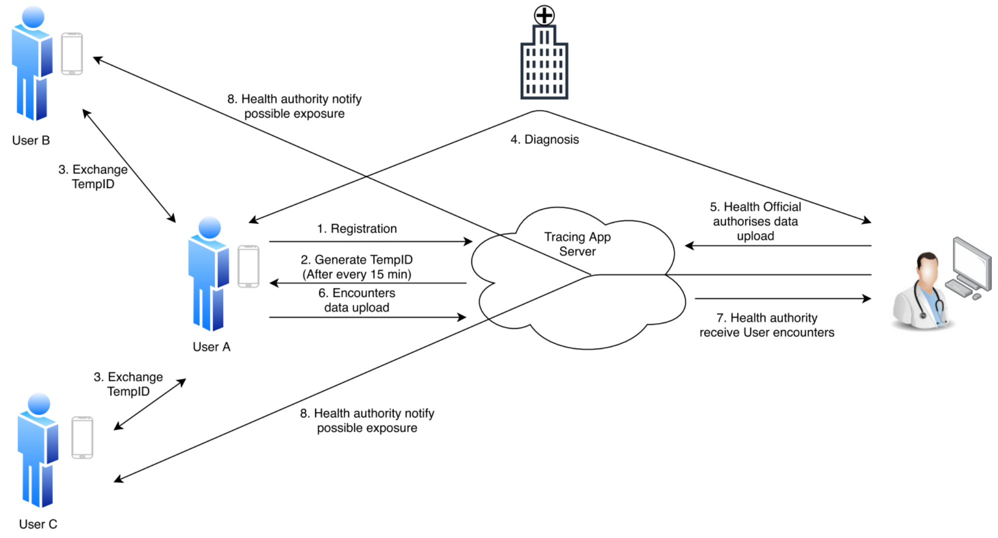
User B
8. Health authority notify possible exposure
4. Diagnosis
3. Exchange TempID
5. Health Official authorises data upload
1. Registration
Tracing App Server
2. Generate TempID (After every 15 min)
6. Encounters data upload
User A
7. Health authority receive User encounters
3. Exchange TempID
8. Health authority notify possible exposure
User C

Figure 1: A centralised tracing app architecture (Ahmed et al., 2020).

user to download the application and provide necessary information to identify the user. The system will then send an OTP (One Time Password) to verify the phone number. After the verification is done, the system will send a temporary ID (TempID) with expiry to the user. After completing the registration processes, the user will start to exchange encounter messages with another user as seen in step 3. This encounter message, which includes their TempID, phone model, and transmit power, are exchanged between one user and another and will be stored locally on their phones. When a user tests positive, they will be asked voluntarily to upload these encounter messages of all other users they had exchanged in the past few days. If the user agrees to upload their encounter messages, they will need to verify an OTP generated by the central system to upload all the stored messages, as in steps 4-6 of the figure. After the COVID-19 positive user uploads the necessary data to the central system, the system will conduct contact matching and build a list of possible infected users. Finally, in step 7 of the figure, health authorities will then get the list containing the contact information of possible infected users for further process.

In summary, there is a central server or system in a centralised model that serves as a core to store users encrypted information, generate temporary IDs, and conduct contact matching for notification process of possibly infected users. Additionally, for this centralised model to function properly, the server is assumed to be trusted.

2.2 Decentralised Model

A decentralised model, in contrast to a centralised model, does not rely on a centralised server to store user information. This model tries to achieve minimum information exchange with the central server and anonymises user identity when communicating with the server (Raman et al., 2021). This model also hopes to reduce privacy infringement and eliminates breaches on the central server (White & van Basshuysen, 2021). One of the most widely used decentralised framework that is adopted by many countries is the Exposure Notification[3] (EN) framework developed by Google and Apple.

[3] Exposure Notification API, as documented on https://www.google.com/covid19/exposurenotifications/, accessed on September 15, 2021.

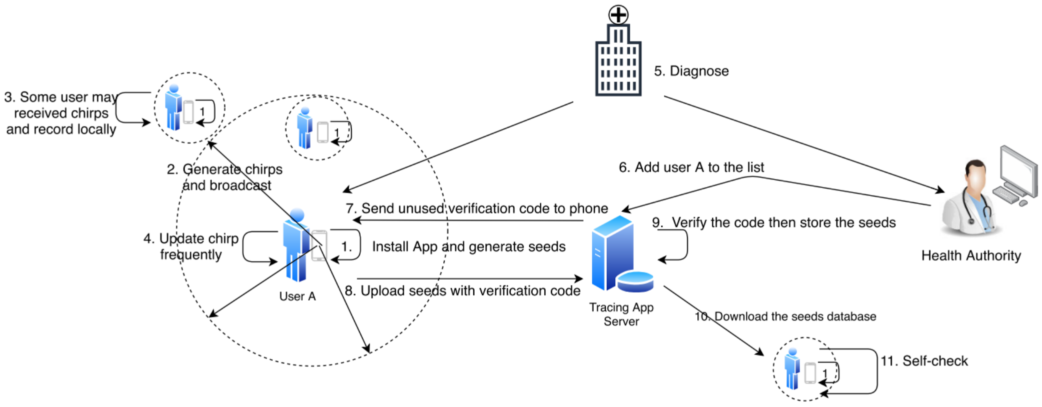
3. Some user may received chirps and record locally
2. Generate chirps and broadcast
4. Update chirp frequently
User A
1.
Install App and generate seeds
7. Send unused verification code to phone
8. Upload seeds with verification code
5. Diagnose
6. Add user A to the list
9. Verify the code then store the seeds
Health Authority
Tracing App Server
10. Download the seeds database
11. Self-check

Figure 2: A decentralised tracing app architecture (Ahmed et al., 2020).

For the decentralised model, as in Figure 2, there is no active registration needed as each local devices generate their temporary IDs periodically. The users will then exchange the locally generated TempIDs with other users they come in contact with. These TempIDs will be stored locally on users' phones and will not be linked to any individual or phone to maintain user anonymity. When a user tests positive for COVID-19, they can voluntarily upload the list of their exchanged TempIDs to the central server. The central server will then broadcast a list containing TempIDs of all positive COVID-19 users. All users will download this list automatically every day and do an automatic check to see if they had stored one of these positive COVID-19 TempIDs. The application will then appropriately notify the user.

The key difference of the decentralised model is the existence of local devices that generate TempIDs, and conduct contact matching for notification process. The sole purpose of the central server is to receive the TempIDs of the positive COVID-19 users and broadcast it to all users.

2.3 Tradeoffs

The differences between centralised and decentralised models create some trade-offs between them. The centralised model concentrates more on the effectiveness of contact tracing. It enables the authorities to check data from the central server to trace and contact the infected users. On the other hand, the decentralised model focuses on preserving the users' privacy by anonymising the data uploaded to the central server (Hernández-Orallo et al., 2020).

The effectiveness of contact tracing applications is correlated with these three measures: precision, utilisation, and speed. Precision is measured based on the ability of the application to detect contacts that may result in possible infection. In the centralised model, by having the ephemeral identifier and the user's contact data, the contact can be quickly identified and alerted correspondingly. However, in the decentralised model, the central server does not store the user's ephemeral identifier. The authorities can only broadcast the ephemeral identifiers to all users of the application (Vaudenay, 2020).

Additionally, the effectiveness of contact tracing applications is based on utilisation, which is the adoption rate of the application (Hernández-Orallo et al., 2020). The decentralised model might seem to get a higher utilisation rate when we look at the fact that it is privacy-preserving by design, which means that it imposes fewer risks on possible breaches. However, this principle does not hold due to privacy concerns of tech-savvy individuals leaking the information. Furthermore, a study by White & van Basshuysen (2021) found that due to its slower contact tracing, a decentralised model would require a utilisation rate of 60% to 80% to be effective. Contrarily, a centralised model would only require around a 50% adoption rate to be effective.

The last measure is based on the speed of contact tracing. A centralised model will be faster due to the authorities being able to trace the user immediately by having the user's data. Contrarily, a decentralised model will depend on the user's eagerness to upload their ephemeral identifiers to the application which then will be broadcasted to other users of the application. Each individual who has the application would have to check the application themselves to see if they came into contact with the infected users. This could cause a longer contact tracing time for all possible contacts and ultimately will cause tracing to be more inefficient (Hernández-Orallo et al., 2020).

2.4 Implications on Privacy

The data collection of both models can have their own privacy implications on users. For a centralised model, it may raise some concerns about the purpose of the data collected, information that can be inferred from the data collected, and subjects who can view the data (Culnane, 2021). On the other hand, these issues are not prevalent in the decentralised model as this model does not collect and store any information that identifies users.

Generally, people are concerned about the purpose of the data collected as many apps collect private data with no relation to stopping the spread of the virus (Culnane, 2021). In some cases, applications require permissions to access contacts, photos, media, files, location data, and camera. Procedures from some governments, such as Australia and Singapore, require sharing the collected data of users with other departments and law enforcements. Consequently, there is also fear on how the data may be used after the pandemic ends (Sharma & Bashir, 2020).

The data on different users' interactions can be used to infer the relationship between the users. This can lead to the creation of a social graph and the user's location pattern (Culnane, 2021). There is also the possibility of users profiling from the data collected, as suggested by Wen et al. (2020), and no clear information on who can view these data. The privacy policy of most apps states that only relevant authorities will have access to it, but no additional information whether other state authorities, law enforcement, or other organisations will be handed the information (Culnane, 2021). There is also the possibility of data breaches done by unauthorised third parties to track the identity of other users (Cho et al., 2020).

Contrarily, the issues stated above do not have a substantial impact on the decentralised model, as the decentralised model does not collect and store users' information (Culnane, 2021). This implies that no data can be processed by relevant authorities or other unauthorised parties. However, decentralised models are also still prone to individual device compromise outside of the data collection process.

3 Implementations of COVID-19 Tracing Apps

Country	App Name	Features	Model
Australia	COVIDSafe	BTS	Centralised
UK (England, Wales)	NHS COVID-19	BTS	Decentralised
Singapore	TraceTogether	BTS, QR Code Scanner	Centralised
Canada	COVID Alert	BTS	Decentralised
Germany	Corona-Warn-App	BTS	Decentralised

Table 1: Examples of contact tracing applications; *BTS: Bluetooth Signals.*

Different countries have opted to develop their tracing apps based on either a centralised or a decentralised model, as seen in Table 1. This section will focus on the implementation of tracing apps developed by Australia and the United Kingdom (UK).

Australia's COVIDSafe[4] is a tracing app based on a centralised architecture. It will ask for the user's details when registering on the app. When the user of the app tests positive, they will be asked to voluntarily upload the TempIDs of the other users they had come in contact with. COVIDSafe keeps the users' data in a central server for contact tracing and notification purposes. UK's NHS COVID-19[5], on the other hand, allows COVID-19 positive users to voluntarily upload their TempIDs to be broadcasted to other users, but the data is not saved after the period of infection has ended.

Both Australia and UK's COVID-19 tracing apps today have some differences when compared to their original release, as discussed by Culnane (2021). Australia now uses the Herald Protocol for their COVIDSafe app after switching from the BlueTrace protocol. These changes are made to achieve significant improvements to bluetooth tracking capabilities and performance of the app. The UK with its NHS COVID-19 app now adopts a decentralised model, after switching from a centralised model at the early stages of the pandemic because of privacy concerns. The current decentralised model adopts the Exposure Notification (EN) framework, which received a better public sentiment for its privacy.

Additionally, Culnane (2021) also discusses the different approaches with the release of COVID-19 tracing apps for both Australia and UK. The UK has released its client and server code to the public since the first version of the app when they implemented the centralised model. This makes it possible for the general public to see how the data is processed and if it infringes on the user's privacy. On the other hand, Australia with its COVIDSafe app only makes the client code open source, while their server code is not publicly available. This approach by Australia needs the users to trust that the authorities will follow existing security and privacy policy. The key takeaway is that releasing source code on both client and server can increase the confidence in the management of privacy and security as the people can review the code on their own (Culnane, 2021). This is different with apps that only release the client source code. Since the public cannot review the security and privacy policy, there is a need for complex action such as passing legislation to maintain the public's trust in the app.

[4] COVIDSafe privacy policy, retrieved from https://covidsafe.gov.au/, accessed on September 13, 2021.

[5] NHS COVID-19 privacy policy, retrieved from https://www.gov.uk/, accessed on September 13, 2021.

4 Security Risks and Countermeasures

Previous sections have shown that both centralised and decentralised models entail their own risks, including from a security aspect. In some cases, these security risks might be a necessary trade-off as some apps try to optimise the contact tracing and notification process. In light of this, there are some countermeasures that can be applied to mitigate the security risks.

As suggested by Hernández-Orallo et al. (2020), the efficiency of tracing apps can be optimised by trading off privacy risks. The decentralised model adopts this notion of privacy-preserving by design, which refers to the principle of inherent safety. Inherent safety means that it eliminates any possible risk without having to apply more safety measures to lower the probability of the risk occurring (Möller & Hansson, 2008). However, decentralised models still do not rule out the possibility of breaches. For instance, there are still some concerns regarding parties identifying infected users by getting the log of additional location information and the corresponding pseudonymised identifiers from the broadcast. In this case, it could expose a bigger privacy risk compared to the health authorities owning the data (Li et al., 2020). Additionally, attacks against decentralised models are usually harder to detect, have a wider scale of the attack, and do not have many generalised countermeasures.

On the other hand, centralised models offer a variety of countermeasures, such as accounting and auditing (Vaudenay, 2020). However, if an attacker breaks into the central database, they can reveal the social graphs from data of multiple users stored in the database. Even though the severity of revealing social graphs is higher than identifying a particular user, the possibility of this kind of breach happening is not as likely as individual device breaches (White & van Basshuysen, 2021). Therefore, for the decentralised model to be preferable, it needs to have a much lower risk of possible breaches (White & van Basshuysen, 2021). Since we have previously discussed that decentralised models also entail risks of privacy attacks, therefore the centralised model works better in capturing the objective of contact tracing application based on the efficiency and speed of intervention.

According to Ranisch et al. (2020), there are two types of ethical risks correlated with contact tracing applications: risks for privacy, and risks for public health if the contact tracing effort turns out to be ineffective. These two risks contradict each other. If there is a high probability of a centralised system being effective, an increase in the level of privacy risks may be tolerable. The same goes for decentralised systems, where there needs to be a much lower privacy risk for the system to be acceptable. Since both of the models impose various risks and the overall ethical risks from the centralised system are less severe, the centralised model is ethically more preferred compared to the decentralised model (White & van Basshuysen, 2021).

5 Measure of Success

To further evaluate the performance of tracing apps, a definition of a successful tracing app is provided. A tracing app is considered to be successful if it serves its purpose to help reduce spread by effectively notifying contacts of a positive case. This can be very crucial since it cancels out the need of state-wide lockdowns because of untraceable local cases. Additionally, non-close contacts are able to go out without posing any risk to the community. There are several factors that can make a successful contact tracing implementation.

In addition to the choice of the system being centralised or decentralised, the speed of intervention and a sufficient uptake of the app are very crucial factors for an effective contact tracing (White & Van Basshuysen, 2021). Speed of intervention emphasises the importance of quickly notifying possible close contacts since individuals are very likely to become infectious shortly after being infected. White & Van Basshuysen (2021) also states that contact tracing will be effective only if the time between infection and isolation can be minimized. Meanwhile, a sufficient uptake means an app needs a large enough percentage of the population to be traced for the app to be effective. There are many ways to achieve this depending on the situation of each country. In this case, it may be important to use an acceptance model to understand the public sentiment towards tracing apps. This will be crucial to determine the optimal steps that are needed to be taken, and ultimately increase the efficiency of tracing apps.

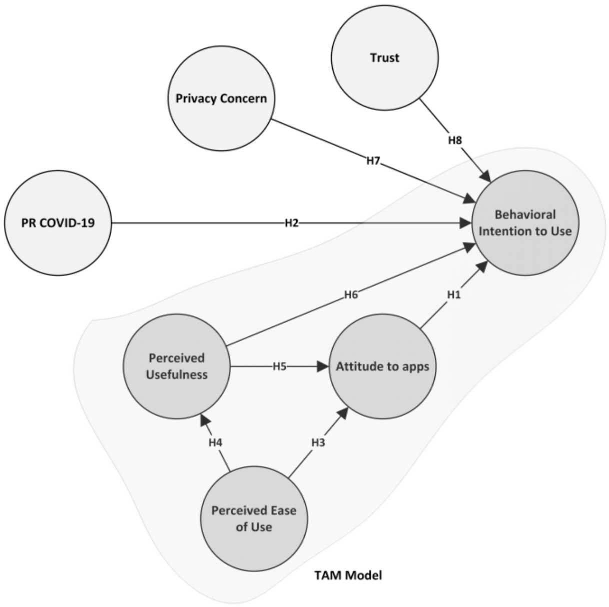

Figure 3: The extended Technological Acceptance Model (TAM) proposed by Velicia-Martin et al. (2021).

People's opinion and sentiment on using tracing apps can be evaluated using the Technological Acceptance Model (TAM). A study by Velicia-Martin et al. (2021) constructed an extended TAM model, as seen in Figure 3, to investigate whether citizens are willing to adopt a contact tracing application in the COVID-19 pandemic. The model includes variables such as perceived usefulness, privacy concern, and trust as important factors to understand people's intention on using tracing apps. The study found most variables to have high explanatory power on people's acceptance, and concludes that a high perceived utility of the app is likely to outweigh the possible privacy concerns if they found the app to have a concrete benefit. In conclusion, understanding the public sentiment towards tracing apps is key to determine the factors that can increase usage and ultimately make tracing efforts more effective.

6 Recommendations

Weighing all privacy, security, and ethical considerations of both centralised and decentralised systems, we have found that the ideal implementation across different countries may be different. Our overall recommendation is that with all other things being equal, a centralised system should always be preferred for a better tracing, and a higher level of privacy risk should be acceptable if it is proven to have a better benefit for the public health. For a country that puts a high importance of privacy, it is important to address the scale of ongoing outbreaks in the country and evaluate if a decentralised tracing model is enough control the spread, as in some cases it is not worth preserving privacy at the cost of a less efficient mitigation of outbreaks. Countries should put their main importance on their public health, and privacy should be a concern when the current tracing efforts are already sufficient contain spreads.

Additionally, at the time of writing, there are no decentralised models that can provide a better tracing efficiency than a centralised system. As sound the technical explanations may seem, this can change in the future as many of current studies are developing decentralised models using state-of-the-art technologies that are always developing. For example, a recent study by Matthews et al. (2020) proposed a decentralised diagnostic tracing system that opposes the current centralised laboratory-based diagnostic modality. Another exploratory study (Hasan et al., 2021) investigates a decentralised architecture based on blockchain technology, which enforces accountability and transparency for user privacy. In conclusion, it is important to always follow the development of current tracing technologies that may optimise benefits for both tracing efficiency and privacy. Furthermore, as a preventive countermeasure to possible pandemics in the future, countries also need to consider tracing systems that do not necessarily require each person to have a device. This will be crucial for containment in developing countries where the use of mobile devices are not as prominent as in developed countries.

7 Conclusion

After exploring both centralised and decentralised tracing models, a centralised system is found to have more efficient tracing through a significantly faster notification process which can be very beneficial for the public health in containing spreads. However, some countries are opting for a decentralised system for the reason of better privacy, which cannot be achieved by a system with centralised data storage. However, this does not entirely cancel privacy risks for a decentralised model since identifiable ephemeral identifiers of positive cases are also broadcasted. Additionally, while centralised systems are prone to central server breaches, decentralised architectures are also prone to other sophisticated types of security attacks, which may be harder to detect and mitigate. Furthermore, we also found that a Technology Acceptance Model (TAM) is helpful to map public sentiment towards using a tracing app, which can be a crucial step since creating a healthy public trust is key to keep a high number of app usage among the population.

Although decentralised systems are developed for better privacy preservation in exchange for a slower tracing process, they still entail their own privacy and security risks. Therefore, a system with a higher chance of being effective — a centralised system — should always be preferred to provide a more efficient tracing despite the presence of higher privacy risks that should be acceptable.

References

Ahmed, N., Michelin, R. A., Xue, W., Ruj, S., Malaney, R., Kanhere, S. S., … Jha, S. K. (2020). A survey of covid-19 contact tracing apps. *IEEE Access*, *8*, 134577-134601. doi: 10.1109/ACCESS.2020.3010226

Cho, H., Ippolito, D., & Yu, Y. W. (2020). *Contact tracing mobile apps for covid-19: Privacy considerations and related trade-offs.*

Culnane, C. (2021). Security and privacy in covid apps. Retrieved from `https://canvas.lms.unimelb.edu.au/courses/108231/files/8694719?module_item_id=3135389`

Hasan, H. R., Salah, K., Jayaraman, R., Yaqoob, I., Omar, M., & Ellahham, S. (2021). Covid-19 contact tracing using blockchain. *IEEE Access*, *9*, 62956-62971. doi: 10.1109/ACCESS.2021.3074753

Hernández-Orallo, E., Calafate, C. T., Cano, J.-C., & Manzoni, P. (2020). Evaluating the effectiveness of covid-19 bluetooth-based smartphone contact tracing applications. *Applied Sciences*, *10*(20). doi: 10.3390/app10207113

Li, T., Jackie, Yang, Faklaris, C., King, J., Agarwal, Y., … Hong, J. I. (2020). *Decentralized is not risk-free: Understanding public perceptions of privacy-utility trade-offs in covid-19 contact-tracing apps.*

Matthews, Q., da Silva, S. J. R., Norouzi, M., Pena, L. J., & Pardee, K. (2020). Adaptive, diverse and de-centralized diagnostics are key to the future of outbreak response. *BMC biology*, *18*(1), 1–5.

Möller, N., & Hansson, S. O. (2008). Principles of engineering safety: Risk and uncertainty reduction. *Reliability Engineering & System Safety*, *93*(6), 798–805.

Raman, R., Achuthan, K., Vinuesa, R., & Nedungadi, P. (2021). Covidtas covid-19 tracing app scale—an evaluation framework. *Sustainability*, *13*(5). Retrieved from `https://www.mdpi.com/2071-1050/13/5/2912` doi: 10.3390/su13052912

Ranisch, R., Nijsingh, N., Ballantyne, A., van Bergen, A., Buyx, A., Friedrich, O., … Wild, V. (2020). Digital contact tracing and exposure notification: ethical guidance for trustworthy pandemic management. *Ethics and information technology*, 1–10.

Sharma, T., & Bashir, M. (2020). Use of apps in the covid-19 response and the loss of privacy protection. *Nature Medicine*, *26*(8), 1165–1167.

Tang, Q. (2020). Privacy-preserving contact tracing: current solutions and open questions. *arXiv preprint arXiv:2004.06818*.

Vaudenay, S. (2020). *Centralized or decentralized? the contact tracing dilemma.* Cryptology ePrint Archive, Report 2020/531. (`https://ia.cr/2020/531`)

Velicia-Martin, F., Cabrera-Sanchez, J.-P., Gil-Cordero, E., & Palos-Sanchez, P. R. (2021). Researching covid-19 tracing app acceptance: incorporating theory from the technological acceptance model. *PeerJ Computer Science*, *7*, e316.

Wen, H., Zhao, Q., Lin, Z., Xuan, D., & Shroff, N. (2020). A study of the privacy of covid-19 contact tracing apps. In *International conference on security and privacy in communication systems* (pp. 297–317).

White, L., & van Basshuysen, P. (2021). Privacy versus public health? a reassessment of centralised and decentralised digital contact tracing. *Science and Engineering Ethics*, *27*(2), 1–13.

White, L., & Van Basshuysen, P. (2021). Without a trace: Why did corona apps fail? *Journal of medical ethics*.