This section aims to provide a basic understanding of mobility data, to serve as a foundation for then exploring ways to support the use of data in the passenger transportation sector.
Framework document: Data-driven transportation systems

This document aims to provide a basic understanding of mobility data, to serve as a foundation for then exploring ways to support the use of data in the passenger transportation sector.

What is data? 1
  What is mobility data? 2
  Who owns, uses, and regulates mobility data? 3

Why is data useful in the mobility sector? 4

What are the key ways in which mobility data can benefit a city? 5

How can a city unlock these benefits? 7

How can a city monitor progress with respect to data collection, sharing, and use? 8

What current data-related policies and initiatives already exist at the national level? 10
**What is data?**

Data, in a broad sense, can refer to any sort of information, both qualitative and quantitative. Data can be measured, collected, and reported by a number of means, and then analyzed to provide insight into a situation. The Government of India’s Personal Data Protection Bill (2018) defines data as “representation of information, facts, concepts, opinions, or instructions in a manner suitable for communication, interpretation, or processing by humans or by automated means.”

**What is mobility data?**

Mobility data encompasses a wide range of data about or related to the interactions and movement of people, goods, and vehicles in the transportation system. This data can be both quantitative and qualitative (e.g. customer satisfaction with a service). Mobility data can describe transportation assets and trips, as well as data that affects and goes in tandem with mobility, such as information about the weather, air quality, and traffic violations.

Mobility data is collected and generated in a number of different manners by a number of different parties. For example, data can be collected using physical infrastructure like sensors and cameras, mobile applications, and surveys; this data is often generated by the movement of vehicles or individual travellers. It may be created and collected by public transit agencies, private companies, and individual citizens.

Some examples of mobility data include:

1) Metro timetables created by a public transit agency
2) Real-time locations of buses gathered by GPS trackers installed in buses
3) Video footage of intersections captured by cameras installed by the city
4) Weather data produced by satellites
5) Starting and ending locations of an individual traveler’s trip, collected by an app-based service provider such as Ola
6) Records of traffic violations kept by the city traffic police

---

Who owns, uses, and regulates mobility data?

Three categories critical to the stakeholder landscape for mobility data are data owners, beneficiaries, and government:

1) Data owners: companies, organizations, and individuals that produce and own datasets.
2) Data beneficiaries: any group or individual that benefits from using mobility data.
3) Government: in addition to owning and benefiting from data, can play a role in enabling interactions between data owners and beneficiaries and protecting their interests.

These categories are not exclusive or have clearly defined boundaries; data can flow from owners to beneficiaries, but it can also flow within each category. Many organizations and individuals function as both a data owner and beneficiary.

Data owners

Data ownership can typically be divided between two categories: public data, which is freely available (though not always accessible); and private data, which is generally kept within an organization. A single data owner may have both private and public datasets. Another key distinction is that of open data—this falls under the category of public data, because it is freely available to the public, but the term specifically refers to data that is typically well-structured, maintained, and published on portal to make it easy to access and use. According to Ministry of Electronics and Information Centre, “a dataset is said to be open if anyone is free to use, reuse, and redistribute it – Open Data shall be machine readable and it should also be easily accessible.”

Data beneficiaries

Within the category of data beneficiaries for passenger mobility, there are three primary sub-groups: cities and governments, travellers, and researchers. Each of these beneficiary groups correlates with a set of use cases (i.e. end goals of collecting and analyzing mobility data), which are outlined in more detail in the Workbook document.

---

3 “Open Government Data,” Ministry of Electronics and Information Technology, Government of India
1) Cities and governments: includes regulators, planners, and operators responsible for system-level design, operations, and policy.
2) Travellers: any individual moving from one location to another.
3) Researchers: any organization or individual conducting research in the area of mobility.

**Figure 1:** Summary of key stakeholders in the data ownership, use, and regulation ecosystem for passenger mobility

**Why is data useful in the mobility sector?**

Data analytics can help to unlock tremendous value in the transportation system by providing governments and organizations with the data they need to create more efficient transportation systems, with optimally designed routes, services, infrastructure, and regulations. Doing so will lead to lower levels of congestion, reduced tailpipe emissions, and less time spent in transit, resulting in communities that are cleaner, safer, better designed, and more economically prosperous.

More and more transit organizations, both public and private, as well as individuals, are collecting significant amounts of transit-related data. The range, scope, and volume of data collection is expanding. This increase in data presents a massive
opportunity to better integrate components of existing transport systems, optimize transit options to users’ needs, and plan and regulate cities to match mobility patterns. The potential value that mobility data can unlock has led some analysts to dub data “the new form of oil”\(^4\) for transport systems.

**What are the key ways in which mobility data can benefit a city?**

Increased collection of and access to data for the mobility sector can provide cities with information to better design their cities and transport systems, which in turn can improve the quality of life for their citizens. Using data analytics in the mobility space has the potential to create more efficient commutes and allow for the optimized design of city infrastructure and regulations. Taken from a societal perspective, this should lead to cities that are:

1) **Cleaner**—using data to optimize commutes and goods transport will lead to fewer cars on the road, which means less tailpipe emissions.
2) **Better designed**—planners armed with historic mobility data can better optimize infrastructure design to meet typical transport patterns, or recognize where areas need to be re-designed to minimize commutes.
3) **Safer**—improved monitoring will lead to quicker emergency response times and better understanding of pain points will allow cities to address underlying causes to reduce accident rates.
4) **More economically prosperous**—less time in transit means citizens have more time to contribute to economic activities.

Overall, the use of data can allow for mobility assets to be better utilized and integrated, and boost economic growth while building cleaner, more liveable communities.

The benefits of using data analytics for mobility can be further examined from the perspectives of various beneficiaries for passenger transport. Looking from the perspectives of travelers, cities, and researchers, additional benefits include:

---

\(^4\) Catapult Transport Systems, “The Transport Data Revolution: Investigation into the data required to support and drive intelligent mobility,” March 2015. [Link](#).
<table>
<thead>
<tr>
<th>Stakeholder-specific benefits</th>
<th>Travelers</th>
<th>Cities (including planners, developers, and regulators)</th>
<th>Researchers</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Visibility of options:</strong> Easier to discover and compare transit options</td>
<td></td>
<td>Transit planning: Using historic traffic patterns to better understand the best corridors to build new public transit routes and non motorized transport infrastructure to meet commuter needs</td>
<td>Improved data analytics capabilities: Greater ability to conduct in-depth analysis in order to draw conclusions about and make recommendations for the mobility system</td>
</tr>
<tr>
<td><strong>Coordination between modes:</strong> Easier to coordinate between different modes of transit required to reach a final destination</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Accessibility:</strong> Smart mobility services, enabled by open data, can meet the diverse needs of travelers and provide easier access to mobility</td>
<td></td>
<td>Improved monitoring: Better understanding of how the mobility system is being used, in order to identify where greater enforcement is required, or where new regulations may be needed to help the system function smoothly</td>
<td></td>
</tr>
<tr>
<td><strong>Better experience:</strong> Integrated, multi-modal transit platforms would increase ease of transport, as well sense of security given that transport providers are tracked and their location is known</td>
<td></td>
<td>Reliability: Better able to manage traffic incidents and use data analytics to build a reliable transport system</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Urban design: Understanding commute patterns and areas of concern in the system would allow for designers to create cities that better support mobility</td>
<td></td>
</tr>
</tbody>
</table>

Table 1: Summary of the stakeholder-specific benefits of mobility data

These benefits map to various use cases for data, which in turn requires different types of data. For example, a traveller needs data that will help him optimize his trip to his destination, which likely only requires real-time data or potentially short-term projections of his transport options when he chooses to depart. In contrast, a city planner who is designing infrastructure for the future would benefit from historic transit data so that she can examine past trends. These use cases are outlined in the Workbook document.
How can a city unlock these benefits?

There are a number of actions a city can take to build a strong data collection and sharing ecosystem in order to unlock the benefits of mobility data and implement a set of data use cases. These actions include developing policy, creating initiatives, and convening the relevant stakeholders in the city to work together. In order to do this effectively, the proper institutional framework must be laid, beginning with the appointment of a City Data Champion to lead all mobility data-related initiatives of the city. Once this individual has been appointed, he or she can lead the process of evaluating the city’s needs and goals, prioritizing use cases for data, and developing initiatives and acquiring the necessary data to support the selected use cases. A high-level overview of this process is summarized in Figure 2; each of these steps, along with additional recommendations for supporting a data-driven transportation system, is outlined in more detail in the Workbook. The Evaluation Metrics document provides more detail on reaching key benchmarks and monitoring continued progress in building a strong ecosystem for data collection, sharing, and use.

Promoting data sharing between parties is key to maximizing the benefits of mobility data. Many public transit agencies, private companies, and individuals are generating and collecting transport data; however, this data tends to be siloed between organizations and individuals, and often recorded using different standards and formats. A city can play a key role in convening the relevant stakeholders and providing a platform for data sharing in order to build an effective data sharing ecosystem and practice to increase the amount of data available for planning, decision-making, and innovation.
How can a city monitor progress with respect to data collection, sharing, and use?

Cities can play a critical role in supporting the development of an effective data collection and sharing ecosystem that involves both public and private entities. In addition to improving their own data collection practices, municipal corporations and Smart City Special Purpose Vehicles (SPVs) can play a key role as a convener to get the many stakeholders involved in the mobility data landscape—public transit agencies, private service providers, travellers, etc.—on board with standardizing and sharing data. Doing so will help unlock the benefits outlined above, by making more, higher-quality data available. As public entities representing the needs of the citizens, the municipal corporations and Smart City SPVs may be the most persuasive and neutral organization to serve this function.

There are a number of steps that a city can take to build and strengthen this data collection, sharing, and use ecosystem. The checklist below provides an introduction to these actions, which are described in more detail in the Workbook and Evaluation Metrics document. While it is difficult to design key performance
indicators (KPIs) in the traditional sense for building a data collection, sharing, and use ecosystem, a city can think of its progress with respect to the steps it has taken and the improvements it is making in strengthening the ecosystem, such as the level of buy-in from various stakeholders and the success of specific use cases the city chooses to implement. The improvement in data collection will in-turn enable monitoring and tracking KPIs in other areas (e.g. electric vehicles and freight efficiency).

The checklist below is comprised of a set of benchmarks aimed at building and strengthening a city’s data collection, sharing, and use ecosystem. Achieving each of the checklist items at a basic level will ensure that the city develops a foundational data capability and capacity; however, many of the checklist items are ongoing, and should be periodically revisited. As an initial step, a City Data Champion should be appointed to hold responsibility for leading all data efforts in the city. The Data Champion should begin by leading the effort to come to a consensus on why the city wants to improve its data collection ecosystem for the transportation sector (this set of documents aims to support that process), and prioritize a few use cases on which to focus. The Data Champion then needs to ensure that sufficient data collection mechanisms exist to support the selected use cases. The Data Champion will also be responsible for ensuring adequate safeguards for data privacy and security and spearheading the development of an Open Data Policy and a city-level data-sharing platform.

The development of a participatory framework will allow key data stakeholders to surface shared challenges, aid in solution development, and provide input on policies to ensure they are in-line with industry trends and standards; additionally, engaging stakeholders will support the development of partnerships to promote data sharing. Finally, efficient investment in mobility data initiatives will support the development of an effective data-driven transportation system.

### Checklist for monitoring progress

<table>
<thead>
<tr>
<th>Does the city have...</th>
<th>Progress</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. An appointed City Data Champion</td>
<td></td>
</tr>
<tr>
<td>2. Clarity and communication around the purpose and value of data collection and</td>
<td></td>
</tr>
<tr>
<td>sharing</td>
<td></td>
</tr>
<tr>
<td>----------------------------------</td>
<td></td>
</tr>
<tr>
<td>3. Defined and prioritized data use cases</td>
<td></td>
</tr>
<tr>
<td>4. Sufficient data collection mechanisms</td>
<td></td>
</tr>
<tr>
<td>5. Appropriate safeguards for data privacy and security</td>
<td></td>
</tr>
<tr>
<td>6. A participatory framework for data stakeholders</td>
<td></td>
</tr>
<tr>
<td>7. An Open Data Policy</td>
<td></td>
</tr>
<tr>
<td>8. A city-level data-sharing platform</td>
<td></td>
</tr>
<tr>
<td>9. Investment in mobility data initiatives</td>
<td></td>
</tr>
</tbody>
</table>

The Evaluation Metrics document provides more detail on how to reach these benchmarks, and monitor continued progress with respect to each as the ecosystem continues to develop and strengthen.

**What current data-related policies and initiatives already exist at the national level?**

The Government of India has several relevant policies in place or drafted that pertain to data (though not specific to mobility data). Numerous states and cities also have their own data policies and guidelines.

Relevant policies at the central government level:

1) **National Data Sharing and Accessibility Policy (NDSAP)**: This policy was approved in February 2012, and applies to all shareable non-sensitive data, in either digital or analog form, that is generated using public funds by various ministries, departments, subordinate offices, organizations, and agencies of Government of India and state governments. The goal of the policy is to promote data sharing and enable access to Gov-owned data for national planning, development, and awareness. NDSAP aims to provide a platform for proactive and open access to data generated by various Gov entities, in machine-readable form through a wide area network, to permit a wider accessibility and usage by the public. In addition to the policy document,  

---

5 “National Data Sharing and Accessibility Policy,” Government of India Department of Science and Technology. Link.
there is a set of implementation guidelines\(^8\) for implementing NDSAP that provides a comprehensive guide for cities to implement open data.

2) **Open Government Data (OGD) Platform in India\(^9\):** Government of India has launched Open Government Data (OGD) Platform (data.gov.in) to support Open Data Initiative for nation-wide data sharing. This initiative was launched as part of NDSAP as well as the Digital India Initiative. OGD platform provides open access to datasets, documents, services, tools and applications collected by various ministries/ departments/ organizations of Government of India for public use. The main goal of OGD is to provide open access to the data generated through public funds and to enhance transparency, accountability, citizen engagement, collaboration, better governance, decision making & innovation. Currently, there are over 2 lakh data points, which have cumulatively been viewed over 19 million times.\(^{10}\)

3) **Personal Data Protection Bill\(^{11}\):** The draft of this act was released in July 2018 by the Ministry of Electronics and Information Technology. The act focuses on the fair and reasonable processing of data. The Draft Act specifies that there must be a clear, specific, and lawful purpose behind data processing and stipulates that only necessary data should be collected. The Act proposes to replace traditional terms such as data controller and data subject (i.e. person whose data is being collected) with data fiduciary and data principal, respectively. The data fiduciary would be responsible for ensuring the personal data quality as well as complete, accurate, and not misleading data processing. Sensitive personal data may be processed on the basis of explicit consent. The Act is currently under review by the Ministry of Electronics and Information Technology.\(^{12}\)

4) **A Free and Fair Digital Economy\(^\text{13}\):** This report was released in draft form in July 2018, under the Chairmanship of Justice B. N. Srikrishna. The report recommends that a data protection law should be set up, that will be responsible for the enforcement and effective implementation of the definition of personal data and sensitive personal data, legal affairs, policy and standard setting, research, and awareness. B.N. Srikrishna’s report

---

\(^7\) “National Data Sharing and Accessibility Policy (NDSAP) – 2012,” Government of India Department of Science and Technology. Link.
\(^8\) “Implementation Guidelines for National Data Sharing and Accessibility Policy (NDSAP),” Open Government Data Division, National Informatics Center. Link.
\(^12\) “Data Protection Framework,” Government of India Ministry of Electronics & Information Technology. Link.
outlines seven key principles for effectively designing a privacy policy. At a high level, the principles are that the policy should be 1) technology agnostic; 2) holistic; 3) include language on informed consent; 4) recommend data minimization; 5) assign controller accountability; 6) structure enforcement; and 7) include deterrent penalties. The draft is currently under review by the Ministry of Electronics and Information Technology.

5) **One Nation One Card**\(^\text{14}\): Government of India is soon to release a one-nation-one-card policy for public transit, which will mandate a single payment card across the country that works for all forms of public transit such as buses, metros, trains, and toll payments. The goal of the card is to provide seamless connectivity across various modes of transport, and promote use of public transport. NITI Aayog made the announcement\(^\text{15}\) in September 2018; the timeline for implementation has not yet been clarified. The government held a contest open to the public for the naming of the card, which closed at the end of August. Delhi has been running a pilot project for a common travel card for metros and public buses since January 2018.

---

\(^{14}\) “National Common Mobility Card,” Digital India. [Link](#).

\(^{15}\) “One-nation-one-card for public soon: NITI Aayog,” The Economic Times. [Link](#).