CENTER FOR DATA INNOVATION

How the EU Can Unlock the Private Sector's Human-Mobility Data for Social Good

By Hodan Omaar | March 28, 2022

Many businesses routinely collect data about the location of consumers, such as where they are when they make a purchase or use a mobile app. Aggregating this information reveals useful insights about human mobility and social interaction. Researchers, governments, and others can use this mobility data, while respecting user privacy, to study and address many pressing societal challenges, such as disease spread, urban functioning, forced migration, climate change, and disaster response. To support these types of applications, EU policymakers should encourage businesses to share mobility data by implementing policies that provide firms with regulatory clarity, financial incentives, and technical resources to give out this type of data.

Private firms face a number of challenges that limit their willingness and ability to share mobility data. The government's role should be to coordinate the behavior of individuals, companies, and researchers toward social good.

INTRODUCTION

Mobility data describes people's movements from one location to another at certain points in time. Population censuses and travel history surveys have historically served as sources of mobility data for researchers exploring population movements. But censuses and surveys are costly to implement, limited in scope and granularity, and ineffective in situations where timely information is needed, such as during conflicts or epidemics. So instead, researchers have increasingly turned toward location data from call records and mobile apps that provide location-based services to fill in the gaps of human mobility patterns at high spatial resolutions, spanning wide temporal periods, and across international borders.

Consider the use of mobility data to track the movement of people from Ukraine after Russia invaded the country in February, which represents the fastest and largest displacement of people in Europe since World War II.[1] Crisis Ready, a collaboration between Harvard University and nonprofit humanitarian organization Direct Relief, has been using mobility data from Meta's (formerly Facebook's) Data for Good program to show where Ukrainian refugees are moving from and to in close to real time, enabling a range of policymakers and response agencies to prioritize the allocation of limited resources to the locations where they are most needed.[2]

Unfortunately, access to novel mobility data is difficult for researchers to obtain because it typically rests in the hands of private firms that face significant legal, financial, and practical challenges to sharing this data.

To address privacy concerns, firms can use tools that limit the risk that an individual's records are uniquely identified in or inferred from a dataset. However, the more specific and less general de-identified mobility data is, the more useful it may be to researchers, which means firms must strike a fine balance in de-identification between the utility of the data and the risk of re-identification. Proper de-identification tools are often costly, which smaller firms may see as a prohibitive cost to data sharing. Those that do share proprietary mobility data may reduce their own competitiveness if other companies use their data to inform their business strategies, and may even incur negative impacts on their reputation if their data is used for unpopular public policies, such as extended lockdowns to address the spread of an infectious disease. But perhaps the greatest challenge the private sector faces to sharing mobility data is the lack of regulatory clarity on what types of data firms should share, whom they should be sharing with, and for what purposes.

To address these challenges and create opportunities for social good, policymakers in the EU should work to encourage the reuse of private sector mobility data. This report offers several recommendations for how EU policymakers can support unlocking private sector mobility data for social good:

- Revise the Data Governance Act (DGA) to allow private sector firms to be listed in national registers of recognized data altruism organizations
- Pilot a common European data space for human mobility that researchers can use for social good
- Issue a Horizon Europe challenge to encourage researchers to use the data space for specific problems in the public interest

- Proactively support experimentation of methods that use mobility data for social good
- Draft a common EU approach for the use of mobility data in emergency contexts to support future public health and humanitarian-crisis exit strategies
- Pursue global partnerships in mobility data

THREE WAYS MOBILITY DATA CAN BE USED FOR SOCIAL GOOD

Mobility data can provide valuable information regarding movement patterns that can help address questions in the public interest, such as those related to public health, urban development, transportation, poverty, migration, and disaster response. While the application areas are broad, the role mobility data plays in helping researchers address social problems can be structured around three distinct types of functions: descriptive, predictive, and prescriptive.

First, mobility data can play a descriptive role, helping researchers quantitively understand and communicate trends and patterns in information, such as through visualizations.[3] Consider one of the earliest and most high-profile use cases of mobility data during the COVID-19 pandemic: a model that visualized the spread of the disease in Wuhan, China, during and after the Chinese New Year holiday.[4] The model used de-identified, aggregated mobility data provided by Baidu, the largest Chinese search engine; air passenger itinerary data provided by the International Air Travel Association (IATA); and case reports from the Chinese Center for Disease Control and Prevention (Chinese CDC). Research showing that a high volume of international travelers left Wuhan for hundreds of different destination cities around the world during the two weeks before the first travel ban was implemented in the city enabled a range of policymakers to prioritize the allocation of limited resources for surveillance to locations where the risks of importation were calculated to be highest.[5]

Second, mobility data can play a predictive role, helping researchers forecast the likelihood of a given event in the future. For instance, using mobility data from Meta's Data for Good program, researchers at the École Polytechnique institution in France developed a machine learning model to forecast up to 14 days in advance the number of COVID-19 cases regions within the EU would likely experience.[6]

Finally, mobility data can play a prescriptive role, enabling researchers to examine the possible consequences of different choices and decide on the best course of action. For example, researchers at the Polytechnic

University of Milan in Italy modeled the daily movements of over 4 million individuals in Italy using data from Meta, and examined the economic outcomes derived from changing the timing and modality of government mobility restrictions to contain COVID-19.[7] They showed that implementing different mobility restriction measures could reduce disposable income by between 10 percent (in the best case) and 40 percent (in the worst case), and that reductions in disposable income are due to non-linear interactions between mobility policies and infection transmission rates.[8]

SOURCES OF MOBILITY DATA

Traditionally, measuring human mobility has relied on data from population and housing censuses, travel history surveys, and air travel data. More recently, call detail records (CDRs), location-based services, and Earth observation data have emerged as novel sources of data on human mobility.

POPULATION CENSUSES

National population censuses systematically count the populations of countries at regular intervals and collect data on their main demographic, social, and economic characteristics.[9] These surveys, which date back to the 18th and 19th centuries in most EU countries, are conducted every 10 years and are one of the most widely available sources of comparable data on international human movement.[10]

While each EU member state collects census data in the way best suited to its administrative practices, EU-wide legislation governs how this data is reported to ensure harmonization that makes data easily comparable. EU legislation provides technical specifications for the census topics and their breakdowns, and requires both statistical data be supplemented with metadata and that member states submit data quality reports.[11] To ensure census data from all member states would be easily accessible and comparable, in 2011, the EU's statistical office together with the statistical offices of all EU member states set up the Census Hub, an online portal that enables users to access data from any member state.

One of the distinct advantages of population censuses is they seek to cover all residents in a country, including undocumented immigrants, unlike many administrative sources, such as population registers, which exclude them by their very nature.[12] Of course, censuses inevitably fail to reach all people, especially those who have a vested interest in avoiding being counted. But because censuses collect detailed demographic and socioeconomic information about each individual, such as their age, sex, level of education, and occupation, researchers can perform in-depth analyses on the causes and consequences of mobility.

The main drawback to population censuses is that they can only really be used to understand long-term trends in human mobility, as these surveys are usually only carried out once every decade. They cannot capture trends in international mobility in a sufficiently timely fashion to support responses to events that occur when conditions change, such as during an acute public health crisis.

TRAVEL HISTORY SURVEYS

Many European countries conduct travel surveys, which are household surveys of individual travel behavior that collect quantitative and qualitative information on how, why, when, and where people travel, as well as factors affecting travel such as the availability of vehicles and driver's licenses, to inform the development of transportation policy. While different countries use different methods to conduct travel surveys, most select a random sample of households or individuals to participate and require them to provide information on their most recent travel behavior. For instance, the annual national travel survey in England includes approximately 16,000 individuals in 7,000 households and requires respondents to self-complete a seven-day written travel diary.[13] Because travel surveys collect comprehensive data on individuals, they can serve as a useful source of information on people's mobility patterns, especially when supplanted with workplace surveys, license-plate surveys, and transit on-board surveys.[14]

However, there are several limitations to using travel surveys. First is nonparticipation. Many of the individuals selected for random sampling simply do not respond, which negatively affects the quality of data collected.[15] Second is recall bias. Because travel surveys require respondents to self-report their travel behavior, the data collected suffers from the systemic error that occurs when respondents do not remember previous events accurately or omit details.[16] Third is that the availability of travel data varies across countries. For instance, a 2013 analysis by the European Commission finds that Cyprus, Hungary, and Latvia do not make household travel data available to anyone; Belgium and Sweden make travel data available only to academic researchers and public authorities; while Finland, Germany, Spain, and the United Kingdom make travel data available to the general public.[17] Last is the inability to easily compare travel patterns across countries. For example, the range and classification of age groups of surveys across countries differs. Some countries such as Germany collect data on all age groups, while other countries only collect data on some age groups, making analysis of mobility patterns across the EU difficult.

AIR TRAVEL DATA

Data on the global flow of air travel passengers is a valuable resource for making inferences about trends in and the types of transnational human mobility. Many reports show that air travel is both directly and indirectly responsible for the spread of infectious diseases, including dengue, severe acute respiratory syndrome (SARS), and most recently COVID-19. Unfortunately, access to complete air travel data remains a challenge for researchers. Current available data sources can be grouped into two categories.

The first category is commercially available datasets from international organizations and private companies. While these datasets can include granular, timely data about the movement of people between two points, they can be very expensive and often have strict limitations on how researchers can use them. For example, the IATA has complete passenger origin and destination data disaggregated by airline for sale in the tens of thousands of dollars, with payment required repeatedly to maintain the most up-to-date data.[18] And Sabre, a private company that collects data directly from the airline industry, has monthly data about air passenger traffic between hundreds of countries and territories worldwide.[19] But it only makes annual country-to-country air travel data publicly available at no cost. Researchers have to pay for access to monthly air travel data.

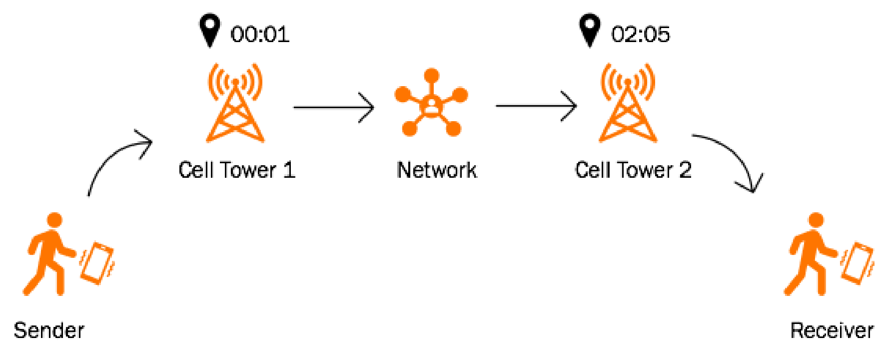
The second category is open-access data resources. Several researchers have developed statistical models that use publicly available data to model passenger flows. For example, U.K. researchers created a passenger flow model in 2013 using airport location data and scheduled route information from FlightStats, a free real-time flight status tracking website, and passenger flow statistics from a variety of sources, including government transportation department websites.[20] However, the problem with existing models is they are limited to predicting annual passenger flows and volumes, which hides seasonal dynamics that are important for understanding global processes such as disease spread and labor migration.[21]

CALL DETAIL RECORDS

CDRs are data mobile network operators (MNOs) collect for billing, monitoring usage, and other purposes. CDRs document the times of calls or messages, the lengths of calls, the callers' and receivers' phone numbers, and their locations.

Every time a person makes a call or sends an SMS text, their phone transmits a signal to the nearest cell tower, which connects them to the mobile phone network their SIM card is subscribed to. Then, the network "pings" the call or text message across towers until it reaches the tower nearest to the receiver, at which point the tower connects to the receiver's SIM. MNOs have logs recording the location of both the starting and ending towers as well as a timestamp of each event.

Figure 1: How CDR data captures location data



When aggregated, CDR data can provide valuable data about human movement. For example, telecommunications company Vodafone partnered with WorldPop, an applied research group at the University of Southampton in England, and the government of Mozambique to inform malaria-elimination strategies in the country. Malaria, a mosquito-borne disease caused by a parasite, is one of the biggest causes of death in Mozambique, with the National Malaria Control Program estimating that the disease was responsible for 29 percent of all hospital deaths among the general population and 42 percent of deaths in children under five years old in 2015.[22] Vodafone provided pseudonymized data on more than 80 billion call records, enabling researchers to derive information about the large-scale flows of people from the usage of cell towers across the country.[23] The researchers then combined this data with malaria incidence maps to identify areas where the disease was most likely to spread. Among other things, they found that districts in southern Mozambique are more connected to each other through human mobility than they are to districts in the north, which suggests that efforts to eliminate malaria from the southernmost provinces could have lasting effects on reducing the risk of cross-border exports of the disease into neighboring countries.[24]

Despite the value CDR data offers, this data suffers from several drawbacks for tracking mobility. First, CDR data tracks the use of a SIM card, not individual people, which presents challenges to obtaining granular information about mobility for certain groups. SIM cards are often associated with individual use because they store information people want consistent access to or prefer not to share, such as a unique identity number (phone number), personal contacts, and security authentication.[25] But sharing SIM cards is not uncommon among refugees and people in developing countries because it is more cost effective. Because patterns of SIM card ownership and usage differ across contexts, CDR data can exclude valuable information about the mobility of certain demographic groups and fail to provide a full picture of the movements of these individuals.[26]

Second, CDR data can underestimate the total distance people travel because it only records location data when people make calls and send texts, discounting those who rarely use their phone to communicate. For instance, CDR data about a delivery driver who uses their phone frequently to communicate with customers may accurately depict their daily movements, but CDR data about a person who chooses to only use their phone very occasionally may not be as reliable a source of information on their movement in time and space.[27]

Most importantly, CDR data can be onerous to obtain, as it is proprietary, which means research groups or humanitarian groups who want to use the data must first set up data sharing agreements with MNOs. According to the International Organization for Migration, these agreements "typically state the legal basis for which the data is being shared, the exact purpose for sharing the data, what the receiving party can and cannot do with the data, who will receive access to what data (role-based access), and the geographic and temporal limits of data sharing."[28] A data sharing agreement to model the spread of malaria, for example, would likely state the need for data sharing over a long period of time and cover large geographic areas, whereas a data sharing agreement to support an immediate humanitarian crisis response from an earthquake might be limited to a shorter time period and cover a smaller area. The significant time and effort needed to set up data sharing agreements mean any collaboration should be set up well in advance of when CDR data is needed. The practicality of setting up these agreements can be further complicated by the relationship between the actors involved. For instance, in Bangladesh, the government has full authority over MNOs' operations, so data sharing agreements with humanitarian groups are easier to set up than in many other countries.

LOCATION-BASED SERVICES

Location-based services are mobile, desktop, or web applications that use the location of a user's device to provide tailored information and personalized services.[29] Examples include search engines that provide geographically relevant search results, camera apps that assign location tags to photos and videos, weather apps that offer localized weather information, and dating apps that pair users with prospective matches in their area.

To use location-based services, users must allow location tracking services on their devices and authorize each app or website to use the location data that their devices collect. Location-based services may gather data through a variety of technologies to precisely determine an individual device's location, including GPS and triangulation from cell towers or Wi-Fi networks. In addition, these services may map IP addresses or user-submitted information to a geographic location to identify a less-precise estimate of a device's location.[30]

Location-based services have seen tremendous growth over the past two decades for several reasons. First, there have been significant advances in enabling technologies, such as low-cost mobile phones and ubiquitous wireless connectivity. Second, the range of applications for these services has expanded from traditional application fields such as navigation and tourism to diverse applications such as location-based social networking, entertainment, personal fitness, dating, gaming, advertising, and search.[31] Finally, there have been rapid advances in the interface technologies and devices for location-based services such as smartwatches, digital glasses, and augmented and virtual reality devices, among others. [32]

One of the unique benefits of mobility data from Internet platforms and applications is a user's data can be collected across all the different devices they use an app on and consolidated to their unique user account. The data is also precise because location is identified using both the device's internal GPS and connected Wi-Fi devices, which makes the data useful for mapping travel routes across time and space. Because it is typically passively collected, data from location-based services avoids both many compliance issues in studies that use GPS trackers as well as the recall bias prevalent in self-reported travel surveys.[33] A 2018 paper published in the *International Journal of Health Geographics* finds that Google Location History (GLH) data passively recorded by Android phones:

> can provide mobility data over periods and at a resolution infeasible from other typical sources of movement information.... GLH data are a greatly underutilized and novel dataset for understanding human movement. While biases exist in populations with GLH data, Android phones are becoming the first and only device purchased to access the Internet and various web services in many middle and lower income settings, making these data increasingly appropriate for a wide range of scientific questions.[34]

As is the case with telecom companies, however, very few Internet and smartphone application companies have made access to the mobility data they collect publicly available. Meta's Data for Good program is an example of one company that has (see box 1), but researchers have few other options regarding comprehensive mobility analytics from location-based services.

Box 1: Meta Data for Good Program Case Study

Meta created its Data for Good program in 2017 with a stated mission of supplying international agencies, nonprofits, and academic researchers with data for humanitarian response. Meta launched this program with its first product, Disaster Maps, which uses aggregated location data to provide real-time information on where disaster-affected populations are located, how they are moving, and whether they have access to cellular networks and power.[35] According to the Internal Displacement Monitoring Centre, the leading international body monitoring conflict-induced internal displacement, using Meta location data to measure displacement patterns in the weeks and months after disasters is a valuable resource to fill data gaps in understanding the dynamics of disaster displacement.[36] For instance, because Meta's Disaster Maps generate daily estimates of movement, humanitarian organizations can better understand the progression of displacement.[37] In addition, because Meta's data provides a matrix of city-to-city movement, those same groups can understand displacement dynamics at a more granular level.[38]

In 2019, Meta introduced a second product, Disease Prevention Maps.[39] These maps use a range of datasets to improve the effectiveness of epidemic response. For example, colocation maps show the probability that people in one area will encounter people in another, helping to indicate where an infectious disease may spread next. And movement-range maps show, at a regional level, whether people are staying within a small area surrounding their homes or visiting many parts of their town, which can provide insights into the adherence to lockdown measures.[40]

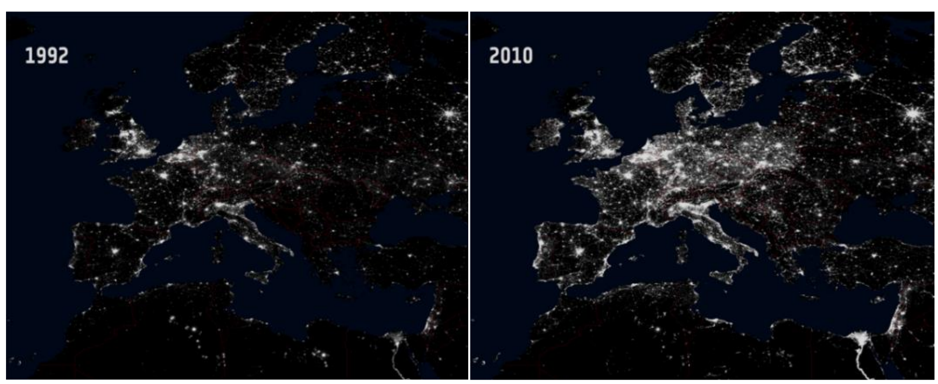
In April 2020, Meta expanded its disease prevention efforts by providing visualizations, datasets, and surveys that help public sector institutions answer questions related to the COVID-19 pandemic. For example, Meta created a social connected index that measures the strength of connectedness between two geographic areas as represented by Meta "friendship ties." Because friendship ties are much more influential in long-distance mobility than in short-range mobility, understanding the impact social network links have on movement can reveal important insights about the emergence of variations in infection rates for diseases such as COVID-19.

EARTH OBSERVATION DATA

Remote sensing technologies that use satellites to detect and monitor different physical characteristics of Earth can provide valuable long-term measures of population densities at geographic scales that are otherwise nearly impossible to obtain. Today, there are more than 150 Earth observation (EO) satellites in orbit that have sensors measuring solar radiation, thermal energy, and light from the Earth's surface. One of the most novel applications of this data is the use of satellite imagery of nighttime lights to understand long-term changes in human settlement patterns. Many researchers have found that nocturnal lighting is a quantifiable indicator of human presence, and measuring the dynamics of nocturnal lighting can help quantify seasonal fluctuations in population sizes.[41]

Figure 2 displays satellite images of Europe at night showing lights from sources in cities and along roads in 1992 and 2010, with brighter regions corresponding to more densely populated regions such as London, Paris, and Rome.[42]

Figure 2: Satellite imagery of night lights in Europe



One benefit of using EO data on anthropogenic illumination to infer human mobility patterns is the data represents the majority of individuals within a population. Unlike data derived from mobile phones, which can underrepresent rural and low-income populations, EO data is highly representative, capturing the light humans produce both when they use electrical power and when they use fire.[43] Another benefit is this data is publicly available from scientific and space agencies such as the European

Space Agency, the U.S. National Oceanic and Atmospheric Administration, and U.S. National Aeronautics and Space Administration.

While EO data does provide data on changes to population densities, it does have one major drawback in that it does not provide insight on individual trip characteristics or motivations. Knowing where people are moving from and to is key to effectively modeling population movement, which informs policies such as those on disease prevention. As a result, EO data is most useful in modeling mobility when it is enriching other, more granular data sources.

BARRIERS TO MOBILITY DATA FROM THE PRIVATE SECTOR

Having as large a range of sources for mobility data as possible is important to improve the prediction accuracy of research models. As a 2021 report from the Joint Research Center, the European Commission's science and knowledge service, explains, "When two complementary datasets are merged or aggregated into a single data pool, the aggregated dataset may produce more insights and economic value than the sum of insights and values of the individual datasets."[44]

However, while the previous examples illustrate the potential to use a variety of sources of mobility data for social good, researchers are largely unable to tap into those that rest in the hands of the private sector because private firms face a number of challenges that limit their willingness and ability to share mobility data. These include privacy and security concerns associated with sharing sensitive location data; the financial costs to de-identify data; the economic opportunity costs that may arise if competitors use their proprietary data for their own gain or data is used in ways that harm their reputation; and most importantly, the lack of regulatory clarity on what types of data to share, whom to share it with, and for what purposes.

PRIVACY AND SECURITY CONCERNS

Human mobility data contains personal information about people's whereabouts, which means firms have to consider the privacy and security risks associated with sharing this data. One oft-cited risk is that aggregated mobility data may be deanonymized and used in ways that undermine individual and group privacy and security, an issue that has become more prominent after Edward Snowden's revelations on the use of CDRs as part of the U.S. National Security Agency (NSA) digital surveillance programs.[45] Privacy issues are top of mind for policymakers all over the world and are one of the greatest obstacles to the use of private sector mobility data.[46]

But even privacy activists acknowledge that when it comes to human mobility data, privacy should be balanced against the collective good. Speaking to Dutch newspaper *De Volkskrant* regarding contact tracing tools for COVID-19, Bas Filippini, chairman of privacy organization Privacy First, said, "I think many people voluntarily want to temporarily give up their privacy by installing an app that tracks them."[47] And in an article published on his organization's website, Filippini noted that "privacy can sometimes be temporarily limited (i.e. not in structural legislation) if strictly necessary in the public interest."[48]

It is important for EU policymakers to remember that European citizens are in favor of having their data used for social-good purposes. An online consultation by the European Commission on the European strategy for data in 2020 finds that 70 percent of respondents believe technology should enable citizens to make their data for available for the public interest, and 83 percent believe they should do so without any direct reward.[49]

Figure 3: Europeans support the reuse of their data for social good

Given the economic and reputational damage firms face from data breaches, however, firms are cautious about sharing de-identified data for secondary use, especially since they (not the firm's customers) are responsible for giving consent to its use.

DIRECT FINANCIAL COSTS

To responsibly share mobility data, companies must invest in costly tools to aggregate and de-identify sensitive information, such as differential privacy algorithms that, by injecting noise to datasets, make it hard to reidentify people with high levels of certainty.[50] While some larger, more resourced firms can afford to invest in such tools, other smaller or mid-sized firms may see de-identification as a prohibitive cost to data sharing, and won't voluntarily share the mobility data they collect unless the business advantage to doing so outweighs the costs.

ECONOMIC OPPORTUNITY AND REPUTATIONAL COSTS

Even though the non-rivalrous nature of data implies that multiple entities can use the same data without the data losing value, the original data collector and user may face economic opportunity costs from others reusing it. Most firms therefore exclude anyone else from reusing the data for any purpose.

The human mobility data private firms collect is largely proprietary and sharing this data can raise competitive concerns. Consider a ride-sharing company that makes the mobility data they collect publicly available. A competitor company could obtain access to this data and learn which areas their customers are requesting rides from and at what times to strategically influence the volume of trade in their favor.[51]

While sharing mobility data for social good can certainly boost the reputation of a company, it can also bring reputational risks if the data is used to support unpopular public policies, such as extended lockdowns to quell the spread of an infectious disease.[52] A decrease in a firm's reputation can then lead to a decrease in demand for its products and services, and thus negatively impact its profits.

LACK OF LEGAL CLARITY

Despite several health emergencies, such as the 2013 Ebola outbreak in West Africa and more recently the COVID-19 pandemic, highlighting the need for legal clarity around the re-use of private sector location data, firms in the EU that want to share mobility data for socially beneficial secondary use operate in a legal and political vacuum. Existing EU laws on data protection and privacy include exemptions for legal directives that support data sharing for the public good. But the EU has yet to establish comprehensive directives that provide regulatory clarity for private firms that want to share mobility data.

The Existing Legal Basis for Sharing Data on Mobility in the EU

The EU's ePrivacy directive, which mandates electronic communication providers to only transmit location data to third parties only after they anonymize the data or obtain prior consent from data subjects, and the General Data Protection Regulation (GDPR) appears to provide the most relevant legal basis for firms sharing and processing location data for social good. Under Article 6 of the GDPR, it is legal for firms to process mobility data if it "is necessary for the performance of a task carried out in the public interest."[53] And according to Article 23, such tasks include those related to "important objectives of general public interest of the Union or of a Member State, in particular an important economic or financial interest of the Union or of a Member State, including monetary, budgetary and taxation matters, public health and social security."[54] In addition, according to Article 15 of the ePrivacy directive, "Member States may adopt legislative measures to restrict the scope of the rights and obligations provided for in … this Directive when such restriction constitutes a necessary, appropriate and proportionate measure within a democratic society."[55]

While the GDPR and ePrivacy directive leave room for laws that sanction the secondary use of personal data for social good, any directives must specifically spell out conditions for the secondary use of data, such as the types of personal data that are subject to be processed, the entities to, and purposes for which, the data may be disclosed, and the processing procedures entities must follow, including measures to ensure lawful and fair processing.

An example of such a directive is Finland's Act on Secondary Use of Health and Social Data.[56] This act, which came into force in May 2019, provides a legal basis for collecting health and social care data from public and private healthcare providers for secondary use. The law covers the entire lifecycle of accessing health and social care data, from the types of data organizations must make available to which entities can apply for access to detailed descriptions of the IT systems for requesting, gathering, accessing, and analyzing the data.[57] Additionally, since January 2020, the entire data sharing process has been overseen by a single entity, the Finnish Health and Social Data Permit Authority (Findata). When a researcher wants pseudonymized data on individual patients or aggregated, anonymized statistics on specifically Finnish patients, they do not need to request data from the healthcare providers directly and instead can apply for access through Findata, which reviews applications on a case-by-case basis. After approval, Findata uses its authority to mandate data controllers provide the relevant data, which it then uploads to its own secure hosting environment. Next, Findata pre-processes the data by linking data entries from different datasets to the personal identification number of every Finnish resident, before pseudo- or wholly anonymizing the data. Once the data is pre-processed, researchers are invited to access and analyze it in Findata's secure IT environment—or, if the data is fully anonymized, the researchers can download it directly. Finally, once the research project is complete or the period for data access expires, Findata deletes all the data it gathered from its IT environment. It also obligates researchers to allow

Findata employees to check any draft publications to ensure full adherence to its privacy and security protocols.

Unfortunately, the EU has not created comprehensive legal clarity for the secondary use of private sector mobility data.

Proposed Laws to Support Data Sharing Do Not Cover For-Profit Firms Sharing Mobility Data

The EU has several initiatives to encourage data sharing, but existing initiatives either exclude for-profit companies from participating or focus on mobility data to create intelligent transportation systems rather than human mobility for social good.

Consider the DGA, the EU's framework for data sharing mechanisms. It introduced the idea of data altruism, defining it as "the consent by data subjects to process personal data pertaining to them, or permissions of other data holders to allow the use of their non-personal data without seeking a reward, for purposes of general interest, such as scientific research purposes or improving public services"[58] In other words, the EU's framework for data altruism encompasses individuals actively consenting to share their personal data for research (where personal data refers to personally identifiable information) and organizations sharing nonpersonal data about their users for socially beneficial purposes. An example of the former is the Personal Genome Project, which enables willing participants to publicly share their genome sequence and health data for use in scientific research.[59] An example of the latter is Entur, a government-owned transportation company in Norway that collects data on public transport, sharing anonymized data about trips people take.[60]

While the DGA is a positive step forward in establishing a framework for organizations to facilitate the collection, sharing, and use of data for altruistic purposes, it excludes for-profit organizations and those that do not perform activities related to data altruism "through a legally independent structure, separate from other activities it has undertaken."[61] But as discussed earlier, when it comes to human mobility data, much of the most granular, representative, and useful data needed to address societal challenges rests in the hands of for-profit companies and is collected primarily for commercial reasons.

It is unclear why the EU excludes for-profit companies in its framework for data altruism organizations, but it may stem from the idea that profit maximization is at odds with social-mindedness. Consider the European Commission's 2019 working paper on corporate social responsibility, which details the progress the commission has made on promoting social entrepreneurship and social innovation. In it, the commisson noted that it is “supporting the development of social enterprises that … prioritize their social impact over profit.”[62] But in the data economy, firms do not have to trade one for the other. When a firm practices traditional philanthropy, the gifts, money, or services they donate are no longer available to the firm to use or give to anyone else. However, when a firm is philanthropic in the data economy by offering its database for socially beneficial purposes, it can still capitalize on the same database for its own commercial interests because data is non-rivalrous. Multiple entities can use the same data simultaneously without the value of the data being diminished.[63] Data can also be excludable, meaning firms can prevent other firms from using it. Consider Meta’s Data for Good program detailed earlier in box 1; Meta only permits select researchers to reuse the mobility data they collect, which addresses any competitiveness concerns that may arise. As such, firms in the data economy can theoretically be simultaneously profit maximizing and altruistic, and the EU should not exclude them from participating in their efforts to encourage altruistic data sharing.

The EU’s other major initiative to encourage mobility-related data sharing established in its Communication on a European Strategy for Data is a European Mobility Data Space (EMDS), a common data pool that aims to facilitate access, pooling, and sharing of data for the development of applications that support intelligent transport systems. For instance, users can access data on road safety conditions, infrastructure and maps, road utilization, and weather.[64] This initiative is aimed at bolstering European competitiveness and has a worthwhile goal, but it is not designed to encourage the sharing of data on human mobility to promote social good.

As a result, researchers who want access to data related to human mobility from private firms rely on ad hoc agreements or data challenges initiated at the will of companies. According to a letter signed in April 2020 by more than 20 leading universities, private companies, nonprofits, and international organizations:

> Although ad hoc mechanisms leveraging mobile phone data can be effectively (but not easily) developed at the local or national level, regional or even global collaborations seem to be much more difficult given the number of actors, the range of interests and priorities, the variety of legislations concerned, and the need to protect civil liberties. The global scale and spread of the COVID-19 pandemic highlight the need for a more harmonized or coordinated approach.[65]

In response to the demand for legal clarity around the use of location data as part of the COVID-19 pandemic response, the European Data Protection

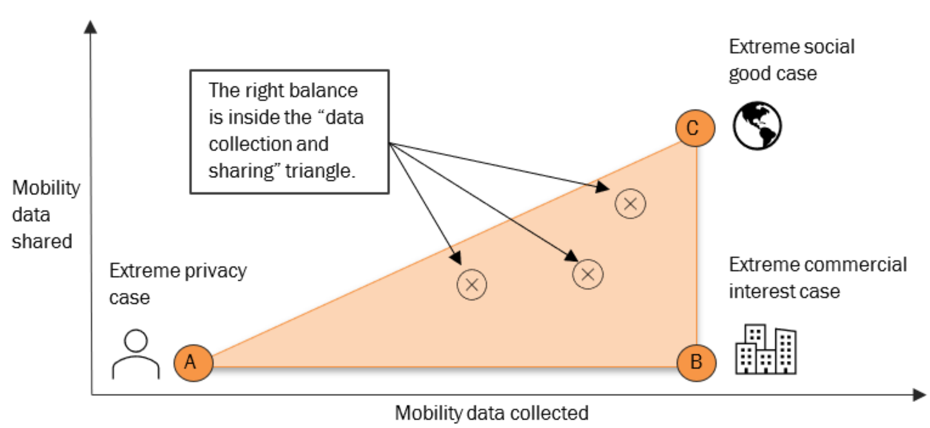
Board (EDPB) published guidelines in April 2020. While the guidelines provide principles for the effective, necessary, and proportional use of location data and contact tracing tools, the EDPB specifies they are only to "support the response to the pandemic by modelling the spread of the virus so as to assess the overall effectiveness of confinement measures [and contact tracing.]" [66] Firms wishing to share data for social good purposes beyond the COVID-19 pandemic have no equivalent guidelines.

THE GOVERNMENT'S ROLE SHOULD BE TO COORDINATE THE BEHAVIOR OF INDIVIDUALS, COMPANIES, AND RESEARCHERS TOWARD SOCIAL GOOD

To maximize social good from reusing privately collected mobility data, the interests and behavior of three different groups must be aligned: individuals whose data is being collected, companies that collect the data for commercial purposes, and researchers and the general public who benefit from accessing the data for social good purposes. But each agent in the data sharing partnership has their own unique priorities, which often compete with one another. This characterizes a non-cooperative environment—and in the absence of a central coordinating entity, can lead to suboptimal solutions for sharing, collecting, and using mobility data.

Figure 4 uses a model introduced by the Data Pop Alliance, a collaborative laboratory focused on big data issues, to delineate a "data collection and sharing" triangle that illustrates the range of possible data sharing solutions of mobility data for social good.[67]

Figure 4: The range of possible mobility data sharing options[68]



Position A defines the extreme individual privacy and security case and would result in no mobility data being collected or shared, which would inhibit the use of mobility data for social good. Position B maximizes the commercial and financial interests of the companies that collect and use mobility data and would result in the maximum amount of mobility data being collected with none of it shared, which would enable firms to offer personalized services to consumers but inhibit the use of data for social good. The final position, position C, maximizes data collection and sharing with regard to societal impact by collecting and sharing as much mobility data as possible, but due to the privacy and commercial rationales laid out earlier, would not be a sustainable position.

The optimal data sharing solution that maximizes the utility of all agents exists within this triangle but differs from case to case depending both on the type of data involved and contextual factors. For instance, since CDR data contains highly sensitive information, privacy and security considerations may have greater weight in a non-acute crisis context. However, during a public health crisis, the expected social benefits of sharing non-anonymized data would change the weight assigned to privacy considerations in favor of the collective good.

The role of government should be to coordinate the behavior of agents to optimize to a collective objective by setting policies and providing resources and incentives that induce them to act cooperatively. To this end, there are several steps the commission could take to pave the way for the responsible reuse of private sector mobility data for public good.

RECOMMENDATIONS

1. Revise the Data Governance Act to allow private sector firms to be listed in national registers of recognized "data altruism" organizations

The EU's DGA, which seeks to facilitate the collection, sharing, and use of data for social good purposes, excludes for-profit companies from its scope. As currently written, the legislation requires all "entities that seek to support purposes of general interest by making available relevant data based on data altruism" be nonprofit entities.[69] But private sector MNOs and app providers collect and hold much of the most valuable data on human mobility, and many seek to provide access to their data for social good. The commission should not assume that nonprofit organizations are necessarily altruistic or that for-profit organizations are not. It should revise the DGA to enable for-profit organizations to be listed in national registers of recognized data altruism organizations and thus encourage private companies to share data on human mobility for social good.

2. Pilot a common European data space for human mobility that researchers can use for social good

In its Communication on a European Strategy for Data, the European Commission outlines its plans to create an EMDS to be a common data pool for mobility data specifically designed to facilitate access to data that supports the development of intelligent transport systems for EU competitiveness, such as data on road safety conditions, infrastructure and maps, and road utilization. It will not be designed to support access to human mobility in order to promote broader types of research in the public interest. The commission should pilot a space specifically for data on human mobility and seek to incorporate a wide variety of data sources to help foster the development of novel research.

The commission has already seen success in such a data space with its development of a common union toolbox of technology and fully anonymized and aggregated mobility data shared by European MNOs to combat COVID-19. Within a few months, 17 MNOs covering 22 EU member states and Norway were transferring data to the commission every day, with an average latency of just a few days, and in most cases covering historical data from February 2020.[70] The commission could expand this effort to include mobility data from location-based apps and services and broaden the use cases to other social good causes.

3. Issue a Horizon Europe challenge to encourage researchers to use the data space for specific problems in the public interest

Over the past decade, a few telecommunications companies have established and provided funding for humanitarian challenges, inviting researchers to come up with innovative solutions in the public interest using the mobility they hold. Consider, for example, Telecom Italia's Big Data Challenge, Orange's Telecom Data for Development Challenge, or Turk Telekom's Data for Refugees Challenge.[71] But such challenges are few and far between. As part of Horizon Europe, the EU's funding program for research and innovation, the commission should issue a call for proposals to develop high-impact solutions to a specific set of social innovation challenges using the data from the data space for human mobility and provide at least €10 million in prize funding.

4. Proactively support experimentation of methods that use mobility data for social good

The use of new data sources for human mobility such as CDRs and data from location-based apps for social and disaster response hold great potential for social good, but their application is still experimental. To date, analysis to build sufficient understanding and trust, or to critically examine models that use this data, has lacked nuance, particularly in fragile and vulnerable contexts. The EU should establish and fund transparent experiments that develop and test mobility-data-based methods for social good and share insights with the research community. For example, the GIZ Data Lab, an outfit of the German development agency GIZ founded in 2019, conducts experiments to "promote the effective, fair, and responsible use of digital data for sustainable development."[72] Its most recent initiative called Data Powered Positive Deviance focuses on exploring the potential for administrative data, satellite imagery, urban data, social media data, and mobility data to identify positive deviants—defined as individuals or groups whose uncommon behaviors and strategies enable them to find better solutions to problems than their peers with the same resources—in a range of diverse contexts.[73] The EU could identify other data-based methods that use mobility data for social good and fund diverse pilot projects to test their efficacy and promote the use of those that work well.

5. Draft a common EU approach for the use of mobility data in emergency contexts to support future public health and humanitarian crisis exit strategies

To respond to the COVID-19 pandemic, many EU member states created new emergency regimes or amended existing emergency tools to enable the use of mobility data to assess social distancing measures and support contact-tracing efforts. However, because no EU-wide systematic emergency framework exists, some member states implemented hasty and ill-thought-out rules that fueled discrimination toward vulnerable populations. For instance, the Bulgarian government allowed authorities to deploy drones with thermal cameras to specifically track and measure the temperature of people in Roma communities, an ethic minority that has faced widespread discrimination in Europe.[74] Despite low rates of confirmed infections among Roma communities, several settlements were placed under strict lockdowns and restricted from entering other parts of society.[75] While the commission did issue recommendations for a pan-European approach to the use of mobility data for COVID-19 exit strategies in April 2020, it came two months after the World Health Organization (WHO) declared the virus a public health emergency of international concern, and a month after many member states had declared a state of emergency and begun adopting their own measures for contact tracing.[76]

Similarly, in conflict situations such as the Russian invasion of Ukraine, private companies have been left to make their own judgment calls on how and when to restrict access to publicly accessible data on human mobility that can support military intelligence in novel ways. For instance, shortly after U.S. researchers had used traffic data from Google Maps to infer a forthcoming Russian military invasion on the Ukrainian border, Google disabled live traffic data in order to protect Ukrainian civilians.[77]

The commission should use these lessons to pre-emptively draft emergency rules and guidance for the use of mobility data in future emergencies.

6. Pursue global partnerships in mobility data

Human mobility is not confined to Europe, so the EU should pursue international partnerships in mobility data in order to understand global mobility patterns. And as European researchers will need access to non-European data the same way non-European researchers will need access to European data, EU policymakers will need to ensure that global partnerships on mobility data do not contain data localization restrictions that prohibit transferring mobility data abroad.[78] Instead, EU policymakers should work to develop consensus among partners on issues such as proper safeguards to balance user privacy and public good for shared datasets and proper vetting of researchers and their projects.

REFERENCES

1. Matthew Luxmoore and Bojan Pancevski, "Refugees Fleeing Ukraine Now Represent Biggest Movement of People in Europe Since World War II," *The Wall Street Journal*, March 5, 2022, https://www.wsj.com/articles/refugees-fleeing-ukraine-now-represent-biggest-movement-of-people-in-europe-since-world-war-ii-11646493910.

2. "Situation Reports," CrisesReady, accessed March 8, 2022, https://www.crisisready.io/resources/situation-reports/.

3. Emmanuel Letouzé et al., "The Law, Politics, and Ethics of Cell Phone Data Analytics" (published as part of the Data-Pop Alliance White Paper Series. Data-Pop Alliance, World Bank Group, Harvard Humanitarian Initiative, MIT Media Lab and Overseas Development Institute, April 2015), http://datapopalliance.org/wp-content/uploads/2015/04/WPS_LawPoliticsEthicsCellPhoneDataAnalytics.pdf.

4. Shengjie Lai et al., "Assessing spread risk of Wuhan novel coronavirus within and beyond China, January-April 2020: a travel network-based modelling study," preprints from medRxiv and bioRxiv, https://doi.org/10.1101/2020.02.04.20020479.

5. Hodan Omaar, "5 Q's for Andrew Tatem, Director of WorldPop about Facebook's Data for Good Program" (Center for Data Innovation, April 2021), https://datainnovation.org/2021/04/5-qs-for-andrew-tatem-director-of-worldpop-about-facebooks-data-for-good-program/.

6. Hodan Omaar, "5 Q's for the Data Science and Mining team at École Polytechnique about Facebook's Data for Good Program" (Center for Data Innovation, February 2021), https://datainnovation.org/2021/02/5-qs-for-the-data-science-and-mining-team-at-ecole-polytechnique-about-facebooks-data-for-good-program/.

7. Alessandro Spelta et al., "After the lockdown: simulating mobility, public health and economic recovery scenarios," *Nature*, October 2020, https://www.nature.com/articles/s41598-020-73949-6.

8. Hodan Omaar, "5 Q's for Fabio Pammoli, Professor of Economics and Management at the Polytechnic University of Milan about Facebook's Data for Good Program" (Center for Data Innovation, December 2020), https://datainnovation.org/2020/12/5-qs-for-fabio-pammoli-professor-of-economics-and-management-at-the-polytechnic-university-of-milan/.

9. Paolo Valente, "Census taking in Europe: how are populations counted in 2010?" *Population & Societies*, no. 456 (2010), https://www.ined.fr/fichier/s_rubrique/19135/pesa467.en.pdf.

10. United Nations (UN), *Handbook on Measuring International Migration through Population Censuses* (New York, UN, 2020), https://unstats.un.org/unsd/demographic-social/Standards-and-Methods/files/Handbooks/international-migration/2020-Handbook-Migration-and-Censuses-E.pdf.

11. Eurostat, *EU legislation on the 2021 population and housing censuses explanatory guidance* (Luxembourg, Eurostat, 2019), https://ec.europa.eu/eurostat/documents/3859598/9670557/KS-GQ-

18-010-EN-N.pdf/c3df7fcb-f134-4398-94c8-4be0b7ec0494?t=1552653277000.

12. United Nations (UN), *Handbook on Measuring International Migration through Population Censuses* (New York, UN, 2020).

13. "National Travel Survey," last modified September 22, 2021, https://www.gov.uk/government/collections/national-travel-survey-statistics#about-the-national-travel-survey-data-and-reports.

14. Robert Griffiths et al., "Travel Surveys," *Transportation in the New Millennium* (2000), http://onlinepubs.trb.org/onlinepubs/millennium/00135.pdf.

15. Ibid.

16. Elizabeth Spencer et al., "Catalogue of Bias Collaboration," https://catalogofbias.org/biases/recall-bias/.

17. Aoife Ahern et al., "Analysis of National Travel Statistics in Europe 2013," *Publications Office of the European Union* (2013), DOI: 10.2788/59474.

18. "Monthly Traffic Statistics," IATA, accessed January 23, 2022, https://www.iata.org/en/publications/store/monthly-traffic-statistics/.

19. "Sabre Market Intelligence," Sabre, accessed January 23, 2022, https://www.sabre.com/products/market-intelligence/.

20. Zhuojie Huang et al., "An Open-Access Modeled Passenger Flow Matrix for the Global Air Network in 2010," *PLoS ONE 8(5): e64317* (May 2013), doi:10.1371/journal.pone.0064317.

21. Liang Mao et al., "Modeling monthly flows of global air travel passengers: An open access data resource," *Journal of Transport Geography* (September 2015), https://cyberleninka.org/article/n/1237443/viewer.

22. "Mozambique Malaria Facts," Severe Malaria Observatory website, accessed December 2021, https://www.severemalaria.org/countries/mozambique#9.

23. "World Malaria Day 2020: Vodafone coronavirus mapping," Vodafone website, April 24, 2020, https://www.vodafone.com/news/digital-society/world-malaria-day-2020-vodafone-fighting-malaria.

24. Jessica R Floyd et al, "Malaria parasite mobility in Mozambique estimated using mobile phone records" (presented at the 24th European Conference on Artificial Intelligence, ECAI 2020), https://ecai2020.eu/papers/1320_paper.pdf.

25. Carleen Maitland and Ying Xu, "A Social Informatics Analysis of Refugee Mobile Phone Use: A Case Study of Za'atari Syrian Refugee Camp," *TPRC 43: The 43rd Research Conference on Communication, Information and Internet Policy Paper* (April 2015), http://dx.doi.org/10.2139/ssrn.2588300.

26. International Organization for Migration (IOM) and Princeton School of Public and International Affairs, *Assessing the Use of Call Detail Records (CDR) for Monitoring Mobility and Displacement* (February 2021), https://www.migrationdataportal.org/resource/assessing-use-call-detail-records-cdr-monitoring-mobility-and-displacement.

27. Ziliang Zhao et al., “Understanding the bias of call detail records in human mobility research,” *International Journal of Geographical Information Science* (January 2016), DOI: 10.1080/13658816.2015.1137298.

28. IOM and Princeton School of Public and International Affairs, *Assessing the Use of Call Detail Records (CDR) for Monitoring Mobility and Displacement*.

29. Haosheng Huang et al., “Location based services: ongoing evolution and research agenda,” *Journal of Location Based Services* (August 2018), https://doi.org/10.1080/17489725.2018.1508763.

30. Testimony of Robert D. Atkinson before the U.S. Senate Committee on the Judiciary Subcommittee on Privacy, Technology and the Law, “The Location Privacy Protection Act of 2014” (ITIF, June 2014), https://www2.itif.org/2014-location-privacy-protection-act.pdf.

31. Huang et al., “Location based services: ongoing evolution and research agenda.”

32. Ibid.

33. Nick Warren Ruktanonchai et al., “Using Google Location History data to quantify fine-scale human mobility,” International Journal of Health Geographics 17(1) (July 2018), DOI:10.1186/s12942-018-0150-z.

34. Ibid.

35. “Disaster Maps,” Meta’s Data For Good website, accessed November 2021, https://dataforgood.fb.com/tools/disaster-maps/.

36. Eugenia Giraudy et al., “Measuring Long-Term Displacement using Facebook Data” (published as part of IDMC's Global Report on Internal Displacement, February 2021), https://research.fb.com/wp-content/uploads/2021/04/Measuring-Long-Term-Displacement-using-Facebook-Data-1.pdf.

37. Ibid.

38. Ibid.

39. “Disease Prevention Maps,” Meta’s Data For Good website, accessed November 2021, https://dataforgood.fb.com/tools/disease-prevention-maps/.

40. “Movement Range Maps,” Meta’s Data For Good website, accessed November 2021, https://dataforgood.fb.com/tools/movement-range-maps/.

41. Brian Dickinson et al., “Inferring Nighttime Satellite Imagery from Human Mobility,” preprint on arxiv (February 2020), https://arxiv.org/abs/2003.07691.

42. “Night lights in Europe,” European Space Agency website, accessed on January 13, 2022, https://www.esa.int/ESA_Multimedia/Images/2012/03/Night_lights_in_Europe.

43. Nita Bharti and Andrew Tatem, “Fluctuations in anthropogenic nighttime lights from satellite imagery for five cities in Niger and Nigeria,” *Nature*,

November 3, 2018, https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6233255/.

44. Seyit Höcük et al., “Economies of scope in the aggregation of health-related data,” JRC Digital Economy Working Paper (September 2021), https://joint-research-centre.ec.europa.eu/publications/economies-scope-aggregation-health-related-data_en.

45. Emmanuel Letouzé et al., “The Law, Politics, and Ethics of Cell Phone Data Analytics.”

46. Craig Timberg and Drew Harwell, “Government efforts to track virus through phone location data complicated by privacy concerns,” *The Washington Post*, March 19, 2020, https://www.washingtonpost.com/technology/2020/03/19/privacy-coronavirus-phone-data/.

47. Remco Takken, “EU member states loosen privacy rules for location data to contain COVID-19,” *Geospatial World*, March 27, 2020, https://www.geospatialworld.net/blogs/eu-member-states-covid-19/.

48. Bas Filippini, “Corona approach irresponsible,” *Privacy First*, March 12, 2020, https://www.privacyfirst.nl/over-ons/columns/item/1172-corona-aanpak-grote-schande.html.

49. European Commission, “Summary report of the public consultation on the European strategy for data” (July 2020), https://digital-strategy.ec.europa.eu/en/summary-report-public-consultation-european-strategy-data.

50. “Differential Privacy,” Harvard University, accessed February 2, 2022, https://privacytools.seas.harvard.edu/differential-privacy.

51. Siddhartha Banerjee et al., “Pricing in Ride-Sharing Platforms: A Queueing-Theoretic Approach,” *EC '15: Proceedings of the Sixteenth ACM Conference on Economics and Computation* (New York: Association for Computing Machinery, June 2015), 639, https://doi.org/10.1145/2764468.2764527.

52. The Open Data Institute (ODI), “The Use of Mobility Data for Responding to the COVID-19 Pandemic” (ODI, 2020), http://theodi.org/wp-content/uploads/2021/03/Data4COVID19_0318.pdf.

53. Regulation 2016/679 (General Data Protection Regulation), Article 6, https://gdpr-info.eu/art-6-gdpr/.

54. Regulation 2016/679 (General Data Protection Regulation), Article 23, https://gdpr-info.eu/art-23-gdpr/.

55. Directive 2002/58/EC (Directive on privacy and electronic communications), Article 15, https://eur-lex.europa.eu/legal-content/EN/ALL/?uri=CELEX%3A32002L0058.

56. Finland’s Secondary Use of Health and Social Data Act, https://stm.fi/en/secondary-use-of-health-and-social-data.

57. Jef Ausloos et al., “Operationalizing Research Access in Platform Governance: What to learn from other industries?” *Algorithm Watch* (June 2020), https://algorithmwatch.org/de/wp-

content/uploads/2020/06/GoverningPlatforms_IViR_study_June2020-AlgorithmWatch-2020-06-24.pdf.

58. Regulation 2929/767 (Data Governance Act), https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A52020PC0767.

59. "The Personal Genome Project," accessed March 2, 2022, https://www.personalgenomes.org/.

60. "Entur," accessed March 2, 2022, https://om.entur.no/reisende/hva-gjor-entur/.

61. Regulation 2929/767 (Data Governance Act).

62. European Commission, "Commission Staff Working document - Corporate Social Responsibility, Responsible Business Conduct, and Business and Human Rights: Overview of Progress" (March 2019), https://ec.europa.eu/docsroom/documents/34482.

63. Hodan Omaar, "No, The Data Economy Is Not A Barter Economy" (Center for Data Innovation, September 2021), https://datainnovation.org/2021/09/no-the-data-economy-is-not-a-barter-economy/.

64. "Mobility Data Space," accessed December 2, 2021, https://mobility-dataspace.eu/#c289.

65. Nuria Oliver et al., "Mobile phone data for informing public health actions across the COVID-19 pandemic life cycle," *Science Advances vol. 6 issue 23* (June 2020), https://advances.sciencemag.org/content/early/2020/04/27/sciadv.abc0764?versioned=true.

66. European Data Protection Board, *Guidelines 04/2020 on the use of location data and contact tracing tools in the context of the COVID-19 outbreak (*April 2020)*,* https://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_20200420_contact_tracing_covid_with_annex_en.pdf.

67. Emmanuel Letouzé et al., "The Law, Politics, and Ethics of Cell Phone Data Analytics."

68. Adapted from Emmanuel Letouzé et al., "The Law, Politics, and Ethics of Cell Phone Data Analytics," 16.

69. Regulation 2929/767 (Data Governance Act).

70. Stefano Maria Iacus et al., "Mobility functional areas and COVID-19 spread," *Transportation* (September 2021), https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8482960/.

71. "Data Collaboratives Explorer," accessed September 2021, https://datacollaboratives.org/explorer.html.

72. "GIZ Data Lab," accessed January 2022, https://www.giz.de/expertise/html/61847.html.

73. "Data Powered Positive Deviance," *GIZ Data Lab blog*, accessed December 29, 2021, https://www.blog-datalab.com/experiments/dppd/.

74. Krassen Nikolov, "Bulgaria extends 'epidemic emergency' and mulls tightening COVID-19 restrictions," *EURACTIV*, March 13, 2020,

https://www.euractiv.com/section/health-consumers/short_news/bulgaria-update-covid-19/?utm_source=EURACTIV.

75. Patrick Kingsley and Boryana Dzhambazova, “Europe’s Roma Already Faced Discrimination. The Pandemic Made It Worse,” *The New York Times*, July 6, 2020, https://www.nytimes.com/2020/07/06/world/europe/coronavirus-roma-bulgaria.html.

76. European Commission, “Coronavirus: Commission adopts Recommendation to support exit strategies through mobile data and apps,” press release, April 8, 2020, https://ec.europa.eu/commission/presscorner/detail/en/ip_20_626.

77. Aaron Gordon and Matthew Gault, “Google Maps Live Traffic Showed the Russian Invasion of Ukraine,” *Motherboard*, February 24, 2022, https://www.vice.com/en/article/xgd7dd/google-maps-live-traffic-showed-the-russian-invasion-of-ukraine.

78. Nigel Cory and Luke Dascoli, “How Barriers to Cross-Border Data Flows Are Spreading Globally, What They Cost, and How to Address Them” (ITIF, July 2021), https://itif.org/publications/2021/07/19/how-barriers-cross-border-data-flows-are-spreading-globally-what-they-cost.

ACKNOWLEDGEMENTS

This report was made possible in part by the generous support of Meta. The Center maintains complete editorial independence for all of its work. All opinions, findings, and recommendations are those of the Center and do not necessarily reflect the views of its supporters. Any errors and omissions are the author's alone.

ABOUT THE AUTHOR

Hodan Omaar is a senior policy analyst at the Center for Data Innovation. Previously, she worked as a senior consultant on technology and risk management in London and as a crypto-economist in Berlin. She has an MA in Economics and Mathematics from the University of Edinburgh.

ABOUT THE CENTER FOR DATA INNOVATION

The Center for Data Innovation is the leading global think tank studying the intersection of data, technology, and public policy. With staff in Washington, D.C., and Brussels, the Center formulates and promotes pragmatic public policies designed to maximize the benefits of data-driven innovation in the public and private sectors. It educates policymakers and the public about the opportunities and challenges associated with data, as well as technology trends such as open data, artificial intelligence, and the Internet of Things. The Center is part of the nonprofit, nonpartisan Information Technology and Innovation Foundation (ITIF).

contact: info@datainnovation.org

datainnovation.org