

# Pandemic data challenges

The worldwide outbreak of COVID-19 has led to great tragedy and poses unprecedented challenges for countries' healthcare systems. Data has become an important instrument in the global fight against the unprecedented spread of the virus. But how will we ensure a return to previous forms of data privacy once the pandemic subsides?

In a bid to help curb a disastrous course of the disease, scientists are coming together to share research findings and data as quickly and collaboratively as possible. Immediate goals are to understand what measures can slow the spread of COVID-19, which people are the most vulnerable and how to treat critical patients under challenging circumstances. In the medium term, data-driven approaches can help crack the virus's biology to develop and test vaccines.

There is an eagerness from many different research communities to get involved, especially as ongoing research projects are put to a temporary halt. One starting point is the [Crowdfight](#) platform where experts propose research projects and anyone can volunteer to help. Another is the COVID-19 data challenges on [Kaggle](#), a well-known platform for sharing data and computational tools. The latest Kaggle challenge calls for researchers to develop tools for mining a new free online database called [CORD-19](#), which collects all peer-reviewed papers and preprints on COVID-19. The database is hosted by the Allen Institute for AI and is updated on a weekly basis. On 6 April, it contained 36,000 full text articles.

Many [publishers](#), including Springer Nature, are making COVID-19 related papers free to access during the duration of the outbreak and are actively encouraging researchers who submit papers on this topic to make their work immediately available via preprint servers. Clear statements regarding the availability of underlying data are also required.

Available data mainly comes from health services, but there is huge potential in harnessing smartphone location and other data for comprehending the dynamics of the pandemic. Such information could help understand and fight the spread of COVID-19 among populations. In this new era of digital pandemics, existing Western sensibilities regarding data privacy are put to the test<sup>1,2</sup>. Germany [adjusted](#) the wording of GDPR-enabling legislation, to allow for processing personal data in the event of an epidemic. Several mobile phone companies have already handed over data, to be analysed at the German national

disease control centre, the Robert Koch Institute. The data are handled in aggregated and anonymized form and assurances are made that this is not about tracking individuals. However, there are still concerns that this action was taken without sufficient analysis of legal and human-rights implications. A group of German data privacy experts published a [statement](#) on 26 March urging governments to act with caution in wielding legal powers in obtaining data, and calling for attention to principles and guidelines on data protection in the current COVID-19 crisis.

Other governments are also busy obtaining citizens' smartphone data, with varying levels of restraint. In Israel, a high-court-enforced limit on government powers was swiftly overruled and access to citizens' mobile location data was made [available](#) for the purpose of retracing individuals' movements. The UK government is in talks with mobile phone companies about access to data, as is the US government. But full transparency in governments' procedures and motivations are needed, as well as reassurances that civil rights are not put second to all else. A [joint statement](#) from civil society organizations around the world, published on 2 April, warns governments that their response to the public health emergency posed by COVID-19 should not result in a new era of invasive digital surveillance. The statement spells out eight conditions that should be met to ensure that digital technologies implemented in the fight against the virus are strictly in line with human rights, among others: expanded monitoring and surveillance powers must be time-bound, and any data-sharing agreements between governments and other entities must be based on law and publicly disclosed.

With present-day technology for collecting and crunching data, there is great promise in using innovative, data-driven tools to fight the virus. But institutionalizing large-scale citizen data collection systems is a dangerous path that could lead to intrusive practices from which it will be difficult to retrace our steps. Michael Ryan, a key advisor for the [World Health Organization](#), [observed](#) during a briefing on 26 March that there is enthusiasm for using digital technologies, but warned that there are

serious data-protection and human-rights principles involved in collecting information on citizens or tracking their movements, and that we must never step beyond the principles of individual freedoms and rights.

An alternative to governments using their powers to get hold of data collected by tech companies is to encourage citizens to voluntarily install applications for non-profit research purposes and donate data from smartphones, fitness trackers and other mobile devices. Citizen science initiatives such as the [COVID Symptom Tracker](#) app, which was launched on 24 March and developed by scientists at Kings College London in collaboration with health company Zoe Global, seem an empowering way for individuals to contribute with their smartphone data. In the United States, in the meantime, Facebook has teamed up with researchers from Carnegie Mellon University Delphi Research Center to develop a [survey](#) that asks Facebook users to voluntarily share information about their health. The company promises that they won't have access to the results of the survey, which will also hide individuals' identities. Another group of scientists and legal experts is developing a decentralized Bluetooth-based [contact tracing system](#) that avoids storing and processing private information in a central server, minimizing the risk of misuse or repurposing of data on individuals. These approaches, however, would require substantial uptake across a population in order to prove beneficial, as well as careful oversight.

The COVID-19 pandemic is a daunting challenge for the world and requires an extraordinary and coordinated response. The crisis calls for us to demonstrate our shared humanity and we must make sure that our efforts to fight the virus are not inadvertently eroding people's rights. □

Published online: 8 April 2020  
<https://doi.org/10.1038/s42256-020-0172-7>

## References

1. Ienca, M. & Vayena, E. *Nat Med.* <https://doi.org/10.1038/s41591-020-0832-5> (2020).
2. Mittelstadt, B., Benzler, J., Engelmann, L., Prainsack, B. & Vayena, E. *Life Sci. Sci. Pol.* **14**, 9 (2018).