Guidance Note on Location Data

For the Working Group on Remote Supervision established by SD VP

THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP

Table of Contents

This Guidance Note was prepared for the Working group on Remote Supervision established by the Sustainable Development Vice Presidency (VP). The note has been prepared by Manuel Gonzalez-Schuler, Dunstan Matekeny, Trevor Monroe, and Nancy Lozano-Gracia, with inputs from Samuel P. Fraiberger, Jessica Grisanti-Bravo Nicholas Jones, and Manuel Tonneau, and under the overall guidance of Abhas K. Jha.

Introduction

"US government, tech industry discussing ways to use smartphone location data to combat coronavirus"

Washington Post's headline (March 18, 2020)

"Google and Apple are announcing a joint effort to enable the use of Bluetooth technology to help governments and health agencies reduce the spread of the virus"

Joint statement by the two tech companies (April 10, 2020)

After the World Health Organization (WHO) declared the COVID-19 pandemic on March 11, 2020, governments around the world imposed drastic mobility restrictions—including travel bans, border closures, and city lockdowns—as they rapidly mobilized medical, financial, scientific, and technological resources to control the spread of the disease. The world's most advanced economies and emerging markets—first in East Asia, then in Europe and North America—heavily relied on cutting-edge technologies as part of their strategy to combat the virus: "from artificial intelligence for medical diagnostics to mobile technology for data collection and contact tracing, technologies associated with the Fourth Industrial Revolution offer efficient and effective ways to cope with the speed, scope, and impact of the COVID-19 pandemic."[1]

Among the many sources of data being used to understand the impacts of the pandemic on many angles, location data stands out as a critical asset providing countries the capacity to track human mobility using cellular networks and Global Positioning System (GPS) technology through smartphones and other electronic

[1] Signé, Khagram, and Goldstein 2020.

devices.[2] This innovative approach has equipped several governments in Australia, China, Israel, Republic of Korea, Taiwan, and Western Europe,[3] to name a few, with the capacity to advance contact tracing, track individuals and populations at risk, isolate positive cases, and better manage the spread of the disease.

With the entry into force of **shelter-in-place orders**, **quarantine rules**, and **social distancing protocols**, life as it used to be has suddenly changed. The World Bank Group (WBG) also moved to remote work on March 13, 2020. This note is part of the efforts under the Sustainable Development Working Group on Remote Project Preparation and Supervision, which focuses on reviewing and proposing approaches to help task teams navigate the new challenges of remote work and adjust to the '**new normal**'. This includes a focus on how to leverage the offerings of new technologies (that is, satellite imagery, artificial intelligence, machine learning, location data, and unmanned aerial vehicles, among others) to improve WBG operations and collaborative practices, and ultimately serve borrowing countries in an effective and equitable manner.

The crisis has provided an incentive for tech and data companies and researchers alike to share information and analysis fast with the objective of informing policy decisions at an unprecedent pace. This openness to sharing information, spearheaded by the pandemic, is likely to change the way that we look at and access data. This provides opportunities for the World Bank teams to think about how they can use such information under the current crisis and beyond.

The purpose of this particular Guidance Note is to add to these efforts by focusing on providing information on what location data is, what data sources are available to gather it, and how these have been used to better understand and diagnose changes on the ground. While the note does not provide direct guidance on preparation and supervision, we hope that the information presented here can help task teams think about how they can leverage location data for their projects. We expect this note—and the data and methods discussed here—to be most useful in informing policy dialogue with clients and inform not only the preparation of new projects. While not directly discussed here, these data and methods could also be valuable as monitoring tools, and task teams could think about how to integrate them into their project monitoring and evaluation mechanisms.

The note is organized in four main sections: (a) it explains what location data is, how it works, and how it is collected; (b) it showcases concrete policy and international development examples—from within and outside the WBG—to illustrate how location data has been used to inform policy dialogue; (c) it provides some guidance for task teams that are considering the use of location data for their project, including information on who within the World Bank teams can help; and (d) it closes by providing an overview of the well-known location data owners/providers while discussing the accessibility of task teams to such data.

[2] World Bank 2020a.
[3] De Vynck 2020.

1. What is Location Data?

Some general definitions excerpted from desk research are provided in Box 1.

Box 1 Defining location data

"Any data processed in an electronic communications network or by an electronic communications service having the capability to indicate the geographical position of the terminal equipment of a user, in particular information regarding the latitude, longitude or altitude of the terminal equipment; the direction of travel of the user; or the time the location information was recorded." IGI Global [4]

"Information about the geographic positions of devices (such as smartphones, tablets, sensors, among others) or structures (such as buildings, landmarks, or attractions). The geographic positions of Location Data is defined by coordinates, and they are commonly expressed in terms of latitude and longitude. Additional attributes such as elevation or altitude may be included and helps data users get a more accurate picture of the geographic positions of the item in question." Quadrant.io[5]

Mobile phone data is probably the most popular type of location data. The term refers to digital footprints of human activities coming from cell phone devices whether they are basic feature phones (low-end devices) or sophisticated smartphones embedded with numerous sensors including GPS (high-end devices). Location data includes two attributes that are critical to carry out data analytics: geographic location and time stamp associated with that location. World Bank, Working Paper.[6]

All definitions of location data refer to three fundamental components: position, time, and movement. They imply three traits of location data: the capacity to determine the geographic position of an electronic device—using the geographic coordinate system; the capacity to track mobility and identify travel direction, including details of the destination and origin; and the capacity to record time and calculate the duration of a commute.

In the past decade, mobile devices (smartphones, tablets, and so on) have had tremendous influence on how we collect and analyze data as well as on what kind of data we have at our fingertips in many areas such as urban planning, agriculture, and environment monitoring. There are two main ways that we can use mobile devices in the space of data. First, we can opportunistically use passively generated data from both smartphones and feature phones. A good example of such opportunistic data includes GPS traces continuously logged by smartphones, which is useful for human mobility studies, among others. Second, mobile devices have also revolutionized data collection with platforms such as Mapillary, Open Data Kit (ODK), and KoBo ToolBox, which have simplified geographic information system (GIS) data capture in the field. Furthermore, the sensors in smartphones can enable us to collect environmental data in ways that were not possible before.

4 IGI Global is a leading international academic publisher committed to facilitating the discovery of pioneering research that enhances and expands the body of knowledge available to the research community (URL).
5 Quadrant Global, "All You Need to Know about Location Data", n.d. Accessed on 2020. Quadrant is a cutting-edge platform designed to allow data professionals to obtain and use high-quality location data and provide companies with location-based business solutions (URL).
6 World Bank 2020a.

1.1 How is location data produced and transmitted?

Contrary to popular belief, mobile devices do not produce or transmit location data by default. Location data is the result of signaling released by other devices. Indeed, mobile devices act as powerful receivers that listen and react to signals sent by other sensors. These signals mostly come from four well-known sources, which basically constitute the modern infrastructure of information and communication technology (ICT) systems that allow people to connect to the internet and communicate with each other: GPS satellites, WiFi networks, beacons (Bluetooth), and cell towers.[7]

Note: Modern smartphones include all these four technologies in one single device (that is, cellular data, WiFi, Bluetooth, and GPS), so it is common to blend multiple sources of location information (Box 2). For instance, if the GPS signal is weak, the smartphone can use the WiFi signal to detect its location (of course this is only possible in places with many WiFi access points). However, feature phones lack GPS sensors and the ability to connect to WiFi; their localization capabilities are strictly dependent on the cellular network infrastructure, whose location data is not accessible from the device itself but requires coordination with the telecommunications company (TelCo).

Figure 1 Modern ICT systems

GPS Satellites	Wi-fi Networks	Beacons (Bluetooth)	Cell Towers
GSP devices can determine positioning and navigation by sensing satellite signaling. They process the signals and calculate the coordinates of the device on earth. They also perform best outdoors with open skies.	are great at providing high-quality location signals particularly indoors. They are also great substitutes when cell towers or GPS are not accessible. They also enhance location accuracy in smartphones when enabled.	are small hardware devices able to sense other devices that come into proximity and transmit low signals through Bluetooth. Although their coverage is quite limited, they are accurate when calculating positioning.	provide the network infrastructure enabling phone calls and messages in mobile phones. Location data can be calculated by triangulation of cell towers when phones are actively used. Accuracy depends on the signal's strength and infrastructure capacity.

Box 2 This is a phone

A smartphone today is so much more than a phone. Yes, it is a telephone. But it is also a camera, a GPS, a survey instrument, a power supply sensor, a seismograph, a microphone, and a compass. All these features packed into this affordable smartphone allow to measure and monitor people's location and mobility. And it is cheap—this particular phone, targeted at consumers in the Latin America and the Caribbean region, is less than US$40 on Amazon.

Source: Compiled by the Development Data Group Analytics and Tools Unit (DECAT).

[7] Ewen 2019; Russo n.d. accessed on 2020.

1.2 How is location data collected?[8]

Surveys. These are traditional surveys like household, employment, or firm surveys carried out on the ground using electronic- or paper-based questionnaires. Information collected through traditional surveys can provide location data inputs such as addresses, zip codes, county, or city of residence of the survey respondents to mention a few.

Cell phone devices. This method mostly relies on communications technology and is conditioned to the territorial coverage of telephone and cellular network infrastructure. By tracing calls and messages (SMS) it can help determine the approximate geographic position of subscribers through triangulation using location identifiers such as mobile country codes, network codes, local area codes, cell IDs, and other cell characteristics. The output accuracy relies on usage, making cell phones undetectable when switched off.

GPS-enabled devices. GPS relies on satellite technology, and it is currently the most used positioning and navigation system in the word. It was developed by the United States military in the 1970s and then made fully available for civilian use in 2000.[9] It provides users with accurate and reliable positioning, navigation, and timing services. GPS technology can be embedded in a variety of portable devices such as smartphones and tablets (through operating systems or apps) and also as sensors that can be incorporated in vehicles, buildings, or any physical structure to track location and movement. Although users also have the capacity to turn on and off their GPS devices enabling/disabling their capacity to record a trajectory, automated data collection is possible through passive usage.

Note: There are two key elements distinguishing mobile- and GPS-based approaches from traditional survey-based approaches for data collection, which can make them more attractive and profitable: data accuracy and cost-effectiveness. For accuracy, we consider travel surveys as an example; people are asked to report on the trips they made, and which routes they took: this is error prone due to people's recall limitations. On the other hand, when using GPS-based approaches, which pick up this information automatically, this problem does not occur. These new approaches are also cost-effective because they do not require hiring of interviewers to collect the data as is often required in low resource regions.

Sensors and cameras. This type of technology is particularly used in transportation and mobility to collect traffic data through sensors and enforce transit rules. Speed sensors can also include cameras that record the vehicle's plate number and help identify drivers through machine learning. Street cameras are also used to inform urban safety interventions: in addition to recording the sequence of unlawful or criminal events, street cameras are effective at helping identify suspects through facial recognition. However, cameras can also provide key location data inputs as they register the name and attributes of the location where events took place (for example, site, street, or neighborhood). When information from multiple street cameras is combined, police can trace the direction, time, and displacement of the suspects across the city.

8 Ewen 2019 (URL); Russo n.d. accessed on 2020 (URL).
9 Russo n.d. accessed on 2020 (URL).

Box 3 Where is location data stored? What you need to know about sources.

Photo: peterhowell

There are different ways in which location data is gathered and managed.

Software Development Kits (SDKs) data. Collected by app developers, SDKs are one of the most reliable location data sources. Information is embedded in the smartphone apps, which use the core location services of a mobile device's operating system (for example, iOS or Android). As previously explained, the device reacts to signaling emitted by different sensors (for example, GPS, WiFi, Bluetooth, accelerometer, magnetometer, and barometer) to accurately determine location. Since information on geographic position is often generated from GPS, the output includes geographic coordinates (latitude and longitude) as well as time stamp. Although data accuracy and precision is high, the volume of data collected can be intermittent: the number of observations per device per day, for instance, varies a lot as it is subject to the user's consent (users not allowing apps to access GPS sensor due to privacy concerns) and activity (interactions with the app and the type of device). Furthermore, for people who do not have smartphones, as is the case with majority of the population in low-income countries, this approach may not provide any data or just provide an amount with a limited coverage.

Carrier data. Gathered by TelCos and phone operators, carrier data is generated by interpreting cellular signals using a process called trilateration. As explained earlier, signals are produced by phone calls and messages made/received by subscribers. Owing to the installed capacity of cellular network, the amount of data gathered is high. However, TelCos usually operate at a national scale, limiting the volume subscribers' datasets to the residents of a particular country. In addition to location data, TelCos own demographic and other personal information from their subscribers. Unlike in the SDK data presented above, where we essentially get the user's exact location based on the GPS signal, in most of the TelCo-based location data such as call detail records (CDRs), the accuracy of the location is low because we can only identify the user's position up to the cellular tower level. In densely populated urban areas where there are more cellular towers, precision of the location can vary from 100 m to 300 m, while in rural areas with few people, the precision can be as low as 3 km.

Bid-stream data. Bid-stream data is collected from advertising servers managed by publishers. Although gathering bid-stream data is quite similar to SDK-based data collection, the two concepts are technically different. When an advertisement is requested and delivered to an electronic device—either through the app in a cell phone or a website on a laptop—there is the option of providing the location of the device in that bid request. Bid requests are actually one of the easiest ways to collect location data. However, the quality might not be the best. Bid-stream data has widely varying levels of accuracy and precision. Another challenge with bid-stream data is that it is intermittent, and some devices have many observations per day, while others are observed only once or twice a month.

2. What Can You Do with Location Data?

Location data can be a powerful tool to diagnose development challenges and inform policy decisions. When used strategically and creatively, it can equip practitioners and policy makers with reliable inputs to intervene in a variety of sectors. The sections that follow reflect on different types of location data interventions in the fields of government and development while bringing up a series of cases to better illustrate what can be done with location data.

2.1 How has location data informed government and development interventions?

The data generated when people use mobile devices has important uses in urban planning, environment monitoring, agriculture, water and sanitation, and disaster risk management, among other sectors. For instance, in urban planning, we need to evaluate a broad range of factors, such as traffic flow and human mobility to formulate effective plans. Let us consider human mobility that is essential in urban planning: unlike using travel surveys to get this kind of data for urban planning, which is expensive to produce and update, mobility data from mobile phones can be easily updated and acquired at a relatively lower cost. The ubiquitous location data available from mobile devices is not only useful for urban planning but it also opens up many opportunities as geographic location presents an avenue to link together other data sources.

In environmental monitoring, one of the important tasks is tracking carbon emissions and transportation has often been pointed as one of the major sources of emissions. In this regard, being able to get people's mobility and commuting patterns enables us to better monitor the environment. In addition, we can use the plethora of sensors that come with smartphones nowadays to perform basic environment monitoring. For example, it is becoming increasingly common for smartphones to come with temperature and humidity sensors, which play an important role in measuring air quality.[10]

[10] Aram, Troiano, and Pasero 2012; Nemati, Batteate, and Jerrett 2017.

In agriculture, smartphones have also become useful. In their paper, Pongnumkul, Chaovalit, and Surasvadi (2015) provide a detailed review of how smartphones have been used in agriculture. They identify four thematic areas: farming, farm management, information system, and extension services. For instance, in farming, which can involve a range of activities from sowing to weed control to fertilizer application, one of the areas where smartphones have been used is disease detection.[11] Another paper from Chouhan, Singh, and Jain (2020) provides a review of computer vision methods employed in smartphones to detect disease[12], while others provide examples of use cases.[13] Even further, this YouTube video provides an appealing example of the use of smartphones in disease detection of cassava in Africa.

In water, sanitation, and hygiene (WASH), mobile phone technologies and location data have strengthened the capabilities of local governments to improve data collection for strategic decision-making, monitor the quality and quantity of water resources, refine forecasting, and support monitoring and evaluation efforts.[14] Research in Sub-Saharan Africa, for instance, suggests that using mobile phone technology to transmit water quality data is positively associated with an increase in effectiveness and sustainability of testing and information flow.[15] Moreover, the WASH sector can benefit from the offerings of mobile technology

[11] Pongnumkul, Chaovalit, and Surasvadi 2015.
[12] Chouhan, Singh, and Jain 2020.
[13] Prasad, Peddoju, and Ghosh 2014; Wu and Chang 2013.
[14] USAID Center for Water Security, Sanitation, and Hygiene 2019.
[15] Kumpel et al. 2015.

through smart metering (to better monitor and maintain utility systems), mobile services (to provide customer support and allow users to report service delivery statuses), and mobile payments (completed remotely and securely). Precise location is a crucial attribute here as it provides utility companies with the right information to assess and monitor the system's performance and also reassures customers that they will be billed for what they actually consume. [This video](#) offers a good snapshot of how mobile technologies can facilitate access to water and sanitation, improve service quality, and ensure delivery in developing countries.

Disaster risk management is probably one of the fields that has benefited most from technological development and greater data availability. When a disaster occurs, good data can help local authorities respond more effectively to the emergency. For example, location data can be quickly processed to provide first responders with the exact position of victims of a particular disaster. This allows emergency teams to prioritize rescue operations, better plan for action, and make informed decisions on the urgency of cases. Precise location technologies have been used to warn citizens on the spread of fires (for example, California wildfires in 2017) and determine whether a person shall be rescued by land, sea, or air (for example, Hurricane Harvey in 2017). Similarly, social media platforms like Facebook play a key role in providing users with the ability to mark themselves as 'safe', share their exact whereabouts, or publish something about their current status in real time. This output is a valuable asset for both emergency teams and worried friends and family members.[16] Moreover, location data can be used by government agencies to issue early warnings or send large-scale messages to reach targeted populations. Twitter is probably the signature platform which disseminates instant news and live notifications in a matter of minutes when tragic events happen (for example, the #AustraliaBushFires in 2019, the #MexicoEarthquake in 2017, or the #ParisAttacks in 2015). Both public agencies and private organizations can leverage the power of connectivity to keep people well-informed, give clear instructions, encourage positive behavior, and eventually manage a crisis.

[16] Elichai 2018.

Photo: Krakow, Poland. Agnes Kantaruk / Shutterstock.com

Reaching the most vulnerable - social development. With shelter-in-place orders and city lockdowns, violence against women has skyrocketed.[17] With the need for urgent and innovative solutions, location data has become a remarkable asset to prevent and/or report sexual harassment and domestic violence through apps. [In Montenegro](), a multi-stakeholder partnership launched 'BeSafe', a mobile app to reach out to women exposed to violence at home. In less than 10 days they had more than 4,000 downloads and were able to determine that most of their users were located in the northern rural region of the country.[18] [In India](), women are turning to smartphone apps to report harassment in public spaces. The data processed helps the police map 'harassment prone' spots in precise city areas.[19] Moreover, organizations like [Safetipin]() work with municipal governments to map cities and provide recommendations for urban upgrading, with the aim of creating safer and more inclusive public spaces for women.

Finally, there are other sectors, beyond sustainable development, which have benefited from the advantages of location data. Although they are not within the scope of work of this Guidance Note, it is worth mentioning some of them as they may provide opportunities for cross-fertilization between sectors:

- **Transportation.** Location data is used to analyze how people use public and private transportation systems; understand mobility patterns; identify congested routes and unserved city areas; assess systems' performance; and inform policy makers' decisions to provide sufficient, high quality, and appropriate mobility services to their citizens.

- **Military and national defense.** National authorities use location data to trace and locate terrorist groups or criminals, perform intelligence plans, support and coordinate foreign operation on the ground, and contain threatens national security.

[17] UN Women 2020.
[18] Gasparikova and Djurickovic 2020.
[19] Nagaraj 2018.

- **Migration and humanitarian aid.** Location data is a good tool to better diagnose, understand, and intervene in situations of massive displacement (for example, exodus or refugee crisis). In addition to quantifying and identifying population groups affected, location data will equip governments in recipient countries with reliable information to mitigate the effects of a humanitarian crisis.

- **Social protection.** While we found no evidence of the use of location data collected from cash transfer programs, the availability of mobile applications for these kinds of programs poses an opportunity to gather additional location information (while ensuring privacy is preserved). Location data is a powerful resource as it allows policy makers to know where the poorer and more vulnerable groups are settled, so they can better target program beneficiaries and ensure effective delivery of aid.

2.2 How have companies, governments, and academia used location data to inform policy and action?

Examples that follow include location data as a fundamental piece in the analyses and interventions described. Most of them are excerpted from academic research, publications, and policy programs carried out in the recent months. They are sorted in four different sections according to the type of immediate purpose they pursue. Further studies and categories will be added once more information becomes available.

(a) Location data to track the spread of a disease like COVID-19 or malaria

Location data is being actively used during the COVID-19 outbreak to track people's mobility through their personal mobile devices: "It is being used to track the transmission chains of confirmed cases or proactive alerts for people to get tested at a nearby facility."[20]

- Sveta Milusheva (2020): **Monitoring travelers to estimate and control the spread of a disease.** Using malaria as a case study and 15 billion mobile phone records across 9 million SIM cards, this paper causally quantifies the relationship between travel and the spread of disease. The estimates indicate that an infected traveler contributes to 1.7 additional cases reported in the health facility at the traveler's destination. It simulated the use of a strategic targeting policy tool to monitor travelers and concluded that targeting informed by mobile phone data could reduce the caseload by 50 percent more than other policy strategies.[21]

(b) Location data to understand human behavior, reactions, and perceptions

- Maloney and Taskin (2020): **Understanding people's reaction to the COVID-19 outbreak.** This paper uses Google mobility data to explore which factors are proving important during the outbreak in the United States and globally. Findings indicate that in the United States much of the decrease in mobility is voluntary, driven by the number of cases reported and proxying for greater awareness of risk. This suggests that much social distancing will happen regardless of the presence of non-pharmaceutical interventions. In poorer countries, evidence shows a limited effect of non-pharmaceutical interventions and no voluntary component, consistent with resistance to abandon sources of livelihood. It also confirms that economic activity is mostly driven (and affected) by the voluntary component, implying that economic recovery will not happen just with the lifting of restrictions.[22]

- Selod and Soumahoro (2020): **Estimating the impact of lockdowns on traffic congestion in major developing world cities.** This paper discusses big data literature applied to urban transportation; provides a taxonomy of big data sources relevant to transportation analyses; and describes how locational data—among others—can be used to measure mobility, externalities, and welfare impacts. Specifically, it aims to estimate the causal impact of stay-at-home orders during the COVID-19 pandemic on traffic congestion in different developing countries cities, particularly Bogota, New Dehli,

[20] Sanchez-Andrade Nuno and Krambeck 2020.

[21] Milusheva 2020.

[22] Maloney and Taskin 2020.

New York, and Paris. They find that lockdown policies drastically reduced daily traffic congestion—as measured by a traffic congestion index—in each of the four selected cities that implemented strict stay-at-home orders during the pandemic.[23]

- Aral et al. (2020): **Deconstructing people's behavior toward mobility restrictions in the United States.** Researchers at the Massachusetts Institute of Technology (MIT) combined information on shelter-in-place and business closure policies (state orders) with mobility trends and social networking data (Facebook) and found that people's behavior toward social distancing and mobility regulations, for example, were not just affected by local restrictions in their community or neighboring states but were also influenced by perceptions from their social network connections. That said, obedience (or violation) of measures taken in a particular state may have an impact on the behaviors of residents of other states, regardless of how proximate or distant they are (for example, New York and Florida). Such findings underline the urgency of a coordinated regional response to the crisis.[24]

(c) Location data to inform investment choices, assess policy effectiveness, and improve service delivery

- Oliver et al. (2020): **Tracing contact and assessing policy effectiveness.** Location data can be used at different stages of the COVID-19 outbreak to meet different goals: (a) in the initial stages, location data provides information that is used for contact tracing, as it allows to track the location and social contact of positive cases; (b) in the rising curve phase it can help assess the efficacy of and compliance toward policies implemented (for example, social distancing protocols or stay-at-home orders) by observing mobility between and within affected municipalities; and (c) at the peak and lowering curve phases, location data could contribute to improve the understanding of how lifting and reestablishing certain measures could translate into behavioral changes and, through that, be used to feed such changes into epidemiological, economic, and policy models. This last piece would allow governments to observe how people react to the loosening of mobility restrictions and the activation of economic recovery measures allowing on-the-fly adjustments to policy decisions and ultimately contributing to finding the optimal policy mix.[25]

- USAID and Global Waters (2018): **Building capacity to increase water systems' functionality in Ethiopia.** Throughout a public-private partnership, 107 sensors were installed in the Afar region to conduct real-time data monitoring of water schemes' operations. Sensors transmitted a daily data flow signal through cell phones or satellites to a central cloud-based remote dashboard managed by the Regional Water Bureau. This technology allowed the bureau to quickly identify where functionality problems were, then plan and prioritize their maintenance intervention, which in turn was tracked through a ticket system. Besides equipping public officials with an effective information management and accountability tool, this technology could also serve as an appealing advocacy instrument for the regional government to address additional financial, equipment, and manpower needs. Lessons learned from this pilot program were shared in a United States Agency for International Development (USAID)-United Nations Children's Fund (UNICEF) initiative in the Somali region to install similar sensor and monitoring systems.[26]

- Hamsteada et al. (2018): **Enhancing parks, transforming streets, and rethinking open spaces.** This paper used geolocated Flickr and Twitter data as a rapid indicator to estimate park visitation in New York City. It analyzed variations in use across 2,143 parks and modeled visitation based on park attributes and amenities, as well as on accessibility and demographic features at the neighborhood level. The authors found that social media activity in parks was positively associated with proximity to public transportation and bike routes, as well with

[23] Selod and Soumahoro 2020.
[24] Aral et al. 2020.
[25] Oliver et al. 2020.
[26] USAID Water Team 2018.

particular park characteristics (for example, water bodies, athletic facilities, and impervious surfaces). They also observed that it was negatively correlated with green space in neighborhoods where minority groups are concentrated. They concluded that parks in high-minority neighborhoods are of lower quality and less accessible compared to parks in low-minority neighborhoods.[27]

The COVID-19 context is a good opportunity to fix this inequality problem. To keep social life and community spirit alive, cities are advancing a transformative urban agenda to rethink and adapt public spaces. This includes showcasing new parks, waterfronts, and plazas where neglected lots or degraded green areas remain; constructing new protected bikeways where car lanes are underutilized or unused; rehabilitating and expanding sidewalks to minimize pedestrians' crowding; and encouraging outdoor dining by authorizing restaurants' use of further public space.[28] As the Italian Mayor put it, the global pandemic seems to be an opportunity "to make our public spaces more functional and more beautiful than they were before the lockdown."[29]

(d) Location data to better target program beneficiaries and vulnerable population groups

- [Aiken et al. (2020):] **Targeting ultra-poor households in Afghanistan.** By combining rich survey data from a poverty alleviation program in Afghanistan with detailed mobile phone logs from program beneficiaries, the authors studied the extent to which machine learning methods could accurately differentiate ultra-poor households eligible for program benefits from other poor households deemed ineligible. Indeed, supervised learning methods leveraging mobile phone data can identify ultra-poor households as accurately as standard survey-based measures of poverty, including consumption and wealth. Moreover, combining survey-based measures with mobile phone data produces classifications more accurate than those based on a single data source.[30]

[27] Hamstead et al. 2018.
[28] Hu and Haag 2020.
[29] Bello 2020.
[30] Aiken at al. 2020.

- **Solidarity Income Program in Colombia.** The National Government of Colombia launched the unconditional monetary transfer program *Ingreso Solidario* in mid-April to support the poorer and most vulnerable households in the country who were not registered in any social assistance program led by the government (that is, *Familias en Acción, Colombia Mayor, Jóvenes en Acción, and Devolución del IVA*). While this program has not used location data of individuals, it has leveraged cell phones to transmit and collect information of the most vulnerable and ensure that they are reached by support programs. In the first stage, the Colombian Government identified program beneficiaries through existing databases (Sistema de Identificación de Potenciales Beneficiarios de Programas Sociales, SISBEN) managed by social development agencies (President's Department for Social Prosperity) and ministries (Labor, Health, and Finance) and reached out to them through SMS to share instructions on how to receive financial support. In the second stage, they assisted families who did not have a bank account and worked closely with financial institutions to carry out a digital banking process for the most vulnerable. Once again, communications were managed using the users' mobile devices. A third phase of the program is currently going on to include further households, in specific municipalities, who have not received the unconditional cash transfer yet.[31]

- A similar program, Renta Basica Bogota (formerly called *Bogota Solidaria en Casa*) is led and carried out by the city government of Bogota to support low-income families living in the most vulnerable urban areas. The program is funded by taxes and donations from citizens and big companies. Direct beneficiaries received groceries, monetary transfers, and other type of social and financial assistance. No sign-up was needed as communications were managed directly with mobile phone subscribers through SMS.[32]

[31] Prosperidad Social – DPS 2020.
[32] Alcaldía Mayor de Bogota 2020.

2.3 How has the World Bank used location data? Case studies from within

Case studies that follow are excerpted from the World Bank's knowledge and lending portfolios. Further studies will be added once more information becomes available from task teams and other WBG departments.

Case #1

Supporting Farmers and Low-income Rural Families in Belize
Country: Belize
Sector: Agriculture and Food

In the framework of the Climate Resilience Infrastructure Project (P127338) the Government of Belize through its Ministry of Agriculture enabled an Emergency Action Plan to support low-income farmer families affected by COVID-19. They collected farmers' data using iPads and a web-based application, which provided information on farmers' location, household characteristics, and individual attributes. The data collected allowed the Government of Belize to better define beneficiaries' criteria and target financial aid recipients. The task teams responsible for this collaboration also used (or are planning to use) drones as well as WhatsApp and cell phone data to perform monitoring and evaluation.

Source: Internal brown bag lunch June 2020.

Case #2

Improving Transit Infrastructure and Commuting Time and in Abidjan
Country: Côte d'Ivoire
Sector: Transport and Mobility

"Mobile phone location data is used to infer origin and destination flows in the city, which are then converted to ridership on the existing transit network. Sequential travel patterns from individual call location data is used to propose new candidate transit routes. An optimization model evaluates how to improve the existing transit network to increase ridership and user satisfaction, both in terms of travel and wait time" (IBM AllAboard 2013). According to United Nations Global Pulse (2013), the research work resulted in a partial solution to the commuting problem: four new bus routes were added to Abidjan's commuting infrastructure, and one route was extended. This should reduce traveling time by 10 percent. United Nations Global Pulse sees such a provisioning use of CDRs as "useful for better urban planning and public transportation."

Source: Haddad et al. 2014.

Case #3

Designing New Bus Routes in Seoul
Country: Republic of Korea
Sector: Transport and Mobility

Using citizens' late night calls and text messages, the city government was able to inform the design of new bus routes in Seoul. It turns out that bus routes were not frequent at night and the Metro was closed, so people decided to use illegal cabs to get to their destinations (home, night clubs, and so on). In consequence, the city decided to offer a new route—the Night Owl—but it did not have information on origins or destinations. They thus partnered with the Korea Telecom Corporation, which provided the Seoul Metropolitan Government with 3 billion call and text records for free. The government used the data to analyze where people were traveling after midnight and was able to choose routes for its new night bus service based on the most heavily trafficked areas. Seoul's Night Owl service reduced annual car trips by 2.3 million, with 7,000 people using the buses nightly.

Source: Internal brown bag lunch n.d.

Case #4

Cleaning Streets Initiative in Los Angeles
Country: United States
Sector: Public Works/Public Health

The Los Angeles Mayor's Office created a 'cleanliness' score to assess the state of city streets. The index is calculated based on four factors: litter, weeds, bulky items, and illegal dumping. The streets are thus given a score and classified in to four categories (that is, clean, somewhat clean, not clean, and encampment presence). The tool has helped the Los Angeles Sanitation and Environment Department identify where the most critical street cases are, prioritize public intervention, and allocate public resources more efficiently. This prioritization has helped the city reduce the number of unclean streets by 82 percent just one year after the process started. Relying on the score has thus allowed the city government to move from a reactive complaint-based system to a more proactive data-driven system, saving the city time and money.

Source: CleanStat 2020.

Case #5

Planning for Water Management in the Netherlands
Country: The Netherlands
Sector: Water

This is what the Netherlands is doing Digital Delta Project. The Netherlands' Government has partnered with the private sector to harness information coming from different sources, including water sensors. It uploads the data into an intelligent dashboard, which is shared with other city agencies, so that everyone has access to real-time data to better plan water management. The costs of water management, including anticipating flooding, droughts, and low water levels, added up to €7 billion per year. With the introduction of the Digital Delta Project, water authorities were able to reduce the cost of managing water by up to 15 percent while getting better prepared to prevent natural disasters and environmental degradation.

Source: Internal brown bag lunch n.d.

Case #6

Uneven Reactions to Lockdown Policies in Jakarta
Country: Indonesia
Sector: Human Mobility and COVID-19

Following Indonesia's nationwide stay-at-home order on March 15, 2020, mobility in Jakarta began to slow down significantly. When a state of emergency was declared on March 23, 2020, it led to an even sharper decline in activity. Using privacy-protected GPS location data, the World Bank analyzed location data from over 80 mobile apps and more than 275,000 app users who consented to sharing their location. They found that, on average, users stayed home 2.5 hours longer than usual, increasing from 13.0 to 15.5 hours per day. When looking at individual wealth categories, they found that low-wealth users only stayed home 11 percent longer than before the stay-at-home order, whereas high-wealth users did the same for 20 percent longer. Likewise, economically advantaged residents reduced the number of neighborhoods they visited by 25 percent, 8 percentage points higher than what was observed for lower-income people (17 percent). What these observations reveal is that lockdown policies affect citizens differently. Vulnerable groups find it more challenging to comply with social distancing measures for a wide variety of reasons, including limited household savings, a weak or nonexistent social safety net, incomes that depend on face-to-face contact, crowded living arrangements, and poor access to basic services.

Source: Fraiberger, Jones, and Lozano 2020.

3. How to Use Location Data Resources in Your Project?

Short guidelines for task team leaders and team members interested in leveraging location data for their projects and operational processes.

3.1 Initial steps

Figure 2 Steps to include location data into your work

Do you really need Location Data?	*What will you do with Location Data?*	*Which Location Data features are you looking for?*	*How will you collect the data you need?*	*Who owns the Location Data you need?*
1	2	3	4	5
Establish wether or not your project requires some sort of Location Data to develop analytical products and/ or complete preparation activities.	Define the object(s) and action(s) that you seek to monitor or measure through Location Data.	Define the scope and type of metrics needed/ desired to carry out your project.	Determine the best Location Data collection mechanisms available.	Identifiy key Location Data owners and providers.

3.2 Two pathways: open data or private sector data

Pathway 1: Public/Open Data	Pathway 2: Private Sector Data
Description. This includes location data that different companies and organizations have made publicly available online (for example, Google Mobility Reports, Apple Mobility Trends Tool [MTT], Facebook's Data for Good Public Datasets, and so on). For the World Bank teams, access to many of these data sources can take place under the existing partnerships built by the data collaboratives.	**Description.** This is location data owned and managed by private companies that is not available online. Yet, it might be accessed or obtained through public-private partnerships, like the Development Data Partnership[a], by meeting specific conditions and a certain cost.
Cost. Overall costs are hard to estimate because they depend on the scope of analysis. Through this route, data comes at no cost to teams. The costs will be limited to staff time and short-term consultant (STC) time for analysis and data manipulation. The teams may need to consider data processing costs. (An example of recent work includes a contract for data analysis with a firm for USD30,000 to process data, create the needed libraries, and provide analysis of mobility patters before and after an earthquake). Similar work could be done by a series of individual consultants with the right skills (for more information on costs of this work, please contact Nick Jones, Sam Fraiberger, or Nancy Lozano).	**Cost.** As mentioned in the case of open data, the costs are hard to estimate because they depend on the scope of analysis. The main difference is that in this case, data costs should be added to total costs. The costs of staff time and STC time for data management, manipulation, and analysis should be included. The teams should consider data processing costs (for example, amazon cloud).
Risk assessment ought to be performed (low/intermediate/high) in light of final product definition, output characteristics, and contextual variables, based on, but not limited to, the following criteria: ▪ **Technical risk.** Data availability and level of detail ▪ **Organizational risk.** Inexperienced staff ▪ **Operational risk.** Alignment with project goals and specific objectives ▪ **Legal risk.** Data property and ownership ▪ **Security risk.** Data privacy and protection ▪ **Business risk.** Sharing data and findings with third parties ▪ **Reputational risk.** Loss of trust in the WBG and partners.	**Risk assessment** ought to be performed (low/intermediate/high) in light of final product definition, output characteristics, and contextual variables, based on, but not limited to, the following criteria: ▪ **Technical risk.** Data availability and level of detail ▪ **Organizational risk.** Inexperienced staff ▪ **Operational risk.** Alignment with project goals and specific objectives ▪ **Legal risk.** Data property and ownership ▪ **Security risk.** Data privacy and protection ▪ **Business risk.** Sharing data and findings with third parties ▪ **Reputational risk.** Loss of trust in the WBG and partners.

Note: a. The current list of data partners include Carto, ClimaCell, Cuebiq, European Space Agency -ESA, Environmental Systems Research Institute -ESRI, Facebook, Google, Grab, Global System for Mobile Communications (GSMA), Indigo, InLoco, LinkedIn, Mapbox, Mapillary, Mobike, Orbital Insight, Premise, SafeGraph, Twitter, Uber, Unacast, Veraset, Waze, WhereIsMyTransport, and X-Mode.

3.3 Whom to call?

Big Data (DECAT) Team: Contact information and capabilities	
Aivin Solatorio	Location/mobility (GPS) Agent-based modelling Natural Language Processing Data engineering High performance computing
Dunstan Matekenya	Location/mobility (CDR/GPS) AI/machine learning Geospatial data science High performance computing
Samuel Paul Fraiberger	Location/mobility (CDR/GPS) AI/machine learning Network data science NLP/text
Brian Blankespoor Benjamin P. Stewart Katie McWilliams	Geospatial data science
STCs - big data team can help identify	Location/mobility (CDR/GPS) High resolution SAE Machine learning Natural language processing Computational social science Data engineering High performance computing
Other staff/teams that may be able to help	
Nicholas Jones	GFDRR, recently worked with location data from Facebook and Cuebiq
Sveta Milusheva	DIME, experience working with CDRs
Holly Krambeck	Location/mobility data for transport and key point person for Data Collaboratives initiative and SD Data Lab
Sarah Antos	Geospatial data science and land

Note: AI = Artificial Intelligence; DIME = Development Impact Evaluation; GFDRR = Global Facility for Disaster Reduction and Recovery; NLP = Natural Language Processing.

If we have missed you and your name should be here, please contact nlozano@worldbank.org to be added.

Box 4 **The Big Data Program will provide data products and services as well as capacity-building activities to support operations through developing capabilities in the following:**

- Mobile analytics using spatial data science to analyze location analytics to deliver new products and services
- Data integration and fusion using methods to carefully combine traditional (census/survey) and alternative data sources (social, satellite, text, and mobile) to produce new data sets and now-casting data products
- Capacity development in data science tools and methods to help clients use and sustain modern data products in operations and country systems (open code, capacity, and communities)
- Privacy-preserving methods differential privacy, encryption, and homographic techniques to preserve privacy.

Some examples of projects that have been supported by DECAT are as follows:

- Guatemala: Mobile poverty mapping (Hernandez et al.)
- Sierra Leone: Working paper in progress. Contact point: Dunstan Matekenya
- Haiti: Urbanization review, presentation on mobile. Contact point: Nancy Lozano; Andrew Whitby
- How good are CDR-based economic measures, project summary (P153635). Contact point: Tariq Khokhar; Andrew Whitby
- Senegal (mostly) + Colombia: OPen ALgorithms (OPAL). Contact point: Sam Fraiberger
- Colombia mobility assessment. Contact point: Felipe Targa; Transport; Dunstan Matekenya

Source: Excerpted from DECTA. 2020. Innovations in Big Data Analytics - FY20 Results Report. Note: DG stands for Data Group.

3.4 Some useful external links with resources

- Flowminder Flowkit supports mobile phone data access, management, and analysis
- University of Tokyo CDR Analysis Toolkit
- Bandicoot open source python toolkit to analyze mobile data
- scikit-mobility - mobility analysis in Python
- OpenCellID cell tower open data
- International Telecommunication Union handbook on Mobile Phone Data for Official Statistics
- International Telecommunication Union methodology guide on big data for official statistics
- OPAL on reconciling mobile open data and security
- GSMA COVID-19 privacy guidelines
- GSMA Mobile Data and Big Data Analytics privacy guidelines
- United Nations Global Pulse - Big Data for Social Good
- United Nations Handbook on privacy preservation techniques
- United Nations Big Data working group conference on mobile

4. Well-known Providers of Location Data and Where to Find Them

Major location data owners/providers can be classified in two different groups: big tech companies and third-party firms. TelCos can be a third key player owning location data; however, they are not covered in detail for the guide. The sections that follow provide an overview about these stakeholders and the role they play in the location data ecosystem.

4.1 The 'big five' tech giants

Google (over 1.5 billion active Gmail users in 2020)[33] is an American technology company specialized in providing internet services and products. It is the most popular search engine worldwide and offers online advertising technologies as well as cloud computing, software, and hardware. Google accesses location data through both its operating system (Android) and the array of apps that the company owns that include GPS technology (for example, Google Maps).

Google uses aggregated and anonymized data from Google Maps. This year, it created **Community Mobility Reports** (CMRs) to show how people's mobility patterns have changed over time in response to policy restrictions that entered into force with the pandemic outbreak. The CMRs exhibit the percentage change in visits to particular places such as retail stores, pharmacies and grocery stores, parks, transit stations, workplaces, and residential places within a particular geographic area. The CMRs are produced for world countries and US regions and can be downloaded as PDF or CSV files.

Apple (1.5 billion active devices in 2020)[34] is an American technology company that is famous for developing software (iOS), personal computers (MacBook and iMac), and innovative technology products (iPhone, iPad, and Apple Watch) and offering online services (Apple TV). Apple also gathers location data through its operating system's functionalities as well as through its tailor-made apps that include GPS technology (for example, Apple Maps).

Like Google, Apple created a **Mobility Trends Tool** (MTT) using data from Apple Maps. The MTT shows the daily percentage change in volume of people driving, walking, or taking public transit in their communities. The MTT provides mobility trends for major cities and 63 countries or regions. The information is generated by counting the number of direction requests made in the app. Datasets are then compared to reflect a change in volume relative to a base date (that is, January 13, 2020).

Note: As of January 2020, Google Android (74.30 percent) and Apple iOS (24.76 percent) jointly possessed almost 99 percent of the global smartphones market share.[35] The total number of smartphone users world-

33 Cnbc.com (URL).

34 Appleinsider.com (URL).

35 Statcounter.com (URL).

wide is projected to grow to 3.5 billion in 2020.[36] This duopoly undoubtedly makes Google and Apple the champions of location data, as they are the only two players having control on both hardware (smartphone) and software (operating systems) components.

Facebook (2.6 billion monthly active users in 2020)[37] is the largest social networking platform worldwide and is the owner of a large portfolio of 82 companies, which includes WhatsApp (2 billion active users globally)[38] and Instagram (1 billion monthly active users),[39] through which it harvests most of the location data it owns.

Facebook designed **Disease Prevention Maps** to help public health organizations close gaps in understanding where people live, how people are moving, and the state of their cellular connectivity, to improve the effectiveness of health campaigns and epidemic response. These datasets, when combined with epidemiological information from health systems, assist not-for-profit organizations in reaching vulnerable communities more effectively and in better understanding the pathways of disease spread by human-to-human contact.

[36] Bankmycell.com (**URL**).
[37] Statista.com (**URL**).
[38] Oberlo.com (**URL**).
[39] Oberlo.com (**URL**).

Amazon (+150 million users and Prime subscribers)[40] is an American technology company and the world's largest retailer and cloud services provider. It is the champion of e-commerce, cloud computing, streaming, logistics, and artificial intelligence. It accesses location data mostly through its mobile app and its cloud services platform (Amazon Web Services [AWS]).

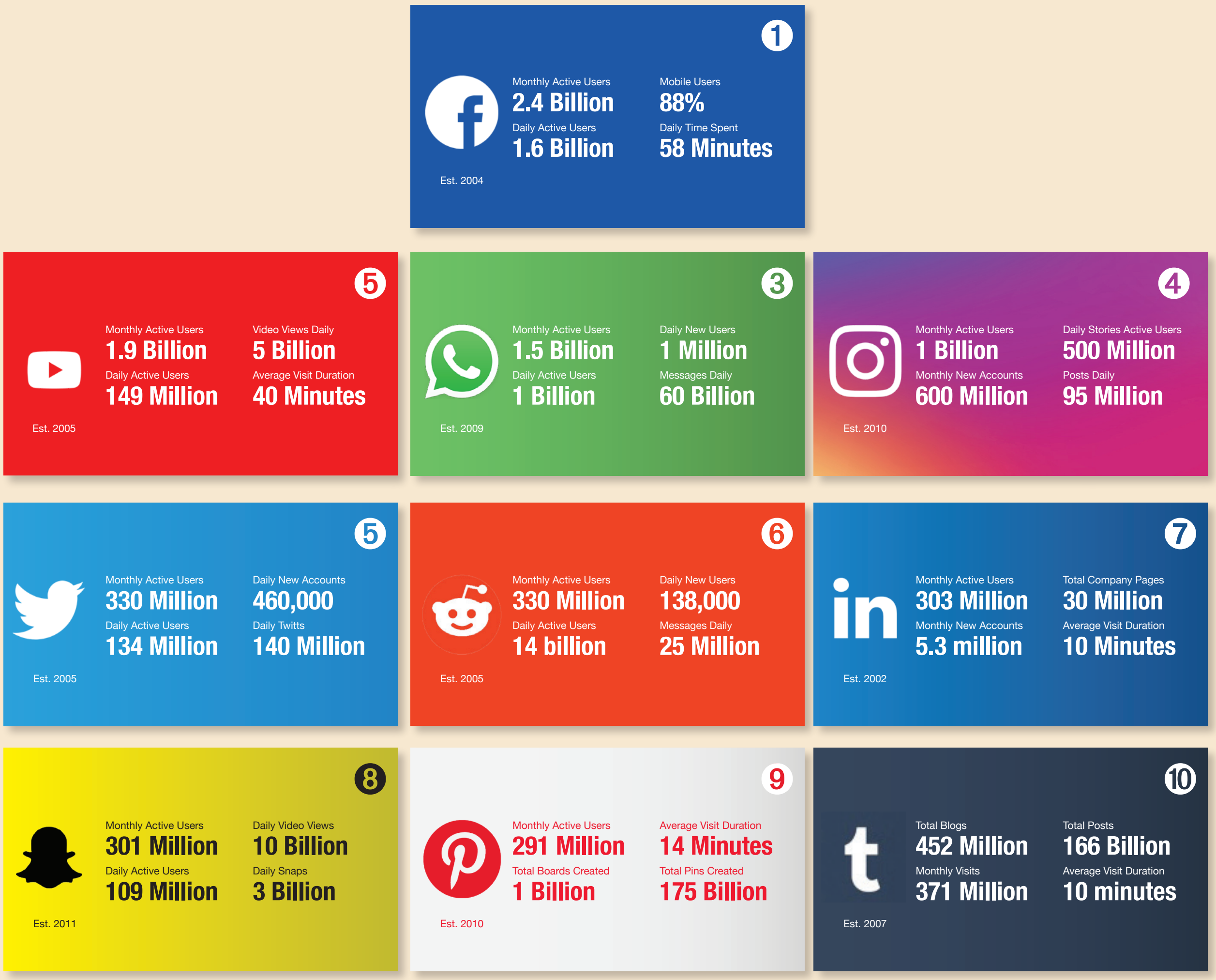
Microsoft (+1 billion devices running Windows 10)[41] is an American technology company and a leading global vendor of software, hardware, consumer electronics, cloud computing, and mobile and gaming systems. It accesses location data through its signature products like Windows 10.

Note: All the big tech companies, apps, and social network platforms report enabling strict privacy policies and cybersecurity protocols to safeguard users' identity (see Box 6 for further details on privacy safeguards).

[40] Oberlo.com (**URL**).
[41] News.microsoft.com (**URL**).

Box 5. Social media platforms and apps are key players in the location data arena.

Apart from the big tech companies, many of these apps rely heavily on the navigation business to provide a particular service—that is the case for Waze, Uber, and Lyft, among others. For many other apps, location tracking provides an extra layer that ultimately enhances user's experience—that is the case for food delivery apps like Uber EATs, DoorDash, and PostMates; grocery delivery apps like Instacart; sports apps like Strava; and even dating apps like Tinder and Bumble. All of them include geographic location as part of their core or complementary features. Infographics (Figure 3) provide a clearer picture of the volume and scale of data that the top-10 social media platforms could own in 2020.[42]



[42] Stout 2019 (URL).

Box 6. Privacy and personal data protection

Personal data is defined as the combined information related to identifiers such as names, ID numbers, location, and online identifiers and also metadata and factors specific to physical, physiological, genetic, mental, economic, and cultural or social identity of a natural person. The sense of urgency for an effective response to the global pandemic, which heavily relies on the use of tech-based solutions, has yet created a complex policy and ethical dilemma: how to ensure the right to health for everyone while preserving the fundamental right to privacy?

Governments are in fact harnessing location data to trace movements of COVID-19 positive cases, establish transmission chains, and estimate how the virus might be spreading; reach out and warn people who have been in close contact with somebody with the virus; track and monitor infected and suspicious individuals who are isolated at home; oversee and enforce compliance to stay-at-home and social distancing orders; and better understand general patterns of people's movement and behaviors and how those change over time. However, the combination of tracing and tracking tools and the massive gathering of this incredible amount of data has raised international concerns about the high risk of infringement of personal privacy.

Despite the effectiveness of such measures and the efforts to anonymize data collected, the truth is that mobile data contains sensitive and revealing details about people's lives, including information on location, behavior, relationships, movements, associations, and activities, which, in the hands of abusive governments, could translate into intrusive surveillance practices, discrimination, and even repression. The right approach thus relies in finding a reasonable balance between an improved data-driven response to the disease and the protection of people's privacy, in the best interest of public health, personal safety, and national security.

However, international organizations such as the United Nations, the WHO, the European Commission, and the World Bank, among many others, have advanced different legal instruments to better address privacy issues and ensure data protection (see the list of COVID-19 Data Protection and Privacy Resources) while proactively responding to the pandemic: "Even without formally declaring states of emergency, States can adopt exceptional measures to protect public health that may restrict certain human rights. These restrictions must meet the requirements of legality, necessity and proportionality, and be non-discriminatory."

To this end, the World Bank has recently enacted a Privacy Data Directive and Policy—that will enter into force on February 21, 2021—which establish a series of concrete principles and practices to better regulate personal data usage and safeguard privacy:

- Use personal data for legitimate purpose(s). That is with the consent and in the best interest of individuals concerned, with the mandate to comply, and consistent with the WBG mission.
- Use personal data in a fair and transparent manner.
- Limit and minimize personal data usage for original purpose(s) only. That is for well-defined archival, statistical, or research use and only in the amount and processing type that is reasonable and necessary.
- Preserve personal data accuracy. Record, update, and maintain data as accurately as possible to achieve original purposes(s) in accordance with this policy.
- Limit personal data storage. Follow records retention and disposition protocols and keep data in a form that permits identification as long as necessary to achieve legitimate purpose(s).
- Properly secure personal data. Use appropriate technical and organizational safeguards to prevent and address unauthorized processing as well as accidental loss, destruction, or damage.
- Share personal data with third parties for legitimate purpose(s) and using the same regard.
- Know your role and keep others accountable. Follow process owner directions, ensure proper records management, report suspected data bleaches, complete trainings and self-assessments, and direct staff on how to process and protect personal data.

Note: For additional information on data privacy at the World Bank, access this [Master Class (January 2020)].

4.2 Third-party firms

Cuebiq is a location intelligence company, which harvests location data from various apps and uses it to optimize marketing strategies for its clients. It works with public, research, academic, and not-for-profit organizations and supports them through their Data for Good program with data-driven initiatives in the areas of natural disaster preparedness and response, urban development and smart city planning, epidemiology and disease prevention, and economic equality and global development.

Some examples of their work include the following:

(a) **Inequality Atlas**—developed with the MIT Media Lab—to analyze how the presence of certain business establishments affect the nature and development of neighborhoods. The tool provides information on the socioeconomic backgrounds of local businesses' visitors (for example, income level) and the time they spend in the stores. It developed an index to estimate how equal/unequal income distribution was across visitors.

(b) **Mobility Insights** with the University of Washington to analyze human mobility before, during, and after Hurricane Harvey hit Houston.

(c) **Hourly updates** on wait times at gas stations across the American Southeast during Hurricanes Irma, Florence, and Michael, so that local authorities could better react to and manage energy shortages and fuel supply.

Cuebiq also created a COVID-19 Support business line, where it provides mobility insights and visitation patterns to better inform businesses and public agents' decisions. Cuebiq has a set of dashboards focused on contact tracing (to measure social distancing and mobility trends), economic reopening (to understanding how consumer visitation is changing), and mobility (to analyze changes in population's mobility trends across the United States at different scales). Other location data-driven analyses offer insights on customers' willingness to leave their homes and travelling behavior (Shelter-in-Place Analysis), how the virus outbreak has affected the mobility gap between wealthier and poorer groups (Social Inequality Analysis), and how new commuters patterns changed and where metropolitan workers are actually travelling (Mobility Flows Analysis).

Unacast provides human mobility insights combining location data, map data, and strategic intelligence to inform clients with the best possible picture of real-world human activity. They serve customers from a variety of industries including retail, real estate, marketing, government, travel, and tourism sectors. Unacast uses location data to gather information on competitors, get context on neighborhoods' activity, uncover the demographics of a place or district, understand mobility trends over time, analyze visitors' preferences and behaviors, and identify new sites offering potential business opportunity.[43]

In the current climate, Unacast has created a COVID-19 Location Data Toolkit offering an array of products such as a Social Distancing Scoreboard to measure and understand the effectiveness of Social Distancing initiatives at the local level (United States only); a Retail Impact Scoreboard to track foot traffic, understand commercial impact, and identify economic recovery signals across different industries and geographies; a Recovery Tool to visualize which, when, and how geographic areas are recovering; and Migration Patterns to monitor from which areas/communities people are leaving and where they end up in response to the outbreak.

Safegraph is a company providing data on business listings, building footprints, visitor analytics, foot traffic counts, and demographic insights for over 6 million points of interest in the United States. Some of its offerings contain neighborhood consumer insights and mobility patterns for every census block group, so that clients learn how people travel between neighborhoods and which neighborhoods attract the most outside visitors. It is also able to use GPS data to determine if a device visited a place, brand, or type of store.

[43] Unacast Solutions, Human Mobility Insights (URL).

Their COVID-19 Toolkit includes dashboards on economic reopening (to analyze foot traffic patterns by state and industry), shelter-in-place (to measure the percent of people staying home all day), foot traffic impact (to observe the changes of aggregated mobility data to point of interests sorted by business categories and brands), and weekly patterns (featuring foot traffic counts and point of interest visitor insights for 5 million places in the United States).

Mapbox describes itself as the 'Location Data Platform' for mobile- and web-based applications. It provides building blocks to add location features like maps, search engines, and navigation into any experience. Mapbox's offerings include (a) mapping and data visualization building blocks to design base maps and data layers, reproduce compelling visualizations of geospatial data, and leverage global street- and address-level data infrastructure; (b) a store locator with building blocks to find stores, places, and people with a unique user experience; and (c) powerful navigation SDKs and APIs[44] to build a complete range of products such as custom routing solutions, voice-guided navigation, and traffic-informed estimated times of arrival, among other features. While providing sophisticated mapping and navigation services to clients, Mapbox ends up harvesting the GPS data. It also collects telemetry data that contains user location.

Note: Through the Development Data Partnership, big tech companies like Google and firms like Mapbox and Cuebiq are sharing their location data with WBG staff for official use under strict conditions. To have access to the data, you have to submit a short proposal which summarizes your use case for the data; once it is approved, the Development Data Partnership team will help you access and use the data with various tools.

The World Bank, through its Development Data Partnership, has recently launched a Mobile Location Data Inventory with information from over 17 data sources for almost 90 developing countries. Data providers include Apple (mobility trends); Baidu (mobility data); Carto (GPS data); Cuebiq (mobility insights); Facebook (co-location maps, movement maps, movement range, population density maps, social connectedness index, and travel patterns); Google (mobility reports); InLoco (movement data); and Mapbox (mobility data).

44 API = Application Programming Interface.

5. References

Research Papers, Journal Articles, and Academic Publications

Aiken, Emily L., Guadalupe Bedoya, Aidan Coville, and Joshua E. Blumenstock. 2010. "Targeting Development Aid with Machine Learning and Mobile Phone Data." Working Paper, University of California Berkley. http://www.jblumenstock.com/files/papers/jblumenstock_ultra-poor.pdf.

Aram, S, A. Troiano, and E. Pasero. 2012. "Environment Sensing Using Smartphone." In *2012 IEEE Sensors Applications Symposium Proceedings*, 1–4. ieeexplore.ieee.org. https://paperpile.com/b/BcmiRH/FG0P.

Bosco, Claudio, and Samantha Wilson. 2019. "Towards High-resolution Sex-disaggregated Dynamic Mapping. Final Report." *Flowminder Foundation*. https://data2x.org/wp-content/uploads/2019/12/TowardsHighResSexDisaggMapping_Flowminder.pdf.

Chouhan, Siddharth Singh, Uday Pratap Singh, and Sanjeev Jain. 2020. "Applications of Computer Vision in Plant Pathology: A Survey." *Archives of Computational Methods in Engineering. State of the Art Reviews* 27 (2): 611–32. https://paperpile.com/c/BcmiRH/IunK.

Haddad, Ryan, Tim Kelly, Teemu Leinonen, and Vesa Saarinen. 2014. "Using Locational Data from Mobile Phones to Enhance the Science of Delivery." *World Bank*, Report Number ACS9644, Washington, DC. http://documents.worldbank.org/curated/en/687441468313509206/pdf/ACS96440REVISE0very00final0Digital0.pdf.

Hamstead, Zoé A., David Fisher, Rositsa T. Ilieva, Spencer A. Wood, Timon McPhearson, and Peleg Kremer. 2018. "Geolocated Social Media as a Rapid Indicator of Park Visitation and Equitable Park Access." *Computers, Environment, and Urban Systems* 72 : 38–50. https://doi.org/10.1016/j.compenvurbsys.2018.01.007.

Holtz, David, Michael Zhao, Seth G. Benzell, Cathy Y. Cao, M. Amin Rahimian, Jeremy Yang, Jennifer Allen, Avinash Collis, Alex Moehring, Tara Sowrirajan, Dipayan Ghosh, Yunhao Zhang, Paramveer S. Dhillon, Christos Nicolaides, Dean Eckles, and Sinan Aral. 2020. "Interdependence and the Cost of Uncoordinated Responses to COVID-19." Working Paper, Massachusetts Institute of Technology. http://ide.mit.edu/sites/default/files/publications/Interdependence_COVID_522.pdf.

Kumpel, Emily, Rachel Peletz, Mateyo Bonham, Annette Fay, Alicea Cock-Esteb, and Ranjiv Khush. 2015. "When Are Mobile Phones Useful for Water Quality Data Collection? An Analysis of Data Flows and ICT Applications among Regulated Monitoring Institutions in Sub-Saharan Africa." *International Journal of Environmental Research and Public Health* 12 (9): 10846–60. https://doi.org/10.3390/ijerph120910846.

Letouzé, Emmanuel. 2015. "Big Data and Development: An Overview." *Data Pop Alliance*, Primers Series. https://datapopalliance.org/wp-content/uploads/2015/12/Big-Data-Dev-Overview.pdf.

Milusheva, Sveta. 2020. "Using Mobile Phone Data to Reduce Spread of Disease.", Policy Research Working Paper 9198, World Bank, Washington, DC. http://documents.worldbank.org/curated/en/143411585572411591/pdf/Using-Mobile-Phone-Data-to-Reduce-Spread-of-Disease.pdf.

Maloney, William, and Temel Taskin. 2020. "Determinants of Social Distancing and Economic Activity during COVID-19: A Global View." Policy Research Working Paper 9242, *World Bank*, Washington, DC. http://documents1.worldbank.org/curated/en/325021589288466494/pdf/Determinants-of-Social-Distancing-and-Economic-Activity-during-COVID-19-A-Global-View.pdf.

Nemati, Ebrahim, Christina Batteate, and Michael Jerrett. 2017. "Opportunistic Environmental Sensing with Smartphones: A Critical Review of Current Literature and Applications." *Current Environmental Health Reports* 4 (3): 306–18. https://link.springer.com/article/10.1007/s40572-017-0158-8.

Oliver, Nuria, Bruno Lepri, Harald Sterly, Renaud Lambiotte, Sébastien Deletaille, Marco De Nadai, Emmanuel Letouzé, et al. 2020. "Mobile Phone Data for Informing Public Health Actions across the COVID-19 Pandemic Life Cycle." *Science Advances* 6 (23): eabc0764. https://doi.org/10.1126/sciadv.abc0764.

Pongnumkul, Suporn, Pimwadee Chaovalit, and Navaporn Surasvadi. 2015. "Applications of Smartphone-Based Sensors in Agriculture: A Systematic Review of Research." *Journal of Sensors*. https://doi.org/10.1155/2015/195308.

Prasad, S, S. K. Peddoju, and D. Ghosh. 2014. "Energy Efficient Mobile Vision System for Plant Leaf Disease Identification." In *2014 IEEE Wireless Communications and Networking Conference (WCNC)*, 3314–19. https://paperpile.com/c/BcmiRH/zXnT+k6P1.

Sacks, Audrey, and Varalakshmi Vemuru. 2020. "Using Big Data to Inform the Response to Vulnerability in Uganda's Settlements and Host Communities." Social Development Practice in the Africa Region (PDF File), World Bank, Washington, DC.

World Bank. 2020a. "The ABCs of Mobile Phone Data." Google Docs, Working Paper. https://docs.google.com/document/d/1Rv662mMzQHX9cRTsX1uYqTTNI-uXEtowlVsFLwISbH0/edit?usp=embed_facebook.

World Bank. 2020b. "Mobility Dashboard for COVID Response - Options Note." https://docs.google.com/document/d/1ayHvqn8NAfe25YCjlL-mboaDreC850n0ki2v6Ufcz-s/edit.

World Bank. 2020c. "Global Responses to COVID-19 in Slums and Cities: Practices from around the World", Internal Presentation, GSG Urban Poverty and Housing.

Wu, Yi, and Klarissa T. T. Chang. 2013. "An Empirical Study of Designing Simplicity for Mobile Application Interaction." In *AMCIS 2013 Proceedings*. aisel.aisnet.org. https://aisel.aisnet.org/amcis2013/HumanComputerInteraction/RoundTablePresentations/10/.

Reports, Briefs, Presentations, and Official Documentation

European Union Commission. 2018. *General Data Protection Regulation: Article 4 - Definitions.* GDPR.eu, https://gdpr.eu/article-4-definitions/.

Office of the United Nations High Commissioner for Human Rights. 2020. *Emergency Measures and COVID-19 Guidance.* United Nations. https://www.ohchr.org/Documents/Events/EmergencyMeasures_COVID19.pdf.

UN Women. 2020. *COVID-19 and Ending Violence Against Women and Girls.* EVAW COVID-19 Briefs. UN Women. https://www.unwomen.org/-/media/headquarters/attachments/sections/library/publications/2020/issue-brief-covid-19-and-ending-violence-against-women-and-girls-en.pdf?la=en&vs=5006.

WBG Board of Directors. "*Personal Data Privacy Directive*," January 23, 2020.

WBG Board of Directors. "*Personal Data Privacy Policy*", February 20, 2020.

News Articles, Blog Posts, and Press Releases

Apple. "Apple Makes Mobility Data Available to Aid COVID-19 Efforts." Newsroom, April 14, 2020. https://www.apple.com/newsroom/2020/04/apple-makes-mobility-data-available-to-aid-covid-19-efforts/.

Aral, Sinan. "The Deadly Cost of Uncoordinated Reopening of our States." The Hill, May 27, 2020. https://thehill.com/opinion/white-house/499727-the-deadly-cost-of-uncoordinated-reopening-of-our-states.

Bello, Lou Del. "How COVID-19 Could Redesign Our World." BBC.com, May 27, 2020. https://www.bbc.com/future/article/20200527-coronavirus-how-covid-19-could-redesign-our-world.

De Vynck, Gerrit. "The World Embraces Contact-Tracing Technology to Fight COVID-19." Bloomberg.com, April 30, 2020. https://www.bloomberg.com/news/articles/2020-04-30/the-world-embraces-contact-tracing-technology-to-fight-covid-19.

Elichai, Amir. "How Big Data Can Help in Disaster Response." *Scientific American Blog Network*, December 13, 2018. https://blogs.scientificamerican.com/observations/how-big-data-can-help-in-disaster-response/.

Ewen, James. "Best Guide to Location Data 2020 - All You Need to Know." Tamoco (blog), September 4, 2019. https://www.tamoco.com/blog/location-data-info-faq-guide/.

Fitzpatrick, Jen, and Karen DeSalvo. "Helping Public Health Officials Combat COVID-19." The Keyword (blog), April 3, 2020. https://www.blog.google/technology/health/covid-19-community-mobility-reports?hl=es.

Gasparikova, Daniela, and Kaca Djurickovic. "'Be Safe' - A Mobile App to Protect Victims of Violence." United Nations Development Programme, May 21, 2020 https://www.eurasia.undp.org/content/rbec/en/home/blog/2020/domestic-violence-app-montenegro-covid19.html.

Hu, Winnie, and Matthew Haag. "Public Spaces Weren't Designed for Pandemics. N.Y.C. Is Trying to Adapt." The New York Times, June 29, 2020, sec. New York. https://www.nytimes.com/2020/06/29/nyregion/nyc-parks-playgrounds-plazas-coronavirus.html.

Human Rights Watch. "Mobile Location Data and Covid-19: Q&A." News, May 13, 2020. https://www.hrw.org/news/2020/05/13/mobile-location-data-and-covid-19-qa.

MIT Initiative on the Digital Economy. "The Cost of Uncoordinated Responses to COVID-19." MIT Initiative on the Digital Economy, May 21, 2020 http://ide.mit.edu/news-blog/news/cost-uncoordinated-responses-covid-19.

Nagaraj, Anuradha. "Safety Apps Map Hotspots of Harassment of Indian Women." Reuters. November 15, 2018, sec. Asia. https://www.reuters.com/article/us-slavery-conference-women-idUSKCN1NK0KA.

Nanos, Janelle. "Every Step You Take: How Companies Use Geolocation Data to Target You—and Everyone around—in Ways You're Not Even Aware Of." The Boston Globe - Business, July 2017. https://apps.bostonglobe.com/business/graphics/2018/07/foot-traffic/.

Prosperidad Social – DPS, "Ingreso Solidario llega a más hogares", n.d. (accessed on March 26, 2021), https://ingresosolidario.prosperidadsocial.gov.co/

Romm, Tony, Elizabeth Dwoskin, and Craig Timberg. "U.S. Government, Tech Industry Discussing Ways to Use Smartphone Location Data to Combat Coronavirus." The Washington Post, Tech Policy, March 17, 2020. https://www.washingtonpost.com/technology/2020/03/17/white-house-location-data-coronavirus/.

Russo, Matthew. "The Definitive Guide to Location Data." Gimbal.com, n.d. (accessed on June 26, 2020), https://gimbal.com/location-data-guide/.

Sanchez-Andrade Nuno, Bruno, and Holly Krambeck. "Facebook and Mapbox: Addressing COVID-19 through Public-Private Data Partnerships - Where Do We Put New Testing Facilities?", March 19, 2020. https://datapartnership.org/updates/covid19-and-public-private-data-partnerships/.

Sanchez-Andrade Nuno, Bruno, and Maksim Pecherskiy. "Google, Apple, and Facebook: Understanding Mobility during Social Distancing with Private Sector Data", Development Data Partnership, May 11, 2020 https://datapartnership.org/updates/understanding-mobility-during-social-distancing/.

Signé, Landry, Sanjeev Khagram, and Julia Goldstein. "Using the Fourth Industrial Revolution to Fight COVID-19 around the World." Brookings (blog), April 28, 2020. https://www.brookings.edu/techstream/using-the-fourth-industrial-revolution-to-fight-covid-19-around-the-world/.

Singer, Natasha, and Choe Sang-Hun. "As Coronavirus Surveillance Escalates, Personal Privacy Plummets." The New York Times, March 23, 2020, sec. Technology. https://www.nytimes.com/2020/03/23/technology/coronavirus-surveillance-tracking-privacy.html.

Stout, Dustin. "Social Media Statistics: Top Social Networks by Popularity." Dustinstout.com (blog), July 8, 2019. https://dustinstout.com/social-media-statistics/.

USAID Water Team. "Mobile Communities in Ethiopia Seek Fixed Solutions to Their Water and Sanitation Challenges." Medium, Jul 24, 2018. https://medium.com/usaid-global-waters/mobile-communities-in-ethiopia-seek-fixed-solutions-to-their-water-and-sanitation-challenges-831380050174.

USAID Center for Water Security, Sanitation, and Hygiene. "Water Currents: WASH & Mobile Technologies." Globalwaters.org, October 29, 2019. https://www.globalwaters.org/resources/assets/water-currents-wash-mobile-technologies.

UN Global Pulse. "COVID-19 Data Protection and Privacy Resources." UN Global Pulse (blog), April 30, 2020. https://www.unglobalpulse.org/policy/covid-19-data-protection-and-privacy-resources/.

Ustaran, Eduardo. "International: The Coronavirus Privacy Dilemma." Data Guidance, March 6, 2020. https://www.dataguidance.com/opinion/international-coronavirus-privacy-dilemma.

Websites, Virtual Events, and Online Platforms

ACM Compass, SIGCAS Computing and Sustainable Societies - Virtual Conference 2020 (**URL**)

Alcaldía Mayor de Bogotá, "Renta Básica Bogotá" (**URL**)

Apple, Mobility Trends Reports (**URL**)

Bankmycell.com, "How Many Smartphones Are In The World?" (**URL**)

Cuebiq, Mobility Insights (**URL**)

Facebook, Data for Good, (**URL**)

Google, COVID-19 Community Mobility Reports (**URL**)

Mapbox, Resources Tab (**URL**)

Microsoft, Story Labs (**URL**)

Quadrant Global. "All You Need to Know About Location Data." (**URL**)

Safegraph.com, COVID-19 Dashboards (**URL**)

Statcounter.com, "Mobile Operating System Market Share Worldwide" (**URL**)

Statista.com, "Number of monthly active Facebook users worldwide 2008–2020" (**URL**)

Oberlo.com, "10 WhatsApp Statistics Every Marketer Should Know in 2020" (**URL**)

Oberlo.com, "10 Instagram Stats Every Marketer Should Know in 2020" (**URL**)

Oberlo.com, "10 Amazon Statistics You Need to Know in 2020" (**URL**)

Unacast Solutions, Human Mobility Insights (**URL**)

Guidance Note on Location Data

For the Working Group on Remote Supervision established by SD VP

THE WORLD BANK
IBRD • IDA | WORLD BANK GROUP