April 2021

A Retrospective:
Protecting Privacy in our COVID-19 Initiatives

TABLE OF CONTENTS

Introduction

It has now been more than a year since the World Health Organization declared the COVID-19 outbreak a global pandemic. Much has been written about the role consumer technology has played in helping people adapt to the social changes COVID has brought about, including the long periods of isolation many have faced as governments sought to curtail the spread of the virus. The countless stories of how, for example, video calling has enabled people to keep in touch with friends and family – and even to celebrate some of life's most important occasions – have reaffirmed the potential for technology to create social value for people around the world.

Less, however, has been written about the other ways that technology companies have informed the public health response to COVID – including through data, research, and other efforts to combat the virus -- and the policy judgments that inform decisions about how to provide that support responsibly. As we reflect on over a year of life under the COVID pandemic, this paper shares the decisions we've made at Facebook about how to assist in the public health response to COVID and what we've learned in providing this assistance. We focus particularly on how we have sought to address the urgent needs of public health authorities and researchers while maintaining our strong commitment to privacy and other fundamental rights.

Section 1 of this paper explains our thinking about how Facebook could most effectively contribute to the COVID response. We discuss how we determined which areas of the public health response to focus on, including our conclusion that our participation should be guided by public health officials, researchers, and other experts. We also decided to prioritize areas where the nature of our services made our contributions particularly impactful – for example, that we'd play to the strengths of our platforms as tools for distributing authoritative information.

After identifying areas of focus, we assessed whether to support particular projects or initiatives. Protecting privacy and human rights is at the core of all our work at Facebook. But as many policymakers and regulators have made clear in providing guidance about the COVID response, privacy is just one of the fundamental rights implicated by

COVID-related challenges. Section 2 discusses how we've sought to protect the rights to privacy and data protection while also honoring other fundamental rights through our assessment processes. It concludes with a proposed framework to carry out these assessments, taking into account potential benefits while also assessing risks, potential harms and safeguards.

To further illustrate the assessment framework we've set out, Section 3 provides a series of case studies. The first is an example of applying the proposed assessment framework to technology to automate contact tracing or exposure notification. The second case study focuses on sensitive data and the tensions evident in identifying the appropriate set of data protection safeguards, such as consent, where the sensitive data has the potential for significant beneficial uses at scale in a public health emergency. The final case study is an in-depth discussion of one key safeguard: privacy-enhancing technologies (PETs) in the context of sharing aggregated data and insights. The application of each of these case studies raises challenging questions, particularly in the public health context.

Although this paper is largely retrospective, we view it as a conversation-starter. We hope it will lead to a broader series of discussions among stakeholders around the globe to understand decision-making in the current emergency and to inform future decision-making should similar challenges arise again. To that end, we conclude each case study with questions for discussion. With greater clarity and consensus around standards governing the questions raised, we see the opportunity to better enable responsible uses of data, including for public health emergencies.

Section 1

Which areas could Facebook assist with the pandemic response?

In early 2020, as COVID-19 spread around the globe, questions quickly emerged from inside and outside the company about how Facebook could assist with the public health response. We received various collaboration requests from public health authorities and the research community.[1] Facebook employees with a variety of experience and skill sets asked how they could assist. And the company was eager to put our data, technology, and talent to work in helping the recovery effort. At this stage, little was known about the virus's eventual impact and many companies were rushing to help. In thinking about the role that Facebook should play, we wanted to maximize the positive impact we could have and — despite the emergency situation — ensure we were upholding our responsibilities around privacy and fairness.

The first question we had to answer was, "where should we focus?" We knew that our considerable resources made it possible to intervene in many different ways. We watched as some companies leveraged their existing products, services, and capabilities to meet the needs of a radically changed world.[2] And we looked on with interest as other companies pivoted to new lines of business entirely – for example, by manufacturing masks and other PPE that were in short supply during the early stages of the pandemic.[3] To decide our path forward, we followed two principles:

1) We would take our cues from the experts leading the pandemic response. While our employees had many ideas about how we could contribute, we recognized at the outset that our decisions should be guided by the things experts in the public health community identified to be in the best interest of individuals and communities as part of the public health response.
2) Rather than explore new product lines, we decided that we could most effectively and most quickly contribute to helping people by playing to the strengths already inherent in our products and among our employees.

Below we've provided examples of how these principles played out in real-time decisions of some areas where we focused our efforts and what we did not pursue.[4]

I. Using social networking technology to help experts connect with communities

Early in the pandemic public health and humanitarian experts identified possible interventions where we could play a valuable role because of our ability to reach so many people on our services. By March of 2021, we had connected over 2 billion people from 189 countries to Covid-19 information from authoritative sources.[5] Not only did we offer services that made it easy for public health officials to easily communicate to their communities, but we also had experience running interventions such as promotions or surveys at a large scale. For instance, before the pandemic, we had surveyed millions of online small businesses biannually on their business challenges and expectations to provide policymakers, research institutions, and nonprofits with insights on the future of business.[6]

We understood from epidemiologists during the early days of the pandemic that surveys, in particular, could assist their efforts in three critical areas: (1) forecasting the spread of the virus; (2) assessing people's understanding of preventive behaviors, such as social distancing, mask wearing, and hand washing; and (3) understanding attitudes and behaviors related to vaccination as vaccines start to become available.[7] We partnered with some of these experts at Carnegie Mellon University's Delphi Research Group, University of Maryland, and the Initiative on the Digital Economy at MIT to create a method of survey collaboration and data sharing that leveraged our strength in enabling communication and their strength in public health research. We promote the surveys to people around the world using Facebook and Instagram. In addition to enabling broad distribution of surveys, we were able to, in aggregate, use information people share with us to mitigate potential bias in sampling. The surveys themselves are run by the research institutions that have expertise in which public health questions can most effectively be answered and conducting analyses of survey responses.[8] This collaboration enabled a setup where we could help with distribution, while only the researchers hosted the survey and collected individual survey responses and they did not share individual responses

with us.[9] The researchers do, however, make the aggregate results of the surveys publicly available.[10] We host public visualizations of this publicly available information to help communicate to a wide audience ranging from individuals curious about disease spread to public health authorities, who may use the information to help prepare responses and allocate resources, as well as to inform messaging tactics and policy decisions at a regional level related to their vaccine rollout efforts.

II. Share aggregated data and insights

Many people share data with Facebook that can, in the aggregate, provide crucial insights for humanitarian work and research. Since 2017, our Data for Good program has partnered with over 450 organizations in nearly 70 countries, and collaboratively built privacy-preserving products, such as aggregated datasets in the form of maps, that provide real-time insights for addressing public health emergencies, spurring economic opportunity, and fighting climate change.[11] In response to COVID-19, we turned to Data for Good tools to help researchers get the information they needed to understand and respond to the pandemic and plan for their recovery. We partnered with dozens of trusted organizations to use Data for Good's Disease Prevention Maps and Movement Range Maps to aid relief efforts. These partners include universities like Harvard School of Public Health in the US, National Tsing Hua University in Taiwan, and University of Pavia in Italy, as well as nonprofits and institutions such as Direct Relief, the Bill & Melinda Gates Foundation, and the World Bank.[12] Our partners established the COVID-19 mobility data network, a global coalition to provide real-time insights from our maps.[13] The maps, which we'll discuss in more detail in Section 3 below, have informed crucial decisions about where to allocate resources, how to educate the public, and understanding the efficacy of non-pharmaceutical interventions and mitigations across Latam, Asia, Europe, and North America.[14] Researchers in Taiwan used the datasets to identify the cities with the highest chance of infection; researchers in Italy analyzed lockdown measures and their impact on income inequality; and public health officials in California and New York reviewed county-level data daily to steer public health messaging.

III. Combine expertise in machine learning and artificial intelligence with large, comprehensive datasets

Machine learning (ML) and artificial intelligence (AI) are important tools to support public health experts around the world in their efforts to keep people safe and informed amid the COVID-19 pandemic. Facebook also has significant expertise in ML and AI. We partnered with academic researchers and other experts globally on a range of initiatives related to COVID-19 that leveraged this expertise. For instance, we worked with academic experts in New York, New Jersey, and Austria to provide localized COVID-19 forecasting models to them, and they in turn used those models to develop and share forecasts with public health authorities and emergency services providers.[15] The information produced by our AI models improved resource planning for in-demand resources, such as hospitals, ICU beds, ventilators, and masks.[16] We then published AI-powered forecasts publicly to help predict the spread of COVID-19 across the entire United States at the county level. These forecasts leverage large, comprehensive datasets, including those discussed above – aggregated data from the Symptom Survey and Movement Range Maps – as well as non-Facebook public data.[17] A critical factor we developed for these forecasts is a new neural autoregressive model that aims to disentangle regional from disease-inherent aspects within these datasets.[18] Our model has the ability to account for relationships among different counties, so, for example, an uptick in one area can have an impact on predictions for adjacent or similar districts.[19]

IV. Refraining from sharing people's precise location

Finally, there were areas we discussed with experts that we decided not to pursue because they did not align well with the strengths inherent in our existing products and therefore would be of limited value to help people. For example, we received requests for user location data from governments seeking to use it to rapidly scale quarantine and physical distinancing measures to better limit COVID-19 spread amongst their population. Even setting aside the significant privacy and data protection issues involved, which we discussed at length with global civil society experts in privacy and data

protection and cover in the first case study below, we simply did not have the kinds of data that would have been useful to achieve the desired goals.

Specifically, we only collect precise location information from users' devices in limited circumstances -- and only with a person's consent to collect the information, which they can turn off or manage via their Location Settings. People consent to provide their precise location information to enable products like Nearby Friends and to see more locally relevant content and ads on our services. Only a portion of the Facebook community consents to such information collection. And of those who have consented to the collection of their precise location information from their device, only some allow us to store that information over time. That is, even though we might initially receive a person's precise location from their device, we only retain that precise location for an extended period if the person also specifically opted-in to that longer term retention. Moreover, when a user shares their precise location with us via the Location Services setting on their device, our app can access location-related data from the device's sensors, including GPS, Wifi and Bluetooth data. However, the precision of the information that we receive can vary widely depending on the setting and the strength of signals available to the user's device. For instance, if a person is in a building that blocks a GPS signal, the operating system will have and will transmit less precise GPS information (and may or may not have other signals such as Wifi available).

Because of the factors above, our location information was of limited benefit for people and communities addressing rapid COVID-19 spread. For instance, even if we could understand two people were inside the same mall, we may not know whether they were in close proximity to each other or even in the same store. For all these reasons, as well as the privacy and data protection considerations discussed in the first case study in Section 3, we decided not to fulfill these types of requests; instead focusing on our existing strengths, such as providing the aggregated data and insights discussed above from our Data for Good program.

Section 2

How could we assist public health efforts while respecting people's privacy and fundamental rights?

The frameworks and laws proposed to address privacy have far-reaching implications with respect to the use of data and technology to support public health aims, and policy choices may prioritize certain fundamental rights over others. Privacy and data protection is often framed as an all-or-nothing proposition, especially in an emergency, with headlines essentially asking individuals or decision-makers to choose between individuals' privacy or their health.[20] This narrative misses a critical conversation about how to best secure multiple aims or fundamental rights concurrently, and can undermine trust in well-designed data protection and information governance to streamline and enhance emergency response systems.

Robust accountability processes and frameworks for balancing and protecting privacy and other fundamental rights were critical for us to build products, form partnerships, and otherwise participate in the public health response to COVID-19. This section details our assessment processes and proposes a framework to carry out these assessments, taking into account potential benefits while also assessing risks, potential harms and safeguards.

I. Assessment processes

At Facebook, the product-development process includes accountability mechanisms designed to incorporate guidance from experts inside and outside the company. With respect to our pandemic-related efforts, we incorporated feedback from experts in the fields of privacy, content policy, health policy, human rights, responsible design, and security.

The Privacy Review process at Facebook is one of the best representations of the shifts we've been making to honor people's privacy in everything we do. We've had a version of Privacy Review at Facebook for many years, but over the past year we've made it even stronger by using a more holistic approach with more extensive technical validation, and more in-depth internal and external consultations. This process focuses not only on making the best decision we can, but also structuring our company processes to consider privacy impacts at the early stages of product development, to document mitigations to any such impacts, and to ensure that mitigations are consistently implemented when a product is built. Facebook's privacy process also focuses on education -- helping to increase capacity to anticipate and address privacy concerns across the company, leading to better future decisions.

In this process, gathering and incorporating internal and external feedback starts at the product ideation phase and continues even after rolling out a new or modified product or service. The process is grounded in principles embodied in legal regimes,[21] data governance frameworks,[22] and context-specific data protection guidance.[23] The process often goes beyond the review of strict data protection and other legal requirements; it also explores policy, ethical and societal implications of our products and services.

II. Assessment framework

To ensure we were appropriately considering people's fundamental rights around data protection while also addressing public interest priorities -- in this case, those associated with the pandemic response -- in our Privacy Review process for COVID-19 related initiatives, we considered and balanced a variety of factors including the benefit of the data or technology, the types and likelihoods of harms that can result from the data or technology, and the feasibility to use safeguards to minimize risk. The following proposed framework draws together many of the factors we considered across a number of our assessments related to addressing the public health emergency. It stands on the shoulders of a wide range of accountability and assessment guidance,[24] and we found the UN Global Pulse's two-phase Risk, Harms and Benefits Assessment Tool to be particularly instructive as we worked through issues related to COVID-19.[25]

Benefits. To ensure that our interventions were effective, we started by identifying the positive impact or problem(s) solved by the proposed product or service. We also

assessed the likelihood that the product or service would be effective in solving the stated problem or achieving the stated positive impact.

Risks. Our Privacy Review process involved analysis of privacy risks that could arise in a number of areas, such as sharing with third parties who then misuse the data for purposes incompatible with the original data collection purpose(s). Additional security, organizational, legal, or reputation risks might also be identified. In addition to the nature of the risk, risks may differ in potential likelihood and severity. Identifying these risks holistically was important to ensuring we could clearly assess whether these risks could be effectively mitigated.

Potential Harms. Harms to individuals or societal harms, encompassing human rights violations, had the potential to arise as a result of the risks identified, such as loss of liberty, loss of autonomy, harassment, physical harm, psychological harm, reputational harm, financial loss, bias or discrimination. As with beneficial value and risks, harms could differ in potential likelihood and severity, and some harms may disproportionately affect certain groups of people.

Safeguards. Safeguards may reduce or eliminate the potential likelihood or severity of risks and/or resulting harms. Safeguards could include things like privacy-enhancing technologies (PETs), transparency, consent, legal and contractual restrictions, or further restricting data collection, use or sharing. The interplay between risks, potential harms, and safeguards is important to understand. For example, if a safeguard reduced the likelihood of a risk but the severity of the resulting harm(s) remained high, the safeguard may have less mitigating impact than if it reduced the severity of the resulting harm(s) but the likelihood of the risk remained high.

Necessity, Proportionality, Lawfulness, and Legitimacy. Drawing from human rights and data protection principles, where there is a legitimate beneficial value and not all risks or harms can be fully mitigated, the product or service should be necessary and proportionate to achieving the beneficial value, as well as lawful and legitimate. This may take into account whether the product or service is the least restrictive means in achieving the beneficial value, i.e. an alternative, less risky or harmful method to achieve the beneficial value is not feasible, as well as counterfactual risks of not acting.

Section 3

Case Studies

While not all factors of the assessment framework proposed above are always relevant or present, they can be generalized across many use cases. To further illustrate the assessment framework proposed above, this section provides a series of case studies. First, we discuss the assessment framework as applied to our decisions about facilitating promotion of automated contact tracing technologies on our platforms. The second case study focuses on sensitive data and the tensions evident in identifying the appropriate set of data protection safeguards, such as consent, where the sensitive data has the potential for significant beneficial uses at scale in a public health emergency. The final case study is an in-depth discussion of one key safeguard: privacy-enhancing technologies (PETs) in the context of sharing aggregated data and insights. The application of each of these case studies raises challenging questions, particularly in the public health context. We explore these challenges and our decision-making within each case study, and we then offer some further questions for discussion to conclude each case study.

I. Applying the overall framework: Technology to automate contact tracing or exposure notification

Historically, contact tracing has been a critical strategy for mitigating the impact of a variety of diseases by slowing the spread of infection.[26] Traditional contact tracing involves public health authorities soliciting medically significant contacts from a person recently diagnosed with a disease. The public health authority then directly notifies people (such as via a phone call) who may have been exposed to the person who tested positive.[27] In the case of COVID-19, public health authorities help the person who was exposed to the virus to quarantine, and, potentially, to get tested, thereby reducing the risk of the exposed person unknowingly spreading COVID-19.[28] In order for this process to work effectively, people must be able to recall all of the contacts who they were with during a specific time period for what the public health authority deems a medically

significant encounter. Today, people around the world carry with them mobile devices with capabilities that can enable understanding of who has been in close proximity to particular individuals or devices.[29] Precise data about where individuals are located and congregated throughout the day therefore has the potential to be useful for notifying people who may have been exposed.[30]

As COVID-19 started to spread to countries around the globe, we received requests that we play a role in providing our data or using our technological expertise to help public health authorities to automate their contact tracing efforts.[31] Many public health experts pointed to the opportunity for mobile phone GPS or Bluetooth data as a way to rapidly scale to meet the growing spread of COVID-19.[32] Those making such requests thought automating contact tracing could benefit public health efforts because it could work without relying on a large, trained workforce within a public health authority to trace infected individuals through traditional methods like telephone calls, and that automated contact tracing would be more effective because it would not rely on infected individuals knowing or remembering who they have come into contact with. This was also true with respect to exposure notification technology developed a bit later in the spring, which traces proximity among infected individuals and notifies the individuals, but does not allow a centralized tracing of possible infection chains by a public health authority.[33]

However, some public health experts and other experts in the technology itself were more cautious about the benefits, noting that the efficacy of such technology was still to be proven and required significant adoption amongst the population of a given region to benefit people within the community.[34] For these reasons, and because mobile OS makers have both more reach and more extensive access to precise location data than we do, we concluded that mobile OS operators were better positioned than Facebook to build technology to automate contact tracing that had a chance of being adopted at the precision and reach necessary to achieve the potential benefits of the technology for people and communities in the midst of addressing rapid disease spread.

Even though we decided not to build such technology, we still had requests to play a role by using our social networking services to promote the work of external public health experts related to contact tracing. For instance, public health authorities requested to use the advertising space we donated on our platforms to promote their automated contact tracing or exposure notification apps.[35] In deciding our response to these

requests, we considered the benefit to people by taking into account whether the apps had been deemed effective by a public health authority, like the WHO or a public health ministry, in helping to contain the spread of COVID-19, and whether they integrated coherently with a country's or state's public health strategy and systems.

We also considered the privacy risks and harms to people that may arise by consulting a number of global privacy and civil liberties experts as well as drawing on global guidance and frameworks, many of which were specific to the context of contact tracing.[36] As an example, privacy experts were concerned about the risk of misuse of data by the entity or entities responsible for the data collected within in an app, particularly in jurisdictions that lacked data protection law addressing the specific context. If several branches of a government had access to location data collected via an app, for instance, that could potentially result in its use for enforcement purposes beyond the scope of addressing COVID-19.[37] Experts highlighted to us the potential for this to result in harms to people, such as discrimination or wider surveillance.[38] They also noted the potential for normalization of surveillance in a region in circumstances beyond the present public health emergency.[39]

We assessed various technical, legal, and organizational safeguards that could mitigate the risks and potential harms above, as well as whether apps met international human rights standards (legitimacy, necessity, lawfulness, and proportionality).[40] For example, we considered whether apps were run by public health authorities with clear intent that any data collected should not fall within the control of law enforcement or security/intelligence authorities as evidenced by technical design or other policy or legal safeguards.[41] As an example of technical design, exposure notification apps allow for data to be collected and stored anonymously and on device, thus preventing an individual's movement to be tracked by others and mitigating concerns about wider government re-use or surveillance. [42] With respect to human rights standards, exposure notification apps therefore seemed necessary and proportionate to achieving the beneficial value. We also chose to facilitate promotion of voluntary apps only.[43]

After conducting case-by-case assessments and receiving evidence of effectiveness, we allowed some public health authorities to use donated advertising space on Facebook to promote their exposure notification apps. A primary factor in our decision-making was that exposure notification apps allow for privacy preserving decentralized processing that

mitigates the significant risks and the potential for harms discussed above.[44] By contrast, we did not facilitate promotion of contact tracing apps allowing a centralized tracing of possible infection chains by a public health authority, which is accompanied by more significant likelihood of risks such as government data re-use or harms arising from wider government surveillance. We recognize that some may disagree with this outcome, either from the perspective that exposure notification is still too risky or potentially harmful as compared to effectiveness, or from the opposite perspective that other automated contract tracing apps have benefits that justify their centralized processing.[45]

While we know there will continue to be disagreement about whether we made the right choices in the apex of the pandemic, we hope this transparency about our assessment framework and decision-making helps to foster continued discussion with the aim of building wider consensus around how we consider and weight these difficult decisions with respect to balancing public health objectives and privacy -- and could help to build greater consensus about the appropriate role of companies like ours in future global emergencies. Given the complexity of the dependencies at play and competing values and societal norms, we believe that this kind of retrospective analysis could be helpful in building alignment around accountability frameworks that could apply to similar choices in the future.

Questions for discussion

1. Does our framework consider the right factors? As applied in this case study, should we have considered different risks and harms or weighed them differently?
2. How do we assess potential benefits in an emergency context where the effectiveness of uses in the public interest may still be uncertain as understanding is developing?
3. The likelihood and severity of the risks and harms was highly speculative, but we had to make a decision with limited information in a relatively short period of time. How do we best gain consensus across domains of expertise such as public health and privacy? Or balance competing views of relative risks and value?

II. High benefits, strong safeguards: Sensitive data, consent and secondary uses for the public health emergency

The traditional approach to privacy protection is notice and consent: tell people what you're planning to do with their data and then ask them if they agree. Consent requirements, particularly explicit consent requirements for sensitive data, tend to focus on gaining permission for specific or narrow data uses. Because sensitive data can give rise to a number of significant privacy risks and potential harms to people, data protection frameworks tend to prioritize individuals' choices in the form of explicit consent. There can also be a number of beneficial purposes for processing sensitive data. Many open questions remain when it comes to processing sensitive data for beneficial purposes such as public health and scientific research. In responding to unforeseen situations like the COVID-19 pandemic, experts have observed that explicit consent, particularly on its own, may not be a feasible or appropriate safeguard of privacy in some circumstances, with broader data governance and ethical considerations also at play.[46] We find there are still important open questions actively being debated – even in jurisdictions with comprehensive data protection law such as the EU – about exactly where to draw the lines for what types of beneficial data uses require consent and what types of secondary uses are appropriate or compatible thus eliminating the need for a separate or additional consent.[47]

The GDPR, for instance, recognizes that the right to data protection must be considered in relationship to its function in society, and assessed alongside other fundamental rights.[48] Such balancing can be seen in the processing of special categories of personal data (similar to what we refer to herein as 'sensitive data,' but not encompassing location information), which allows for processing without explicit consent where necessary for reasons of public interest in the area of public health, such as protecting against serious cross-border threats to health, provided there is a basis for such processing in EU or Member State law.[49] Similar processing of sensitive data necessary for scientific or historical research is allowed provided there is a basis in EU or Member State law which is proportionate to the aim pursued, respects the essence of the right to data protection, and provides for suitable and specific measures to safeguard the fundamental rights and the interests of the data subject.[50] These exceptions to the general prohibition under the GDPR against processing special categories of personal data evidence a balancing between data protection and the potential benefits of this data. However, because such

provisions rely on a basis for processing in EU or Member State law, they are implemented inconsistently across jurisdictions, giving rise to challenges in achieving the full potential benefits in areas such as public health and scientific research.[51]

Our COVID-19 survey efforts provide a practical example of this tension with respect to the potential for beneficial value of sensitive information and its protection in data processing. In Section 1 above, we described how we used our social networking technology to help experts connect with communities, in particular promoting to people on Facebook and Instagram symptom and preventive behavior surveys conducted by academic researchers to help understand and forecast the spread of COVID-19, as well as to gain insight on COVID-19 vaccine attitudes.[52] These surveys may collect sensitive data about a person's health, such as potential COVID-19 symptoms they are experiencing or if they've been immunized, or a variety of demographic information such as age, gender, socioeconomic status, political views, race or ethnicity. The beneficial uses of the sensitive data collected in these include: (1) informing correlation or causation factors among relevant groups and their beliefs or behaviors related to the virus, such as mask wearing or vaccine hesitancy; (2) understanding differences in impact of the disease and predicting future impact for relevant groups; and (3) assessing bias, discrimination, and fairness concerns with respect to relevant groups.[53] All of the above, in turn, may improve and help target public health interventions related to the public health emergency.

While health or demographic data may be accessible to healthcare providers, others such as researchers, healthcare authorities, or private sector actors often lack such data, making it difficult for them to spot and serve populations that may be experiencing disproportionate harms.[54] As the spread of COVID-19 progressed, for instance, healthcare providers began to observe that certain races and ethnicities experienced disproportionate impacts from the virus.[55] Once these disparities became clear, it was important for the academic researchers conducting the symptom and preventive behavior surveys to collect this data from survey respondents to inform responses to these trends.[56] Similarly, as vaccine development and distribution progressed, it was important to collect information about vaccine attitudes and behaviors to inform messaging tactics and policy decisions at a regional level related to vaccine rollout efforts.[57]

Where processing sensitive data, the question then becomes appropriate safeguard for the beneficial purposes set out. Many open questions remain, particularly where seeking to protect against cross-border threats to health such as the pandemic, whether explicit consent is an appropriate safeguard of privacy and whether in some circumstances, but certainly not all, it may limit the scale of collection or sharing of information.[58] As one example, in promoting to people on Facebook and Instagram the preventive behavior and symptom surveys conducted by academic researchers, we provided transparency that the survey would not be on Facebook and we would not receive individual survey responses. If they proceeded, the researchers also separately sought their explicit consent to take the survey. By September 2020, more than 30 million people around the world had taken these surveys,[59] with relevant findings continuing to be published as of the date of this paper.[60]

By contrast, the insights and learnings from our Data for Good maps used in COVID-19 response would not have been possible if we had to seek a separate, new consent to use the location data with each new map and crisis. Implementing such a safeguard would have limited the maps to being forward looking only, which is very limiting during a disaster when you need data as quickly as possible -- and where one of the most important comparisons is between historic population behavior and behavior during the emergency. It would also increase friction resulting in low and unequal distribution of participants, which may reduce the value of the maps, especially as an input in forecasting the spread of COVID-19 or informing regional responses. A lot of the value was in creating forecasts at smaller geographic levels such as county, where the patterns in the location data are complex and rapidly evolving. Such forecasting relies on large, comprehensive data sets for statistically significant results within a given region and time period to help ensure and enhance accuracy of predictions. We therefore use other safeguards of privacy, as discussed in the last case study below.

Recent guidance from EU bodies points to many of the key outstanding questions in this context. For instance, the European Data Protection Board (EDPB) published Guidelines on the processing of data concerning health for the purpose of scientific research in the context of the COVID-19 outbreak repeated the language of the GDPR that categorizes scientific research, by default, as a compatible further use not requiring separate consent to process; however, the EDPB deferred to forthcoming guidance with respect to secondary uses of data concerning health due to the “horizontal and complex nature”

of the topic.[61] The EDPB recently released guidance on the application of the GDPR to health research. It highlighted many of the complications that researchers face under GDPR when conducting studies that span multiple Member States, such as a lack of clarity on how researchers can obtain "broad consent" to data processing and on what "additional safeguards" allow for processing data for scientific research purposes and for using data previously collected.[62] Additional guidance on these questions is expected.[63] Further, the European Medicines Agency (EMA) recently published a discussion paper noting the potential benefits from secondary uses in enabling larger data sets that are key to many public health interventions and strategies.[64] However, there is no consensus on the set of standards to enable this secondary use and, in particular, how to approach consent in such a context.[65] The EMA is also preparing question and answer guidance on the application of EU data protection rules to the secondary use of health data in medicines development, evaluation and supervision.[66]

Questions for discussion

1. What is the appropriate instrument (e.g., industry codes, regulatory guidance, etc.) for ensuring consistency in approach across different stakeholders operating in different jurisdictions about processing sensitive data for promoting public health and scientific research?
2. What regulatory structures and/or substantive data protection provisions sufficiently enable sensitive data to be obtained and used for beneficial purposes, and in what circumstances?
3. How might we foster trust in collective or societal objectives that do not require consent? What is a robust accompanying set of safeguards (e.g. transparency, independent ethical review, privacy-enhancing technologies,, etc.)
4. Is the practicability of seeking consent relevant? Why or why not? If so, what factors should prove impracticability?
5. What are the high-risk uses of data during the COVID-19 pandemic giving rise to privacy harms to individuals or society that should always require consent (i.e. discrimination, prediction of an individual's health condition, government repurposing for non-disease surveillance)?

III. Privacy-enhancing technologies as a safeguard: Sharing aggregated data and insights

This case study will zoom in on one of the most critical and commonly debated safeguards with respect to ours and others' work related to COVID-19 – privacy-enhancing technologies (PETs).[67] We found the EU's data protection guidance on COVID-19 particularly instructive with respect to this safeguard, although, as we will discuss, even within it many important and challenging questions remain unresolved.[68] To illustrate these questions, we'll focus on the sharing of aggregate datasets through our Data for Good program.

As discussed in Section 1, we offer datasets in the form of maps on population and movement that researchers and non-profits use to understand the coronavirus crisis, informing disease forecasting efforts and protective measures during the pandemic. Disease Prevention Maps and Movement Range Maps are aggregated datasets that, alone or when combined with epidemiological information from health systems, help public health organizations close gaps in understanding where people live, how people are moving, and the state of their cellular connectivity, in order to improve the effectiveness of health campaigns and epidemic response. For instance, Taiwanese researchers were among the first to implement a new dataset that was produced with data from Disease Prevention Maps on the colocation probability of people from different regions to understand the rates at which people may be coming into contact.[69]

Our work on Disease Prevention Maps and Movement Range Maps demonstrate the complexity of questions surrounding PETs. In our experience, for some datasets, it may be practical and feasible to create a differentially private or highly aggregated public datasets that achieve their intended purposes for pandemic response. In other instances, even though the data remains aggregated, the objectives of health researchers or humanitarian organizations may require another layer of detail or the dataset may be smaller. There may still be significant beneficial value for individuals and society in sharing aggregated, de-identified or pseudonymous data. PETs can be accompanied by legal and organization safeguards that further help mitigate privacy risk. The discussion below will illustrate how we have applied this approach to Disease Prevention Maps and Movement Range Maps. We do find there is a limit to the data we can share or disclose without

risking re-identification of individual information, and there is a need for greater research and development around PETs that can enable more data sharing.

We release Disease Prevention Maps and Movement Range Maps in a format to help prevent re-identification, while preserving insights that are useful in responding to crises.[70] These datasets are aggregated in a way that protects the privacy of individuals by using techniques like spatial smoothing to create weighted averages and avoid using map tiles where very few people live.[71] We further apply a differential privacy framework for publicly available Movement Range Maps, which concern movement of people over time.[72] Differential privacy minimizes risk of re-identification of individual data with the help of possible additional information — even information we cannot anticipate now.[73] Applying a differential privacy framework takes into account the sensitivity of the data set and adds noise proportionally to ensure with high probability that no one can re-identify users.[74]

We share more detailed, protected datasets, such as Disease Prevention Maps, only with our network of trusted partners that are accompanied by additional legal and organizational safeguards, such as data use agreements to stipulate clear guidelines that ensure responsible data practices.[75] One example is a type of Disease Prevention Map called co-location maps, which reveal the probability that people in one area will come in proximity with people in another, helping illuminate where COVID-19 cases may appear next.[76] The research partners enrolled in the Data for Good program still only have access to aggregate information from Facebook.[77] That is, the sharing of these protected datasets still rely on PETs via de-identification and pseudonymization techniques.[78] For example, we will aggregate data points over a given period of time and a given geographic region.

Entities sharing data should always try to use the strongest PETs available that still meet the intended uses and purposes of the dataset. Yet ambiguity tends to arise in how to implement this approach in practice, as well as where and what further legal and organizational safeguards can be coupled with PETs for a robust approach. For instance, in its Guidelines 04/2020 on the use of location data and contact tracing tools in the context of the COVID-19 outbreak, the EDPB emphasized “when it comes to using location data, preference should always be given to the processing of anonymised data rather than personal data.”[79] The EDPB acknowledges that rendering location data

anonymous is highly complex, but suggests that options for effective anonymization of mobile phone datasets do exist.[80] In another example, a recent op-ed in Science, signed by doctors, epidemiologists, disease modeling experts, and data privacy scholars, underscored the need for PETs coupled with legal and organization safeguards for effective COVID 19 response and the significant beneficial value of sharing such datasets.[81]

Questions for discussion

1. What are reactions to the EDPB's framework for location data? How might we apply to other types of data more broadly?
2. What is the appropriate framework for assessing feasibility and effectiveness of "reasonable" or "robust" anonymization vs other types of de-identification and pseudonymization for a given purpose?
3. What additional legal and organizational safeguards have proved efficient and effective in facilitating the sharing of mobility data to respond to the crisis?
4. How could we create incentives for more research into PETs, to address the desire for more granular data releases that are currently not feasible without compromising data protection?

What's Next?

The unprecedented COVID-19 pandemic has brought consumer technology companies such as Facebook together with public health authorities, researchers, and many more to support communities in responding to the emergency. The GDPR and other laws already contemplated balancing the need for data protection with other fundamental rights, and the pandemic has demonstrated both the importance of this approach and that crucial and practical questions remain. This paper and the conversations that will follow it are intended to lay out key issues and start to address hard questions about how consumer technology companies can assist these vital efforts with data and technology that can be implemented in a privacy-protective way to maximize the benefits while mitigating the risks. Such collaboration has the potential to benefit everyone, from individuals to health and humanitarian organizations to governments. We hope that transparency in our decision-making and practices, as well as better consensus around accompanying rules of the road, can help foster trust in critical and potentially life-saving efforts.

Endnotes

[1] See, e.g., Mostashari, F. and Frieden, T., Technology companies can help fight Covid-19, CNN (13 May 2020), https://www.cnn.com/2020/05/13/health/coronavirus-tech-companies-fight/index.html.

[2] See, e.g., Wade, M. and Bjerkan, H., Three Proactive Response Strategies to COVID-19 Business Challenges, MITSloan Management Review (17 April 2020), https://sloanreview.mit.edu/article/three-proactive-response-strategies-to-covid-19-business-challenges/; Power, R. How 5 Businesses Are Adapting To Life In A Pandemic, Forbes (26 April 2020), https://www.forbes.com/sites/rhettpower/2020/04/26/how-5-businesses-are-adapting-to-life-in-a-pandemic/?sh=7f0395104f5a; and McKinsey Digital, How six companies are using technology and data to transform themselves (12 August 2020), https://www.mckinsey.com/business-functions/mckinsey-digital/our-insights/how-six-companies-are-using-technology-and-data-to-transform-themselves.

[3] See, e.g., Hufford, A., New Manufacturers Jump Into Mask Making as Coronavirus Spreads, The Wall Street Journal (21 March 2020), https://www.wsj.com/articles/new-manufacturers-jump-into-mask-making-as-coronavirus-spreads-11584792003; and Miller, N., How factories change production to quickly fight coronavirus, BBC Worklife (13 April 2020), https://www.bbc.com/worklife/article/20200413-how-factories-change-production-to-quickly-fight-coronavirus.

[4] These are only a few of the many areas and projects we explored. For instance, we also took a number of measures to address COVID-19 related misinformation in accordance with our platforms' Community Standards, and we built new features like Messenger Rooms to help people connect during social distancing. For a more comprehensive list of our products, services, and programs related to COVID-19, see Jin, Kang-Xing, Keeping People Safe and Informed About the Coronavirus, Facebook Newsroom (18 December 2020), https://about.fb.com/news/2020/12/coronavirus/; Jin, Kang-Xing, Reaching Billions of People With COVID-19 Vaccine Information, Facebook Newsroom (8 February 2021), https://about.fb.com/news/2021/02/reaching-billions-of-people-with-covid-19-vaccine-information/; and Mark Zuckerberg Announces Facebook's Plans to Help Get People Vaccinated Against COVID-19, Facebook Newsroom (15 March 2021), https://about.fb.com/news/2021/03/mark-zuckerberg-announces-facebooks-plans-to-help-get-people-vaccinated-against-covid-19/.

[5] Id.

[6] Facebook Data for Good, Future of Business Survey, https://dataforgood.fb.com/tools/future-of-business-survey/.

[7] Zuckerberg, supra note 4. See also Jin, Kang-Xing. and McGorman, L., Data for Good: New Tools to Help Health Researchers Track and Combat COVID-19, Facebook Newsroom (6 April 2020) https://about.fb.com/news/2020/04/data-for-good/; and Zuckerberg, M., How Data Can Aid the Fight Against COVID-19, Washington Post (20 April 2020), https://www.washingtonpost.com/opinions/2020/04/20/how-data-can-aid-fight-against-covid-19/.

[8] Id. Any questions about this research - including with respect to review by Institutional Review Boards - can be sent to covid19symptomsurvey@fb.com.

[9] Id.

[10] Id.

[11] Facebook Data for Good, Annual Report 2020, https://dataforgood.fb.com/wp-content/uploads/2021/01/Facebook-Data-for-Good-2020-Annual-Report-1.pdf.

[12] Id.

[13] COVID-19 Mobility Data Network, https://www.covid19mobility.org/.

[14] Supra note 11.

[15] Facebook AI, Using AI to help health experts address the COVID-19 pandemic (20 Oct. 2020), https://ai.facebook.com/blog/using-ai-to-help-health-experts-address-the-covid-19-pandemic/.

[16] Id.

[17] Id.

[18] Id.

[19] Id.

[20] See, e.g., Petersen, M., Tracking who gets vaccinated is vital for public health, but it's raising privacy concerns, Los Angeles Times (28 December 2020), https://www.latimes.com/business/story/2020-12-28/covid19-vaccine-privacy-personal-data; Singer, N. and Sang-Hun, C., As Coronavirus Surveillance Escalates, Personal Privacy Plummets, New York Times (23 March 2020), https://www.nytimes.com/2020/03/23/technology/coronavirus-surveillance-tracking-privacy.html.

[21] See, e.g., EU General Data Protection Regulation (GDPR), Brazil Lei Geral de Proteção de Dados (LGPD), and the forthcoming Personal Data Protection legislation in India.

[22] See, e.g., the Organisation for Economic Co-operation and Development (OECD) Recommendation of the Council on Health Data Governance, https://www.oecd.org/health/health-systems/health-data-governance.htm; Recommendation on the Protection and Use of Health Related Data, UN Special Rapporteur on the Right to Privacy - Task Force

on Privacy and the Protection of Health-Related Data (6 November 2019), https://www.ohchr.org/Documents/Issues/Privacy/SR_Privacy/FINALHRDDOCUMENT.pdf; the World Health Organization (WHO) Data Principles, https://www.who.int/data/principles.

[23] See, e.g., Amnesty International, Covid-19 Response and Rebuilding Principles (April 2020), https://www.amnesty.org/download/Documents/POL3022612020ENGLISH.PDF; Azarmi, M. and Crawford, A., Use of Aggregated Location Information and COVID-19: What We've Learned, Cautions about Data Use, and Guidance for Companies, Center for Democracy and Technology (CDT) (29 May 2020), https://cdt.org/wp-content/uploads/2020/05/2020-05-29-Use-of-Aggregated-Location-Information-and-Covid-19.pdf; Gillmore, D.K., Principles for Technology-Assisted Contact-Tracing, American Civil Liberties Union (16 April 2020), https://www.aclu.org/report/aclu-white-paper-principles-technology-assisted-contact-tracing; and Principles for Protecting Civil Rights and Privacy During the COVID-19 Crisis (12 June 2020), http://civilrightsdocs.info/pdf/policy/letters/2020/CDPWG%20Principles_for_Protecting_Civil_Rights_and_Privacy_While_Combatting_Coronavirus.pdf.

[24] See, e.g., and Citron, D.K. and Solove, D.J., Privacy Harms (9 February 2021), https://ssrn.com/abstract=3782222; Future of Privacy Forum (FPF), Unfairness by Algorithm, Distilling the Harms of Automated Decision-Making (December 2017), https://fpf.org/wp-content/uploads/2017/12/FPF-Automated-Decision-Making-Harms-and-Mitigation-Charts.pdf; IAPP, Building Ethics into Privacy Frameworks for Big Data and AI, https://iapp.org/media/pdf/resource_center/BUILDING-ETHICS-INTO-PRIVACY-FRAMEWORKS-FOR-BIG-DATA-AND-AI-UN-Global-Pulse-IAPP.pdf; National Institute of Standards & Technology (NIST), Privacy Risk Assessment Methodology (February 2019), https://www.nist.gov/itl/applied-cybersecurity/privacy-engineering/resources; and UN Office for the Coordination of Humanitarian Affairs (OCHA), Data Responsibility Guidelines (Working Draft) (March 2019), https://data.humdata.org/dataset/2048a947-5714-4220-905b-e662cbcd14c8/resource/c7053042-fd68-44c7-ae24-a57890a48235/download/ocha-dr-guidelines-working-draft-032019.pdf.

[25] UN Global Pulse, Risks, Harms and Benefits Assessment, https://www.unglobalpulse.org/policy/risk-assessment/.

[26] See, e.g., Aschwanden, C., Contact Tracing, a Key Way to Slow COVID-19, Is Badly Underused by the U.S., Scientific American (21 July 2020), https://www.scientificamerican.com/article/contact-tracing-a-key-way-to-slow-covid-19-is-badly-underused-by-the-u-s/.

[27] See, e.g., Centers for Disease Control and Prevention, Contact Tracing (25 February 2021), https://www.cdc.gov/coronavirus/2019-ncov/daily-life-coping/contact-tracing.html; and World Health Organization, Contact tracing in the context of COVID-19: Interim guidance (1 February 2021), https://apps.who.int/iris/rest/bitstreams/1332668/retrieve.

[28] See, e.g., Feretti, L. et al., Quantifying SARS-CoV-2 transmission suggests epidemic control with digital contact tracing, Science (08 May 2020), https://science.sciencemag.org/content/368/6491/eabb6936.

[29] Id.

[30] Id.

[31] We highlight some of the most common benefits, risks, potential harms, safeguards, and human rights standards in the discussion that follows, but it is not comprehensive of all factors considered. For a fuller discussion, see, e.g., Dubov, A. and Shoptawb, S., The Value and Ethics of Using Technology to Contain the COVID-19 Epidemic, The American Journal of Bioethics 20(7) (18 May 2020), https://doi.org/10.1080/15265161.2020.1764136.

[32] See, e.g., Bashir, A. et al., Applicability of Mobile Contact Tracing in Fighting Pandemic (COVID-19): Issues, Challenges and Solutions, Computer Science Review (28 September 2020), https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3683404; Mello, M.M. and Wang, C.J., Ethics and governance for digital disease surveillance, Science (29 May 2020), https://science.sciencemag.org/content/368/6494/951; and Oliver, M. et al., Mobile phone data and COVID-19: Missing an opportunity?, Computers and Society (27 March 2020), https://arxiv.org/abs/2003.12347.

[33] Exposure notification apps refer to apps that use encrypted identifiers to connect individual users to each other. This is a decentralized technology that is used to warn users in the case of contact with infected individuals ("exposure notification"), but it does not allow for a centralized tracing of possible infection chains like automated contract tracing technology. See Hecht-Felella, L. and Mueller-Hsia, K. Rating the Privacy Protections of State Covid-19 Tracking Apps, Brennan Center for Justice (5 November 2020), www.brennancenter.org/our-work/research-reports/rating-privacy-protections-state-covid-19-tracking-apps.

[34] See. e.g. Braithwaite, I. et al., Automated and partly automated contact tracing: a systematic review to inform the control of COVID-19, Lancet Public Health 2(11) (19 August 2020), https://www.thelancet.com/journals/landig/article/PIIS2589-7500(20)30184-9/fulltext; Ivers, L.C. and Weitzner, D.J., Can digital contact tracing make up for lost time?, Lancet Public Health 5(8) (16 July 2020), https://www.thelancet.com/journals/lanpub/article/PIIS2468-2667(20)30160-2/fulltext; Kretzschmar, M. E. et al., Impact of delays on effectiveness of contact tracing strategies for COVID-19: a modelling study, Lancet Public Health 5(8) (16 July 2020), https://www.thelancet.com/journals/lanpub/article/PIIS2468-2667(20)30157-2/fulltext; Soltani, A., Calo, R., and Bergstrom, C., Contact-tracing apps are not a solution to the COVID-19 crisis, Brookings (27 April 2020), https://www.brookings.edu/techstream/inaccurate-and-insecure-why-contact-tracing-apps-could-be-a-disaster/.

[35] Our role, in this instance, was to provide the donated advertising space to promote the apps. We did not collect or receive any information from the apps themselves.

[36] See, e.g., supra note 23. See also Global scientists and researchers, Joint Statement on Contact Tracing (19 April 2020), https://giuper.github.io/JointStatement.pdf; Mitchell, S.S.D., "Warning! You're entering a sick zone": The construction of risk and privacy implications of disease tracking apps, Online Information Review (14 Oct. 2019), https://www.emerald.com/insight/content/doi/10.1108/OIR-03-2018-0075/full/html; Sharma, T. and Bashir, M., Use of apps in the COVID-19 response and the loss of privacy protection, Nature Medicine 26 (26 May 2020), https://www.nature.com/articles/s41591-020-0928-y; and UK scientists and researchers, Joint Statement on Contact Tracing App NHSX (29 April 2020), https://drive.google.com/file/d/1uB4LcQHMVP-oLzIIHA9SjKj1uMd3erGu/view.

[37] Id.

[38] Id.

[39] Id.

[40] We also considered whether the sponsoring government had a recent record of atrocity crimes or severe human rights violations.

[41] Id. This can be difficult to ensure. For instance, the Singaporean government recently admitted that data from its contact tracing app could also be accessed by the police for use in criminal investigations, despite previous assurances that the data would never be used for any purpose other than virus tracking. See, e.g., Chandran, R., Singapore to limit police access to COVID-19 contact-tracing data, Reuters (2 February 2021), https://www.reuters.com/article/us-singapore-tech-lawmaking-idUSKBN2A20ZI; and Han, K., COVID app triggers overdue debate on privacy in Singapore, Al Jazeera (10 February 2021), https://www.aljazeera.com/news/2021/2/10/covid-app-triggers-overdue-debate-on-privacy-in-singapore.

[42] Supra note 33.

[43] See, e.g., Letter from Andrea Jelinek, Chair EDPB to Olivier Micol, Head of Unit European Commission DG for Justice and Consumers (14 April 2020), https://edpb.europa.eu/sites/edpb/files/files/file1/edpbletterecadvisecodiv-appguidance_final.pdf. In addition to providing individuals with choice, voluntary apps were in greater need of promotion in order to advance adoption by the population to reach effectiveness. See, e.g., Valentino-DeVries, J., Coronavirus Apps Show Promise but Prove a Tough Sell, New York Times (7 December 2020), https://www.nytimes.com/2020/12/07/technology/coronavirus-exposure-alert-apps.html.

[44] Supra note 33.

[45] For instance, France's data protection authority, Commission Nationale de l'Informatique et des Libertés (CNIL), performed two reviews of StopCovid, the contact-tracing app backed by the French government. StopCovid uses a centralized contact-tracing protocol called ROBERT, which isn't an anonymous system — it relies on pseudonymisation. Nonetheless, the CNIL found that ROBERT tries to minimize data collection as much as possible and that its focus on exposed users instead of users who are diagnosed COVID-19-positive "protects the privacy of those persons." See CNIL, Deliberation 2020-056 on a draft decree relating to the mobile application called "StopCovid", request for opinion n ° 20008032 (25 May 2020). https://www.legifrance.gouv.fr/cnil/id/CNILTEXT000041940832?init=true&page=1&query=ROBERT+&searchField=ALL&tab_selection=cnil&timeInterval=.

[46] See, e.g., Center for Information Policy Leadership (CIPL), Data Protection in the New Decade: Lessons from COVID-19 for a US Privacy Framework (August 2020), https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/data_protection_in_the_new_decade_-_lessons_from_covid-19_for_a_us_privacy_framework.pdf; Klar, R. and Lanzerath, D., The ethics of COVID-19 tracking apps – challenges and voluntariness, Research Ethics (5 August 2020), https://doi.org/10.1177%2F1747016120943622; Schwartz, P., Illusions of consent and COVID-19 tracking apps, International Association of Privacy Professionals (19 May 2020), https://iapp.org/news/a/illusions-of-consent-and-covid-tracking-apps/Zwitter, A. and Gstrein, O.J., Big data, privacy and COVID-19 – learning from humanitarian expertise in data protection, Journal of International Humanitarian Action 5, 4 (18 May 2020), https://doi.org/10.1186/s41018-020-00072-6.

[47] See, e.g., EDPB Document on response to the request from the European Commission for clarifications on the consistent application of the GDPR, focusing on health research (2 February 2021), https://edpb.europa.eu/our-work-tools/our-documents/other/edpb-document-response-request-european-commission-clarifications_en, pp. 7-9; European Commission, DG Health and Food Safety, Assessment of the EU Member States' rules on health data in the light of GDPR (February 2021), https://ec.europa.eu/health/sites/health/files/ehealth/docs/ms_rules_health-data_en.pdf, pp. 77-79; European Data Protection Supervisor, A Preliminary Opinion on data protection and scientific research (6 January 2020), https://edps.europa.eu/sites/default/files/publication/20-01-06_opinion_research_en.pdf, pp.18-20.

[48] GDPR, Recital 4, Data Protection in Balance with Other Fundamental Rights, https://gdpr-info.eu/recitals/no-4/. See, e.g., Bradford, L., Aboy, M., and Liddell, K., COVID-19 contact tracing apps: a stress test for privacy, the GDPR, and data protection regimes, Journal of Law and the Biosciences, 7(1) (28 May 2020), https://doi.org/10.1093/jlb/lsaa034.

[49] GDPR, Article 9(2)(g) and (i).

[50] GDPR, Articles 9(2)(j) and 89(1); See also supra note 48.

[51] See, e.g., CIPL supra note 46 (p. 5); DIGITALEUROPE, DIGITALEUROPE recommendations on health data-processing (9 April 2020),

https://www.digitaleurope.org/resources/digitaleurope-recommendations-on-health-data-processing/ ; Future of Privacy Forum, Can Gdpr Work for Health Scientific Research? (2018), https://fpf.org/wp-content/uploads/2018/12/Can-GDPR-Work-for-Health-Scientific-Research-Report.pdf; and Peloquin, D. et al., Disruptive and avoidable: GDPR challenges to secondary research uses of data, European Journal of Human Genetics 28 (2020), https://www.nature.com/articles/s41431-020-0596-x.

[52] As we mentioned above, we do not host the surveys nor collect survey participant responses, and we only have access to public, aggregated survey data provided by the universities.

[53] See, e.g., Choi, K. et al., Data linking race and health predicts new COVID-19 hotspots, The Conversation (20 May 2020) https://theconversation.com/data-linking-race-and-health-predicts-new-covid-19-hotspots-138579; Golestaneh, L. et al. The association of race and COVID-19 mortality, The Lancet, 25 (1 August 2020) https://www.thelancet.com/journals/eclinm/article/PIIS2589-5370(20)30199-1/fulltext; Klugman, K.P., Younger ages at risk of Covid-19 mortality in communities of color, Gates Open Research (26 June 2020), https://gatesopenresearch.org/articles/4-69.

[54] For instance, in the US, Senator Warren introduced bicameral legislation in April 2020 to require the federal government to collect and report coronavirus demographic data, including race and ethnicity. https://www.warren.senate.gov/newsroom/press-releases/senator-warren-introducing-bicameral-legislation-to-require-federal-government-to-collect-and-report-coronavirus-demographic-data_--including-race-and-ethnicity; Similarly, in the UK, the Health Protection (Notification) (Amendment) (Coronavirus) Regulations 2020 requires test providers to report demographic data, including ethnicity. See https://www.legislation.gov.uk/uksi/2020/1175/made.

[55] See, e.g., Farah, W. and Saddler OBE, J., Perspectives from the front line: The disproportionate impact of COVID-19 on BME communities, NHS Confederation (December 2020), https://www.nhsconfed.org/-/media/Confederation/Files/Publications/Documents/Perspectives-from-the-front-line_FNL_Dec2020.pdf; Ford, T. et al., Race gaps in COVID-19 deaths are even bigger than they appear, Brookings (16 June 2020) https://www.brookings.edu/blog/up-front/2020/06/16/race-gaps-in-covid-19-deaths-are-even-bigger-than-they-appear/; and Johns Hopkins University & Medicine, Maps & Trends: Racial Data Transparency, Coronavirus Resource Center, https://coronavirus.jhu.edu/data/racial-data-transparency.

[56] An important question for future projects like this is whether this sensitive information should be collected from the start so that disparities can be more rapidly identified and addressed -- and, if so, what protections are appropriate to guard against the misuse of this data -- or if this type of collection should only happen when there is some level of confidence that a disparate impact exists or will exist.

[57] The Delphi Group at Carnegie Mellon University in partnership with Facebook, COVID-19 Symptom Survey, Topline Report on COVID-19 Vaccination in the United States (12 March 2021),

https://dataforgood.fb.com/wp-content/uploads/2021/03/CMU_Topline_Vaccine_Report_20210312-1.pdf.

[58] See supra note 47. Even with explicit consent, individuals may de-prioritize their own privacy in an emergency in favor of competing values, such as security, health, or safety. Moreover, it may create difficulty in measure for equity or bias depending on the representativeness of those consenting. See, e.g., Crawford, K. and Finn, M., The Limits of Crisis Data: Analytical and Ethical Challenges of Using Social and Mobile Data to Understand Disasters, GeoJournal, 80(4), (1 August 2015), https://link.springer.com/article/10.1007/s10708-014-9597-z; and Greenwood, F. et al., The Signal Code: A Human Rights Approach to Information During Crisis, p. 44, https://hhi.harvard.edu/sites/default/files/publications/signalcode_final.pdf.

[59] Farr, C., Over 30 million people told Facebook if they had the coronavirus or wore masks — and now it will be used for science, CNBC, (4 September 2020), https://www.cnbc.com/2020/09/04/over-30-million-people-told-facebook-if-they-had-covid-or-wore-masks.html

[60] See, e.g., supra note 57.

[61] EDPB, Guidelines 03/2020 on the processing of data concerning health for the purpose of scientific research in the context of the COVID-19 outbreak (21 April 2020), p. 10, https://edpb.europa.eu/our-work-tools/our-documents/ohjeet/guidelines-032020-processing-data-concerning-health-purpose_en.

[62] Supra note 47.

[63] The European Strategy for Data published in February 2020 calls for this guidance, noting "...[t]he GDPR has created a level playing field for the use of health personal data, [but] fragmentation remains within and between Member States and the governance models for accessing data are diverse." European Commission, Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions: A European Data Strategy (19 February 2020)), p. 29, https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1593073685620&uri=CELEX%3A52020DC0066.

[64] The paper notes the need to draw from larger datasets across EU member states, stating: "Data collected in one country alone is often insufficiently powered to answer many of the public health questions that national authorities face, e.g. for rare disease exposed, rare outcomes or public health emergencies. Large data sets also provide a greater degree of precision and accuracy, and combining data across Member States provides information on national variations, on the effectiveness and impact of different public health interventions and strategies on larger numbers of patients, and may provide complementary types of information, which is important for regulatory and policy decision-making, both at Union and national levels." European Medicines Agency, The General Data Protection Regulation: Secondary Use of Data for Medicines and Public Health Purposes - Discussion

Paper for Medicines Developers, Data Providers, Research-Performing and Research-Supporting Infrastructures (2020), http://www.encepp.eu/events/documents/Discussionpaper.pdf; Heads of Medicines Agencies and European Medicines Agency, HMA-EMA Joint Big Data Taskforce Phase II report: 'Evolving Data-Driven Regulation' (2019), https://www.ema.europa.eu/en/documents/other/hma-ema-joint-big-data-taskforce-phase-ii-report-evolving-data-driven-regulation_en.pdf.

[65] For example, a recent legal analysis of Ireland's Health Research Regulations, which were passed after GDPR, found their consent requirements "short-circuits the pandemic exemptions set out in GDPR." Kirwan, M., Mee, B., Clarke, N. et al. What GDPR and the Health Research Regulations (HRRs) mean for Ireland: "explicit consent" – a legal analysis. Irish Journal of Medical Science (30 July 2020), p. 6, https://link.springer.com/article/10.1007/s11845-020-02331-2.

[66] European Medicines Agency, Questions and Answers (Q&As) on data protection and the secondary use of personal data for medicines development and public health purposes (29 September 2020), https://www.ema.europa.eu/en/documents/presentation/questions-answers-data-protection-secondary-use-personal-data-medicines-development-public-health_en.pdf.

[67] See, e.g., Top Strategic Technology Trends for 2021: Privacy-Enhancing Computation, Gartner (12 January 2021), https://urldefense.com/v3/__http:/www.gartner.com/document/3995508?ref=sendres_email&refval=77482551__;!!Bt8RZUm9aw!qp-4MxU_LAD0xDLSArYbrFHaszoVNgpi2CBwk0R6Od9JNkhz0XIyYjP-biI$ (subscription required).

[68] See, e.g., EDPB, Guidelines 04/2020 on the use of location data and contact tracing tools in the context of the COVID-19 outbreak (21 April 2020), https://edpb.europa.eu/our-work-tools/our-documents/ohjeet/guidelines-042020-use-location-data-and-contact-tracing-tools_en.

[69] For more examples, see supra note 11.

[70] McGorman, L., Privacy Matters: Data for Good, Facebook Newsroom (3 June 2020), https://about.fb.com/news/2020/06/privacy-matters-data-for-good/.

[71] Herdağdelen, A. et al., Protecting privacy in Facebook mobility data during the COVID-19 response, Facebook Research (3 June 2020), https://research.fb.com/blog/2020/06/protecting-privacy-in-facebook-mobility-data-during-the-covid-19-response/.

[72] Id.

[73] Id.

[74] Id.

[75] Supra note 70.

[76] Id.

[77] Id.

[78] Id.

[79] Supra note 68, p. 5.

[80] Id, p. 6. The EDPB guidance says anonymization methods should be assessed in accordance with a "reasonability test" that "must take into account both objective aspects (time, technical means) and contextual elements that may vary case by case (rarity of a phenomenon including population density, nature and volume of data)." (p. 5) Processing to meet the reasonability test includes considering location datasets as a whole, not just anonymizing individual data on its own, as well as processing data "from a reasonably large set of individuals using available robust anonymization techniques." (p. 6) Further, "evaluating the robustness of anonymisation relies on three criteria: (i) singling-out (isolating an individual in a larger group based on the data); (ii) linkability (linking together two records concerning the same individual); and (iii) inference (deducing, with significant probability, unknown information about an individual)." (pp. 5-6)

[81] Buckee, C.O., et al., Aggregated mobility data could help fight COVID-19, Science (10 April 2020), https://science.sciencemag.org/content/368/6487/145.2.full?fbclid=IwAR1jcZK1fxtO4. The article spoke to the value of aggregated mobility data in this context: "A map that examines the impact of social distancing messaging or policies on population mobility patterns, for example, will help county officials understand what kinds of messaging or policies are most effective. Comparing the public response to interventions, in terms of the rate of movement over an entire county from one day to the next, measured against a baseline from normal times, can provide insight into the degree to which recommendations on social distancing are being followed."