Towards a Trustworthy Coronavirus Contact Tracing App

Marion Oswald
Commentary, 4 May 2020

The use of a coronavirus contact tracing app has not yet been demonstrated to be trustworthy, in terms of its purpose, reliability, effectiveness or potential harmfulness. Furthermore, the binary nature of its output must be addressed if trustworthiness is to be achieved.

Data-driven responses to coronavirus are being developed at speed, with smartphone contact tracing apps proving one of the most controversial. Questions around efficacy and the wider implications of these apps moved to Parliament on 28 April, at an oral evidence session held by the Commons Science and Technology Committee.

This session, which included evidence from Matthew Gould, Chief Executive of NHSX – the unit responsible for setting national policy and developing best practice for the National Health Service (NHS) technology, including digital and data innovation – was perhaps the first public forum to provide concrete information about the in-development NHS contact tracing app.

While privacy questions were raised, it was the issue of trust that appeared to be the recurring theme behind much of the parliamentary discussion. Will, and should, the public trust, and therefore use, the technology? Gould commented that, in order to achieve significant levels of download, 'the message needs to be, if you want to keep your family and yourselves safe … the app is going to be … an essential part of the strategy for doing that … it will require us to earn and keep the trust of the people'.

So this implies that trust is still to be achieved. In my view, the use of a contact tracing app has not yet been demonstrated to be trustworthy. By this, I mean a whole system including the people within it – not just the technological element – that can be relied upon to do what it is supposed to do, and to

What is the system (of which the contact tracing app is part) supposed to do? This sounds a simple question but it is not. An NHSX blog gives a basic explanation:

'Once you install the app, it will start logging the distance between your phone and other phones nearby that also have the app installed using Bluetooth Low Energy … If you become unwell with symptoms of COVID-19, you can choose to allow the app to inform the NHS which, subject to sophisticated risk analysis, will trigger an anonymous alert to those other app users with whom you came into significant contact over the previous few days. The app will advise you what action to take'.

But this only deals with how the app works in a limited sense. Still to be determined is the overall purpose of the system of which the app is part.

Gould acknowledged that use of an app only makes sense as part of an integrated strategy including manual contact tracing (especially of those not using the app) and increased testing, all working towards a consistent aim. This could include the use of data to analyse virus spread, but again, exactly what information, by whom and for what purposes (statistical or other) has yet to be laid out in detail.

What will individuals receive in return for using the app? This will not necessarily take the form of individual benefit – rather a contribution to a public health outcome. This outcome – and the uncertainties as to whether it can be achieved – need to be made clear to the public to facilitate understanding and debate. This could move the public conversation away from (in my opinion unhelpful) disputes around 'centralised' and 'decentralised' contact tracing approaches, to instead determining exactly what action and data analysis are necessary and proportionate to support a long-term strategy. Ultimately, strategy is the state's responsibility, but as the Nuffield Council for Bioethics argues, to date the government's strategy appears 'massively simplified', and lacks detail 'of what principles or values are informing the decisions about how to proceed with the "exit"'.

CAN THE SYSTEM BE RELIED UPON?

I would suggest that we do not know. NHSX says that 'millions of us are going to need to trust the app and follow the advice it provides.' As yet, there are few convincing reasons why we should. Discussing models on which the NHSX app will apparently be based, Professor Christophe Fraser from the Big Data Institute in Oxford proposed that the app will operate with self-diagnosis, conceding that this would mean 'more people receiving notifications as a result of false warnings' but that the 'effect of suppressing the epidemic more quickly outweighs the risks in waiting for a test before the notification'. However, the 'over-70s have not been factored in' – they have to 'remain at home'. How will these limitations and inaccuracies be mitigated by the overall virus-suppression strategy? An approach that is dependent on everyone over 70 staying at home for an indefinite time would not – from their perspective of the over 70s – 'work' for them.

As explained during the Committee evidence session, the app will turn potentially inaccurate Bluetooth contact measurements into a 'binary decision' as to whether to advise an individual to self-isolate. Notification will happen when measurements hit a certain threshold. What this threshold should be – bearing in mind that the app will lack 'knowledge' of individual context – is an opinion-based judgement. How will it be made?

Furthermore, a binary decision hides the uncertain nature of the app's output, and varying levels of infection risk that the measurements may suggest, depending on individual circumstances. The app will have no idea, for instance, whether a person was wearing a mask. My work with Alexander Babuta on algorithms and data analytics in policing has recommended that:

NHSX expects the public to trust the app, but a binary output appears unlikely to provide enough explanation for this to happen. The human user should be given the information that they need to make risk judgements themselves, including the uncertainties involved in the app's output and the options available to them.

DEMONSTRATING EFFECTIVENESS

The system must be shown to do what it is supposed to do in order for people to trust it. There will be a data protection impact statement, according to the evidence session, but details of the proposed model will not be published before local testing is completed. There will be an ethics advisory board for the app, but its full membership, terms of reference and commitment to transparency are yet to be confirmed.

There are missed opportunities here: first, to have incorporated the legal, ethical and societal input from the start (the app was said to have been in development since March). As Chair of the West Midlands Police and Crime Commissioner and West Midlands Police data ethics committee, I know that an approach which incorporates input from people with a variety of perspectives at early stages of a project is more effective and therefore more trustworthy. Second, to invite challenge from diverse perspectives, other disciplines, the 'rebels', to ensure that the direction is not set by those who vehemently agree. Third, to expand the discussion away from one-dimensional definitions of privacy and individual data protection preferences, to the wider question of how the state will defend individuals from potential detriments that they may not even be aware of.

MINIMISING HARM

How to ensure that the system does nothing it should not? Although important, the answer to this does not just revolve around data security and anonymisation. This element of trustworthiness requires us to predict the bad things that might happen – something lawyers are trained to do! If we can anticipate unintended uses – such as the app being demanded for employment, property occupation or access to services – we must put in place ways of stopping these things happening, by a combination of law (a team of UK academics have suggested safeguards; Australia has a Determination pending formal legislation), regulation, oversight, enforcement and technical design.

As the Chair of the Commons Committee, Greg Clark, concluded last week, 'fundamental questions' remain; answering these will require a willingness to listen and respond to diverse perspectives. Achieving trustworthiness requires a series of system-wide graduated steps, not a binary process.

The views expressed in this Commentary are the author's, and do not represent those of RUSI or any other institution.

BANNER IMAGE: A visualisation of a coronavirus infection. Courtesy of Adobe Stock

AUTHOR

Marion Oswald
Associate Fellow

Marion is the Vice-Chancellor's Senior Fellow in Law at the University of Northumbria, an Associate Fellow of the Royal United Services... read more

Related

Threat

Commentary, 28 July 2020

Tom Keatinge

The illicit finance challenge faced by the UK and highlighted in the recent parliamentary report into Russian operations threatens to be overwhelming. Urgent and radical action is required.

Tags: Centre for Financial Crime and Security Studies, UK

Combat Air Choices for the UK Government

Occasional Papers, 28 July 2020

Justin Bronk

In light of the upcoming Integrated Review, this Occasional Paper outlines the combat air choices facing policymakers.

Tags: Occasional Papers, Military Sciences, UK, UK Defence

The UK's Labour Party: The Long March to Regaining Trust and Electability on Security Policy

Commentary, 27 July 2020

Mike Gapes

The UK's chief opposition party has overhauled its foreign and security policies. But some of the biggest policy choices are yet to be made.

Tags: UK

1 2 3 4 5 6 7 8 9 ... NEXT › LAST »

defence and security community.

Individual

RUSI LIBRARY

The collection is dedicated to developing our knowledge of war and sharing theoretical approaches to modern military thinking... read more

The Library is now closed until further notice due to the Coronavirus

SUBSCRIBE TO OUR NEWSLETTER

Receive updates on RUSI's research initiatives, publications and events, with highlights of commentary and analysis.

email address

Subscribe

SUPPORT RUSI

Noted for its quality, RUSI's analysis is driven by an ethos of accuracy, objectivity and policy relevance.

Donate