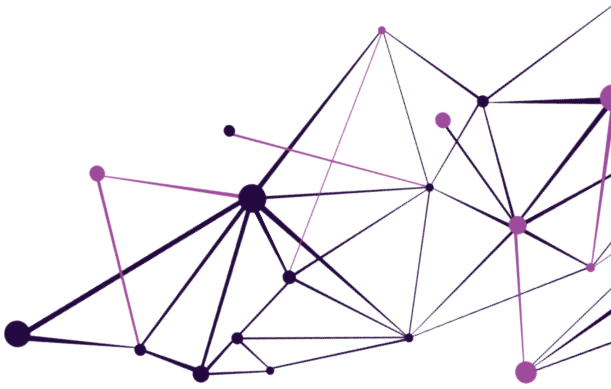
RUNNING AN ETHICS AUDIT

Case Study: Digital Contact Tracing

EXECUTIVE SUMMARY

Since the onset of the COVID-19 pandemic, the tech industry has been hard at work developing potential solutions to help fight the virus. One of the more prominent proposals has been for the use of Digital Contact Tracing (DCT), the technical application of manual contact tracing methods. DCT quickly became a topic of hot debate, as privacy and trust concerns were thrust into the spotlight as never before. Suddenly, ethical principles that had been discussed theoretically at length had a clear actionable case study plastering headlines as society began to understand the importance of ethics in emerging technology and the risks that unchecked innovation holds.

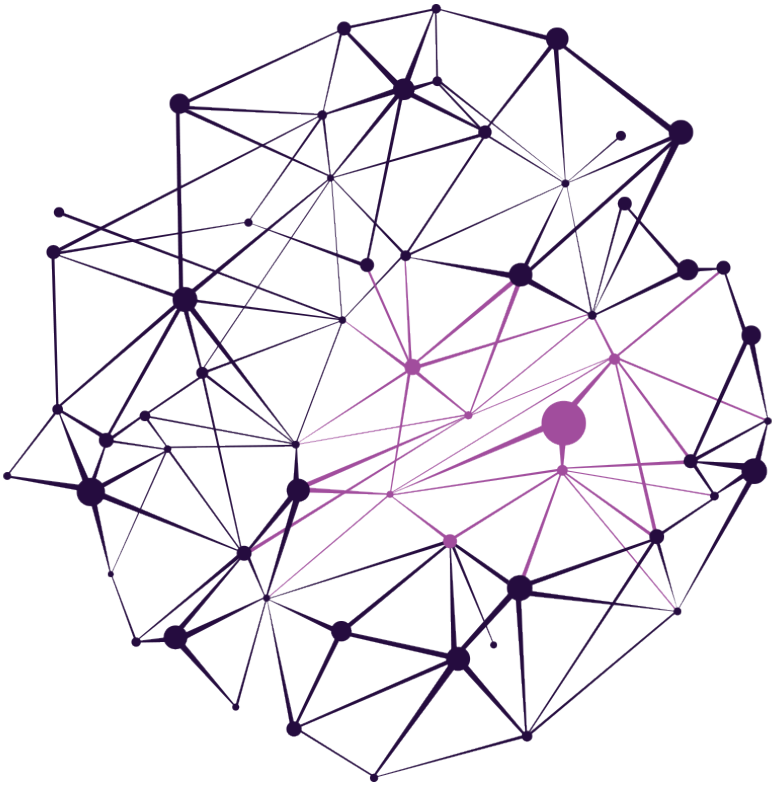
Although there is now a clear call for the incorporation of ethical principles in DCT applications, the path to such is not certain. There is no rule book for how to implement DCT while still respecting human rights and ethics, no clear cut guide laying out each step to take. However, this does not mean that we are without a solution. As we look to roll out DCT applications to help fight the coronavirus, it is vital that these applications undergo due diligence ethics audits if we are to respect human dignity.



The purpose of this paper is to illustrate, from a high level perspective, what it would look like to run a preliminary ethics audit on a DCT application, with the hope that it may provide some guiding structure to those currently working on similar applications.

To demonstrate such, we will be evaluating a hypothetical Bluetooth Digital Contact Tracing application against the seven principles laid out by the European Union's High-Level Expert Ethics Guidelines. For each of the seven principles we will highlight the essential considerations that must be made in order to embed the principle into practice in terms of the chosen technology. These considerations are not only applicable to DCT, but also exemplify the high-level process any COVID-19 tech should be undertaking.

WHAT IS DIGITAL CONTACT TRACING (DCT)?



It is generally accepted that people who have come into close contact with an infected individual are at higher risk of contracting and spreading the disease. Contact tracing is a manual method that involves interviewing an infected person about their activities so as to closely identify and monitor the people they came into contact with for the purpose of reducing the spread of the infection. Contact tracing is not a novel concept, in fact it was used during the ebola epidemic to effectively control outbreaks, as well as playing a major role in the eradication of smallpox.

Manual Contact Tracing is time and labour intensive. In light of COVID-19, efforts have been made to digitise the process by creating Digital Contact Tracing applications. DCT aims to automate the process by utilizing the potential of smartphones to gather contact and location data.

DCT apps use built-in GPS or Bluetooth (BLE) technology to track users. GPS tracking is location-based, as it records where and when an individual has been in certain locations. BLE, on the other hand, does not necessarily record location information and instead relies on the phone's ability to detect when it has come into close proximity with another phone. This offers privacy advantages but limits the traceability of those who may have come into contact with the same surface. Currently, all apps developed for COVID-19 DCT use Bluetooth due to its stronger privacy protection and lower battery consumption.

HOW TO APPROACH THIS REPORT

In the following sections we will be drawing on the ethical principles as defined by the EU High Level Expert Group's Ethical Guidelines. We will be considering Privacy and Security, Human Agency and Autonomy, Transparency, Fairness, Accountability, Technical Robustness and Safety, and Societal Wellbeing.

Each section will begin with a brief working definition for the principle. Following the definition, we will explore a handful of the considerations that must be made in light of the specific principle. The purpose of these considerations is to illustrate the mental thought process that must be undertaken when implementing the chosen ethical principle into the technology. Finally, we will end with commentary from Ethical Intelligence's perspective on the matter.

As you read, keep in mind this is a high level case study meant solely to illustrate how one would begin to go about running an ethics audit on a digital contact tracing application.

PRIVACY & SECURITY

Privacy and security as principles closely related to the prevention of harm. According to the EU, this is a fundamental right when regarding AI systems advocating for better data governance that ensures the quality and integrity of the data used.

CONSIDERATIONS

When looking to implement Privacy and Security in regards to DCT, it is important to evaluate the risks of data storage via the use of centralised or decentralised networks. Both networks use bluetooth signals to track when smartphones are close to one another, alerting a user to a person infected with COVID-19. The centralised model sends the data to a remote server where the information is processed and matches can be made with other contacts to alert potential COVID-19 infection. On the other hand, the decentralised model enables the user more autonomy with their data by storing information and matching contacts on personal devices. In other words, centralised networks upload user data into one single server point, whereas decentralised use multiple sources. Although the centralised network allows for analysis of public health, decentralised networks have gained popularity more recently due to their ability to distribute data among different sources. The evident advantages to a decentralized network are the reliability of the system, the scale and most importantly, privacy. Consequently, the risk to privacy is far less than with one single point that is present with a centralised server.

This leads into the key consideration for privacy in the context of DCT of data handling. There is widespread uncertainty of who will have control of this sensitive data, as well as where exactly will it be stored and for how long. Trust is imperative if DCT applications are to be adopted, and trust can only be achieved if the individuals using the application believe their privacy is being respected and security of their data is upheld.

EI'S PERSPECTIVE

When considering proper data collection, storage and usage in the context of DCT, personal privacy is a critical concern in Western cultures. In light of an individual's entitlement to privacy of personal data, government mandate of the application would come into direct violation of individual privacy as defined by the European Union. Furthermore, it appears that decentralised systems respect this notion of individual privacy. However, it is important to consider that the case for centralised systems may be strengthened if the public was willing to trade a certain level of privacy for analysis of the movement of the virus. This of course could only be done if the controller of the central system was highly trusted by the public. Either way, each system requires that the individual be privy and have some measure of control over how their data is being collected and handled, and for what purposes it is being used.

AGENCY & AUTONOMY

Agency and autonomy are closely related terms, however they can be distinguished as such; Agency is the capacity of individuals to act independently and to make choices of their own free will (Ability), Autonomy is the capacity to be one's own person, free from manipulative or distorting external forces (Freedom).

CONSIDERATIONS

Though similar, it is useful to consider both these terms individually, as the perception of the technology prior to utilisation, the safeguards that exist while in use, and the freedom to change our mind, with our data no longer utilised, are vital when using DCT technology. Major considerations include the ease of use of the technology, obtaining consent and how the data is ultimately utilised.

These considerations translate into concrete questions such as the ability to discontinue the service. Is it as simple as deleting an app? Or will it require 'un-subscription' once used, alongside burdensome bureaucracy complicating the process. Will the consent process be many pages in length, effectively dissuading those who would otherwise embrace such technology? Other considerations include the lag in time between an infection event and informing all those involved, to protect the anonymity of people involved in such events, perhaps by a standard hourly alert giving either the all-clear or a warning alert, if one is warranted. The use of DCT to track individuals based on the data obtained from their associations would need to be counterbalanced through proper data governance arrangements, giving confidence to users that data is safely encrypted. Users should be alerted when crossing borders that their application will either no longer function or that an application that works across borders will likely be sending their data to a different collection site, if centralised, and be under different legal protection, so the user can then decide how to proceed. The alert may need to be given prior to travel, especially if the nation being visited has a different approach to privacy.

EI'S PERSPECTIVE

It is vital for consent to be obtained transparently and that a thorough, but simple and concise, explanation is given to users who may not be IT literate. The utilisation of the data and rights of access must be defined. The application should either be limited to a sovereign state or a mutual agreement about data use obtained from all countries where it will be utilised and this clarification provided to users, with any update giving them an opportunity to delete the application/technology and their data to no longer be used. The state and/or companies must not remove individual autonomy and require or pressure individuals into using the technology, no limitations on daily living or transport should be made and the choice should be purely to benefit the individual and wider society, if the individual chooses to do so. Importantly the information of an AI system should be intelligible for review by a competent human authority, but not identifiable. The ability of nominated third parties to delete or withdraw consent must be considered should the individual become incapacitated or die and be unable to change their consent. Death should imply automatic discontinuation of the tracking service and this is a further reason why deletion of the app/add on should constitute automatic withdrawal of consent for DCT.

TRANSPARENCY

Data, systems and AI business models must all exhibit a certain level of transparency by clearly communicating decision making factors to the relevant stakeholders. Furthermore, any human interaction with an AI system should be informed in so much that the human is aware they are interacting with an AI and understands the system's capabilities and limitations.

CONSIDERATIONS

Transparency is essential to building trust in technology, and trust in turn is necessary for the required adoption of any DCT application. In order to implement transparency in a way that builds trust, significant considerations must be made in terms of mission creep. Essentially, as a DCT application is developed, it may require deeper data points than originally projected in order to accurately track when people come into contact. It is important to consider what data the DCT application will need to collect and how this data collection will be communicated in a clear and effective manner to the user. Furthermore, if higher accuracy is needed, will the scope of data be expanded, or will the operating systems instead need to be updated? An application versus an operating system are two very different manners of enabling contact tracing, how will the difference between the two be communicated to stakeholders so that users understand fully when they are being traced and when they are not? Personal information is constantly leaked, often unbeknownst to users, due to the ability of systems to infer information, such as mental health metrics from patterns of music consumption. It is vital that any developer of a DCT application considers the limitations that need to be set on the data collected, or in other words, clearly defines the scope of the data and system itself. Moreover, if these systems are being deployed on a national level by government bodies, it is important to consider clear timelines and end dates to the application. Contact tracing may be acceptable during a world pandemic, but in normal times it is likely to be rejected and therefore must have clearly defined timelines, not vague end dates.

EI'S PERSPECTIVE

Essentially, transparency considerations in light of DCT will boil down to clearly defining scope of the system and data collected, as well as how it will be communicated to the relevant stakeholders. This means understanding and setting limitations on what data will be collected, how long it will be stored for, and what information will be inferred from that data. From there, these limitations must be communicated clearly to any potential user, graphics and simple text being the best way to do so. Finally, a user must be informed when they are being traced as well as how to discontinue the tracing.

FAIRNESS

The principle of fairness is meant to ensure that benefits and costs are equally distributed and that both individuals and groups are not subjected to unfair bias, discrimination or stigmatisation.

CONSIDERATIONS

In the case of DCT, while the technology itself is not biased, there may be unintended consequences in its deployment and adoption that may lead to disadvantaging lower income families and older generations, groups which are already the most vulnerable when it comes to COVID-19. DCT requires a minimum of 60% adoption rate in order to be effective. Countries like Singapore, where DCT was launched earlier in the pandemic, have seen only a ~20% adoption rate and have consequently made the app mandatory in certain locations because of this low voluntary adoption rate. The need for widespread adoption, coupled with common privacy concerns among the public, means governments will likely want to incentivise its use by linking it to possible benefits. For example, it is possible to limit certain social benefits to those using DCT or requiring it for access to public spaces. This act of resource delegation in accordance to DCT application usage is essentially mandating the adoption of the technology. One of the key considerations here is the fact that mandating DCT adoption can negatively impact those who are already in vulnerable economic positions by potentially limiting their access to education, work, and other resources, as well as risking stigmatisation.

But it is not only economic standing that must be considered in terms of fairness. Adults over the age of 50 are significantly less likely to have smartphones, even in advanced economies where smartphones are ubiquitous. Similarly, underdeveloped nations have low adoption of smartphones with only 45% of adults in developing countries owning smartphones according to Pew research. Those without a smartphone, or smartphones for which updates are not supported, are essentially invisible to DCT and unable to access resources limited to DCT usage if it were to be implemented.

EI'S PERSPECTIVE

DCT has the potential to broaden inequality where its adoption is linked to incentives such as welfare or access to public spaces, as it can exacerbate pre-existing issues. The tendency for vulnerable groups to be left out of datasets and studies has been widely documented, and DCT-based datasets, which can be used to study the disease, will inevitably suffer from the same drawbacks. If any form of access to resources or personal freedom is to be attached to the usage of DCT applications, fairness of the application is placed at significant risk.

ACCOUNTABILITY

Mechanisms, such as audits, should be utilised to ensure accountability for both an AI system and its outcomes. If a system is auditable, then its algorithms, data and design are all open to critical assessment. Most importantly, the outcomes must be subject to analysis and rectification if significant harm is caused.

CONSIDERATIONS

The first accountability consideration deals with who will be inputting the code into the app that will signal its owner has been infected. There are two potential answers here, either the medical professional will be able to trigger the signal through a code limited to their own acces, or the individual will be given the code that triggers the signal to enter into their application. This is an important distinction as it dictates whether the responsibility lays with the medical professional or individual. Whoever is responsible is then in turn accountable in the case that the code is used but the individual was not infected, or the individual was infected but the code was not used.

Although significant considerations must be made in terms of understanding the algorithms, data and design of DCT applications to ensure responsible innovation, the potential outcomes of such applications on personal accountability is a consequential point often overlooked. There is a danger for an individual utilising a DCT application to no longer feel at risk or accountable for following additional health measures. This is because the usage of a DCT app can create a false sense of security, as the user may assume they have not come into contact with an infected individual due to lack of notification from the application when in reality this is not necessarily true.

EI'S PERSPECTIVE

Who is accountable for any imaginable scenario must be determined prior to the issuing of a DCT application. In other words, it must be clarified if it is the responsibility of the medical professional or the individual to identify via the app an infection, and that a clear record is kept in case an infection code is falsely used or not used when it should have been. Furthermore, if DCT applications are deployed, then the application must clearly and consistently communicate the importance of continuing to follow health and safety measures, as well as emphasize the fact that the application does not make an individual immune to infection.

ROBUSTNESS & SAFETY

In order for an AI system to be robust and safe, it must be ensured that the system is resilient, accurate, and reliable. This is essential to minimizing intentional harm.

CONSIDERATIONS

Much of the discussion about the role DCT should play in public health policy has operated with an idealized vision of the function of DCT. In other words, it is taken as a premise that we can deploy a smartphone app that can effectively automate contact tracing. However, the actual technical and epidemiological correctness of the DCT model is central when we want to consider whether DCT can be used in a way that is fair, and as part of policies that make ethically and legally justifiable claims on individuals. If there will be social and individual costs and benefits to be distributed by way of policies involving DCT, determinations about their justness and fairness depends on having a clear view of the efficacy we can realistically expect from the technology.

If we had reason to believe that DCT apps can reliably replicate the role of manual contact tracing, which has a long history of use and is broadly regarded as ethically permissible, then we might focus ethical attention to problems such as data retention, privacy, and protection from off-label use. We would focus on the unique aspects of the DCT as a delivery mechanism. Call these mechanism-problems, to mark the sense in which we are comfortable with the function of the technology, but want to be sure it is implemented in a way that is ethically justifiable. But in fact DCT is not a replacement for manual contact tracing, and most policy proposals generate a new role for DCT, such as requiring testing. However mere proximity is not a reliable predictor of COVID-19 transmission because BLE is not a reliable indicator of our proximal contact with others, and DCT does not operate with strong enough non-causal signals to generate strong inferences about transmission risk. Consequently, the actions we might attach to its predictions can only be weakly justified.

EI'S PERSPECTIVE

It may be reasonable to use DCT as a method to ration limited testing resources, and require mandatory testing on proximity indication. This is a function that is better aligned with the kind of model DCT can give us. We can then ask mechanism questions about that function - maybe the nature of DCT app usage would cause some groups to be over or under tested. The primary reason efficacy matters for the ethical analysis of DCT is to help us reason about prediction-to-action policies, by allowing us to interrogate the precise nature of the predictions. A secondary reason sits more comfortably alongside some of the implementation concerns that have dominated policy discussions, especially those regarding scope-creep of the app, and misuse. If we can reasonably foresee that DCT cannot play the role that a policy expects it to, for example, by yielding reliable inferences about who has been infected with COVID-19, then we have good reason to worry that its implementation will be changed. For example, a change could be broadening its data collection and relaxing privacy protections, in order to fit it to the policy role. This could justify an analysis that includes technically informed conjecture about the sorts of alterations to the technology that will have to be made to perform the stated policy function.

SOCIETAL WELLBEING

Societal wellbeing is the quality of life or social welfare, measured not only on an economic, but on mental and physical health basis. The health of the institutions that govern a society should also be considered.

CONSIDERATIONS

If the threshold for adoption is a large factor for the technology to have a successful impact, then it will either be quite a few years before it proves effective or the form (smartphone application etc.) will need to be modified to ensure that no one is left behind, should they wish to participate. In addition, the shape this will take will very much depend on the engagement of individuals. Insufficient information about the technology is available to a lay audience and therefore this may induce anxiety, leading to societal rejection of the technology, prior to seeing the potential benefits of wide-scale adoption. If the approach is one where there is an expectation from every member of society to participate, such as through tracking of health service, national insurance or social security numbers, then this will lead to a backlash due to individual privacy concerns. Such an enforcement of technology use will need to be critically evaluated, as a loss of credibility and public confidence would be devastating as well as antithetical to the principle of autonomy. Aside from disease status, the contact tracing app could easily be modified to monitor activity and exercise levels, punishing or rewarding individuals for such actions, as well as possibly by social media platforms to allow their users to know when a threshold number of their close friends had gathered to socialise without sending individual invitations. The range could be from relatively trivial, to serious medical usage and for each iteration the necessary implications of use and misuse to privacy, mental health and societal wellbeing would need to be made.

EI'S PERSPECTIVE

Prior to adoption of DCT a state/nation must carefully consider if adequate safeguards are in place, informing the public of how their data will be used and stored, while emphasising the long-term benefits of the technology on society, not just during a crisis. This will be key in ensuring the confidence of the public, with provable examples in the form of successfully targeted healthcare actions, such as social distancing or disease epicentre/flare mapping, where targeted resource allocation is thought to make a substantive difference to the situation. An educational approach will be most helpful, to make everyone in society, from employees to leaders, a stakeholder and a willing, benefiting participant, without creating a bubble where some individuals would be excluded or disadvantaged. A proper public consultation to determine if most individuals are supportive of implementing DCT is vital to allay fears of an undemocratic pronouncement while preserving societal wellbeing.

FINAL REMARKS

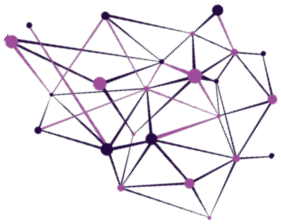
Although an ethics audit should be run as a preliminary step during the development stage of any potentially influential tech application, it is essential to emphasize that it is only one step of many. Underlying ethical issues are not solved in an afternoon of debate, instead they require monitoring through the entire lifecycle of the technology. By beginning with an ethics audit, initial issues can be identified and either solved or marked for monitoring; it is like laying a strong foundation for a house. A technology application can only develop to its full strength if it has begun on a robust foundation.

The purpose of this report was to highlight the fact that controversial applications, such as DCT, must go through a rigorous ethical analysis in order to uncover any hidden risks inherent in the systems. Moreover, this report was designed to illustrate the thought process required to undergo such an ethical analysis, as responsible technology is only made possible by proper consideration.

We would like to reiterate that we only covered a handful of the necessary considerations, as the emphasis for this report was on the thought process. This is solely a case study, whereas an actual ethics audit would be more in depth, including action plans, rigorous testing of the technology and analysis of stakeholder sentiments.

CONTRIBUTING AUTHORS: Andrew Buzzell
Amanda Curry
Michael McAuley
Olivia Gambelin
Oriana Medlicott
Vikas Sharma

CONTRIBUTING EDITOR: Fiona Melzer



ETHICAL INTELLIGENCE

bringing the human back into the equation

Ethical Intelligence (EI) empowers clients to create human-centric technology through ethical innovation. EI is translating the knowledge of AI Ethics from academia into practical solutions for industry. Working with a network of experts ranging from ethicists to AI programmers, EI educates and equips companies with the tools necessary to embed ethics into business and tech processes to mitigate ethical risk and gain the competitive edge in a volatile market.

info@ethicalintelligence.co

Ethical Intelligence Associates

www.ethicalintelligence.co

@ethicalai_co

Ethical Intelligence Associates, Limited
C/O Wright, Johnston & Mackensy
The Capital Building
Edinburgh, EH2 2AF
https://www.ethicalintelligence.co