MDC00003202108
Issued: 2021-08

MOBILITY DATA SHARING ASSESSMENT: OPERATOR'S MANUAL

A resource for mobility data sharing

INTRODUCTION

In support of its vision to support the effective, responsible sharing of mobility data to support safe, equitable, and livable streets for all, the Mobility Data Collaborative™ (MDC) has identified the need to improve and coordinate understanding among all parties around foundational policy and legal issues to support mobility data sharing.

The MDC engaged the Future of Privacy Forum™ (FPF) to work with MDC members to develop a practical set of resources for organizations to ensure appropriate legal and policy review of mobility data sharing initiatives, as well as assess the fairness of these initiatives. Collectively, these resources make up the Mobility Data Sharing Assessment (MDSA):

- An infographic that provides a visual overview of the MDSA process.
- A tool that provides a practical, customizable, and open-source assessment framework.
- An operator's manual [this document] that provides detailed instructions, guidance, and additional resources to assist organizations as they complete the tool.

The MDSA is designed to be part of a flexible and scalable process and to support transparent and accountable decision-making about how and when to share mobility data. Since there is no "one-size-fits-all" approach to privacy and data protection, the length and complexity of the MDSA process will depend on the nature of the data sharing initiative and the risk tolerances of the organizations involved.

The MDSA is most useful for mobility companies partnering with public agencies, public officials requesting mobility data from service providers, local departments of transportation contemplating sharing with other public agencies, companies that want to license mobility data, or any organization considering sharing mobility data for academic research.

Just as the Mobility Data Collaborative™ *Guidelines for Mobility Data Sharing Governance and Contracting* apply to all parties involved in mobility data sharing, the MDSA is meant to be used by all organizations that either share mobility data or receive mobility data. The MDSA provides operational guidance and is consistent and interoperable with leading industry frameworks, such as the *NIST Privacy Framework* [1]. The MDSA was also designed to be technology-neutral and can be used for all data sharing methods.

Although privacy risk cannot be completely eliminated, the goal of this process is to help organizations maximize the benefits of mobility data sharing initiatives and minimize the privacy risks to individuals and communities. The MDSA does not provide legal advice; organizations must comply with all local laws and should consult with legal counsel.

This initial version of the MDSA is intended for any organization or group that wants to share or receive mobility data related to shared ground-based transportation, such as ride-hail, e-scooters, and public or shared transit. The MDSA was also designed for ongoing data sharing initiatives (e.g., research projects, permit requirements, commercial licenses, or voluntary partnerships); it is not meant for one-off data sharing requests or internal data operations. Future versions of the MDSA might include questions or guidance appropriate for assessing other modes of transportation, such as personal vehicles, delivery robots, or air-based mobility services.

The goal of the MDSA is to enable responsible data sharing and to equip organizations with an open-source, interoperable, customizable, and voluntary framework with guidance to help reduce barriers to sharing mobility data. Each organization will have a different operational privacy maturity level and organizational risk tolerance; in addition, the MDSA represents a high bar for privacy assessments, thus organizations are encouraged to customize the Tool and use what works, leave what does not, and continue improving internal mobility data privacy and security practices.

HOW TO USE THIS OPERATOR'S MANUAL

The MDSA operator's manual is designed to accompany the MDSA tool and follows the same structure while providing additional privacy guidance and resources relevant for mobility data sharing initiatives involving mobility data generated by the personal use of shared ground-based transportation, such as ride-hail and e-scooters. This operator's manual includes "**Practice Tips**" throughout, which are meant to provide advice to organizations completing the MDSA tool while referencing the operator's manual.

The MDSA operator's manual can be used…

- As a "how-to" resource and guide when filling out the MDSA tool.
- As the starting resource for organizations to build upon and customize as they are building out their privacy programs.
- As a communication resource to facilitate conversations about privacy internally and with partners, stakeholders, or communities.
- As an educational resource for employees and when discussing the scope of data sharing arrangements.
- To help organizations operationalize privacy and embed privacy into the design of mobility data sharing initiatives from the outset.
- To help organizations assess privacy and equity considerations *before* the decision to share mobility data has been made.

The MDSA operator's manual is useful for…	The MDSA operator's manual should be consulted…
▪ Data providers and recipients that work together to complete the MDSA tool or for organizations to do on their own. ▪ Public and private organizations at all stages of maturity that are interested in sharing or receiving mobility data.	▪ Before requesting or providing data, ideally. ▪ Retroactively or while the data sharing initiative is in progress, if not beforehand. ▪ When substantial changes to an existing data sharing initiative are contemplated.
The MDSA operator's manual covers…	**The MDSA operator's manual will be most effective when…**
▪ Mobility data, which is all information related to an activity, event, or transaction generated by either the operator or user of a digitally enabled mobility vehicle or service.[1] ▪ Data generated by personal use of shared ground-based transportation, such as ride-hail and e-scooters.	▪ Organizations think strategically and long-term about mobility data sharing. ▪ Consulted by collaborative, multi-disciplinary teams reflecting multiple points of view (e.g., privacy, legal, business, product, engineering, research, community, and other expertise). ▪ Organizations leverage the insights and resources in this operator's manual and the infographic to complete the MDSA tool.

TABLE OF CONTENTS

[1] Organizations should treat mobility data as personal information unless it can be demonstrated to be non-personal (e.g., data is de-identified or relates to an operator or object and not any particular individual).

1. MOBILITY DATA SHARING IN CONTEXT

An essential first step in assessing any mobility data sharing initiative is to put it in context. All stakeholders in the initiative should understand the environment in which the mobility data sharing will occur, including what the initiative is, what it intends to achieve, and who the relevant stakeholders are. Clearly documenting each of these components will help organizations better communicate their initiatives, as well as streamline their privacy and ethical data use assessments.

1.1 Backdrop

Organizations should create a high-level overview of the mobility data sharing initiative that they want to assess. This overview should identify key aspects of the initiative, organize key documents into one place, and record key details about the initiative that will be useful when reviewing the initiative at a later date.

Describe the initiative. Organizations should—in writing—briefly describe the data sharing initiative and its key components so that assessors and others can quickly understand the initiative's scope. The suggested information to include here is: (1) the name of the initiative; (2) a list of all partners involved in the initiative and their roles (e.g., funders, sponsors, or promoters); (3) any start and end dates (e.g., for pilot projects or partnerships); (4) any specific products or services being used (e.g., a particular scooter or bike share service or transit analytics product); and (5) any funding, reporting, or other deadlines.

Identify any data sharing agreements or other key documentation. It is important for partners in a data sharing initiative to have a written agreement (e.g., a data sharing agreement, partnership agreement, MOU, or other contracts) that clearly sets expectations, defines roles and responsibilities, and identifies key terms and conditions. Key provisions in these types of agreements include, but are not limited to:

- A license or permission to use mobility data for a particular objective.
- Any restrictions on how mobility data can be used (e.g., territorial or time limitations, exclusivity requirements, retention and deletion requirements, commercialization, etc.).
- Any warranties or other assurances (e.g., about the data provider's rights to the mobility data, that appropriate consent to share has been acquired where necessary, etc.).
- Any allocation of liability for contract breaches, data breaches, and other liabilities.
- Any confidentiality and non-disclosure requirements.
- The term and duration of the agreement.
- The governing law of the agreement and how the parties will resolve disputes.

Identify any regulatory bodies or other entities that might oversee the initiative, or the organizations involved in the initiative. Organizations should know which entities have regulatory authority over the sharing of mobility data for their initiative. For example:

- If a mobility data sharing initiative involves U.S. residents or companies, those organizations should understand relevant requirements and enforcement actions from the U.S. Federal Trade Commission, any relevant state Attorney(s) General, and any relevant self-regulatory bodies.
- If a mobility data sharing initiative involves U.S. municipalities, those organization should understand relevant requirements and enforcement actions in state and local laws.

- If a mobility data sharing initiative involves European Union (EU) residents, companies, or public authorities, those organizations should understand relevant requirements and enforcement actions from their particular member state's data protection authority, courts, or other supervisory bodies.
- If a mobility data sharing initiative involves a federally funded academic institution, it may require approval by an institutional review board or other ethical review body.
- In other jurisdictions, organizations should be aware of national and regional laws that may apply.

Record administrative information about the data sharing assessment (i.e., the MDSA tool). Organizations who undertake privacy assessments of their data sharing initiatives should clearly document the assessment process, including the date the assessment was completed and when it is scheduled for re-review, if applicable. It is also helpful to include contact information for the individuals or teams who completed the assessment and those who approved the initiative, in case there are questions in the future.

1.2 Objective

It is essential that organizations be able to clearly articulate what they intend to achieve through their initiatives and how sharing mobility data will help accomplish those goals. These objective(s) are the baseline against which privacy assessments like the MDSA tool evaluate the impacts of a particular data sharing initiative. Without a clearly defined objective or purpose for sharing mobility data, a privacy assessment will not be effective. It may be helpful to think of the objective statement as an "elevator pitch": articulate the objective in such a way so that everyone within the organization—as well as individuals using mobility services and the general public—can understand it.

- **Practice tip**: As a best practice, organizations are strongly encouraged to assess each objective or goal separately, so that the privacy and ethical safeguards described throughout this operator's manual can be tailored to maximize the value of the data being shared while minimizing risks to privacy.
- **Practice tip**: Once the objective has been defined, organizations must consider what mobility data will be shared to achieve the objective. (See 3.1.)

Be as specific as possible. Organizations should be as specific as possible when explaining what the initiative intends to accomplish (e.g., "improve transportation equity" is not specific, while "address transportation inequity by increasing the availability of last-mile micromobility devices within designated economic opportunity zones in City X" is much more specific). Organizations are advised against "bundling" objectives (e.g., "the initiative will share mobility data to evaluate changes in congestion along a major corridor and to enforce compliance with micromobility permits in another zone''), as this will make it more difficult to assess the specific benefits and privacy risks associated with the data sharing and to tailor safeguards appropriately.

Identify outcomes and decisions to be made. To the extent possible, organizations should also document the specific outcomes and decisions that will be informed by the mobility data that is being shared. For example, if organizations are sharing data to help manage and monetize public curb space for delivery and ride-hail drop-off/pick-up, organizations could include details about the factors that will go into the decision about which curb spaces to monetize. Likewise, organizations could include whether the initiative's success may lead to new fee structures and policies for curb spaces in the future. Organizations should remember that being able to effectively measure the success or impact of the data sharing initiative is often itself an important objective (e.g., performance metrics such as increased number of trips to transit hubs, shorter duration of trips, etc.).

Consider existing use cases and resources. Organizations that are struggling to articulate their objectives may want to take advantage of resources and real-world use cases of mobility data sharing compiled by new mobility organizations and peers.[2] Organizations are also strongly encouraged to engage with the communities and stakeholders who will be impacted by the initiative, to define and prioritize the key problems that mobility data may help solve, as well as to identify the best types and sources of mobility data to achieve those goals.

No one-size-fits-all. There is no one-size-fits-all method for defining an objective and organizations may need to go through several rounds of iterating in order to best refine their objectives. Organizations might also articulate their objective(s) differently for different audiences. For example, how an organization describes its objective for conducting a Mobility Data Sharing Assessment may be much more formal and detailed than how an organization describes that same objective to the public. (See 3.3.)

1.3 Stakeholders

Mobility data sharing initiatives have the potential to impact a wide variety of individuals, organizations, and communities. In order to effectively assess the privacy and ethical impacts of a particular initiative, organizations must understand who the full set of stakeholders are.

Identify all data provider(s) and recipients(s). Organizations should be as specific as possible and:

- Identify the specific departments, teams, projects, or other organizational subdivisions that will provide or receive the mobility data as part of the initiative. Even within the same organization, departments may have very different partnerships, missions, or objectives and key results by which they are motivated.
- Identify whether the data provider and/or recipient are part of any coalitions or collaboratives relevant to the sharing of mobility data. Some mobility data sharing initiatives may have only one provider and one recipient, while others may include multiple providers/recipients (e.g., a transportation data hub) or even other third parties (e.g., a major employer or insurer who helps promote or subsidize the use of micromobility services generally but does not actually share or receive mobility data). Organizations should also note whether there is any type of anticipated onward data sharing by any party.
- Identify whether this is an open-ended data sharing initiative with no specific data recipient (e.g., the initiative's results will be published as public open data, or the organization will develop a new service or analytic product that uses mobility data). In these cases, identify the anticipated customers or users of the data/product/service.

Identify all impacted individuals (data subjects). Organizations should clearly identify the types of individuals[3] whose information may be shared as part of the initiative. This would involve individuals who provide mobility data directly (e.g., drivers or riders), but also those whose activities and information may be captured more indirectly (e.g., passengers).

[2] For example, refer to NUMO's Micromobility and Your City Use Case and Metrics platform, or the Open Mobility Foundation's Use Case Database.

[3] These individuals are sometimes referred to as "data subjects" (e.g., under the General Data Protection Regulation (GDPR)) or "consumers" (e.g., under the Mobility Data Collaborative™ *Guidelines for Mobility Data Sharing Governance and Contracting*).

Identify all other reasonably identifiable impacted communities. Organizations should also clearly identify the people and groups who may be impacted by the initiative, but whose information is not being directly shared. Depending on the types of mobility data to be shared or received, there may be a wide range of stakeholders to consider, including both individuals (e.g., bystanders, pedestrians, commuters, residents of a particular neighborhood, or members of a historically marginalized community) and organizations (e.g., public transit operators, other mobility service providers, transportation or mobility nonprofits, state or federal transportation efforts, or unions).

2. LAWFULNESS: CAN YOU SHARE THE DATA?

One of the first things that an organization must determine is whether or not the data sharing in question is lawful. For the purposes of the MDSA tool, in order for an organization's sharing of mobility data to be lawful, it must not be prohibited by any applicable laws, contractual agreements, privacy notices, or other binding commitments. Organizations should identify and document all laws and obligations applicable to their particular initiative and ensure that mobility data sharing is permitted. Because this is a dynamic and quickly changing environment, organizations are also encouraged to regularly monitor the legal landscape for changes.[4]

2.1 Privacy Laws and Obligations

Organizations should identify all of the privacy and data protection obligations applicable to the sharing of mobility data and document any specific requirements or provisions that may impact the initiative. Organizations may also find it helpful to document requirements related to other types of processing (e.g., collection, retention, deletion), but it is not strictly necessary. In some circumstances, data sharing may also be permitted only as long as certain conditions are met (e.g., notice and consent, necessity); organizations should clearly identify these conditions and clearly describe how they will be met.

Identify and document applicable laws and obligations. Organizations should think broadly about the types of laws and obligations that might apply to sharing mobility data. For example:

- In the EU, commercial organizations and public agencies are subject to the General Data Protection Regulation (GDPR) as well as member state-specific data protection and mobility/transportation laws.[5]
- In all jurisdictions, including the U.S., legal requirements or other obligations may arise from a wider range of sources, including, but not limited to:
    - The data being shared (e.g., sectoral laws applicable to location, health, financial, biometric data, or other sensitive personal information).
    - Transparency or reporting requirements applicable to different types of organizations (e.g., public records laws, regulatory reporting requirements).
    - Specific state or federal laws that permit data collection and sharing by public transportation authorities (e.g., the National Transit Database).
    - Local government regulations and policies (e.g., permit or license requirements, procurement processes, surveillance ordinances, privacy principles, or data governance policies).

[4] The MDSA does not provide legal advice; organizations must comply with all local laws and should consult with legal counsel.
[5] Organizations should consult with legal counsel about the GDPR's application, as the GDPR has extraterritorial application when some factors are present and some U.S.-based organizations may trigger compliance obligations.

  - State or federal laws relating to privacy and consumer protection (e.g., FTC Act, state unfair and deceptive acts and practices laws, tort law, Electronic Communications Privacy Acts).
  - Anti-discrimination laws (e.g., Title VI of the Civil Rights Act of 1964, Americans with Disabilities Act).
  - State or federal constitutional provisions (e.g., 4th Amendment).
  - Federal or other laws that place restrictions on data recipients and/or providers when mobility data is shared across jurisdictional borders (e.g., GDPR, Wiretap Act in the U.S.).
- Data sharing initiatives may also be bound by other commitments made by the organizations involved (e.g., contracts, public privacy notices, self-regulatory programs, consent decrees/settlements, etc.).
- In practice, organizations may also find it helpful to document any internal notes about these obligations (e.g., whether a legal interpretation was sought by in-house or outside counsel and what, if any, guidance was provided; important enforcement actions under a particular law; the date that a contractual obligation expires; etc.).

Conditional mobility data sharing obligations. If any of the applicable laws and obligations allow mobility data sharing only under certain conditions, organizations should clearly describe how the data sharing initiative meets those conditions. For example, some laws or obligations may allow the sharing of mobility data only with the explicit consent of the individual, when "necessary" for the accomplishment of certain statutory purposes, during a public health emergency or other exigent circumstances, with the approval of an institutional review board or other ethical review body, or with a valid court order or other legal processes.

2.2 Changing Legal Landscape

In some circumstances, there may be few, or no, laws or obligations applicable to a particular initiative. As the legal and policy landscape surrounding mobility data sharing is dynamic and quickly evolving, organizations will need to actively monitor the legal landscape and identify any potential changes to federal, state, or local laws, regulations, or other legal actions[6] that may have an impact on the organizations or data involved in the initiative.

Emerging mobility data sharing obligations or prohibitions. Organizations should briefly describe their plans to monitor new or changing laws and obligations throughout the initiative's lifecycle. Organizations are encouraged to consider new developments arising from the sources described in 2.1 and to pay special attention to new regulatory guidance, enforcement activities, and case law, which can often move more quickly than legislation. There may also be specialty privacy and mobility groups or services available to help organizations keep abreast of policy developments and trends.[7]

[6] For example, new legislative amendments, regulatory guidance, enforcement activities, or judicial interpretations on existing law.

[7] For example, privacy trends and trackers are available from the International Association of Privacy Professionals, Future of Privacy Forum™, global law firms, and others, while industry trends and trackers are available from organizations such as the Mobility Data Collaborative™, Open Mobility Foundation, New Urban Mobility Alliance, North American Bike Share Alliance, or National Association of City Transportation Officials.

3. FAIRNESS: SHOULD YOU SHARE THE DATA?

"Human mobility, much like a physical fingerprint, is highly unique and can be used to find a person across mobility datasets [2]."

Organizations should also ensure that mobility data is shared in a way that is fair to individuals and communities. Fairness requires organizations to consider whether they should share mobility data, even when it is technically legal to do so.[8] Stakeholder trust is essential to ensuring that data sharing initiatives achieve long-term success. Moreover, organizations will not be able to sustain that trust if (1) mobility data is shared or used in ways that are irrelevant or disproportionate to the objective(s); (2) if the quality of mobility data is poor or not fit for purpose; (3) if the initiative runs contrary to individuals' and communities' reasonable expectations of privacy; or (4) if fairness, equitable, and anti-discrimination considerations are not addressed.

3.1 Relevancy and Proportionality

In order for mobility data sharing to be fair, the mobility data must be relevant and proportionate to the initiative's objectives. However, "mobility data" is not a monolith: there are many different types of mobility data, some of which create greater privacy or fairness concerns than others.[9] This means that organizations should break down each and every type of mobility data to be shared, consider its key attributes, and document why the data being shared is relevant and proportionate to the initiatives' objective(s). Given its sensitivity from a privacy perspective, organizations should give special consideration to location data throughout this process.

- **Practice tip**: **Use and adapt existing tools**, **such as the MDSA tool**. Organizations can use the table found in the MDSA tool Section 3.1 to record details about the types of mobility data to be shared in a particular initiative. Organizations are encouraged to customize this table to their specific initiatives, such as by adding or editing key attributes.
- **Practice tip**: **Consult diverse experts**. Some of the key attributes of mobility data (e.g., level of identifiability) may require a case-by-case evaluation by legal, technical, and other experts. Organizations are strongly encouraged to supplement the guidance found in this operator's manual with other relevant expert advice.

[8] The MDSA takes a broad approach to fairness, which is reflected in 3.1, 3.2, 3.3, and 3.4. Note, however, that "fairness" is also a requirement of many global privacy laws, e.g., U.S. Federal Trade Commission Act: Section 5: Unfair or Deceptive Acts or Practices [79] regulates "unfair" acts and practices and GDPR Article 5: Principles Relating to Processing of Personal Data [80] requires data to be processed "fairly." Also refer to Information Commissioner's Office, Guide to the GDPR, Principle (a): Lawfulness, fairness and transparency.

[9] Organizations can find additional information about "mobility data" since the concept is used throughout the MDSA in Appendix A and about location data and re-identification in Appendix B.

3.1.1 The Key Attributes of Mobility Data

To properly assess whether the mobility data[10] to be shared is relevant and proportionate to the data sharing objective, organizations should first consider key attributes of mobility data that are included in the data sharing initiative: (1) data type, (2) data source, (3) format or method of data transfer, and (4) any other sensitive attributes.

Data type. Organizations often talk about sharing "mobility data" as if it is a uniform entity. However, in reality, organizations usually share many different types of data as part of their initiatives, each of which may have unique privacy implications. Some common examples of mobility data types include:

- Trip data (e.g., origin/destination point).
- Telemetry data (e.g., route information).
- Operational data about vehicles and devices (e.g., battery life, vehicle IDs).
- Number of rides requested and completed.
- Trip lengths (distance and time).
- Spatial data (e.g., single locations, trajectories, direct and indirect location based on proximity).
- Temporal data (e.g., real-time and non-real time).
- Fleet size.
- Number of vehicles in operation.
- Wait times.
- Disability and accessibility related data (e.g., Americans with Disabilities (ADA) trips).
- Trip prices.
- Vehicle identification number (VIN).
- Driver or rider ID.
- User demographics (e.g., gender, age, income, zip code).
- User surveys or other qualitative inputs (e.g., customer satisfaction surveys, comment or complaint forms, user experience interviews).
- Metadata (e.g., data about data and characterizes mobility data to make it easier to retrieve, interpret, or use the data).
- Transaction or payments data (e.g., voucher enrollment, fare costs, etc.

[10] Mobility data is all information related to an activity, event, or transaction generated by either the operator or user of a digitally enabled mobility vehicle or service. For the purposes of this MDSA, it includes data generated by personal use of shared ground-based transportation, such as ride-hail and e-scooters.

Data source. Organizations may also generate or obtain different types of data from different sources. It is important that organizations understand the context in which data originated [3], as it may impact many aspects of the privacy assessment process. Some useful questions to ask are: How reliable is the data for this objective? How aware were individuals that their data was being collected? Are there specific contractual limits on what this data can be used for?. Sources of different types of mobility data might include:

- Data collected **directly** from individuals, such as surveys or applications.
- **Observations** about individuals, such as loyalty cards, enabled location sensors on personal devices, traffic sensors, or CCTV.
- Information **generated** or **derived** by an organization, such as credit or trust ratios or average travel times.
- Information that is **inferred** by using analytics to find correlations, such as fraud or credit scores or other risk analysis.
- Information **purchased** or **obtained** from a third party.

Format or method of sharing. Different types of mobility data may also be shared in different formats or by different methods, which may have different implications for privacy or security. Some examples of technical delivery modes [4, pp. 77-81] for mobility data include:

- **Wireless**: This mode typically involves setting up a wireless network and connecting devices to that network without any hardware cable material (e.g., Bluetooth, wireless LAN, wi-fi).
- **Remote access** or **virtual private network** (VPN): A VPN uses a network connection to enable remote access data on devices or servers that are connected to the network at the time.
- Database accessible by **application programming interface** (API): This mode enables applications to exchange data and functionality easily and securely.
- **Email**: An easy to use mode of transfer and security concerns can be alleviated by using password protected files and adding public key cryptography mechanisms. This mode does not support sharing large files.
- **Removable storage media** (e.g., a USB flash drive).
- **Wire transfer**: This mode typically involves a fiber-optic hardware cable (e.g., a LAN cable or ethernet cable) directly connecting two data storage sites.
- **Distributed ledger technology** (e.g., blockchain).

Identifiability. Organizations should also be careful to distinguish between data that is considered identifiable and data that is not.[11] There are very different privacy implications when an organization shares or receives data that can directly or indirectly identify a particular individual and when it shares or receives data that cannot be reasonably linked to a particular individual or device (e.g., statistical data about congestion zones or corridor analysis). Some factors [5] to consider include:

- **Direct identifiers** can be used to identify a particular person without additional information (e.g., rider names, home address, or transaction information).
- **Indirect (or quasi-) identifiers** do not identify a specific individual alone, but can be aggregated and "linked" with other information to identify a particular person (e.g., precise trip or location data, work address, or demographic information). Even a vehicle or device's technical data (e.g., data about the wear and tear on vehicle parts) could indirectly identify a particular person if cross-referenced with other data. [6]
- Data that has been **de-identified**, **aggregated**, **or anonymized** such that it cannot be reasonably linked to a particular individual or device (e.g., statistical data about congestion zones or corridor analysis).

Other sensitive attributes. Some types of data may have additional sensitive attributes that deserve special consideration. While organizations should tailor these attributes to their own circumstances, some other types of data that may warrant additional assessment include:

- Location data (see 3.1.2).
- Physical biometrics (e.g., fingerprints, eye gaze, face detection or recognition).
- Behavioral biometrics (e.g., speed or braking patterns).
- Equity or socioeconomic data (e.g., riders' use of low-income discount or equity programs, use of mobility data for pricing regulation).
- Data that may be regulated in other sectors (e.g., financial or transaction data, health data, children's data).

3.1.2 The Key Characteristics of Location Data

The ability to use and share location data is at the heart of all mobility data sharing initiatives[12]. However, the ability to understand when and how individuals move from one location to another is as valuable to governments, researchers, and businesses as it is revealing of individuals' activities and interests [7]. Even when location datasets are stripped of identifying details, they can reveal information about individuals when combined with other data (including behaviors, attitudes, and even identities). Given the potential sensitivity of mobility data, and the importance of location privacy to individuals and communities,[13] mobility data sharing initiatives should pay special attention to location data. Some key characteristics to consider include:

[11] Determining how identifiable a particular data element is will require a case-by-case legal and technical analysis and may vary depending on jurisdiction. See Appendix B for more information about identifiability and re-identification.
[12] See Appendix B.
[13] Location privacy is about individuals' right to move in public spaces with the expectation that under normal circumstances their location will not be systematically and secretly recorded for later use [77] [9]. Location privacy is not absolute; it depends on the individual's expectation of "normal circumstances" and how the information is collected and used. Because an individual's reasonable expectations and the way location data is collected and used changes over time, context, nuance, and using a case-by-case analysis are crucial when assessing the privacy implications of location data.

Precision. Sometimes location data can be quite precise or granular (i.e., can pinpoint an object or individual to a specific block or intersection); other times it may be imprecise or coarse (i.e., can only place an object or individual in a particular city or country). To what extent location data is considered precise can vary based on sector and jurisdiction.[14] However, the more precise and accurate the location data, the more revealing it tends to be [7, p. 4] [8].

Density. The density (e.g., urban, suburban, or rural) and zoning (e.g., residential, commercial, industrial) of the location data to be shared is another important spatial consideration. Data about high-density urban or commercial locations, for example, typically carries lower privacy risk than rural or residential areas. However, it is important for organizations to also address potential outliers within otherwise relatively homogeneous zones (e.g., a patch of single-family homes in an otherwise commercial district).

Immediacy. The speed with which location data is shared between organizations can also have privacy implications. Data shared in real-time or near-real-time is often considered higher risk than historical location data (e.g., data shared at a day's, week's, or month's delay), in part because it may enable organizations to intercept or engage with an individual while they are in movement. It can also be more challenging to protect data shared in real-time rather than in non-real time [9].

Continuous versus snapshot sharing. Organizations may also share location data in different frequencies, including in limited, discrete batches (i.e., one-off snapshots) or on a more continuous basis (i.e., periodic or on-demand data). More frequent and continuous location data is typically more challenging from a privacy perspective, because it is more likely to generate the spatial "bread crumbs" that can reveal sensitive information about individuals' movements and may require the use of more sophisticated privacy-enhancing tools [10] [11].

Time and location. The combination of time (temporal elements) with location data can create a greater privacy risk than either time or location data alone. Adding temporal data to a set of location points can reveal more precisely when and how an individual made it from point A to point B to point C. Including the temporal element can also provide additional context and allow for new inferences compared to location data alone (e.g., the same trip from home to work at 8 am versus 8 pm might reveal what shift an individual works). Research has shown that only a few timestamped location points are needed to uniquely identify most individuals; accordingly, location data shared without temporal elements typically carries lower risk than location data combined with specific times [7] [12].

3.1.3 Making the Connection: Mobility Data and Objectives

From a legal, policy, and practical perspective, it is important for organizations to assess whether mobility data is adequate, relevant, and limited to what is necessary to achieve a specific objective. This is a case-by-case question, and depends upon the objective to be achieved, as well as the nature and sensitivity of the mobility data at issue. If using the MDSA tool, organizations should justify why the data elements to be shared in a particular initiative are relevant and proportionate to their objectives.

[14] Determining how precise a particular location data element is will require a case-by-case legal and technical analysis, and may vary depending on jurisdiction. For example, the National Advertising Initiative (NAI) considers latitude and longitude coordinates with two or fewer decimal places to be imprecise [78], while the California Privacy Rights Act of 2020 considers "precise geolocation" to be "any data that is derived from a device and that is used or Intended to be used to locate a consumer within a geographic area that Is equal to or less than the area of a circle with a radius of one thousand, eight hundred and fifty (1,850) feet, except as prescribed by regulations" [18]. See more about how "location data" is defined under existing privacy laws in FPF Policy Brief: Location Data Under Existing Privacy Laws [65].

Relevant. Mobility data is relevant for a data sharing initiative if there is a rational link between the data and the initiative's specific objective [13]. In order to assess whether the right amount and type of mobility data is being shared, organizations must first be clear about why the mobility data is needed (i.e., their objective; see 1.2).

- Example: If the objective of the data sharing initiative is to enforce a cap (or limit) on how many e-scooters are deployed in a particular operating area, then relevant data might include: the number of e-scooters permitted/in operation per provider, the boundaries of geographical zones, and timestamps for any e-scooters entering/exiting the zone. On the other hand, data about rider demographics, pricing, or trip duration may not be relevant for that particular objective.
- Example: If the objective of the data sharing initiative is to measure whether a new mobility product is helping reduce racial and socioeconomic gaps in access to transportation services within a particular neighborhood, then relevant data might include demographic information about individuals who use the service. On the other hand, precise location data from all trips taken by riders citywide may not be rationally connected to that particular objective.

Proportionate. Proportionate means that organizations should strike a balance between the potential privacy impact of the mobility data (given its data type, source, method/format of sharing, identifiability, and other sensitive attributes) and the importance of the objective to be achieved [14] [15].

- Example: A transportation planning or product development objective in a dense urban corridor might be achievable with aggregate data instead of precise route and location data with timestamps.
- Example: In emergency situations, such as a wildfire, a city may need to know in real time the precise location of all mobility devices in a certain geographic area to effectively deploy emergency services.

Privacy-friendly alternatives. Organizations should also consider whether there are privacy-friendly alternatives to the mobility data they had planned to share or receive that would be effective for achieving their objectives. If there are alternatives, but they are not suitable for a particular initiative, it is helpful to document why and to periodically reassess decisions, due to changes in technologies and data sources.

- Example: A differentially private dataset (a dataset that describes the patterns of groups within the dataset while withholding information about individuals in the dataset) might be appropriate for some use cases, such as urban planning or regional coordination; however, it may not be suitable for others, such as when mobility data is used to make decisions with legal consequences or to understand impacts on small populations.[15]

3.2 Data Quality and Retention

In addition to relevance and proportionality, organizations must also assess the quality of the mobility data to be shared or received and whether it is fit for its purpose. If the data is not sufficiently accurate, complete, or representative to achieve a particular objective, then it will likely not satisfy stakeholders' notions of fairness and privacy. It is also important that organizations consider how long they will need to retain mobility data in order to achieve an objective. An organization or initiative that plans to keep mobility data indefinitely is not embodying fair and responsible data practices.

Accurate. Organizations should ensure that the mobility data they share or receive is accurate, reliable, and up to date. This includes assessing the degree to which the data maps to the initiative's objectives, and that there are processes in place to ensure that inaccurate or outdated information is caught and corrected. Some key considerations include:

[15] For example, differential privacy offers robust privacy protection in the U.S. census, but obscures small populations, with unintended impacts on equity and research [81].

- **Nature of the data source**. How an organization collects, derives, or obtains mobility data can impact the data's accuracy, timeliness, and what the data is permitted to be used for. (See 3.1.1 and 3.1.2.)
    - Example: If a company's goal is to monitor road conditions and identify potholes or other repair needs, then relying solely on data from on-vehicle cameras or motion sensors, which can have a high number of false positives [16], may not be the best fit.
    - Example: If a city's goal is to enforce micromobility permit requirements for approved operators, then using aggregated trip data obtained from a third-party aggregator or analytics derived by a commercial location analytics service, which might obscure individual trip and operator activity or come with contractual restrictions, may not be the best fit.
    - Example: If a researcher's goal is to understand route decisions made by bikeshare riders in urban, suburban, and rural spaces, then relying solely on the devices' global positioning system (GPS) data, which may be less technically accurate in dense urban canyons than in other environments, may not be the best fit.
- **Ongoing monitoring and correction**. Because data sharing initiatives are dynamic, organizations should invest in resources (e.g., training, tools, and policies) that help to evaluate the accuracy of mobility data on an ongoing basis. In the private sector, organizations may be required to provide individuals with access to copies of their personal data and mechanisms to request that inaccurate or outdated information is corrected or deleted [17] [18].[16]

Completeness and representativeness. Organizations must also assess whether the initiative's objective can be accomplished with the mobility data intended to be shared and whether that data is sufficiently representative of the community to allow accurate conclusions to be drawn. Data that is incomplete or unrepresentative can lead organizations to false conclusions and have negative impacts on individuals and communities. In these cases, it may be necessary for additional information to be gathered or shared to ensure that the initiative is able to achieve its objective. Some key considerations include:

- **Who or what may be missing**. For various reasons, mobility data may have gaps or missing information (e.g., perhaps a mobility service was offered in only one language and was not widely adopted in other communities; perhaps there was a power outage and data collection was interrupted for a few days). Organizations should examine the data they intend to share from different perspectives and, in light of their intended objectives, consider the implications of attributes like demographic representativeness, market share, and geographical representativeness.
    - Example: An organization may wish to use data about the frequency, duration, and distance of ride-hail trips to evaluate the impact of COVID-19 stay at home orders. However, while ride-hail data may be precise enough for clear measurement, the population of ride-hail users (typically younger, wealthier, and digitally literate) is not sufficiently representative—on its own—to draw accurate conclusions about the effectiveness of the policy intervention.[17]
- **Invite fresh eyes to identify blind spots**. Because all teams have biases and blind spots, it can be difficult and even impossible to evaluate mobility data for completeness or representativeness from all perspectives and in all circumstances. Organizations should not be afraid to reach out to knowledgeable stakeholders and experts to help identify potential blind spots.

[16] Refer to Information Commissioner's Office, Right of access; Article 15, GDPR; California Consumer Privacy Act (CCPA).
[17] For example, refer to studies measuring mobility to monitor travel and physical distancing interventions during COVID-19 using smartphone GPS traces [82] [83].

- **Data provenance and metadata**. Data provenance (a record trail accounting for the origin of the data with an explanation of how and why it got there) and metadata can be useful tools when assessing the completeness and representativeness of mobility data. Organizations should also consider whether mobility data has been consistently collected or measured in the same way across all records intended to be shared and should note where there are inconsistencies and what impact this could have on data quality.

Retention. In order to ensure that mobility data is shared and used in a fair and privacy-protective manner, it is important for organizations to set limits on how long they will retain mobility data in identifiable forms. If an organization cannot describe how long it will need mobility data (e.g., until the end of a pilot period; until a statute of limitations expires; until no longer necessary to carry out the purpose for which it was shared), then it may be impossible for organizations to truly ensure the data remains fit for purpose. Some key considerations include:

- **Data may be needed for longer periods of time, depending on the objective**. Organizations should be able to articulate how long they need to retain mobility data to achieve their objectives, but there is no one-size-fits-all retention period that they must adhere to. Data does not always need to be deleted immediately, but it should not be retained indefinitely. (See 5.4.)
- **Insights based on mobility data**. Organizations should be thoughtful about whether there are ways to retain important insights or analysis gleaned from mobility data, without retaining the mobility data itself. This may help decrease privacy risks, as well as operational and storage costs.

3.3 Transparency

Transparency is a fundamental aspect of privacy and fairness. In order to assess transparency, organizations should consider the perspectives of the individuals whose data is being shared and the perspectives of the other communities who may be impacted by the data sharing initiative.[18] There may be heightened expectations regarding transparency and engagement for certain organizations, such as government agencies. (See Section 5.)

  - **Practice tip**: All participants in a data sharing initiative should collaborate to make the initiative's work more open and transparent to the public. It is also important for organizations to be open and transparent to each other throughout the initiative.

Individual engagement. Organizations should consider whether the individuals whose data is being shared would be surprised or offended to learn that their mobility data had been shared as part of the initiative. If individuals are not aware of the sharing, or the sharing is not consistent with their reasonable expectations of privacy, then organizations have likely not been suitably transparent.[19] Some key considerations include:

- **Reasonable expectations of privacy**. Organizations should look to the mobility data's source to understand the reasonable (subjective and/or objective) expectations that individuals had when their data was first processed. User interviews and research on cultural and local norms may also provide additional insight [19]. Organizations should also keep in mind that individuals' expectations of privacy can change depending on the design and availability of mobility services offered.[20]

[18] For example, refer to Draft NISTIR 8312, Four Principles of Explainable Artificial Intelligence.
[19] Organizations should also consider the best ways to be transparent about the results of a data sharing initiative (e.g., through open data portals and annual reports). For example, refer to LADOT, Year One Snapshot, A Review of the 2019-2020 Dockless Vehicle Pilot Program; or refer to Sidewalk Labs' Digital Transparency in the Public Realm, which brought together participants for co-design sessions and released the materials and a summary of the inputs afterwards.
[20] When it comes to assessing the individual's expectation of privacy, some commentators believe that it no longer makes sense to differentiate between a smartphone and mobility devices or services, such as e-scooters or ride-hail [96].

- **Freedom of information**. Because there are heightened transparency laws and norms for governments, agencies should understand whether to what extent information about the initiative or the mobility data itself can or must be shared through open data portals or public records requests. (See 5.5.)
- **Privacy notices**. All organizations' websites should include accessible, easy-to-find information about the mobility data sharing initiative, including the initiative's objectives, what types of mobility data will be shared, how long mobility data will be retained, for what purposes mobility data may be used, and other applicable safeguards. While organizations tend to have a single detailed privacy notice, it is often more helpful to take a different approach and some privacy laws require different types of notices.[21] Instead of a single, detailed notice, consider more dynamic forms of notice, such as:
    - Layered notices, which are helpful when a full privacy notice may be long and complex. Organizations can layer their notices by producing both a condensed notice (with key highlights up front) and a longer, complete notice with all of the legal requirements.
    - Contextual notices, which allow organizations to highlight any purposes or uses of data that would not be obvious to the individual or reasonably expected based on the context.
    - "Just-in-time" notices, which appear at the time an individual accesses a feature (such as before sharing precise location data or contact information with a mobile app), instead of only at the time they sign up for the service. The notice should have relevant and focused privacy information delivered at the time mobility data is collected.
    - Privacy or data protection dashboards, which provide an interactive interface for individuals to change their privacy settings in real-time.
    - Visual notices, such as "nutrition label" models or public signage where passive data collection may occur [20] [21].
    - Technical specifications and developer terms, which may provide a window into the technical aspects of data sharing. (However, organizations should note that developer terms and platforms such as GitHub are not easily navigated by the general public.)

Community engagement.[22] It is also important for organizations to engage with the community at large, who may be just as impacted by the initiative's efforts as those whose data is actually shared. For example, an initiative that plans to allocate sidewalk space to scooter parking should consider the impact on nearby residents, businesses, and pedestrians as well as scooter users. Organizations that fail to be open and transparent with all stakeholders may struggle to create social license and legitimacy or build public trust. Some key considerations include:

- **Reasonable expectations of privacy**. Cities and communities are often mosaics representing individuals from many walks of life, with different backgrounds, interests, and priorities. Organizations should not assume that everyone in a community shares the same privacy expectations and preferences, or that communities in different cities or countries are similar to each other; instead, organizations should take the time to understand and consider each impacted communities' specific perspective. Where there may be conflicting views, it becomes even more important that decisions about how mobility data is shared and used are made transparently.

[21] Organizations should be mindful of privacy laws that prescribe certain types of notice. For example, refer to the California Consumer Privacy Act, "notice at collection" [91] [103].

[22] Community engagement is a process of working collaboratively with groups of people who are affiliated by geographic proximity, special interest, or similar situations to address issues affecting the wellbeing of those people to create social license and legitimacy and build public trust [47] [102].

- **Meaningful engagement and empowerment**. To the extent possible, organizations should also provide opportunities for communities who will be impacted by the initiative to have a meaningful role in its development and decision-making, especially around how mobility data will be shared and used. Organizations may also want to provide ways for impacted communities to provide input to the data sharing initiative, such as identifying new data sources, unintended consequences, or opportunities for future work.
- **Inclusive engagement**. Organizations should put special effort into engaging any marginalized communities[23] that may be impacted by the initiative. Organizations should thoughtfully consider whether their outreach and communications channels are accessible and welcoming for every individual.[24]

3.4 Ethics, Equity, and Anti-Discrimination

There may be situations where sharing mobility data is technically lawful, but still has the potential to negatively impact individuals and communities. Organizations should be careful to assess whether their data sharing initiative raises ethical or equity considerations either intentionally (by advancing anti-discrimination goals) or unintentionally (by using biased data, creating disparate impacts, or allowing function creep).

- **Practice tip**: Data ethics is an evolving field that considers the impact that data-driven activities can have on individuals and communities. Formal ethical reviews[25] are often not required as part of mobility data sharing initiatives; nevertheless, organizations navigating mobility data sharing initiatives with significant ethical considerations might find it valuable to obtain expert guidance or review (either internal or external).[26]

Ethical considerations for individuals or communities. Many decisions to share mobility data are a matter of identifying which laws, regulations, frameworks, or agreements apply. However, sometimes these frameworks may not specifically address all concerns that individuals and communities may have about how and why mobility data is shared, resulting in an initiative being labeled "creepy" or unethical. For example, a mobility data sharing initiative using only aggregated data to inform transportation policy might have a low privacy impact, but could still negatively impact individuals and communities if the data were biased in favor of one population over another. Consulting research on ethical impacts, establishing or following existing ethical principles,[27] creating a mobility data ethics checklist,[28] or conducting outreach to individuals and communities (before the data sharing initiative is approved and has started) can help organizations better understand what is appropriate (or not) for their particular initiatives.

[23] Marginalized communities are those that experience discrimination and exclusion because of unequal power relationships across economic, political, social, and cultural dimensions [84].

[24] This means considering more than physical accessibility requirements. For example, a university may offer physically accessible rooms, but may feel inaccessible and exclusive to individuals who have not attended a university. Other considerations for culturally diverse and inclusive spaces, platforms, and materials include language and literacy; time and location; physical and digital accessibility; incentives and appeal; art and storytelling; and power dynamics.

[25] For example, The Common Rule requires any university-based and federally funded research on human subjects to be reviewed by Institutional Review Boards (IRB); 45 CFR § 46, 4 [94].

[26] For example, organizations may establish their own internal review committees (e.g., Facebook/Google Review Board) or avail themselves of private review committees or expert ethicists (e.g., EDSRC). For example, refer to Future of Privacy Forum™, *FPF Ethical Data Use Committee Will Support Research Relying on Private Sector Data*, May 5, 2021 [95].

[27] For example, some indigenous populations have a code of ethics for researchers that may require consent and approvals before data is collected [85].

[28] Some example questions are: Is data science the right tool to meet the objective? How does the organization or team consider individuals and communities who will ultimately be impacted by the initiative? Were the systems and tools used to collect the data biased against any groups? Are particular stakeholders empowered or disempowered as a result of this initiative? Could the sharing of the data be reasonably expected to cause tangible harm to any individual's well-being?

Anti-discrimination goals. While data sharing initiatives should be free of bias and inaccuracies that may lead to discrimination against individuals and communities, organizations may have anti-discrimination goals as the very objective for mobility data sharing.[29] Attempts to create fairness and ameliorate inequities may sometimes be perceived as bias, and organizations should be transparent about these intended outcomes. Organizations may also wish to undertake broader strategic planning efforts to identify and prioritize anti-discrimination goals outside of a specific mobility data sharing initiative.[30]

- Example: Because equal access to mobility services by all demographics is a challenge, some organizations may share and use mobility data to design spaces and services that prioritize anti-racism and anti-discrimination [22] [23] [24].
- Example: A city may receive race-based data to monitor and prevent racial disparities and systemic discrimination and to develop evidence-based policies to address problems [25] [26] .

Bias and disparate impacts. Organizations must understand to what extent the mobility data being shared and the decisions it informs as part of their initiative may hide, reflect, or reinforce systemic biases or have a disproportionate impact on certain groups or communities.[31] Organizations should proactively identify and address potential sources of bias in mobility data itself and in the analytic models and decisions that it informs. Some key considerations include:

- **Sources of bias**. There are many ways bias can find its way into mobility data, including, but not limited to:
  - Incomplete or unrepresentative data. For example, data that has been crowdsourced from social media platforms or newer transportation modes may not be representative of all demographics, such as older persons or recent migrants [27]. If this data were used to generate insights about the larger population or provide a service at scale, it might inadvertently favor younger populations or those who may already have more transportation options.
  - Historical data. Data collected in the past may no longer accurately reflect reality. For example, surveys and analytic models about average commute times that were designed decades ago by car-centric organizations may not accurately account for new mobility modes (including e-scooters, bikes, and ride-hail options).
  - Through algorithm development. In addition to considering data sources, bias can arise from algorithmic design decisions. Bias can emerge when optimization strategies and the choice of algorithm do not account for the expertise of the subject matter experts or for the affected communities (e.g., if there is a lack of diversity among the programmers designing the training sample) [28].
- **Targeted and personalized services**. There may also be situations where a data sharing initiative intends to impact a particular group or community, such as by offering special services or incentives to transit riders or carpoolers. Sharing mobility data to develop products and services for certain groups is not necessarily unethical, but organizations should be careful to consider whether this sort of segmentation may lead to discriminatory outcomes (e.g., segregation of services based on certain characteristics) and whether individuals would consider personalization based on things like race or income level as unfair.

[29] For example, mobility data sharing initiatives meant to address the challenges of providing mobility services in rural communities and to older adults [86].
[30] For example, by creating a gender action plan, adopting anti-racism data standards, or engaging external experts to conduct an equity assessment of an organization's data sharing practices [25, pp. 7-8] [88] [87].
[31] Organizations should be aware that, in some cases, activities that have disparate impacts against protected classes of people may trigger legal liability under civil rights laws. (See Section 2.)

- **Tools and strategies for identifying and mitigating bias**. There is a growing field of research and tools (as well as law) aimed at helping organizations identify and mitigate systematic bias in data and disparate impacts. Organizations should consider how to incorporate these new resources into their workflows, including: specialized algorithmic impact assessments (AIAs) to help organizations or agencies assess and mitigate the impacts associated with deploying an automated decision system [29]; making data sources and algorithms open source so that independent researchers can provide feedback on whether the bias or inaccuracy is systematic and the potential implications [27]; and providing mechanisms for individuals who are substantially affected by algorithmically driven decisions to appeal or request human review over those decisions.[32]

Function or scope creep. Finally, organizations should be mindful of function or scope creep when sharing mobility data. Function creep is the expansion of a system or technology beyond its original purposes [30] and often occurs gradually over time. Allowing substantial creep of scope would eventually cause the data sharing initiative to fall out of line with what was learned and communicated during stakeholder engagement and may also push the initiative outside of any consent obtained and agreements signed. Although function creep does not always result in negative impacts, it is contrary to privacy and ethical principles and can undermine public trust in organizations and initiatives. Some key considerations in assessing function creep include:

- **Clear objectives**. Having a well-defined and properly scoped objective before sharing mobility data can help to mitigate against function creep [31]. When organizations define their mobility data sharing objectives, it is often a good idea to have foreseeable and compatible purposes in scope as well. This is especially relevant where algorithms are applied to big mobility datasets to explore new or unanticipated correlations within mobility patterns.
    - Example: If a researcher is studying traffic collisions to advance pedestrian safety goals, it may be foreseeable that eventually her work could expand to include near-misses as well. If this is considered and addressed before the initiative launches, it is not likely to create function creep. However, if that expanded work were to require new types of mobility data with significantly different privacy impacts (e.g., real-time location data in addition to historical location data), a new Mobility Data Sharing Assessment may be required.
    - Example: If a city DOT is receiving mobility data to enforce micromobility service operators' compliance with city permits, it may not be foreseeable that changing political priorities by city council would redirect the agency's focus to enforcing rider compliance with traffic laws or other ordinances. Unless there are opportunities for public debate and discussion prior to the change, this likely constitutes function or scope creep and a new Mobility Data Sharing Assessment would need to be conducted.
- **Process safeguards**. The best way for organizations to safeguard against function creep is to institute clear processes. For example, organizations are encouraged to develop lists of all of the foreseeable ways that mobility data may be used within their data sharing initiatives—including if there are any uses that are clearly out of bounds from the outset (e.g., selling mobility data to law enforcement or immigration authorities). Organizations can also provide opportunities for public notice and discussion before a new partner joins an initiative or data is shared for new purposes and apply privacy by design and privacy by default safeguards to safeguard against function creep.

[32] Note that in some jurisdictions, human review and appeals may be required by law. For example, refer to Articles 15 and 22, GDPR [104].

4. IMPACTS: DO THE BENEFITS OUTWEIGH THE PRIVACY RISKS?

Once organizations have determined that it is both lawful and fair for mobility data to be shared, they must next determine whether the benefits of the data sharing outweigh the privacy risks. In order to do this, organizations must identify the anticipated privacy risks and anticipated data benefits of their mobility data sharing initiative, consider any safeguards or controls that can be applied to mitigate those risks, and ultimately determine if the initiative achieves an appropriate balance of privacy risks and data benefits given the circumstances. The approach in this section builds upon risk management frameworks from the U.S. National Institute for Standards and Technology (NIST) [1] [32] and is designed to be flexible, interoperable, and technology-neutral.

- **Practice tip**: Privacy risks can be more subjective than objective, and tend to be:[33]
  - Incremental. As mobility data sharing grows over time, so does the likelihood of a data breach, a successful re-identification attack, or a discriminatory impact on vulnerable or historically marginalized communities.
  - Inequitable. Privacy risks may accrue unevenly throughout society. If not addressed in advance, some community members may reap the benefits of mobility data sharing, while others bear the privacy burden.
  - Non-obvious. Certain privacy risks are more impactful or more likely to occur for particular groups, and can be overlooked by program designers who have not specifically incorporated those individuals' inputs. For example, publishing information about a ride-hail trip that ends at a busy bus terminal will create a lower privacy risk than publishing a trip that ends at a rural single-family home.
  - Intrusive. Privacy is closely tied to feelings about self-control and autonomy, and its real—or perceived—loss can leave people feeling vulnerable, exposed, and without control of their own lives. In these situations, individual and community behavior can be chilled, relationships harmed, and trust lost.

4.1 Evaluate the Privacy Risks

Mobility data—especially precise location or trip information—is often considered sensitive and identifiable information, and thus may pose a significant risk to individual privacy.[34] In particular, there are widely acknowledged privacy risks to certain types of data (e.g., information about individuals' religious, financial, or health status, etc.) and types of people (e.g., vulnerable individuals, such as children).[35] Determining privacy risk is a highly contextual process. It depends on many factors, including the nature of the data at issue, the purposes for which data will be used, and any applicable technical, legal, and organizational safeguards and controls. Organizations will need to determine the anticipated privacy risks arising from each mobility data sharing initiative on a case-by-case basis.

[33] Adapted from FPF's *Nothing to Hide Toolkit* [47].
[34] "Mobility data is among the most sensitive data currently being collected. Mobility data contains the approximate whereabouts of individuals and can be used to reconstruct individuals' movements across space and time… While in the past, mobility traces were only available to mobile phone carriers, the advent of smartphones and other means of data collection has made these broadly available." [12]
[35] See Appendix B for more information about the privacy risks related to location data.

Identify and describe anticipated privacy risk considerations. Some key considerations for organizations as they work to identify potential privacy risks include:

- **Specific risks**. Rather than reinvent the wheel, organizations may find it helpful to consult structured lists from scholars and standards organizations to develop a more consistent, comprehensive understanding of potential privacy impacts.[36] However, organizations should be careful to focus their Mobility Data Sharing Assessments on risks that arise specifically from the sharing of mobility data. Two of the most prominent concerns expressed by communities, advocates, and academics regarding mobility data sharing are the risks of (a) re-identification,[37] and (b) data misuse by data recipients.
- **Differences by sector**. There are certain privacy risks that are unique to the type of organization involved. For example, government agencies are often subject to transparency obligations or data sharing commitments that private companies are not, such as Public Records or Freedom of Information Acts, Open Data commitments, or intra- and inter-governmental data sharing commitments. On the other hand, companies may have their own unique obligations to share data, such as during mergers and acquisitions or to comply with regulatory reporting requirements. Academics, too, may face their own requirements to share data in support of research reproducibility goals or as part of institutional review processes.

Qualitatively measure privacy risks. It is challenging to measure privacy risks and organizations are encouraged to take advantage of existing guidance by data protection authorities, scholars, and other privacy experts. For example, EU and Canadian authorities have each published detailed guidance and examples of "high risk" data processing relevant for both public and private sector organizations and have also published lists of "No-Go Zones" [33] [34] [35]. See Table 1 for a summary preliminary risk spectrum.

[36] For example, NIST has developed a Catalog of Problematic Data Actions and Problems [89] and leading privacy scholars such as Daniel Solove and Danielle Citron have developed a Taxonomy of Privacy Harms [90].
[37] See Appendix B for more.

Table 1. Preliminary risk spectrum 9
(Adapted from Office of the Privacy Commissioner of Canada, OPC's Guide to the Privacy Impact Assessment Process, Preliminary risk assessment and Article 29 working party of EU data protection authorities (WP29) published guidelines with nine criteria which may act as indicators of likely high risk processing [36] [37])

Lower Risk	Higher Risk
Involves limited personal data.	Involves a large amount of personal data or data processing on a large scale.
Involves the personal data of a few individuals.[38]	Involves the personal data of many individuals.
Does not involve the personal data of vulnerable or marginalized populations.	Involves the personal data of one or more vulnerable or marginalized populations.
Does not involve sensitive personal data, location data, or trip data.	Involves sensitive personal data, location data, or trip data.
The context is not sensitive.	The context is sensitive.
Has a minimal impact on individuals or communities.	Has a major impact on individuals or communities.
Involves a one-time or short-term initiative.	Involves a longer-term initiative.
Does not involve profiling, evaluation or scoring of individuals.	Involve profiling, evaluation, or scoring of individuals.
Does not involve automated decision-making with legal or similar significant effect.	Involves automated decision-making with legal or similar significant effect.
Does not involve systematic monitoring (of a publicly accessible area on a large scale).	Involves systematic monitoring (of a publicly accessible area on a large scale).
Does not involve an innovative use or application of new technologies or organizational solutions.	Involves an innovative use or application of new technologies or organizational solutions.
Is not likely to prevent an individual from exercising a right or using a service or contract.	Is likely to prevent an individual from exercising a right or using a service or contract.

4.2 Evaluate the Benefits

Mobility data is especially beneficial information to governments, companies, and researchers as they seek access to support improvements to transportation infrastructure and develop transportation policy. Mobility data can be used to help better understand and regulate mobility systems at a local level, and develop and deploy new mobility services, including micromobility, contactless delivery options, and electric and automated vehicles.

Determining the benefits of a mobility data sharing initiative is a highly contextual process. Organizations will need to consider a range of factors, including: the nature of the benefits, the range of beneficiaries, and the size and scope of the benefits, as well as the likelihood that the benefits can be achieved. Organizations will need to determine the anticipated benefits arising from each mobility data sharing initiative on a case-by-case basis.

[38] On the other hand, note that the re-identification and personal information exposure is higher, qualitatively, for small samples.

Identify anticipated or intended data benefits. Some key considerations for organizations as they work to identify potential data benefits include:

- **Objectives and use cases**. Organizations should already have clearly defined objectives by this stage in the assessment process (see 1.2). In addition to the resources found within this operator's manual, organizations can look to business plans, strategic reports, mission statements, company goals, and use case libraries to further flesh out their goals.
- **The nature and scope of the benefit**. The benefits of a data sharing initiative may be narrow and focused (e.g., expanding last-mile mobility services to an underserved community or measuring the performance of a particular policy initiative) or wide-ranging (e.g., eliminating pedestrian deaths or incentivizing the use of electric vehicles). The benefit may also be short-term (e.g., a 6-month limited pilot) or longer-term, requiring years or decades to be achieved (e.g., climate-change goals). Organizations should be open and transparent about how the initiative's progress will be measured—especially where the benefits will not be obvious in the short-term.
- **The range of potential beneficiaries**. Organizations should consider what benefits the data sharing initiative might have not only to the individuals whose mobility data is shared and to the organizations that are sharing the information, but also to other organizations, communities, and society at large. Organizations should aim to not only be transparent about the initiative's intended benefits, but which groups are expected to experience those benefits.

Qualitatively measure data benefits. It can be equally challenging to measure data benefits as it is to measure data risks. Some key considerations include:

- **Evidence that benefits will occur**. Ideally, organizations will have already gathered evidence about the likelihood that the initiative will achieve its objectives (see 1.2). For example, organizations engaging in a new mobility data sharing initiative might develop specific performance metrics, commission an independent report, or look to industry research or forecasts to determine what success will look like. Organizations may also want to develop their own reporting tools and lessons learned to use as benchmarks for future mobility data sharing initiatives.
- **Culture-specific preferences and priorities**. Since every city and community is different, mobility data sharing assessments must take culture-specific differences into account as they evaluate the likelihood and impact of a particular benefit. Some societies may place a high value on individual benefits, while others give greater weight to community values [38]. Organizations should be mindful of these differences, but aware that the status quo may change over time [39] [40].

4.3 Weigh the Risks Against the Benefits

Finally, organizations must determine whether the mobility data sharing initiative achieves an appropriate balance of data benefits and privacy risks. Organizations should consider the privacy risk tolerance of the stakeholders involved, the overall benefit of the mobility data sharing, and the operational resources available to mitigate privacy risks. Structured assessment methods, such as MDSA tool Section 4 can help organizations visualize these components. Organizations may need to apply additional safeguards and repeat this assessment process several times until the desired balance of benefits and risks is achieved.

Holistic assessment of the benefits and risks. Ultimately, organizations will need to consider privacy risks and data benefits holistically. Some key considerations include:

- **No definitive rules**. There are no definitive rules on what degree or probability of benefit is needed to overcome presumptions against creating privacy risk. For example, a mere assertion that a product or service can be improved by sharing mobility data is likely not sufficient justification to impose privacy risks on individuals; yet proof that a benefit will occur beyond any doubt is not a reasonable standard either. Instead, organizations will need to determine the right balance in the context of their particular initiative.
- **Risk tolerance and values**. Even within a single mobility data sharing initiative, organizations and other stakeholders may have different risk tolerance and values, given their organizational missions, structures, and legal requirements. It is important for organizations to understand that privacy risk will never be zero, and to be transparent about what they consider to be an acceptable level of privacy risk. Organizations can also seek guidance or feedback from the community and the public, as well as from privacy commissioners or data protection authorities, regarding acceptable levels of privacy risk [36] [41].
- **Equity**. Organizations should also consider that the benefits and risks of a mobility data sharing initiative may accrue unevenly throughout society. If this is not addressed in advance, some community members may reap the benefits of mobility data sharing while others bear the privacy burden [42].

Additional mitigations or safeguards. Where organizations determine that the residual privacy risk is too high in relation to the benefits of their initiative, they are encouraged to apply additional safeguards to mitigate the risks and repeat the assessment until a satisfactory balance has been reached. Organizations should be realistic about what safeguards are practical for a given data sharing initiative, taking into account both the operational costs to implement the safeguards effectively and whether any safeguards require a tradeoff between privacy protection and data quality[39] (see Section 5).

- Example: When an initiative poses low overall privacy risk and offers substantial overall benefits, organizations may be satisfied that they have struck an appropriate balance. Even in these situations, however, organizations may want to further mitigate the privacy risks or increase their transparency and accountability through other safeguards (e.g., additional internal training or transparency reports).
- Example: When an initiative's overall benefits and risks are equal but there remains a high residual privacy risk (i.e., the risk remaining after controls have been applied), organizations will typically apply additional safeguards to further mitigate the privacy risk and repeat the assessment process (e.g., de-identifying sensitive fields, restricting access to mobility data, or completing data audits).
- Example: When an initiative's overall privacy risks substantially outweigh its overall benefits, organizations are strongly recommended to apply additional safeguards to further mitigate the privacy risk and repeat the assessment process (e.g., permanently deleting sensitive data fields, auditing vendors' and partners' data practices, or consulting independent privacy and security experts).

[39] For example, there is a data utility and privacy risk tradeoff and the question about how to preserve user privacy while still yielding useful information and insights is the subject of much research [7].

Countervailing public policy goals or legal considerations. In some circumstances, a mobility data sharing initiative that poses a moderate or high privacy risk may still be worth pursuing where it is supported by compelling legal or public policy considerations (e.g., where sharing mobility data is required by law, or is necessary for government transparency, equity, or public safety). Organizations that intend to rely on these grounds to justify a particular mobility data sharing initiative must be transparent and accountable in their decision-making and should clearly document why less privacy-invasive alternatives are not suitable to meet a specific need.

5. CONTROLS: KEY PRIVACY SAFEGUARDS FOR DATA SHARING

There are a wide variety of technical, organizational, and legal safeguards and controls to mitigate privacy risks. This section highlights eight key safeguards that are especially important and relevant to sharing mobility data. However, this list is not comprehensive or prescriptive,[40] and organizations must determine for themselves which safeguards to apply to maximize the benefits of mobility data while minimizing privacy harms.

Practice tips:

- **Determine what safeguards to apply on a case-by-case basis**. Organizations should consult a range of experts (including technical, legal, business, and other stakeholders) to determine the appropriate set of safeguards for a particular mobility data sharing initiative. What is appropriate will depend on the nature and sensitivity of the mobility data involved, the purpose and context of the sharing, and the capabilities and resources of the organizations involved.
- **Implement a combination of technical**, **legal**, **and organizational safeguards**. Relying on only one type of safeguard may leave initiatives vulnerable to privacy or security threats. Organizations should use a combination of technical, legal, and organization controls to most effectively protect mobility data.
- **Implement safeguards at all stages of the data lifecycle**. While a Mobility Data Sharing Assessment specifically addresses mobility data sharing, it is important for organizations to consider safeguards applicable throughout the entire data lifecycle, from when mobility data is collected, used, processed, shared, and stored, until it is ultimately disposed of.

5.1 Data Minimization

Organizations may share mobility data for a variety of important public purposes, including to provide mobility services, enforce safety regulations, and advance equitable and sustainable transportation policy measures. Nevertheless, the most effective way to reduce privacy risk (as well as data storage and compliance costs) is to not obtain and store unnecessary data. "Data minimization" means that initiatives should only share and use mobility data that is necessary to fulfill their specific objectives, and that the data they share is adequate, relevant, and limited to what is necessary in relation to the objective.

For examples of recommended safeguards, see MDSA tool Section 5. For discussions of relevancy and proportionality, data quality, and fairness, see Section 3.

[40] For a more comprehensive list of privacy safeguards and controls, refer to NIST SP 800-53B *Control Baselines for Information Systems and Organizations* [106].

Appropriate data minimization must be determined on a case-by-case basis. Key considerations for data minimization include:

- **No objective standards**. Determining whether mobility data is adequate, relevant, and limited to what is necessary for a particular initiative is highly contextual and must be assessed on a case-by-case basis. Sometimes this may mean organizations decide to share less mobility data, or to share less identifiable mobility data. Organizations are encouraged to consult external experts when deciding what mobility data to share, and to develop internal guidelines and standards to inform future initiatives.
- **Avoid over-minimization**. While data minimization is a foundational privacy principle and an important safeguard, organizations should not be afraid to share mobility data for beneficial purposes. Organizations should consider their objectives for sharing mobility data and whether not sharing certain types of mobility data could defeat those goals or even result negatively impact certain individuals or communities. For example, if race or socioeconomic data is not included in a mobility dataset, organizations may not be able conduct bias audits or identify potential disparate impacts [43].
- **Privacy by Design**. Organizations should also leverage technical, organizational, and design tools to achieve appropriate data minimization and encourage privacy by design and default [44]. For example, organizations can minimize data sharing by implementing opt-in versus opt-out consent in consumer apps, randomly rotating persistent device identifiers, purging identifiable data after a short period of time, or obtaining non-identifiable mobility data by default.

5.2 Transparency

Organizations should be clear, open, and honest with impacted individuals and communities about how mobility data will be shared and used. Transparency is critical both before mobility data is shared and throughout the data sharing initiative. Even when one organization in an initiative does not have a direct relationship with individuals and communities, it can support and encourage transparency efforts by the others.

For examples of recommended safeguards, see MDSA tool Section 5. For discussions of transparency and fairness, see Section 3.

Meaningful transparency. Organizations should be mindful of privacy laws that prescribe certain types of notice and should consult with legal counsel where applicable. Some key considerations for transparency include:

- **Informed individuals**. Organizations must provide individuals whose data will be shared as part of a mobility data sharing initiative with enough information to understand the initiative's objectives, potential privacy impacts, and safeguards. Individuals should be able to make an informed decision about whether they want to use a mobility service or engage with an organization that participates in the mobility data sharing initiative, including understanding how their data will be shared and with whom it will be shared [45].
- **Informed communities**. Other impacted communities must also be provided with tools and opportunities to understand the initiative's goals and mobility data sharing activities, and to have their perspectives heard and considered. For example, through online and physical education campaigns and (virtual) town halls.
- **Project lifecycle**. Transparency is important throughout the entire life of the data sharing initiative. Key moments for notice about privacy practices include before an initiative is launched, before mobility data is obtained from an individual, before mobility data is shared, whenever there is a material change to how mobility data is used or shared, and after the initiative concludes.

5.3 Consent and Social License

Where feasible, and as required by law, organizations should obtain meaningful consent from individuals prior to sharing their mobility data. Where individual consent is not possible, as is often the case for mobility data sharing initiatives, organizations should seek to create a "social license" or approval to share mobility data. Some common methods of obtaining social license include community advisory communities, focus groups, crowdsourcing, or public workshops.

For examples of recommended safeguards, see MDSA tool Section 5. For discussions of community engagement and fairness, see Section 3.

Obtaining meaningful consent. Meaningful consent is specific, informed, and freely given. This means that individuals should be asked for their consent to specific data sharing activities, not vague, open-ended, or bundled purposes. Organizations should be considerate of power imbalances (such as between governments and residents), "take it or leave it" scenarios, or coercive user interfaces, all of which can cut against truly voluntary consent. Organizations can also implement processes and tools to respect individual choices (e.g., automated mechanisms for withdrawing consent) and to validate that those individuals have given consent (e.g., receipts or metadata) [46]. Organizations that receive mobility data from third parties should, where applicable, verify that the original party obtained appropriate consents to collect and share the data.

Where consent is not feasible. Obtaining consent can be very challenging in mobility data sharing initiatives, where typically only one party will have a direct relationship with individuals. Organizations should ensure that the initiative is in compliance with applicable privacy laws requiring consent, even when they cannot obtain consent directly. However, where consent is not legally required or feasible, organizations should consider alternative ways of obtaining permission to share mobility data. Some key considerations include:

- **Social license**. When data sharing initiatives cannot feasibly ask every individuals' permission to share mobility data, they might instead look for more general societal permission to operate.[41] Organizations seeking to establish social license for a mobility data sharing activity should create inclusive and meaningful opportunities for public engagement, including opportunities for individuals and impacted communities to substantively contribute to decision-making about privacy protections and appropriate data uses [47].
- **Other authorization**. In some cases, particularly when mobility data is obtained by government entities, consent may not be required if there is a clear legislative mandate or other legal basis for mobility data to be used and shared (e.g., to advance public health and safety goals, or to protect critical infrastructure).
- **User controls**. Even where specific consent is not required or feasible, organizations might be able to provide individuals with privacy choices, such as through controls or settings built into mobility devices, apps, or on organizations' websites.

[41] Social license exists when a project has ongoing approval within the local community and other stakeholders, ongoing approval or broad social acceptance, and—most frequently—ongoing acceptance [92] [93].

5.4 Retention and Disposal

Another important privacy principle is that mobility data should be retained for no longer than necessary to fulfill an initiative's stated objectives, unless otherwise required by law, after which it should be reliably deleted or de-identified.

A retention policy should be the maximum amount of time to keep the data, not the minimum. Retention policies should be specific and relative (e.g., 6 months after the trip ends or 1 year after the user unsubscribes). "As long as necessary" to fulfill undefined or general purposes, for example, is not a robust retention period.

For examples of recommended safeguards, see MDSA tool Section 5. For discussions of community data retention and fairness, see Section 3.

Setting a retention period. Organizations should not take a one-size-fits-all approach to setting retention periods, but should instead consider how long mobility data will be needed to accomplish the initiative's objective(s), whether mobility data must be in identifiable form for the entirety of that time period, and whether there are other contextual factors. More often than not, this means that organizations will have multiple retention periods relevant to a particular initiative.

- Example: If a city agency receives mobility data and uses it as the basis of an administrative or judicial enforcement action against a mobility service provider, it may need to retain that data in its original or "raw" (i.e., unobfuscated) form until any appeals or relevant statutes of limitations have expired.
- Example: If a company receives mobility data and uses it to forecast regional transit trends, it may only need to retain identifiable mobility data for as long as it takes to analyze or aggregate it.
- Example: If a public agency receives mobility data, it may need to retain it in accordance with a statewide record retention schedule.

Disposing of data that is no longer needed. A retention policy should also include details about how mobility data will be handled after it is no longer needed. Organizations should delete or de-identify mobility data using tools proportionate to the sensitivity of the data [48], or contract with professionals if they do not have the appropriate tools to safely dispose of sensitive information on site. It is important that organizations ensure mobility data is disposed of everywhere, including data held by service providers or other members of the initiative, data in backup systems, and data on physical media (e.g., USB drives).

5.5 Limitations on Public Disclosure and Other Onward Transfer of Data

Any disclosure of mobility data to the public or other unrelated entities, such as through open data programs, public records requests, memoranda of understanding (MOUs), administrative subpoenas, warrants, or civil discovery should be carefully controlled to mitigate risks that individuals will be re-identified or have sensitive personal information about them revealed. Mobility data shared with one organization for a particular objective should not be readily accessible to other entities, such as law enforcement or immigration enforcement, without specific controls.

For examples of recommended safeguards, see MDSA tool Section 5.

Open data. Open data initiatives offer many benefits to the public, researchers, and businesses, but if they are not properly implemented, they can reveal sensitive and personal information about individuals. Before organizations publish mobility data openly, they should carefully assess whether the data poses a risk of re-identification and either apply appropriate privacy-enhancing safeguards or refrain from publishing.[42] Organizations should also regularly review any mobility data that is already available on open data platforms for potential privacy concerns.

Public records. Although some public records laws shield identifiable mobility data from public disclosure, in many places they do not. This is due to outdated conceptions of what constitutes "personal information." Public records requests to public agencies holding mobility data have already led to concrete privacy harms for individuals and should not be underestimated [49]. To the extent permitted by law, public agencies should withhold identifiable mobility data from public disclosure. Public agencies should also specifically address mobility data in their internal data classification and public records systems, so that staff understand when and to whom such information can be released.

Law enforcement and other government requests for data. Public and private organizations that receive mobility data should ensure that it is not transferred onward to other entities. Mobility data can be very powerful and revealing, and it is especially important that law enforcement, immigration authorities, and other government agencies provide due process. Organizations who receive requests for mobility data from such entities should consider publishing regular transparency reports discussing the types of requests received and whether data was shared.[43] Organizations can also develop internal guidelines on how and with whom mobility data will be shared, in what formats (e.g., identifiable versus de-identified, summary statistics versus trip-level data), under what conditions (e.g., subject to a warrant), and how staff should determine whether there is a potential privacy risk to individuals or groups of people if mobility data were released.

5.6 Third Party Management

Organizations often share mobility data with third parties, such as vendors, processors, or service providers, to support routine business operations and to maximize the data's utility. Contractual and technical restrictions should be placed on third parties to limit inappropriate secondary use and re-identification of individuals.

For examples of recommended safeguards, see MDSA tool Section 5.

Due diligence and supervision. Organizations should conduct due diligence when selecting vendors and service providers and should be satisfied with the level of privacy protection that these parties will provide before any mobility data is shared. Organizations should also periodically (e.g., annually) review vendor contracts and audit vendors' data practices in order to ensure that mobility data is not being used or handled in inappropriate ways. In the event there is a breach of contract or misuse of data, appropriate remedial actions must be taken.

Contractual restrictions. Organizations should only share mobility data with vendors who agree to implement privacy and security controls appropriate to the sensitivity of the mobility data and to only use mobility data to perform their contracted services. It is crucial that contractual and technical restrictions are put in place to restrict any secondary uses of mobility data or attempts to re-identify individuals. Administrative and technical measures should also be used to audit and enforce these restrictions.

[42] For information on how to complete an open data risk assessment, refer to Kelsey Finch, City of Seattle Open Data Risk Assessment, FPF (January 2018).
[43] Refer to LADOT Data Protection principles, committing to publishing an annual transparency report [48].

5.7 Data Security

Before organizations decide to share mobility data, it is essential to implement data security controls to protect the confidentiality, integrity, and availability of mobility data using appropriate technical, administrative, and physical safeguards. The sensitivity of mobility data may make it a particularly attractive target for criminals, malicious actors, and other prying eyes, including employees who may try to exceed their authorized use for personal gain [50] [51] [52].

For examples of recommended safeguards, see MDSA tool Section 5.

Comprehensive data security programs or strategies. Organizations should coordinate their mobility data sharing practices with their Chief Information Security Officers (CISOs), Chief Technology Officers (CTOs), and other senior security leaders to ensure mobility data is appropriately secured. This might include leveraging existing data security standards and frameworks; developing comprehensive data loss strategies;[44] regularly reviewing and updating data handling guidelines to ensure they reflect current best practices (e.g., password management, two-factor authentication) [48] [53]; engaging independent security researchers; and periodically testing the effectiveness of the key administrative, technical, and physical safeguards protecting mobility data.

Incident management. Organizations must also be prepared to respond to a situation where mobility data is compromised. A well-resourced incident response system includes (1) procedures to identify, manage, and resolve incidents and breaches; (2) defined responsibilities and accountable members of the team; (3) a process to determine the required actions and escalation procedures; (4) a process for complying with applicable breach laws and regulations; (5) a process for employees or third parties with remediation or penalties; and (6) periodic review of incidents and testing on systems, with remediation as needed. Mobility data sharing initiatives should determine in advance which organization(s) involved will lead incident response measures and, where applicable, notify individuals whose data has been breached.

5.8 Use of Privacy Enhancing Technologies

Privacy-enhancing technologies (PETs), such as data de-identification or secure data enclaves,[45] mitigate the privacy risks associated with the identifiability of mobility data, while supporting more robust data sharing and access. PETs are particularly important safeguards in situations where mobility data may be made available to the public, such as through open data platforms or public records requests.

For examples of recommended safeguards, see MDSA tool Section 5. For discussions of location data and re-identification, see Appendix B.

Apply PETs on a case-by-case basis. Organizations must apply PETs on a case-by-case basis to determine the best combination of contractual, organizational, and technical controls for the circumstances. PETs are an active research field, with frequently evolving technical literature and policy guidance, and organizations are encouraged to consult disclosure control experts for specific advice [54] [55] [56]. Some key considerations include:

- **Tradeoffs**. PETs typically involve obfuscating or perturbing data in some way, which means that organizations must make a tradeoff between data quality and privacy: as the amount of privacy protection goes up, the amount of analytic utility in the data goes down [7] [57].

[44] This might include taking into consideration mobility data that resides on devices (notebooks, personal computers, portable storage devices, mobile devices), data in transit, and data backups (including servers, databases, backup media, and storage platforms) [4].
[45] A secure data enclave is a secure, centralized service for researchers that work with sensitive research data.

- **Location data**. Certain types of mobility data, such as location data, require special consideration and expert guidance, especially as the research around de-identification methods and re-identification attacks develops. Organizations sharing location datasets should endeavor to keep on top of research on the spatial and temporal aspects of location data, as well as location privacy protection mechanisms [9] [58] [59].
- **Release methods**. Different PETs may be appropriate depending on the context in which mobility data will be shared. For example, if mobility data will be shared in a controlled environment, traditional statistical disclosure control techniques like redaction can be buttressed with administrative and legal controls (e.g., access controls, training, and contractual prohibitions against re-identification). On the other hand, formal methods for measuring and controlling disclosure risk (e.g., differential privacy, synthetic data) should be used for any mobility data that is released publicly and without additional legal constraints [60].
- **Process and oversight**. Organizations cannot approach PETs with a "set-it-and-forget-it" mindset. Instead, organizations must provide appropriate internal training and education regarding re-identification risk; monitor the effectiveness of PETs over time; and consider how to meet transparency obligations and expectations (including, in some cases, publishing the methodology of any PETs used to safeguard mobility data).

6. MOBILITY DATA SHARING IN PRACTICE

6.1 Ongoing Monitoring and Accountability

After mobility data has been shared, it is important that organizations have a system in place for ongoing monitoring of both the initiative and the mobility data that was shared. It is also crucial that someone at the organization be responsible and accountable for the sharing initiative.

Monitoring and compliance. Accountability does not happen on its own; instead, organizations need to implement clear processes to ensure that policies and procedures are being followed, privacy safeguards remain effective, and the initiative is living up to its privacy commitments. Organizations can support monitoring and compliance through both technical tools (e.g., data audits, access logs, security testing) and organizational controls (e.g., contract reviews, personnel training, penalties for non-compliance). Policies and procedures should be reviewed at least annually to ensure that they are consistent with emerging laws, regulations, and best-in-class privacy and security safeguards. Additionally, organizations can implement internal and external mechanisms to encourage transparency and accountability (e.g., ombudspersons, anonymous complaints tools, bug bounties, or whistleblower protections).

Resourcing and expertise. Organizations that wish to strengthen or streamline their mobility data sharing initiatives must invest in privacy resources and expertise. Privacy teams can provide return on investment to both public and private sector organizations, when given the appropriate leadership, tools, and training [61] [62]. For example, organizations should designate a senior official with appropriate authority to implement and oversee privacy protections across the organization or initiative, to help ensure more consistent and confident decision-making about privacy. A privacy team can also ensure that all personnel who handle confidential or identifying mobility data receive appropriate training, monitor partners and vendors, conduct privacy risk assessments, engage with stakeholders, and support the organization's strategic goals.

7. CONCLUSION

7.1 Navigating the Road Ahead

The MDSA process is an important first step on a journey towards responsible mobility data sharing, not the end of the road. Organizations should also use the MDSA process as an opportunity to build a concrete foundation to support future mobility data sharing. Some recommendations include:

- Leverage the MDSA process to strengthen relationships and institutional knowledge around mobility data and privacy within the organization, with partners, customers, and communities.
- Use the MDSA questions and documentation to develop or adapt contract terms, data sharing agreements, or privacy impact assessments to the mobility data context.
- Publish MDSAs (or a summary version) to promote greater transparency and accountability or to engage stakeholders in discussions about the importance of mobility data sharing.
- Build on the MDSA process to develop additional resources and best practices.

The MDSA process is intended to help organizations share mobility data in responsible and privacy-respecting ways while meeting their compliance, planning, research, or innovation goals. The MDSA operator's manual is only one piece of the toolkit: organizations are also encouraged to take advantage of the additional insights and resources in the accompanying MDSA tool and MDSA infographic and to stay up to date on privacy developments for mobility data by visiting www.mobilitydatacollaborative.org and fpf.org.

8. ABOUT THE MOBILITY DATA COLLABORATIVE™

The Mobility Data Collaborative™ serves as a neutral forum for cross-sector collaboration. Its goal is to convene leading mobility partners from public and private sectors to develop a framework of best practices to support effective and secure mobility data sharing.

Vision: Support the effective sharing of mobility data with public agencies to support safe, equitable, and livable streets for all.

Mission: Fostering a collaborative, cross-sectoral, and productive forum that reconciles public agencies' policy and regulatory goals with industry capabilities and consumer interests. This includes establishing a data-sharing framework that provides public benefit while protecting consumer privacy.

9. CONTACT INFORMATION

To learn more about the Mobility Data Collaborative™, please visit www.mobilitydatacollaborative.org.

Contact: mobilitydatacollaborative@sae-itc.org

10. ABOUT FUTURE OF PRIVACY FORUM™

The Future of Privacy Forum™ is a nonprofit organization that serves as a catalyst for privacy leadership and scholarship, advancing principled data practices in support of emerging technologies. FPF brings together industry, academics, consumer advocates, and other thought leaders to explore the challenges posed by technological innovation and develop privacy protections, ethical norms, and workable business practices. More information about FPF can be found at www.fpf.org.

11. ACKNOWLEDGEMENTS

The Mobility Data Collaborative would like to acknowledge Chelsey Colbert and Kelsey Finch at the Future of Privacy Forum™ for coordination, facilitation, and writing of the Mobility Data Sharing Assessment.

The Mobility Data Collaborative would also like to acknowledge the contributions of the member organizations during the development of the Mobility Data Sharing Assessment.

- City of Tallahassee
- Lyft
- Miami-Dade County
- Uber

Additional acknowledgement is extended to John Verdi, Vice President of Policy, FPF; Sara Jordan, Senior Counsel, FPF; Stacey Gray, Senior Counsel, FPF; Marcus Dessalgne, FPF Law and Policy Fellow; and Kathryn Earles, FPF Law and Policy Intern. Their contributions helped shape the MDSA operator's manual.

12. REFERENCES

12.1 Applicable Documents

The following publications were referenced during the development of this document. Where appropriate, documents are cited.

12.1.1 SAE Publications

Unless otherwise indicated, the latest issue of SAE publications shall apply. Available from SAE International, 400 commonwealth Drive, Warrendale, PA 15096-0001, Tel: 877-606-7323 (inside USA and Canada) or +1 724-776-4970 (outside USA), www.sae.org.

MDC00001202004 Guidelines for Mobility Data Sharing Governance and Contracting

MDC00003202108 Mobility Data Sharing Assessment Infographic

MDC00003202108 Mobility Data Sharing Assessment Tool

12.1.2 Other Documents

[1] National Institute of Standards and Technology, U.S. Department of Commerce, "NIST Privacy Framework," January 2020. [Online]. Available: https://www.nist.gov/privacy-framework. [Accessed 15 February 2021].

[2] A. Farzanehfar, F. Houssiau and Y.-A. de Montjoye, "The risk of re-identification remains high even in country-scale location datasets," 12 March 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2666389921000143.

[3] M. Abrams, "The Orgins of Personal Data and its Implications for Goverance," 21 March 2014. [Online]. Available: https://ssrn.com/abstract=2510927 or http://dx.doi.org/10.2139/ssrn.2510927.

[4] Infocomm Media Development Authority of Singapore (IMDA) and Personal Data Protection Commission (PDPC), "Trusted Data Sharing Framework," 2019. [Online]. Available: https://www.imda.gov.sg/-/media/Imda/Files/Programme/AI-Data-Innovation/Trusted-Data-Sharing-Framework.pdf.

[5] S. L. Garfinkel, "De-Identification of Personal Information (NISTIR 8053)," October 2015. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf.

[6] European Data Protection Board (EDPB), "Guidelines 01/2020 on processing personal data in the context of connected vehicles and mobility related applications, Version 2.0," 9 March 2021. [Online]. Available: https://edpb.europa.eu/our-work-tools/our-documents/guidelines/guidelines-012020-processing-personal-data-context-connected_en.

[7] D. Calacci, A. Berke, K. Larson and A. Pentland, "The Tradeoff between the Utility and Risk of Location Data and Implications for Public Good," 11 December 2019. [Online]. Available: https://arxiv.org/pdf/1905.09350.pdf.

[8] S. Gray, C. Colbert, P. Sanderson, K. Ringrose and S. Jordan, "Future of Privacy Forum: Closer Look at Location Data Privacy and Pandemics," 25 March 2020. [Online]. Available: https://fpf.org/blog/a-closer-look-at-location-data-privacy-and-pandemics/. [Accessed 1 February 2021].

[9] B. Liu, W. Zhou, T. Zhu, L. Gao and Y. Xiang, "Location Privacy and Its Applciations: A Systematic Study," IEEE Access, vol. 6, pp. 17606-17624, 2018.

[10] National Association of City Transportation Officials (NACTO) and International Lawyers Association, "Managing Mobility Data," April 2019. [Online]. Available: https://nacto.org/wp-content/uploads/2019/05/NACTO_IMLA_Managing-Mobility-Data.pdf. [Accessed 1 March 2021].

[11] C.-Y. Chow and M. F. Mokbel, "Privacy of Spatial Trajectories In: Computing with Spatial Trajectories," New York, NY, Springer, 2011, pp. 109-141.

[12] Y.-A. de Montejoy, C. A. Hidalgo, M. Verleysen and V. D. Blondel, "Unique in the Crowd: The privacy bounds of human mobility," Scientific Reports, vol. 3, p. 1376, 2013.

[13] Information Commissioner's Office, "Principle (c): Data minimisation," [Online]. Available: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/principles/data-minimisation/. [Accessed 1 March 2021].

[14] European Data Protection Supervisor, "Necessity & Proportionality," [Online]. Available: https://edps.europa.eu/data-protection/our-work/subjects/necessity-proportionality_en. [Accessed 1 March 2021].

[15] European Data Protection Supervisor, "The EDPS quick-guide to necessity and proportionality," [Online]. Available: https://edps.europa.eu/data-protection/our-work/publications/factsheets/edps-quick-guide-necessity-and-proportionality_en. [Accessed 1 March 2021].

[16] Wired, "Potholes and Big Data: Crowdsourcing Our Way to Better Government," 2018. [Online]. Available: https://www.wired.com/insights/2014/03/potholes-big-data-crowdsourcing-way-better-government/. [Accessed 1 March 2021].

[17] European Commission, "EU data protection rules," May 2108. [Online]. Available: https://ec.europa.eu/info/law/law-topic/data-protection/eu-data-protection-rules_en.

[18] Bryan Cave Leighton Paisner, "California Privacy Rights Act of 2020," 2020. [Online]. Available: https://ccpa-info.com/california-consumer-privacy-act-full-text/. [Accessed 1 February 2021].

[19] G. Wu, "How do you find research participants and improve your response rate?," Code for Canada, 31 March 2021. [Online]. Available: https://codefor.ca/blog/how-do-you-find-research-participants-and-improve-your-response-rate/?s=09. [Accessed 1 May 2021].

[20] P. G. Kelley, J. Bresee, L. F. Cranor and R. W. Reeder, "A "Nutrition Label" for Privacy," 15-17 July 2009. [Online]. Available: https://cups.cs.cmu.edu/soups/2009/proceedings/a4-kelley.pdf. [Accessed 1 March 2021].

[21] "Digital Transparency in the Public Realm," DTPR, [Online]. Available: https://dtpr.helpfulplaces.com/. [Accessed 1 March 2021].

[22] G. Bourke, "Equality in mobility matters now more than ever," 2 July 2020. [Online]. Available: https://www.projectsbyif.com/blog/equality-in-mobility-matters-now-more-than-ever/.

[23] G. Bourke, "What privacy preserving techniques make possible: for transport authorities," 30 July 2020. [Online]. Available: https://www.projectsbyif.com/blog/what-privacy-preserving-techniques-make-possible-transport-authorities/.

[24] G. Bourke, "What privacy preserving techniques make possible: for mobility providers," 6 August 2020. [Online]. Available: https://www.projectsbyif.com/blog/what-privacy-preserving-techniques-make-possible-mobility-providers/.

[25] Toronto Transit Commission, "TTC Status Update - Ombudsman Recommendations," 2019 September 2019. [Online]. Available: https://www.ttc.ca/About_the_TTC/Commission_reports_and_information/Commission_meetings/2019/September_24/Reports/19_TTC_Status_Update_Ombudsman_Recommendations.pdf.

[26] M. Heller, "Why the next step for antiracism is transportation," World Economic Forum, 22 April 2021. [Online]. Available: https://www.weforum.org/agenda/2021/04/transport-us-antiracism.

[27] Open Data Institute, "Helping organisations navigate ethical concerns in their data practices," September 2017. [Online]. Available: https://www.scribd.com/document/358778144/ODI-Ethical-Data-Handling-2017-09-13.

[28] N. Turner Lee, "Detecting racial bias in algorithms and machine learning," Journal of Information, Communication and Ethics in Society, vol. 16, no. 3, pp. 252-360, 2018.

[29] Government of Canada, "Algorithmic Impact Assessment v0.9," 22 March 2021. [Online]. Available: https://canada-ca.github.io/aia-eia-js/.

[30] B.-J. Koops, "'The Concept of Function Creep', 13 Law, Innovation and Technology (1) (Forthcoming)," 3 March 2021. [Online]. Available: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3547903.

[31] "Article 29 Data Protection Working Party, Guidelines on consent under Regulation 2016/679, adopted 29 November 2017, revised 11 April 2018," European Commission, Brussels, 2018.

[32] National Institute of Standards and Technology, "SP 800-63 Digital Identity Guidelines," U.S. Department of Commerce, [Online]. Available: https://pages.nist.gov/800-63-3/.

[33] International Association of Privacy Professionals (iapp), "EU Member State DPIA Whitelists, Blacklists and Guidance," [Online]. Available: https://iapp.org/resources/article/eu-member-state-dpia-whitelists-and-blacklists/. [Accessed 1 March 2021].

[34] Office of Privacy Commissioner of Canada, "Guidance on inappropriate data practices: Interpretation and application of subsection 5(3)," May 2018. [Online]. Available: https://www.priv.gc.ca/en/privacy-topics/collecting-personal-information/consent/gd_53_201805/. [Accessed 1 March 2021].

[35] European Commission, "Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for purposes of Regulation 2016/679 wp248rev.01," 4 October 2017. [Online]. Available: https://ec.europa.eu/newsroom/article29/items/611236.

[36] Office of the Privacy Commissioner of Canada, "Expectations: OPC's Guide to Privacy Impact Assessment Process," March 2020. [Online]. Available: https://www.priv.gc.ca/en/privacy-topics/privacy-impact-assessments/gd_exp_202003/. [Accessed 1 March 2021].

[37] European Commission, "Guidelines on Data Protection Impact Assessment (DPIA) (wp248rev.01)," 4 October 2017. [Online]. Available: https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=611236. [Accessed 1 March 2021].

[38] Future of Privacy Forum, "Big Data: A Benefit and Risk Analysis," 11 September 2014. [Online]. Available: https://fpf.org/blog/big-data-a-benefit-and-risk-analysis/. [Accessed 1 March 2021].

[39] International Association of Privacy Professionals (iapp), "Why China's cultural attitudes toward privacy may be in flux," 8 September 2016. [Online]. Available: https://iapp.org/news/a/why-chinas-cultural-attitudes-toward-privacy-may-be-in-flux/. [Accessed 1 March 2021].

[40] International Association of Privacy Professionals (iapp), "Do You Care About Chinese Privacy Law? Well, You Should," Janauary 2015. [Online]. Available: https://iapp.org/news/a/why-chinas-cultural-attitudes-toward-privacy-may-be-in-flux/. [Accessed 1 March 2021].

[41] Office of the Commissioner of Canada, "Organizational Structure," [Online]. Available: https://www.priv.gc.ca/en/about-the-opc/who-we-are/organizational-structure/. [Accessed 1 March 2021].

[42] Automotive News, "Podcast: Anthony Foxx on building the long road to transportation equity (Episode 96)," 9 May 2021. [Online]. Available: https://www.autonews.com/shift-podcast-about-mobility/anthony-foxx-building-long-road-transportation-equity-episode-96. [Accessed 20 May 2021].

[43] M. Andrus, E. Spitzer, J. Brown and A. Xiang, ""What We Can't Measure, We Can't Understand": Challenges to Demographic Data Procurement in the Pursuit of Fairness," 21 January 2021. [Online]. Available: https://arxiv.org/pdf/2011.02282.pdf.

[44] A. Chavoukian, "Privacy by Design: The 7 Foundational Principles, Implementation and Mapping of Fair Information Practices," May 2010. [Online]. Available: https://www.ipc.on.ca/wp-content/uploads/resources/pbd-implement-7found-principles.pdf. [Accessed 2021 March 1 2021].

[45] Information Commissioner's Office (ICO), "Right to be informed," [Online]. Available: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/right-to-be-informed/. [Accessed 1 March 2021].

[46] Office of the Privacy Commissioner, "Guidelines for obtaining meaningful consent," May 2018. [Online]. Available: https://www.priv.gc.ca/en/privacy-topics/collecting-personal-information/consent/gl_omc_201805.

[47] Future of Privacy Forum, "NOTHING TO HIDE: Tools for Talking (and Listening) About Data Privacy for Integrated Data Systems," October 2018. [Online]. Available: https://fpf.org/wp-content/uploads/2018/09/FPF-AISP_Nothing-to-Hide.pdf. [Accessed 1 March 2021].

[48] National Institute for Standards and Technology, "Security and Privacy Controls for Information Systems and Organizations," U S Department of Commerce, 10 December 2020. [Online]. Available: https://csrc.nist.gov/publications/detail/sp/800-53/rev-5/final.

[49] A. Hern, "New York taxi details can be extracted from anonymised data, researchers say," The Guardian, 27 June 2014. [Online]. Available: https://www.theguardian.com/technology/2014/jun/27/new-york-taxi-details-anonymised-data-researchers-warn.

[50] S. Gurman, "AP: Across US, police officers abuse confidential databases," AP News, 28 September 2016. [Online]. Available: https://apnews.com/699236946e3140659fff8a2362e16f43/ap-across-us-police-officers-abuse-confidential-databases.

[51] D. Maass and J. Gillula, "What You Can Learn from Oakland's Raw ALPR Data," Electronic Frontier Foundation, 21 January 2015. [Online]. Available: https://www.eff.org/deeplinks/2015/01/what-we-learned-oakland-raw-alpr-data.

[52] A. Goldman and M. Apuzzo, "With cameras, informats, NYPD eyed mosques," AP in the News, 23 February 2012. [Online]. Available: https://www.ap.org/ap-in-the-news/2012/with-cameras-informants-nypd-eyed-mosques.

[53] L. Cranor, "Time to rethink mandatory password changes," Federal Trade Commission (FTC), 2 March 2016. [Online]. Available: https://www.ftc.gov/news-events/blogs/techftc/2016/03/time-rethink-mandatory-password-changes.

[54] "Federal Data Strategy: 2020 Action Plan," US Office of Managemen and Budget, CDO Council and General Services Administration, 14 May 2020. [Online]. Available: https://strategy.data.gov/2020/action-plan/.

[55] Information Commissioner's Office, "Blog: Building on the data sharing code - our plans for updating our anonymisation guidance," 19 March 2021. [Online]. Available: https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2021/03/building-on-the-data-sharing-code-our-plans-for-updating-our-anonymisation-guidance/ .

[56] "Article 29 Data Protection Working Party, Opinion 05/2014 on Anonymisation 0829/14/EN WP216," 10 April 2014. [Online]. Available: https://ec.europa.eu/justice/article-29/documentation/opinion-recommendation/files/2014/wp216_en.pdf.

[57] H. Zang and J. Bolot, "Anonymization of location data does not work: a large-scale measurement study," in Proceedings of the 17th annual international conference on Mobile computing and networking (MobiCon '11), Las Vegas, 2011.

[58] A. Boutet, S. B. Mokhtar and V. Primault, "Uniqueness Assessment of Human Mobility on Multi-Sensor Datasets," 8 December 2016. [Online]. Available: https://hal.archives-ouvertes.fr/hal-01381986/document.

[59] Google AI Blog, "New Insights into Human Mobility Privacy Preserving Aggregation," 12 November 2019. [Online]. Available: https://ai.googleblog.com/2019/11/new-insights-into-human-mobility-with.html.

[60] United States Census Bureau, "2020 Census Data Products: Disclosure Avoidance Modernization," [Online]. Available: https://www.census.gov/about/policies/privacy/statistical_safeguards/disclosure-avoidance-2020-census.html.

[61] R. Herold, "12 Reasons Why Data Privacy Protection Brings Business Value," CPO Magazine, 3 September 2018. [Online]. Available: https://www.cpomagazine.com/blogs/privacy-intelligence/12-reasons-why-data-privacy-protection-brings-business-value/.

[62] A Beacon Group white paper, co-sponsored by Cisco, "Privacty Gains: Business Benefits of Privacy Investment," 2019. [Online]. Available: https://www.cisco.com/c/dam/en_us/about/doing_business/trust-center/docs/privacy-gains-business-benefits-of-privacy-investment-white-paper.pdf.

[63] National Association of City Transportation Officials (NATCO), "Big Data, Big Questions: Refining the Managing Mobility Data Framework," 21 June 2019. [Online]. Available: https://nacto.org/2019/06/21/big-data-big-questions-applying-the-managing-mobility-data-framework/.

[64] S. Garfinkel, J. M. Abowd and C. Martindale, "Understanding Database Reconstruction Attacks on Public Data," 28 November 2018. [Online]. Available: https://queue.acm.org/detail.cfm?id=3295691.

[65] S. Gray and P. Sanderson, "FPF Policy Brief: Location Data Under Existing Pivacy Laws," Future of Privacy Forum, December 2020. [Online]. Available: https://fpf.org/wp-content/uploads/2020/12/FPF_Guide_Location_Data_v2.2.pdf.

[66] S. Gray, "FPF: If you can't take the heat map: benefits & risks of releasing location datasets," Future of Privacy Forum, 14 December 2020. [Online]. Available: https://fpf.org/2018/01/31/if-you-cant-take-the-heat-map-benefits-risks-of-releasing-location-datasets/.

[67] K. Tezapsidis, "Uber Releases Open Source Project for Differential Privacy," Medium, 13 July 2017. [Online]. Available: https://medium.com/uber-security-privacy/differential-privacy-open-source-7892c82c42b6.

[68] S. L. Garfinkel, "De-Identification of Personal Information (NISTIR 8053)," 2015. [Online]. Available: https://nvlpubs.nist.gov/nistpubs/ir/2015/NIST.IR.8053.pdf.

[69] G. S. Nelson, "Practical Implications of Sharing Data: A Primer on Data Privacy, Anonymization, and De-Identification," 2015. [Online]. Available: https://support.sas.com/resources/papers/proceedings15/1884-2015.pdf.

[70] Federal Trade Commission, "Protecting Consumer Privacy in an Era of Rapid Change, Recommendations for Businesses and Policymakers," March 2012. [Online]. Available: https://www.ftc.gov/sites/default/files/documents/reports/federal-trade-commission-report-protecting-consumer-privacy-era-rapid-change-recommendations/120326privacyreport.pdf.

[71] A. Pyrgelis, C. Troncoso and E. De Cristofaro, "Knock Knock, Who't There? Membership Inference on Aggregate Location Data," in 25th Network and Distributed System Security Symposium (NDSS 2018), 2017.

[72] V. Talla, "Documents Reveal ICE Using Driver Location Data From Local Police for Deportations," ACLU Northern California, 13 March 2019. [Online]. Available: https://www.aclunc.org/blog/documents-reveal-ice-using-driver-location-data-local-police-deportations.

[73] F. Xu, Z. Tu, Y. Li, P. Zhang, X. Fu and D. Jin, "Trajectory Recovery From Ash: User Privacy is NOT Preserved in Aggregated Mobility Data," in Proceedings of the 26th International Conference on World Wide Web, 2017.

[74] A. Tockar, "Riding with the Starts: Passenger Privacy in the NYC Taxicab Dataset," 15 September 2014. [Online]. Available: https://agkn.wordpress.com/2014/09/15/riding-with-the-stars-passenger-privacy-in-the-nyc-taxicab-dataset/.

[75] J. K. Trotter, "Public NYC Taxicab Database Lets You See How Celebrities Tip," 23 October 2014. [Online]. Available: http://gawker.com/the-public-nyc-taxicab-database-that-accidentally-track-1646724546.

[76] P. Golle, "Revisiting the uniqueness of simple demographics in the US population," in WPES'02: Proceedings of th 5th ACM workshop on Privacy in electronic society, 2006.

[77] A. J. Blumberg and P. Eckersley, "On location privacy, and how to avoid losing it forever," Electron. Frontier Found, vol. 10, pp. 1-7, 3 August 2009.

[78] Network Advertising Initiative, "Network Advertising," February 2020. [Online]. Available: https://www.networkadvertising.org/sites/default/files/nai_impreciselocation2.pdf.

[79] U.S. Federal Trade Commission, "Federal Trade Commission Act Section 5: Unfair or Deceptive Acts or Practices," [Online]. Available: https://www.federalreserve.gov/boarddocs/supmanual/cch/ftca.pdf.

[80] GDPR Resources & Information, "Article 5: Principles Relating to Processing of Personal Data," [Online]. Available: https://www.gdpr.org/regulation/article-5.html.

[81] A. R. Santos-Lozada, J. T. Howard and A. M. Verdery, "How differential privacy will affect our understanding of health disparities in the United States," 2021. [Online]. Available: https://www.pnas.org/content/pnas/117/24/13405.full.pdf.

[82] A. Chouldechova and A. Coston, "Mobility Data Used to Respond to COVID-19 Can Leave Out Older and Non-White People," Carnegie Mellon University, Heinz College, [Online]. Available: https://www.heinz.cmu.edu/media/2021/March/mobility-data-used-to-respond-to-covid19-can-leave-out-older-and-nonwhite-people.

[83] A. Coston, N. Guha, D. Ouyang, L. Lu, A. Chouldechova and D. E. Ho, "Leveraging Administrative Data for Bias Audits: Assessing Disparate Coverage with Mobility Data for COVID-19 Policy," in Conference on Fairness, Accountability, and Transparency (FAccT 2021), Virtual Event, Canada, 2021.

[84] "National Collaborating Centre for Determinants of Health," [Online]. Available: https://nccdh.ca/glossary/entry/marginalized-populations. [Accessed 1 March 2021].

[85] J. Daley, "San People of South Africa Issue Code of Ethics for Researchers," SmartNews, Smithsonian Magazine, 23 March 2017. [Online]. Available: https://www.smithsonianmag.com/smart-news/san-people-south-africa-issue-code-ethics-researchers-180962615/. [Accessed 1 March 2021].

[86] J. Debussche, J. César and I. De Moortel, "Big Data & Issues & Opportunities," Bird & Bird, April 2019. [Online]. Available: https://www.twobirds.com/en/news/articles/2019/global/big-data-and-issues-and-opportunities-discrimination. [Accessed 1 March 2021].

[87] S. Hymon, "Metro releases Understanding How Women Travel report," The Source, 19 September 2019. [Online]. Available: https://thesource.metro.net/2019/09/19/metro-releases-understanding-how-women-travel-report/. [Accessed 1 March 2021].

[88] Government of Ontario, "Anti-Racism Data Standards - Order in Council 897/2018, Data Standards for the Identification and Monitoring of Systemic Racism," [Online]. Available: https://www.ontario.ca/page/anti-racism-data-standards-order-council-8972018.

[89] National Institute of Standards and Technology, "NIST Privacy Engineering Program, Resources," U S Department of Commerce, [Online]. Available: https://www.nist.gov/itl/applied-cybersecurity/privacy-engineering/resources. [Accessed 1 March 2021].

[90] D. K. Citron and D. J. Solove, "Privacy Harms," GWU Legal Studies Research Paper No. 2021-11, GWU Law School Public Law Research Paper No. 2021-11, 2021.

[91] State of California Department of Justice, "California Consumer Privacy Act (CCPA)," [Online]. Available: https://oag.ca.gov/privacy/ccpa. [Accessed 23 January 2021].

[92] Learning for Sustainability, "Social license to operate [10]," [Online]. Available: https://learningforsustainability.net/social-license/. [Accessed 1 March 2021].

[93] "Ethics Explainer: Social license to operate," The Ethics Centre, 23 January 2018. [Online]. Available: https://ethics.org.au/ethics-explainer-social-license-to-operate/.

[94] "45 CFR 46: 2018 Requirements (Common Rule)," U.S. Department of Health & Human Services (hhs.gov). [Online]. Available: https://www.hhs.gov/ohrp/regulations-and-policy/regulations/45-cfr-46/index.html. [Accessed 1 February 2021].

[95] Future of Privacy Forum, "FPF Ethical Data Use Committee Will Support Research Relying on Private Sector Data," 5 May 2021. [Online]. Available: https://fpf.org/blog/fpf-ethical-data-use-committee-will-support-research-relying-on-private-sector-data/.

[96] M. Tokson, "Inescapable Surveillance, University of Utah College of Law Research Paper No. 398," Cornell Law Review, 2020.

[97] Google, "GTFS Realtime Overview," Google Transit APIs, 22 October 2020. [Online]. Available: https://developers.google.com/transit/gtfs-realtime.

[98] North American Bikeshare Association, "NABSA/General Bikeshare Feed Specification (GBFS)," [Online]. Available: https://github.com/NABSA/gbfs.

[99] A. Pyrgelis, C. Troncoso and E. De Cristofaro, "What Does The Crowd Say About you? Evaluating Aggregation-based Location Privacy," in Proceedings of Privacy Enhancing Technologies, 2017.

[100] L. Sweeney, "Simple Demographics Often Identify People Uniquely," Carnegie Mellon University, Pittsburgh, 2000.

[101] C. Y. Ma, D. K. Yau, N. K. Yip and N. S. Rao, "Privacy Vulnerability of Published Anonymous Mobility Traces," in IEEE/ACM Transactions on Networking, 2013.

[102] Clinical and Translational Science Awards Consortium Community Engagement Key Function Committee TaskForce on the Principles of Community Engagement, "PRINCIPLES OF COMMUNITY ENGAGEMENT, SECOND EDITION (NIH Publication No. 11-7782)," National Institutes of Health, U S Department of Health and Human Services, Washington, D.C., 2011.

[103] Hogan Lovells, "CPRA countdown: Updated transparency obligations and opt-out rights," 9 May 2021. [Online]. Available: https://www.engage.hoganlovells.com/knowledgeservices/news/cpra-countdown-updated-transparency-obligations-and-opt-out-rights.

[104] EUR-Lex, "Regulation (EU) 2016/679 of European Parliament and the Council of 27 April 2016 (Document 02016R0679-20160504)," 27 April 2016. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A02016R0679-20160504&qid=1532348683434.

[105] City of Los Angeles Department of Transportation (LADOT), "LADOT Data Protection Principles," 12 April 2019. [Online]. Available: https://ladot.lacity.org/docs/ladot-data-protection-principles.

[106] National Institue of Standards and Technology, "SP 800-53 Rev. 5 Security and Privacy Controls for Information Systems and Organization," [Online]. Available: https://csrc.nist.gov/publications/detail/sp/800-53/rev-5/final.

13. ABBREVIATIONS

AIA	Algorithmic Impact Assessment
EU	European Union
FPF	Future of Privacy Forum™
GBFS	General Bikeshare Feed Specification
GTFS-RT	General Transit Feed Specification Realtime
GPS	Global Positioning System
LADOT	City of Los Angeles Department of Transportation
MDSA	Mobility Data Sharing Assessment
MDC	Mobility Data Collaborative™
NIST	U.S. National Institute for Standards and Technology
PETs	Privacy-Enhancing Technologies

APPENDIX A. UNDERSTANDING MOBILITY DATA

Mobility Data
For the purposes of the MDSA, **Mobility Data** means all information related to an activity, event, or transaction generated by either the operator or user of a digitally enabled mobility device, vehicle, or service. Mobility data includes personal data, non-personal data, and data that falls into a grey area where a contextual analysis is required.

The MDSA focuses on identifying and evaluating the privacy and ethical considerations of sharing mobility data. The ecosystem of companies, governments, researchers, and other organizations interested in sharing mobility data is evolving quickly, and privacy is already well-established as a key principle [10] [63]. Nevertheless, "mobility data" does not yet have a universally accepted definition.

From a privacy practitioners' perspective, mobility data is typically a broad concept, capturing all data. Some types of mobility data could be sensitive and identifiable, and thus carry high privacy risk (e.g., detailed records of a person's daily travel history). Other types of mobility data could be far less sensitive or identifiable, and instead carry very low privacy risk (e.g., aggregate price of transit fares over time). Transportation officials or researchers, on the other hand, might use "mobility data" to refer only to data that is personally identifiable, and instead use terms like "public data" or "transportation data" to refer to information that carries low or no privacy risk (e.g., statistical information or information about transportation systems). For example, some transportation experts consider data generated by the General Transit Feed Specification Realtime (GTFS-RT)[46] or General Bikeshare Feed Specification (GBFS)[47] as non-personal and outside the scope of "mobility data."

This MDSA uses a broad definition of "mobility data" in order to help organizations identify and address potential privacy and ethical considerations comprehensively. This definition includes data beyond just location or geospatial information, and includes information from and about individuals, devices, organizations, and other impacted communities as well. In many cases, mobility data that does not appear personal can nevertheless reveal sensitive details about an individual or group. It is also true in many cases that data that is considered non-personal or de-identified today could be used to re-identify individuals or reveal sensitive information in the future [64]. Given this, organizations are encouraged to err on the side of caution and to also adopt a broad definition for the purposes of assessing the privacy and ethical considerations of sharing mobility data.

[46] GTFS Realtime is a feed specification that allows public transportation agencies to provide real time updates about their fleet to application developers [97].

[47] "The General Bikeshare Feed Specification, known as GBFS, is the open data standard for shared mobility. GBFS is intended to make information publicly available online; therefore, information that is personally identifiable is not currently and will not become part of the core specification [98]."

APPENDIX B. UNDERSTANDING LOCATION DATA AND THE RISK OF RE-IDENTIFICATION

Location data is considered "personal information" under most, if not all, privacy laws around the world, when the data **relates to an identifiable person** [65]. Location data is personal information when it is sufficiently precise, accurate, and/or persistent (collected over time) to identify a person with reasonable specificity. This means, for example, that precise location data involving buildings, landmarks, or factory sensors is usually not personal information. Similarly, aggregated data (information about movements of large groups) is not considered personal information. In some cases, precise location data tied to an individual or a device can be "de-identified" or "anonymized" through technical approaches so that it is no longer considered identifiable.[48]

Many organizations are interested in gaining access to "anonymous" or "anonymous and aggregated" location data, to observe population-level trends and movements. While in some cases this is possible, it is very challenging to make any dataset of individual precise location data truly "anonymous." [8] Even if unique identifiers are used instead of names, most individual's behavior can be easily traced back to them—for example, from the location of their home (where a device "dwells" at night) [12]. At times, even highly aggregated data about patterns of large groups of people (such as high-level heat maps) can inadvertently reveal information. In 2017, an interactive "global heat map" of movements of users of the Strava fitness app inadvertently revealed the locations of deployed military personnel at classified locations [66].

These challenges are not insurmountable, but organizations should be very careful not to overpromise, and should treat location datasets as private, sensitive information. This means location data—even in de-identified or aggregated form—should be subject to administrative, technical, and legal controls to ensure it remains protected and limited in who can access it and for what purposes.

The Role of De-Identification

De-identification can be an important and powerful tool for organizations to reduce privacy risk before sharing data publicly or with other organizations. Organizations typically use technical or statistical techniques to reduce the identifiability of information in a dataset, for example by redacting fields that directly identify individuals (e.g., names, email addresses) and obfuscating, encrypting, or perturbing the fields that can indirectly identify individuals (e.g., demographic data, location data, user or device IDs). More sophisticated approaches may inject random "noise" into a dataset in such a way that individual data points become inaccurate, but the dataset as a whole remains statistically valid. Organizations can also strengthen their de-identification processes by combining technical approaches with legal and administrative safeguards, such as contractual commitments not to re-identify data or access restrictions on de-identified data.

The state of the art for de-identification techniques is advancing rapidly, and more formal approaches that offer greater and more lasting privacy protection, such as differential privacy or synthetic data, are beginning to be applied at scale in the mobility space [67].

[48] Organizations should note that terms like "de-identified" and "anonymized" may have different meanings in different jurisdictions. There may also be differences in how these terms are used by laypersons, by lawyers, and by technical or statistical experts. This MDSA uses the term "de-identification" with the understanding that sometimes de-identified information can be re-identified, and sometimes it cannot.

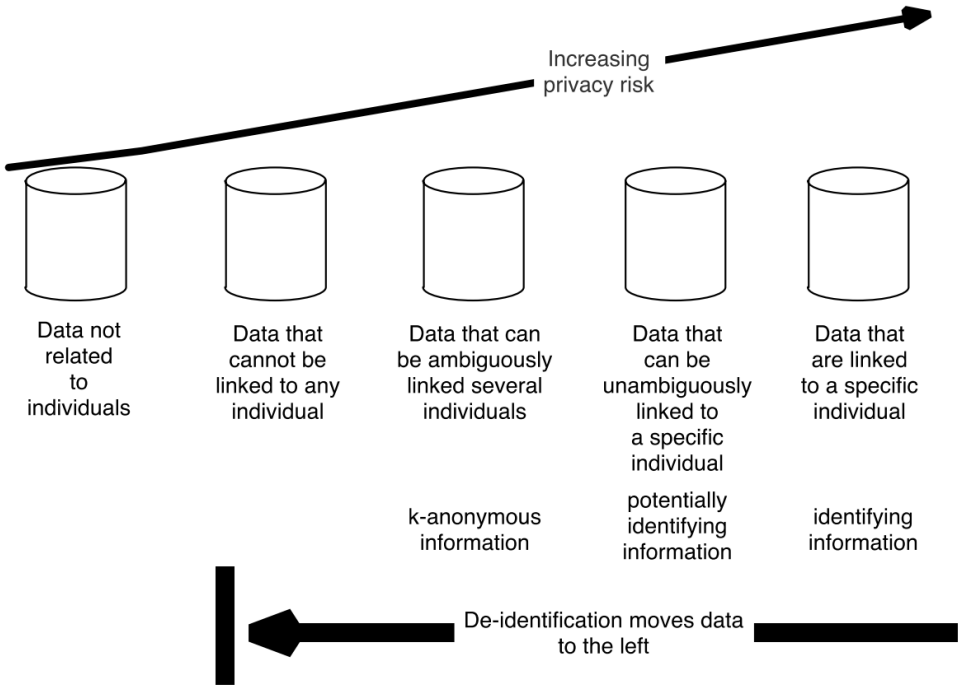


Figure 1. Data identifiability spectrum [68, p. 6]

De-identification is a powerful tool; however, it is not a silver bullet and data identifiability exists on a spectrum (see Figure 1[68, p. 6]). Because de-identification relies on manipulating data to be less identifiable, it inevitably impacts the reliability and utility of that information for analysis. Organizations must determine, on a case-by-case basis, what is an acceptable tradeoff between how reliable and precise the information they share should be and the extent to which privacy is protected in that data. As organizations strive to appropriately balance risk and utility in location datasets, they must consider qualitative and contextual factors as well as quantitative ones, including "the interests of who is using the data (e.g., public good versus private companies) and the people detailed in the datasets who are most at risk [7, pp. 15-17]." See Figure 2 for a visual of the privacy and data utility tradeoff.

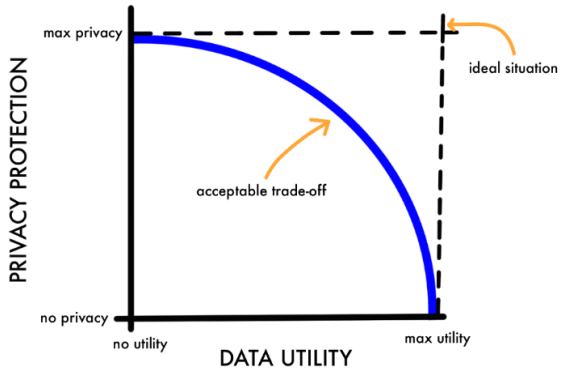


Figure 2. Data privacy versus data utility [69, p. 18]

Evaluating Re-Identification Risk

One of the primary privacy risks arising from location data is the risk that individuals may be re-identified. That is, that their identity or other information about them may be revealed from seemingly non-identifiable datasets.

Re-identification is the process of attempting to reverse-engineer de-identified data to reveal personal information. There is a sizable body of technical scholarship available about the risk of re-identification of location data, and as new sources of mobility data become more broadly available it will likely remain an area of significant research interest. Organizations are strongly encouraged to remain abreast of research and developments.

Determining re-identification risk requires specialized expertise, and must be determined on a case-by-case basis. Experts will consider factors such as: the nature of the original dataset, the de-identification techniques applied, the technical skill of the attacker, the available resources, and the availability of additional (or "auxiliary") data that can be linked with the de-identified data [5].

Re-identification: Types of Disclosure Risk

Organizations often believe that for data to be re-identified it must be possible to link a named individual to the data. This is not always the case; in fact, it may be possible to learn enough from a dataset to violate someone's privacy even without knowing their name. Technical experts typically consider three categories of "disclosure" risk: identity disclosure, attribute disclosure, and inferential disclosure [5, pp. 12-14].

- Identity disclosure is when it is possible to link a specific data item to a specific individual with a high probability.
- Attribute disclosure is when it is possible to determine that an attribute described in the dataset is held by a specific individual with high probability.
- Inferential disclosure is when it is possible to make an inference about an individual (typically a member of a group) with high probability, even if the individual was not in the original dataset.

Applicable Legal Standards

Organizations should carefully consult their local privacy laws to determine which types of disclosure risk must be addressed by law. Different jurisdictions may also apply different tests or standards for determining when data is considered identifiable. For example, in the U.S., companies must be able to show a level of "justified confidence" that data is no longer "reasonably linked or linkable to a particular consumer or device" by applying appropriate technical measures, making public commitments not to attempt to re-identify individuals, and requiring the same commitment of all downstream data recipients [70]. Meanwhile, in the EU, data protection authorities consider three tests for determining if data has been appropriately anonymized [56]:

- Singling out, which is the "possibility to isolate some or all records which identify an individual in the dataset."
- Linkability, which is the "ability to link at least two records concerning the same data subject or a group of data subjects (either in the same database or in two different databases)."
- Inference, which is the "possibility to deduce, with significant probability, the value of an attribute from the values of other attributes."

The Challenges of De-Identifying Location Data

Location data is a type of information that experts consider especially challenging to appropriately de-identify. Geospatial information has strong identification powers: just a few observations of a person's location and time can be highly identifying, even in a dataset that is generalized and statistically "noisy" [68, pp. 26-28, 37-38]. Moreover, there is often a significant amount of auxiliary data that could be used to link de-identified location data to other personal information, including public land use and utilities records, photographs and videos of popular locations, cameras on public streets, 311 or complaints records, geo-tagged social media posts, online check-ins, and smartphone and other connected device GPS traces [12].

Research on location data and re-identification has demonstrated that, for example:

- An attacker with some prior knowledge about an individual's mobility profile can use aggregate information to improve their knowledge or localize the individual. Membership is accurate when groups are small, and when individuals have regular habits, they are easier to classify correctly than those with sporadic movements.[49]
- Ascertaining whether an individual is present in aggregate location time-series becomes a privacy threat if the aggregates relate to a group of individuals that share a sensitive characteristic [71].
- The privacy of individuals is unlikely to be preserved even in country-scale location datasets [2] and some individuals from marginalized communities may be more vulnerable to re-identification risks or negative impacts from inference attacks [72].
- The temporal element is important because temporal-only mobility traces can represent a mobility footprint where researchers are able to highly discriminate between individuals; the uniqueness of individuals is improved when temporal information is combined with spatial information.[50]
- Recent research has shown that in ideal scenarios, an individual's "trajectories" can be recovered with up to 91% accuracy from aggregated location data that was collected from mobile applications [73].

Research on location data has also shown that it is especially susceptible to inferential disclosure (the "possibility to deduce"), which may result in harm to groups or classes of individuals, even those who do not appear in the dataset [5]. For example, location data can be used to predict individuals' income, home and work location, their sleep schedule, gender and age, their personality, their friends, and where they socialize [2]. Some locations may also be particularly sensitive because of what might be inferred about those who visit or spend time there (e.g., certain hospitals, schools, religious centers, nightclubs, abortion clinics, dispensaries, military bases, or political organizations and events).

[49] See the first evaluation of membership inference in the context of location data [99].

[50] In addition, researchers showed that the temporal information improves the capacity to uniquely identify individuals by 14% on average compared to only considering spatial information. "Measuring the uniqueness does not mean reidentification of users. Indeed, pseudo-anonymised mobility traces themselves do not disclose the identity of a user. However, using external knowledge can lead to infer the identity of users [58, p. 2]."

Demonstrations of Re-Identification

Organizations should not discount the difficulty of appropriately de-identifying mobility data, especially location data. Re-identification is not a hypothetical concern, but rather a real possibility that can have significant negative impacts on individuals and organizations. Organizations should be especially cautious before releasing any data publicly, even in de-identified or aggregated forms. Consider the following case studies:

Case Study: Re-Identification Using Mobility Traces [5]
Individuals and vehicles can be identified by their "mobility traces" (a record of locations and times that the person or vehicle visited). In one study, mobility trace data from a sample of 1.5 million individuals was processed, with time values being generalized to the hour and spatial data generalized to the resolution provided by a cell phone system (typically 10 to 20 city blocks). The researchers found that four randomly chosen observations of an individual which put them at a specific place and time was sufficient to uniquely identify 95% of the data subjects [12]. Spatial and temporal points for individuals can be collected from a variety of sources, including purchases with a credit card, a photograph, or Internet usage. A similar study found that 30 to 50% of individuals could be identified with ten pieces of additional information [101].
Case Study: Re-Identification of Pseudonymized Taxi Ride Data
In 2014, the New York City Taxi and Limousine Commission released a dataset containing a record of 173 million New York City taxi trips from 2013. The Commission replaced the taxi medallion numbers and driver license numbers with a one-way cryptographic hash. The data did not include the names of the taxi drivers or riders, but it did include a 32-digit alphanumeric code that could be readily converted to each taxi's medallion number.[51] Users of the dataset discovered the hash algorithm and were able to reverse the pseudonymization by iterating through all possible medallion numbers and license numbers, determining the cryptographic hash of each, and replacing the hash with the original number [5] [74]. This dataset also raised privacy risks for the passengers. Someone discovered that they could find time-stamped photographs on the internet of celebrities entering or leaving taxis in which the medallion number was clearly visible. With this information and other publicly available information on gossip websites, the person was able to discover the other end-point of the ride, the amount paid, and the amount tipped for two of the 173 million taxi rides [74]. A reporter at the Gawker website was able to identify another nine [75].
Case Study: Re-Identification Through Linkage Attacks Using Indirect Identifiers
Another way to re-identify a dataset that has been de-identified is through a linkage attack [5]. One of the most widely publicized linkage attacks was by Latanya Sweeney, who reidentified the medical records of a Massachusetts governor. Massachusetts was distributing a research dataset containing de-identified insurance reimbursement records of employees that had been hospitalized and to protect their privacy names were stripped from the dataset. However, the employees' date of birth, zip code, and sex was preserved to allow for statistical analysis. Because Sweeney knew that the governor had recently been treated at a Massachusetts hospital, she was able to re-identify the governor's records by searching for the "de-identified" records that matched the governor's date of birth, zip code, and sex. She accessed this data from the Cambridge voter registration list for $20. Sweeney generalized her findings, arguing that up to 87% of the U.S. population could be uniquely identified by their 5-digit zip code, date of birth, and sex based on the 1990 census.[52] Follow-up work by privacy researcher Phillip Golle computed a re-identification rate of 62% using the year 2000 census [76].

[51] The 32-digit code was the MD5 cryptographic hash of the taxi medallion number. The MD5 algorithm cannot be inverted, but there are such a small number of possible taxi medallion numbers that it was straightforward to use a brute force attack and hash them all. The problem could have been avoided if the Taxi and Limousine Commission had used a keyed hash and not released the key, if the medallion numbers had been encrypted instead of hashed (and the encryption key had not been released), or if the TLC had created a randomly generated code for each medallion number.

[52] Birthday, zip code, and sex are indirect identifiers [100]. Refer to pages 20-22 of NIST, *De-Identification of Personal Information* for suggested methods for de-identifying indirect identifiers [68, pp. 20-22].

APPENDIX C. ADDITIONAL RESOURCES

General Resources

Australian Government, Best Practice Guide to Applying Data Sharing Principles, March 15, 2019, https://www.pmc.gov.au/sites/default/files/publications/data-sharing-principles-best-practice-guide-15-mar-2019.pdf.

European Commission, Experts say privately held data available in the European Union should be used better and more, Report: Towards a European strategy on business-to-government data sharing for the public interest, https://digital-strategy.ec.europa.eu/en/news/experts-say-privately-held-data-available-european-union-should-be-used-better-and-more.

Information Commissioner's Office, Data Sharing Guide, https://ico.org.uk/for-organisations/data-sharing-a-code-of-practice/.

Personal Data Protection Commission Singapore, Trusted Data Sharing Framework, https://www.imda.gov.sg/-/media/Imda/Files/Programme/AI-Data-Innovation/Trusted-Data-Sharing-Framework.pdf.

Open Mobility Foundation, Privacy Guide for Cities, https://github.com/openmobilityfoundation/governance/blob/main/documents/OMF-MDS-Privacy-Guide-for-Cities.pdf.

D'Agostino, Mollie, Pellaton, Paige, Brown, Austin, Mobility Data Sharing: Challenges and Policy Recommendations, 2019, https://escholarship.org/uc/item/4gw8g9ms.

Information Commissioner's Office, "Accountability and governance," https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/.

Information Commissioner's Office, Data sharing code: the basics, https://ico.org.uk/media/for-organisations/documents/2618790/data-sharing-code-the-basics.pdf.

Mobility Data Collaborative, https://mdc.sae-itc.com/.

Netherlands National Ombudsman, The Citizen is not a Dataset, https://www.nationaleombudsman.nl/nieuws/onderzoeken/the-citizen-is-not-a-dataset.

NACTO, https://nacto.org/wp-content/uploads/2019/05/NACTO_IMLA_Managing-Mobility-Data.pdf and https://nacto.org/wp-content/uploads/2017/01/NACTO-Policy-Data-Sharing-Principles.pdf.

NUMO, Leveraging Data to Achieve Policy Outcomes, Explore Use Cases for Micromobility Data, https://policydata.numo.global/.

Office of the Privacy Commissioner of Canada, "Walk the Talk and Show It: Demonstrable Accountability for Data Protection," https://www.priv.gc.ca/en/opc-news/speeches/2012/sp-d_20120417_pk/.

Populus, A Practical Guide to Mobility Data Sharing and Cities, May 2020, https://www.populus.ai/white-papers/mobility-data.

Stantec and ARA, Preparing for Automated Vehicles and Shared Mobility: State-of-the-Research Topical Paper #1, Models for Data Sharing and Governance for Automated Vehicles and Shared Mobility, September 21, 2020, http://onlinepubs.trb.org/onlinepubs/AVSMForum/products/1-NCHRP_Data_Sharing_and_Governance_Final_10-28-20v2.pdf.

SUM4ALL, Sustainable Mobility: Policy Making for Data Sharing, https://www.sum4all.org/data/files/policymakingfordatasharing_pagebypage_030921.pdf.

Fairness

Berke, Alex, Calacci, Dan, Larson, Kent, Pentland, Alex (Sandy), "The Tradeoff Between the Utility and Risk of Location Data and Implications for Public Good," arXiv.org, 2019, https://www.media.mit.edu/publications/the-tradeoff-between-the-utility-and-risk-of-location-data-and-implications-for-public-good/.

Diverse Voices - Tech Policy Lab, http://techpolicylab.org/diverse-voices.

Debussche, Julien, César, Jasmien, De Moortel, Isis, Big Data & Issues & Opportunities: Discrimination, April 2019, https://www.twobirds.com/en/news/articles/2019/global/big-data-and-issues-and-opportunities-discrimination.

European Data Protection Supervisor, The EDPS quick-guide to necessity and proportionality, January 2020, https://edps.europa.eu/data-protection/our-work/publications/factsheets/edps-quick-guide-necessity-and-proportionality_en.

Future of Privacy forum, "A Closer Look at Location Data: Privacy and Pandemics," March 2020, https://fpf.org/blog/a-closer-look-at-location-data-privacy-and-pandemics/.

Future of Privacy Forum, Nothing to Hide, https://fpf.org/wp-content/uploads/2018/09/FPF-AISP_Nothing-to-Hide.pdf.

Gemma Galdon Clavell, AIES '20: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, Auditing Algorithms: On Lessons Learned and the Risks of Data Minimization.

Information Commissioner's Office, Big data, artificial intelligence, machine learning and data protection, https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf.

Information Commissioner's Office, Data minimisation and privacy-preserving techniques in AI systems, https://ico.org.uk/about-the-ico/news-and-events/ai-blog-data-minimisation-and-privacy-preserving-techniques-in-ai-systems/.

Metcalf, Jacob & Crawford, Kate, Where are Human Subjects in Big Data Research? The Emerging Ethics Divide (2016), http://papers.ssrn.com/abstract=2779647.

Microsoft Research, Operationalizing the Legal Principle of Data Minimization for Personalization, https://www.microsoft.com/en-us/research/publication/operationalizing-the-legal-principle-of-data-minimization-for-personalization/.

Office of the Privacy Commissioner of Canada, "A Framework for the Government of Canada to Assess Privacy-Impactful Initiatives in Response to COVID-19," April 2020, https://www.priv.gc.ca/en/privacy-topics/health-genetic-and-other-body-information/health-emergencies/fw_covid/.

Open Data Institute, The data ethics canvas, https://theodi.org/article/data-ethics-canvas/.

Racial Equity Toolkit: An Opportunity to Operational Equity, https://www.racialequityalliance.org/wp-content/uploads/2015/10/GARE-Racial_Equity_Toolkit.pdf.

Upturn, Data Ethics, Investing Wisely in Data at Scale, September 2016, https://www.upturn.org/static/reports/2016/data-ethics/files/Upturn_-_Data%20Ethics_v.1.0.pdf.

Yves-Alexandre de Montjoye, César A. Hidalgo, Michel Verleysen, and Vincent D. Blondel, "Unique in the Crowd: The privacy bounds of human mobility," March 2013, https://www.nature.com/articles/srep01376.

Safeguards

Ann Cavoukian, The 7 Foundational Principles Implementation and Mapping of Fair Information Practices, Information & Privacy Commissioner, https://www.ipc.on.ca/wp-content/uploads/resources/pbd-implement-7found-principles.pdf.

Federal Trade Commission, A Guide for Business, https://www.ftc.gov/tips-advice/business-center/guidance/start-security-guide-business.

Federal Trade Commission, Privacy By Design and the New Privacy Framework of the U.S. Federal Trade Commission, https://www.ftc.gov/sites/default/files/documents/public_statements/privacy-design-and-new-privacy-framework-u.s.federal-trade-commission/120613privacydesign.pdf.

Francis, P., Probst Eide, S., and Munz, R. (2017). Diffix: high-utility database anonymization. In Annual Privacy Forum (Springer), pp. 141-158.

Gemma Galdon Clavell, AIES '20: Proceedings of the AAAI/ACM Conference on AI, Ethics, and Society, "Auditing Algorithms: On Lessons Learned and the Risks of Data Minimization," February 2020, https://dl.acm.org/doi/10.1145/3375627.3375852.

ICO, Data minimisation, https://iapp.org/media/pdf/resource_center/ICO-data-minimisation.pdf.

Information and Privacy Commissioner of Ontario, De-identification Guidelines for Structured Data, June 2016, https://www.ipc.on.ca/wp-content/uploads/2016/08/Deidentification-Guidelines-for-Structured-Data.pdf.

ISO/IEC 27701, https://www.iso.org/obp/ui/#iso:std:iso-iec:27701:ed-1:v1:en.

Microsoft Research, Operationalizing the Legal Principle of Data Minimization for Personalization, July 2020, https://www.microsoft.com/en-us/research/publication/operationalizing-the-legal-principle-of-data-minimization-for-personalization/.

Mir, D.J., Isaacman, S., Cáceres, R., Martonosi, M., and Wright, R.N. (2013). Dp-where: Differentially private modeling of human mobility. In 2013 IEEE international conference on big data (IEEE), pp. 580-588.

NIST Privacy Framework, https://nvlpubs.nist.gov/nistpubs/CSWP/NIST.CSWP.01162020.pdf (+ crosswalks to other laws and standards: https://www.nist.gov/privacy-framework/resource-repository/browse/crosswalks).

NIST SP 800-53B Control Baselines for Information Systems and Organization, https://csrc.nist.gov/publications/detail/sp/800-53/rev-5/final.

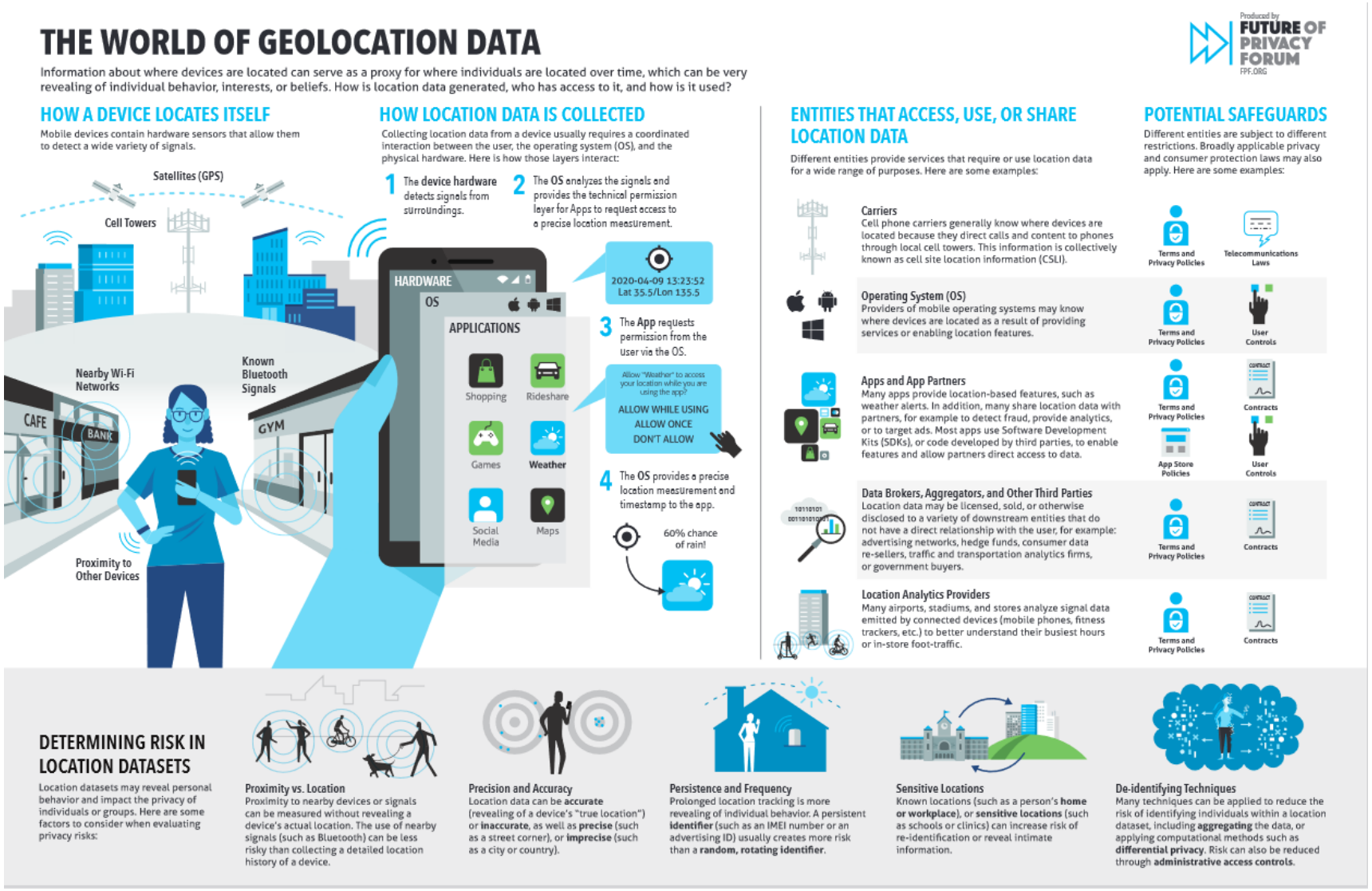
Oehmichen, A., Jain, S., Gadotti, A., and de Montjoye, Y.-A. (2019). Opal: high performance platform for large-scale privacy-preserving location data analytics. In 2019 IEEE International Conference on Big Data (Big Data) (IEEE), pp. 1332-1342.

Office of the Privacy Commissioner of Canada, OPC's Guide to the Privacy Impact Assessment Process, Risk Analysis by Privacy Principle, https://www.priv.gc.ca/en/privacy-topics/privacy-impact-assessments/gd_exp_202003/.

Office of the Privacy Commissioner of Canada, Personal Information Retention and Disposal: Principles and Best Practices, June 2014, https://www.priv.gc.ca/en/privacy-topics/business-privacy/safeguards-and-breaches/safeguarding-personal-information/gd_rd_201406/#fn2-rf.

Office of the Privacy Commissioner of Canada, Privacy Tech-Know blog: Privacy Enhancing Technologies for Businesses, https://priv.gc.ca/en/blog/20210412.

APPENDIX D. INFOGRAPHIC: UNDERSTANDING THE WORLD OF GEOLOCATION DATA



APPENDIX E. INFOGRAPHIC: A VISUAL GUIDE TO PRACTICE DATA DE-IDENTIFICATION