DATA POLICY FOR BHOPAL CITY

The Abstract

Recognizing the challenges and opportunities of urbanization in India, MoHUA, Government of India (GoI), launched the Smart Cities Mission (SCM) in June 2015. Since the launch of the mission, 100 smart cities have been identified and cities have started executing civic infrastructure and ICT projects as per the mission guidelines. Cities have conceptualized projects that enable them to do more with less, enhance their operational efficiency, and deliver timely and quality services to citizens. In order to do more with less, cities are looking for innovative technologies, affordable and sustainable interventions to manage the demand of exponentially growing population in cities to bring in efficiency and optimization in governance and to provide better quality of life to its inhabitants.

With the launch of the Smart Cities Mission in 2015, India started paving a new pathway towards transforming urban management with the power of digital technologies. While the Mission has opened a floodgate of opportunities for cities in applying ICT (Information & Communication Technology) solutions, there is also a greater awareness of the need for building on the 'city-as-a-platform' concept.

The concept recognizes the value of enhancing engagement among all four stakeholders of the quadruple-helix model—Government, citizens, academia, and industry, along with improvements in the internal workflow and decision-making processes of city Governments. In this context, the need for city Governments to take 'digital leadership' has become more pronounced.

This essentially calls for building an enabling ecosystem supported by a robust system of data acting as a backbone. Therefore, making cities 'DataSmart' is key in realizing the full potential of technology interventions and innovation ecosystems in cities. The envisaged outcome of becoming DataSmart is to bring greater efficiency, accountability, and transparency in city governance decisions while fostering civic engagement, co-creation, and innovation in problem-solving.

Eventually, this will lead to the creation of a thriving data marketplace as a sustainable model for smart city solutions for businesses.

To establish this new paradigm of Data Governance, it is critical that cities create data policies that balance privacy, legal and public benefit considerations. At the same time it must define the contours of collaboration between various Governmental/non-Governmental entities on data sharing and access for the intelligent use of data in addressing complex urban challenges.

Table of Contents

Glossary

BSCDCL	Bhopal Smart City Development Corporation Limited
BMC	Bhopal Municipal Corporation
SPV	Special Purpose Vehicle
ICCC	Integrated City Command and Control Centre
CCDCSC	Common Cloud based Data Centre for Smart Cities
ITMS	Intelligent Traffic Management System
ICT	Information and Communication Technology
G2G	Government-to-Government
G2C	Government-to-Citizen
G2B	Government-to-Business
G2A	Government to Academia
IoT	Internet of Things
AI	Artificial Intelligence
ML	Machine Learning
API	Application Programming Interface
OGD	Open Government Data
GoI	Government of India
MeitY	Ministry of Electronics & Information Technology
NIC	National Informatics Centre
NDSAP	National Data Sharing and Accessibility Policy
SaaS	Software as a Service
CDO	City Data Officer
CDP	City Data Policy

SCDA	Smart City Data Alliance
MDO	Mission Data Officer
DC	Data Champion
MIS	Management Information System
NGO	Non-Governmental Organization
CA	Chartered Accountant
e-Gov	e-Governance
CSV	Comma separated Values
XML	Extensive Markup Language
JSON	JavaScript Object Notation
REST	Representational State Transfer
SOAP	Simple Object Access Protocol
XLS	Excel Spreadsheet
ODS	Open Document Formats for Spreadsheets
RDF	Resources Description Framework
KML	Keyhole Markup Language
GML	Geography Markup Language
URL	Uniform Resource Locator
OCR	Optical Character Recognition
UDD	Urban Development Department
MoHUA	Ministry of Housing and Urban Administration
DAM	Data Analytics and Management
DeitY	Department of Electronics & Information Technology
SCADA	Supervisory Control and Data Acquisition
GIS	Geographic Information System
ETL	Extract-Transform-Load

Introduction

Cities are crucial units of local governance within a nation. The development and growth of a nation are influenced to a large extent by its cities. The Smart Cities concept relies on fostering a balanced confluence of two megatrends: Rapid Global Urbanization and Digital Transformation through the Industry revolution. These trends have consequences on our efforts to improve liveability for citizens, enhance human capital and transform the relationship between Government, Civil Society, and Market Players in an environmentally, sustainable and inclusive manner. While Indian cities are the location of its financial institutions and economic drivers, they are also centres of learning and creativity, offering a rich matrix of life to their citizens.

Information is the biggest power of the modern times. It is imperative for the empowerment of communities that cities put out information regarding their functioning, their public services, their governance systems, achievements and failures in the public domain, thereby, empowering their citizens through the access to information. Citizens can collaborate with government easily and with increased frequency; both within their community and beyond, forming stronger groups and exchanging ideas and building new collaborations.

City Governments deal with large number of issues like mobility, management of water, waste water and solid waste, safety and security services, energy, housing, education and health amongst many others. These issues are highly complex in nature and require integrated approaches to resolve. Functions of city governments are organized into multiple departments, agencies and networks. These departments, agencies, networks, work in vertically integrated structures and are each responsible for performance of some functions integral to the working of the city. Besides the departments of governments, private sector organizations, corporates, community organizations, research and academic institutions also play a large role in the functioning of cities, through provision of infrastructure, services, research, co-creation and valuable feedback. All government/ non-government organizations/ individuals are custodians of different types of datasets that is generated through their operations. Since these organizations work as vertically integrated structures, a lot of the data so produced remains in silos within their organizations. In order to solve the myriad complex issues faced by cities, it is vital that data locked in such silos be unlocked and shared amongst these entities.

The future of Governance is data-driven and Indian cities are beginning to adopt this change in their functioning. Bringing data in ‘focus’ ensures a move towards outcome-based planning in governance. This helps us realistically assess the gaps between the outcomes and the desired goals.

Efficient sharing of data among data owners and inter-and-intra governmental agencies along with data standards and interoperable systems is the need of the hour. Hence, there was a need to formulate a policy on Data Sharing and Accessibility which could provide an enabling provision and platform for proactive and open access to the data generated through public funds available with various ministries/departments/organizations of State government and Government of India.

City Data Policy is the first significant step in the direction to provide conceptual clarity over accessing and sharing protocols over city data. A data policy is essential to understand the contours of data sharing, privacy, security and ownership in the context of the city.

Data policy is also needed to define the contours of collaboration between various governmental/ non-governmental entities on sharing and access of data.

Current Assessment

Bhopal City View under Smart City Projects

Bhopal is among the first 20 cities selected in first round of smart cities challenge under Government of India‟s (GoI) smart cities mission (SCM) to implement the smart city proposal (SCP). In this context, Bhopal has incorporated a special purpose vehicle (SPV) – Bhopal Smart City Development Corporation Limited (BSCDCL) (the "Authority") to plan, design, implement, coordinate and monitor the smart city projects in Bhopal. BSCDCL is a company incorporated under Indian Companies Act 2013 with equal shareholding from Madhya Pradesh Urban Development Company Limited (MPUDCL) on behalf of Government of Madhya Pradesh (GoMP) and Bhopal Municipal Corporation (BMC).

Smart City Vision & Mission for Bhopal

Vision: Transforming Bhopal, a City of Lakes, Tradition & Heritage into a leading destination for Smart, Connected and Eco Friendly Communities focused on Education, Research, Entrepreneurship and Tourism.

Mission: People will be happier, healthier, smarter and more prosperous. The city will be more livable, green, clean, sustainable, resilient and competitive.

Government of Madhya Pradesh has embarked on an ambitious journey of developing various cities of the state as Smart Cities. This initiative includes 07 cities identified as part of the Smart City Mission of Government of India and state identified cities over and above these 07 cities. The 07 cities selected as part of Smart City Mission are as follows.

- Indore, Bhopal, Jabalpur, Ujjain and Gwalior, Sagar and Satna

Moving forward towards the provision and implementation of the various initiatives under Smart City Mission Bhopal has incorporated numerous projects. Some of the major projects are as –

<<Bhopal Smart City ICT Projects >>

Integrated Command and Control Centre

Through this project, BSCDCL designed, developed, implemented and operates & maintains:

- Common Cloud based Control and Command Application.
- A Cloud based Data Centre for all 07 smart cities of the state
- A Cloud based Disaster Recovery Centre for all 07 smart cities of the state
- The Integrated Control and Command Centre (ICCC) at each of the 07 cities with city based
- Controls and Analytics.

Common Cloud based Data Centre for Smart Cities (CCDCSC) is a common platform where all the information from various sources like city operation centers and applications is stored. All the information collected here, is analyzed for better planning of the smart cities using integrated analytical layer / BI engine. These insights / trends are helpful in managing incidents across the state and individual city and aid in better planning for the development and delivery of smart city projects.

CCDCSC is a cloud based Data Center based out of any location within India. CCDCSC id connected with various city level ICCCs and various applications of the city from where feeds are received (except video feeds). It hosts common command center application platform for all 7 cities. It also hosts other common applications like integrated analytical layer / BI engine. Eventually all the smart components / applications deployed in the cities will be integrated with the common platform layer for managing smart city operations.

Objective: BSCDCL envisions the planned CCDCSC and ICCC to fulfil following objectives:

- Single source of truth‖ for the city‘s civic functions
- To enable Integrated Operations, Informed Decision Making Real Time Data Analysis and to become “Proactive than reactive” are the major benefits of this project.
- Platform with the ability to receive, intelligently correlate & share information to better predict outcomes
- Act as City‘s emergency and disaster management platform
- Ability to integrate multiple text, voice, data, video and smart sensors communication interfaces
- Ability to integrate and correlate online and offline interactions
- Capabilities to support GIS based incidents visualization
- Future proof - based on Modular, Open, Configurable architecture with capabilities to integrate innovative new applications

- Intelligent and Intuitive work-flow management
- Advanced historical records management and archiving capabilities
- Advanced industrial grade cyber security features

Integration of various IT systems of different stakeholders with the objective of enhancing safety, security and providing better public services in the cities is helping in following:

1. To provide assistance to citizen at the time of emergencies
2. To provide facilities of Ambulance, Police Van, Fire Brigade to the citizens
3. To effectively manage Traffic and Roads and support police to maintain Law and Order
4. Disaster Management
5. Environmental Control/ Pollution Control
6. Efficient user of public resources like electricity and water
7. Efficient and timely delivery of public services.
8. Better health and education services

There are a number of functions and systems that are currently being managed / will be managed out of the Integrated Control and Command Center. Depending on the type of systems and functions, they are being monitored and/or Controlled from the Command Center Application, and have the option of sharing a feed to another agency as required via the platform. This integrates all the City Systems procured under the Smart City Mission.

- E-Governance & City ERP
- DIAL 100
- DIAL 108
- Traffic Management System
- Safe City Cameras Feed
- Emergency Response and Disaster Management
- Met Department
- Smart Parking
- Public Bike Sharing
- Smart Pole
- Smart Lighting
- Solid Waste Management Services
- Intelligent Transport Management System
- Municipal Corporations Call Centre

- Municipal Corporations Services Portal
- City GIS Platform
- City Application
- Water Management System
- Crowdsourcing Data
- Fire Brigade Control System
- Solar Roof Top
- ABD area Utilities
- ABD area Lighting
- ABD area Metering
- ABD area Surveillance
- ABD area building management systems.

Public Bike Sharing

Bhopal Smart City Development Corporation Limited introduced a Public Bike Sharing System to provide a low-cost, environmentally friendly mobility option to city residents under PAN city project of Smart City Mission. Public bike sharing is a flexible system of personalized public transport. Cycles are available in a closely spaced network of semi-automated stations. Users can check out cycles at one station and return them to any other station in the network.

Public Bike sharing is a key element in a city's strategy to expand the use of sustainable transport modes. Cycle sharing is expected to boost the use of public transport by providing crucial last-mile connectivity to the BRT system, thereby expanding the catchment areas for the region's transit systems.

The system will be integrated with the BRT system through the location of its stations and the ITS system. By encouraging a shift to sustainable modes, the Public Bike Sharing System will reduce dependency on automobiles, reduce traffic congestion, vehicle emissions, and demand for motor vehicle parking. In addition, the system will expand the health and wellness benefits of bicycle transport to new users. Finally, the system will support the transformation of streets to become environments where pedestrians and bicyclists feel safe and comfortable. The stations would cover the "catchment area" of the BRT. This will ensure that the first and last mile connectivity to people living in the catchment area is provided for. Importance will be given to place stations near important commercial, cultural, educational, administrative, and residential and tourist attraction points in the catchment area.

BENEFITS AND IMPACTS:-

PBS has diverse benefits by increasing the Non-Motorised Transport. It decreases the use of privately owned vehicles decreasing the pollution of the city and thereby increases the environmental and air quality.

As PBS is designed to solve the First and Last mile connectivity, it improves the convenience of the people being the most reliable mode of transport. PBS also enhances improved road safety and plays an important role in building up a healthier society.

PBS acts as feeder to Existing BRT system increasing the efficiency of BRT. On other hand it decreases over Crowding of the Public Transport by increased use of PBS for small and medium trips.

OUTCOMES:-

- 250 Smart cycles
- 86 docking stations
- 24/7 Operations
- More than 40,000 registered users
- Fully automated
- Unmanned
- Central control system
- Redistribution
- Cashless

Smart Pole and Intelligent Street light project

Bhopal has launched first of its kind public private partnership (PPP) based smart poles and intelligent street lights project in India under the Smart Cities Mission. The project is innovatively structured in such a manner that it delivers bundled smart services to citizens and maximum values for money to the city authority.

There are two parallel part of this project:

1. Smart pole installation across the city: 400 Smart pole to be installed
2. LED street light installation: 20,000 LED street lights to be installed by replacing the conventional sodium lamps and mercury lamps.
3. Smart Pole:

Smart Pole offers multiple services to the city residents. The integral parts of the smart pole are as below:

- Smart pole has telecom tower infrastructure to match with city aesthetic and ready to accommodate upcoming technology as 4G and 5G
- Energy efficient and remotely controllable LED Street Lights
- Wi-Fi hotspot services for the city
- Surveillance cameras for safety and parking violation detection
- Environmental Sensors to monitor Air quality, temperature and humidity
- Electric Vehicle charging points to promote use of electric vehicles in the city
- Mobile based application with functionality of SoS.
- Centralizes Command and Control centre for monitoring the implementation of smart solutions.
- Optical fiber for better bandwidth to the Wi-Fi users/providing backhaul to telecom operators.

Dedicated Mobile Application:-

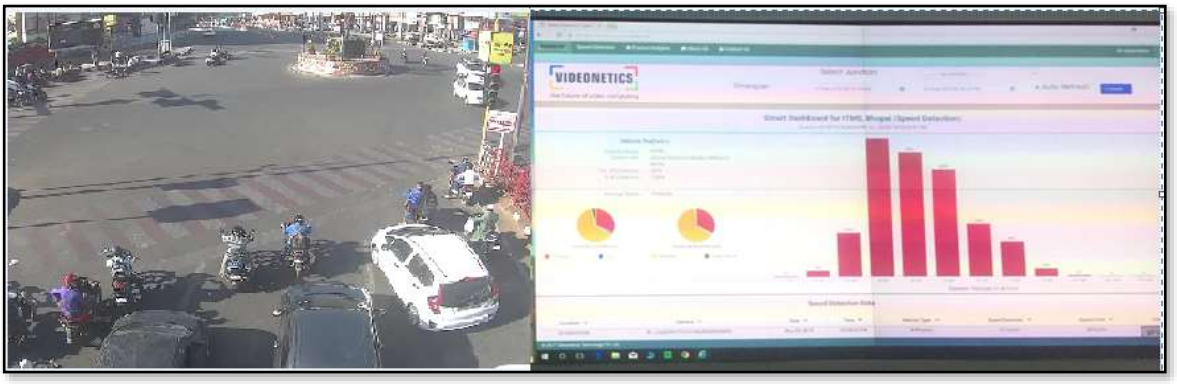
- Complaint register , Redressal Status Tracking, Geotagged Poles.
- Wi-Fi Access
- Environment Parameters- Nox, Sox, Humidity, Temperature etc.

Outcome:

- 20700 Nos. Smart LED Street lights installed at various locations under BMC Limits.
- Individual and group control of street lights through command centre.

Smart Street light Dashboard

Intelligent Traffic Management System



As part of smart city initiatives Bhopal Smart City Development Corporation Limited (BSCDCL)has taken up project for Smart traffic management – Integrated Traffic Management System (ITMS) which provide greater information to the authorities to proactively manage the ongoing traffic situation, enhance traffic discipline through effective enforcements and increase road safety by preventing accidents and allow citizens to make informed travel choices. With ITMS, the overall burden of traffic on the roads will reduce, and will result in enhanced reliability for the citizens and local businesses improving overall productivity levels.

The ITMS services include strengthening of four key areas of traffic management &road safety within the city:

1. Develop a sense of order on the roads by efficiently prosecution of traffic violators and traffic law enforcement
2. Enhance Situational Awareness of existing traffic conditions on real time basis
3. Develop ability to assimilate and Analyze Real Time Traffic Information and historic trends to support decision making on traffic management strategies
4. Create linkages to support Information sharing through traffic controllers, Variable Message Signs, Public Address (PA) systems etc.
5. Ensure long term Capacity Building through training and support for traffic police staff

Key Components of systems:-

- ANPR systems
- Red Light Violation Detection System
- Speed Violation Detection System
- P A System
- E- Challan System
- Establish dash board at Smart city command Center and control room along with TMC

Outcomes:

- 22 Junctions(14 RLVD, 7 Overspeed , 1 surveillance)
- Number of cameras- 260+
- More than 50ooo challans generated
- Revenue collection – 80 lacs

Smart - Road sweeping machines

Conventional Road Sweeping

Road Sweeping Machine

- Live Video Feeds of the Road Sweeping Machines.
- Real Time Data of the Litter Filled , Water Tank status at Command and Control Center

- Dedicated 3 Vehicles to cover the 200 km lane length covering the main roads of Bhopal for the Dry/Wet Cleaning of Main Roads using Smart Multipurpose Machines.
- Road sweeping , litter picking capacity of 5cum, water jetting with capacity of 5000 litre per day

Underground Smart Dustbins

Bhopal, a rapidly developing city of India was facing the problem of rising amounts of waste generation due to urbanization. The solution was Installation of Modern Fully Underground Waste Collection System by Using Smart Bins for Collection of Municipal Solid Waste at Designated Locations in Bhopal City. The aim was to install bin levelling system on each bin and provide a software solution for tracking the levelling of waste filled in underground bins at various locations in the city on real time basis. The software solution is mobile compatible and provides the optimal route management solution with a provision for alerting the operator for emptying a specific bin by keying in the codes. The task of these special sensor-fitted bins is multi-fold.

- Smart Bins with a capacity of 3cum for collection of Municipal Solid waste at 100 designated locations in the Bhopal City
- About 60-70 tons of garbage being collected on a daily basis
- GPS based tracking of the Bins and Vehicles used to transfer the Waste Collected, in Level Sensors.

- Movement of the Redistribution vehicle based on the Route Map based on the Sensor Data

Intelligent Waste Management

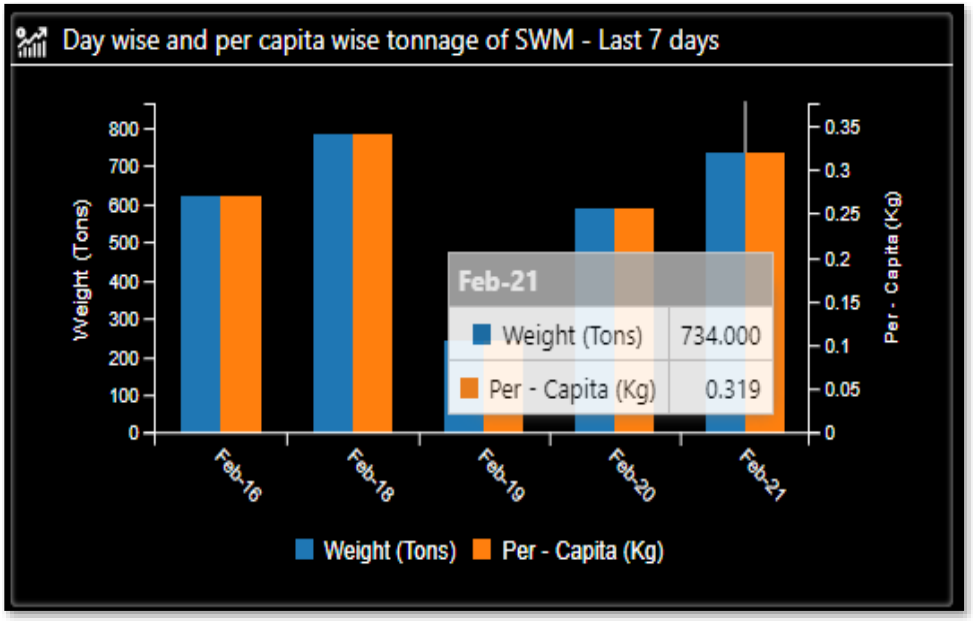
Exploring following technology options for better solid waste management like all SWM vehicles fitted with GPS to optimize solid waste management (Obtain vehicle utilization details and deviations details), underground bins with RFID, installation of weight sensors on transfer stations, obtain citizen feedback on door to door waste collection through the Bhopal Plus city app, biometric location based attendance of all SWM workers, door to door waste collection monitoring through QR based DDN scanning.

- Average Waste Generation per day in Bhopal City – 800 Tonnes
- Average Waste Collected per day through SWM Vehicles – 750 Tonnes
- Underground Dustbin Collection Status
- Vehicle Wise Data Collection :-

➢ No of Vehicles Tracked– 667

a) Type of Vehicles - Tipper, Dumper, Mini Garbage, Refuse Compactor, Dumper Placer, Mini Auto, Truck, JCB, Animal Squad, Sewage etc.
b) No of Trips
c) Total Capacity of Vehicles
d) Garbage Collected (in Tons)

Live vehicle tracking of SWM vehicles to ensure their reachability to all assigned location



Daily trend of garbage collected from entire city and per capita garbage generation

Digital Door Numbering

BSCDCL with an aim to introduce a Unique Smart Addressing Solution for the Urban Dwellings in the city of Bhopal (BMC Limits) based on a Geo-spatial solution, depicting an Alpha Numeric Smart Address code for each property/establishment in Bhopal city with the information of main road, sub road/lane, Landmarks building and the floor rolled out the USASUP (Unique Smart Sequential addressing Solution for Urban Properties / Establishments) project.

- This will ensure the benefits of providing navigation details and address location services to the dwellings/properties within the jurisdiction of BMC and also help see spatial

depiction of information related to electricity, water connection, property tax and fee collection management.

- This project is envisioned to create a unique standardized alpha numeric smart address code for every property/establishment in the jurisdiction of Bhopal municipal limits with the information about the property/establishment, the objective is to formalize door numbers into a usable scientific door numbering system that creates a standard format cross the state. The benefits of this project are immediate and long term - impacting day-to-day communication, governance and above all, create a basic infrastructure for public good.
- The properties shall be affixed with USASUP plate/sticker. To develop this platform and solution, an onsite digital door to door survey is required to be carried out for capturing the Geo position and other attributes

Government Buildings

Commercial Properties

House

- Smart Unique Id will help in interconnectivity of departments such as Property tax, Public amenities etc.
- Future tool for door to door service monitoring
- Enhancement of Citizen Communication owing to a dedicated address system in place.

Smart School for three cities in MP- Bhopal, Sagar, Satna

With this project BSCDCL entails at modernizing the government schools of Bhopal, Sagar and Satna with comprehensive upgrading of school facilities which includes (but not limited to) – equipping schools with technology based learning resources, modern classroom experience for the school children, School Management Solutions, Virtual Learning Environment, Digitalization of educational content, Learning Management Solutions and Library Management Solutions.

The project aims at increasing digital literacy amongst students, enhancing proficiency in academic subjects, Capacity building amongst teachers, automate process of school management (like student admission, examination management, course management) etc.

Bhopal Water Utility Management using SCADA

Bhopal Smart City Development Corporation Limited (BSCDCL) is in process to implement the existing Bhopal City Water Utility Management System (BWUMS) with GIS integrated Supervisory Control, Real time Leakage Detection System and Data Acquisition along with the centralized Control Centre. This system will monitor real time water quality and leakages in the water distribution network (Water tanks, supply lines, etc.).

Bhopal Smart City has now reached to the state where it can actually contemplate to become Data smart and move forward towards the new chapter of Data Driven Governance which can further facilitate a platform and structure of data sharing between different authorized entities to ensure data-centric decision making and outcome-based planning in governance.

Data Driven Governance

Data Governance is the exercise of decision making and authority for data-related matters.

It's a system of decision rights and accountabilities for information related processes, executed according to be agreed-upon models which describe who can take what actions with what information, and when, under what circumstances, using what methods.

Governed data is data that is trusted and understood and for which someone is accountable for both the data itself and for addressing issues about the data.

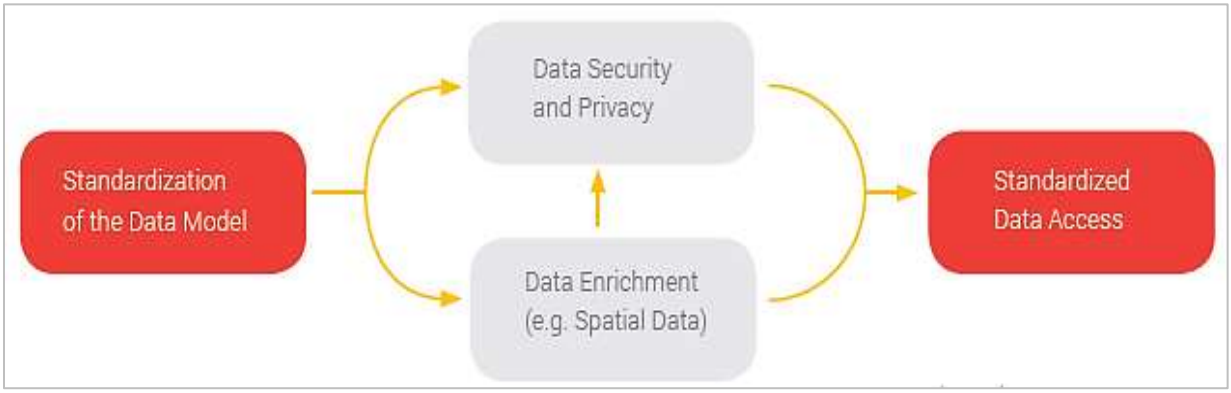


Figure 1: Functional architecture of Data Governance for a Smart City

1. Objectives of Data Driven Governance

- **Empowerment of citizens:** When cities are open about how they function, connect with their communities through various platforms during the development lifecycle of myriad projects, they become true proponents of 'Open Government'. Such cities constantly try to build trust with their citizens and engender a collective conscience amongst their communities around important issues facing their present and future. Citizens can collaborate with government easily and with increased frequency; both within their community and beyond, forming stronger groups and exchanging ideas and building new collaborations.

- **Promotion of Data Sharing and Exchange:** Efficient governance requires ready availability of relevant data. Unfortunately, data is locked up in various systems with different data owners. Open Data initiatives, data sharing, and exchange platforms will assist in facilitating G2G, G2C and G2B data sharing and exchange of data for effective decision making in real time.

- **Promotion of Multi-disciplinary research on Civic Issues:** Local data could unlock research on civic issues like transport, traffic and solid waste. Multidisciplinary researchers may provide different perspectives or solutions on civic issues to the city administration.

- **Co-Creation and Open Innovation and Civic Engagement:** City Governments will be able to work with entrepreneurs, industry, and academia to promote participation in governance, co-creation and open innovation. This will enable greater civic engagement through directed partnerships and collaborations with external and parastatal government agencies, institutions, communities, academic, research, policy and civil society organizations.

- **The emergence of Innovative Technologies:** Data is fuel for the development of solutions based on emerging technologies like IoT, Artificial Intelligence (AI), Machine

Learning (ML), Blockchain, etc. Data platforms will help cities become data ready to kick-start innovation in emerging technologies.

- **Enhancement of Transparency and accountability:** Smart cities intend to deliver reliable services to their citizens through various Smart Solutions. Implementation of the DataSmart Cities Strategy will lead to enhanced transparency and accountability among its citizens and communities by making reliable data available through data platforms. It will help build trust between the city Government and citizens.

2. Key Strategic Partners

In the process, Governments at all levels will be benefitted:

City Administration-
Municipal bodies can strategize, prioritize and manage service delivery to citizens.
Data enables the identification of visual and analytical insight to manage cities more efficiently.

State Government
State Government can utilize the resources effectively to where they are most needed. Appropriate State-level policies can be rolled out. States may choose to expand this framework to other major cities in addition to smart cities.

Central Government
For the Central Government, an open data platform helps to facilitate critical issues of national importance (e.g., pollution or drinking water availability) across multiple schemes. National policies can be formulated keeping in mind the needs of the population which will be captured precisely through the use of analytical capabilities. Incentive plans are being considered for the Smart Cities based on need and performance.

Non-Government Agencies
Entrepreneurs, Industry, and Academia are empowered with the availability of the required civic data. They can become partners in co-creation and open innovation to design cost-effective and contextual solutions to address civic issues.

3. Ensuring Data Security and Privacy

Managing security and privacy of data is crucial to building and maintaining trust between ecosystem participants and thus will be a critical element of this data policy.

Data collection, sharing and analysis must be ring-fenced by a privacy first approach to guarantee protections for residents and users. The usage rules for data elements must specify for what purposes the data can or cannot be used.

For management of Privacy and Security of Data, it is recommended that all data access must be through Application Programming Interface (API) calls to ensure appropriate security controls.

Any kind of data sharing should comply with existing standards and certifications for data privacy and security.

For open data, it is recommended that direct access to data be prohibited and use of APIs mandated. Data dissemination should be only to authenticated and authorized stakeholders (both internal and external) through data fiduciaries.

4. Data Categorization

Data will be categorized into two broad categories:

Personal Data:

Personal data means data consisting of information which is related to an individual who can be identified from that information (or from that and other information in the possession of the data users), including any expression of opinion about the individual but not any indication of the intention of the data user in respect to that individual. 'Data' is defined as information recorded in a form in which it can be processed by equipment operating economically in response to instructions given for that purposes.
Note: Personal Identifiable Information cannot be published by the City on Data Platforms under any data sets. Datasets must be anonymized before publishing.

Non-Personal Data:

Non-personal data also refers to anonymous information/data, namely information which does not relate to an identified or identifiable natural person, or personal data rendered anonymous in such a manner that the data subject is not or no longer identifiable. Anonymization means excluding any personal identifiers from data sets.

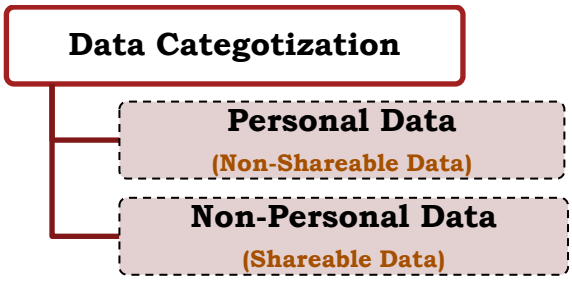


Figure 2: Data Categorization

4.1 Data Classification

Personal and Non-Personal Data will be classified into the following categories:

Classification	Class	Definition
Level 1	Public	Data available for public consumption and use.
Level 2	Internal Use	Information which can only be disclosed to municipal corporation employees for managing operations or delivery of public services on a day to day basis.
Level 3	Sensitive	Data regulated by any city/ State/Central law or regulation like privacy law etc.
Level 4	Protected	Data which needs to be protected e.g. Identity of citizens and disclosure /notification needs to be issued by the municipal corporation in case of any breach or loss of data.
Level 5	Restricted	Data which could lead to a threat to life or loss of public assets or critical infrastructure and are accessible only through a prescribed process of registration and authorization by respective departments / organizations.

4.2 Data Sets and Feeds –

Data sets/feeds must be prioritized over following criteria:

i. Data sets which promotes co-creation and open innovation.
ii. Data/Information which are frequently requested by stakeholders.
iii. Data/Information which leads to academic research on civic issues.
iv. Data/Information which helps in improving operational efficiency of City administration.
v. Data/Information which leads to enhanced transparency and accountability.
vi. Data/Information which leads to citizen convenience.
vii. Data/information which promotes multidisciplinary research using insights or analysis.
viii. Data/Information which leads to institutionalization of a culture of data driven governance.
ix. Data/Information which leads to better situational awareness.
x. Data/Information which leads to mandatory disclosures.
xi. Data/Information which forms the basis of reports/research funded by public funds by city administration.
xii. Data/Information which improves inter-department coordination.

5. Open Government Data (OGD)

In pursuance of the NDSAP- Policy notified by the GoI in March 2012, MeitY through NIC has set up the Open Government Data (OGD) Platform India - https://data.gov.in/ (link is external) to provide open access by proactive release of the data available with various ministries/ departments/ organizations of GoI.

Government collects processes and generates a large amount of data in its day-to-day functioning. But a large quantum of government data remains inaccessible to citizens, civil society, although most of such data may be non-sensitive in nature and could be used by public for social, economic and developmental purposes.

These data need to be made available in an open format to facilitate use, reuse and redistribute; it should be free from any license or any other mechanism of control. Opening up of government data in open formats would enhance transparency and accountability while encouraging public engagement. The government data in open formats has a huge potential for innovation building various types of Apps, mash-ups and services around the published datasets.

Notification of NDSAP mandates government departments to proactively open up data. NDSAP in India is applicable to all entities of Government Setup.

The OGD Platform is now available as Software as a Service (SaaS) model. It is envisaged that Ministries/Departments will release datasets on proactive/auto consumption basis through Application Programming Interfaces (APIs)/Web Services i.e. in line with the principles of Open by Default from all e-Government Service Applications particularly from Digital India initiative of the Government.

The main features of OGD platform include Single point access to open datasets, Responsive Web Layout design, Enhanced Visualization Platform, Better User Experience and efficient discoverability of resources, Cataloguing of similar resources, APIs, Embedding Catalogues, Widgets to share filtered set of data catalogues, catalogues subscription, community participation through Forums, Blogs, Infographics, Visualizations and much more.

Government Open Data License has been recently approved to ensure that the data sets released are not misused or misinterpreted (for example, by insisting on proper attribution), and that all users have the same and permanent right to use the data.
Smart Cities endorse Government Open Data License to ensure that published data is not misused or misinterpreted by its users.

5.1 Foundational Pillars of Open Data Initiative –

People:

People are at the core of the successful implementation of any strategy. Resources will need to be allocated and a robust Institutional Governance need to be set up for multi-layered management of the DataSmart Cities Strategy and capacity building.

Process:

Data Governance processes need to be institutionalized. A robust data strategy requires compliance and considerations to policies such as National Data Sharing and Accessibility Policy (NDSAP), Personal Data Protection Bill (2018) and City Data Policies. Policies go hand in hand with processes for identification, publishing, sharing and procuring of data sets and data feeds.

Platforms:

A DataSmart Cities Strategy platform can be defined as a set of Digital Infrastructure components needed for the management, analysis and use of data for a data-led governance. These would include support of external platforms like Open Government Data (OGD) platform needs to evolve over the lifecycle of DataSmart Cities' roadmap. Platforms also need to support Standards for programming interfaces, models, ontology and file formats.

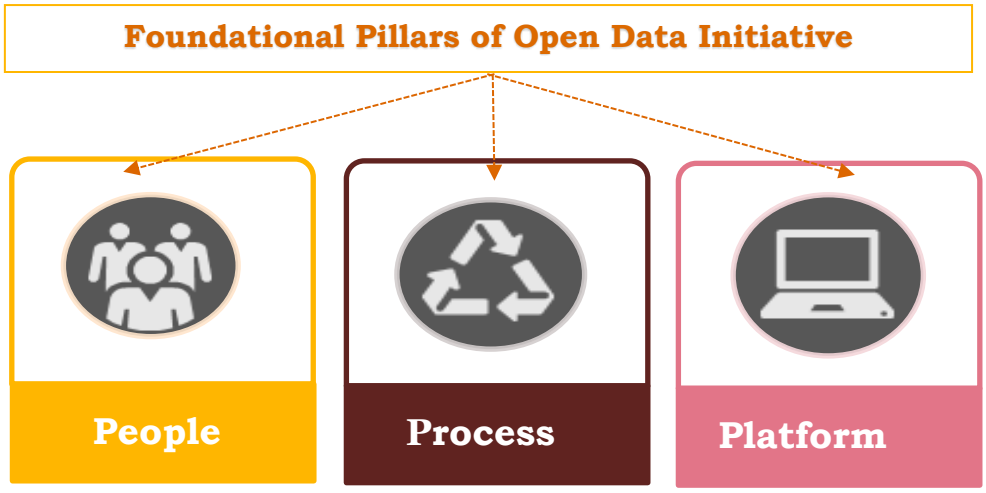


Figure 3: Foundational Pillars of Open Data Initiative

5.2 Open Government Data (OGD) Platform India

OGD Platform India has been set up at https://data.gov.in to provide collated access to Resources (datasets/apps) under Catalogs, published by different government entities in open format. It also provides a search & discovery mechanism for instant access to desired datasets. OGD Platform also has a rich mechanism for citizen engagement. Besides enabling citizens to express their need for specific resource (datasets or apps) or API, it also allows them to rate the quality of datasets; seek clarification or information from respective Chief Data Officer (formerly known as Data Controller).
OGD Platform has a strong backend data management system which can be used by government departments to publish their datasets through a predefined workflow. They can also create and/or view the visualizations created for their datasets. They shall also have a dashboard to see the current status on their datasets, visualizations, usage analytics as well as feedback and queries from citizens at one point.

OGD Platform also has a Communities component which facilitates forming of communities around datasets, domain of interest such as agriculture, education, health, or it could be app developers' community or even of data journalists. This shall give first hand input to development community for building new components, apps. It shall also give input to departments as what kind of datasets is more useful and accordingly prioritize the release of the datasets.

The data which are contributed to the Open Government Data (OGD) Platform India have to be in the specified open data format only. The data have to be internally processed to ensure that the quality standard is met i.e. accuracy, free from any sort of legal issues, privacy of an individual is maintained and does not compromise with the National security.

While prioritizing the release of datasets, one should try to publish as many high value datasets. Grouping of Related Resources (Datasets/Apps) should be planned and are to be organized under Catalogs.

NDSAP recommends that data has to be published in open format. It should be machine readable.

6. OGD Portal for Bhopal (Bhopal DataHub)

A lot of data exists in silos in different departments. Even if the culture and policy considerations are overcome, there is often a technical incompatibility between these silos, which prevents the effective sharing and exploitation of data across a city. Common programming interfaces, data representations formats and data models are necessary to achieve data interoperability. Robust open data portals, secure and intuitive data exchange platforms need to be created for effective sharing and management of data in the city.

<<

>>

The key features of the City Open Data Portal are detailed as below:

- **Open Source Driven** – Developed completely using Open Source Stack, facilitating cost saving in terms of software and licenses and also provisioning community participation in terms of further development of product with modules of data visualization, consumption, APIs to access datasets etc.

- **Metadata** – Resources (Datasets/Apps) shall be published along with standard metadata along with controlled vocabularies on government sectors, jurisdictions, dataset types, access mode etc. Besides facilitating easy access to datasets, this shall be extremely useful in the future for federation/integration of data catalogs.

- **Social Media Connect**–To support wider reach and dissemination of datasets, anyone can share the information about any dataset published on the platform with his/her social media pages on a press of a click.

- **Citizen Engagement** – The Platform has also a strong component of Citizen Engagement. Citizens can express their views as well as rate the datasets w.r.t three aspects (Quality, Accessibility and Usability) on the scale of 5. They can also embed the Resources (Datasets/Apps) in their blogs or web sites. Facility to contact the Chief Data Officers is also available on the Platform.

- **Community Collaboration** – Citizens with specific interest can build communities and discuss online. OGD Platform facilitates the communities to open up online forums, blogs and discussions around various datasets, apps available on the platform. It also provides a platform to express and discuss the kind of Datasets, APPs & APIs they would like to have. It shall also give input to departments as what kind of datasets is more useful and accordingly prioritize the release of the datasets.

- **Government Open Data License - India** – In the Task Force meeting held on September 10, 2015, the following decisions on Commercial use of Open Data has been taken: i) A committee to be formed under the Chairmanship of JS (Legal) & Chief Data Officer, D/o Legal Affairs with members from concerned departments to review and recommend the license to be associated with the data being published on the OGD Platform as well as its use for commercial purposes.

<<

>>

6.1 Evolution Roadmap of the platform

The technology platform will start through a readily available open-data platform and evolve eventually to a mature data market place. The figure below depicts the evolution of data platforms, as the data maturity increases in cities:

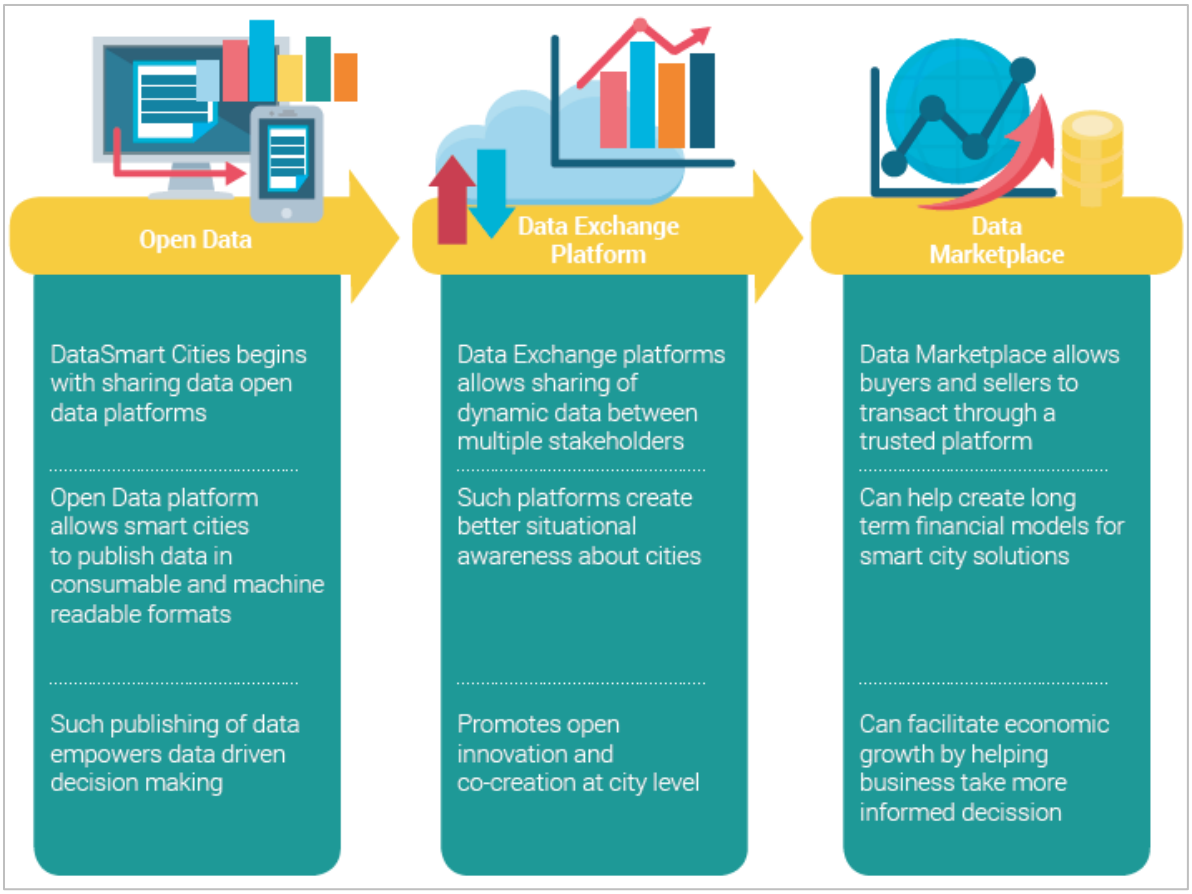


Figure 4: Evolution Roadmap of the Open Data Portal

The following table provides the purpose and the outcome of the platforms at the various stages of the roadmap:

Model	Purpose	Outcome
Open Data Platform	• To create a better understanding of cities • To empower communities through data • To provide insights for data-driven decision making • To strengthen research around cities • To provide free and open static and dynamic datasets covering historical data in a consumable format	• Increased transparency and accountability • Greater trust on Government • Enhanced G2G, G2B and G2A (Academia) collaboration • Leads to social audit and open Government • Increased public participation • Improved resource or asset visibility • Better decision making thereby leading to more efficient and cost-effective solutions. • Deepen open innovation and co-creation

Model	Purpose	Outcome
Data Exchange Platform 	• To allow stakeholders to publish and consume the data via a secure platform • Many to many relationships among stakeholders can be established • Platform act as Data Broker to create partnerships between data producers and consumers • Free and open exchange • Controlled exchange via a platform through API metering	• Better decision making thereby leading to more efficient and cost-effective solutions. • Enhanced G2G, G2B, and G2A collaboration • Helps foster data-driven decisions by diverse players in the urban economic ecosystem • Development of vibrant app ecosystem • Leads to advanced research in academic and research institution • Deepen open innovation and co-creation.
Data Marketplace 	• Marketplace to sell and buy data via a secure platform • One to one relationship among publisher and consumer • Compliance of legal framework around data • Data exchange and payment through the data portal • Central and decentralized architecture.	• Evolution of the marketplace would help strengthen the urban economic ecosystem • Enhanced G2G, G2B, G2A, and B2B collaboration • Helps cities develop new business models • Empowers communities through the sharing of data • Promotes the development of emerging technologies like Artificial Intelligence (AI), Machine Learning (ML), and Blockchain

6.2 How the Data Hubs/ Observatories can contribute

a. Assess performance-

- Track status of achieving service delivery goals in cities
- Track compliance with city master plan and aid in iterations
- Measure outcomes of schemes, projects, plans, and policies
- Track performance through Ease of Living Index, Municipal Performance Index and Sustainable Development Goals

b. Aid in day-to-day decision making

- Response to emergencies, disasters, grievances
- Service delivery, traffic management
- Identifying strategic priorities, fix problems, reduce inefficiencies

c. Aid in preparation of plans (land use, infrastructure etc.)

- Inform city managers to spot trends in city growth, better forecast impacts and, in turn, inform city plans
- Inform regulatory instruments
- Facilitate big data analysis for better decision making

7. Institutional Governance Structure

The primary goal of any institutional governance mechanism is to deliver on the set rules which are derived from the principles of data governance, i.e., integrity, transparency, auditability, accountability, stewardship, checks and balances, standardization, and change management.

7.1 Key Principles

There are two broad categories of actors who are responsible for delivering and enforcing these rules. These are:

a) Actors who have authority for decision rights, and
b) Actors who have accountability for execution

The aspiration to build a 'city-as-a-platform' model would mean the rise of a collaborative governance structure with multiple nodes, amongst which one node would assume centrality in decision-making.

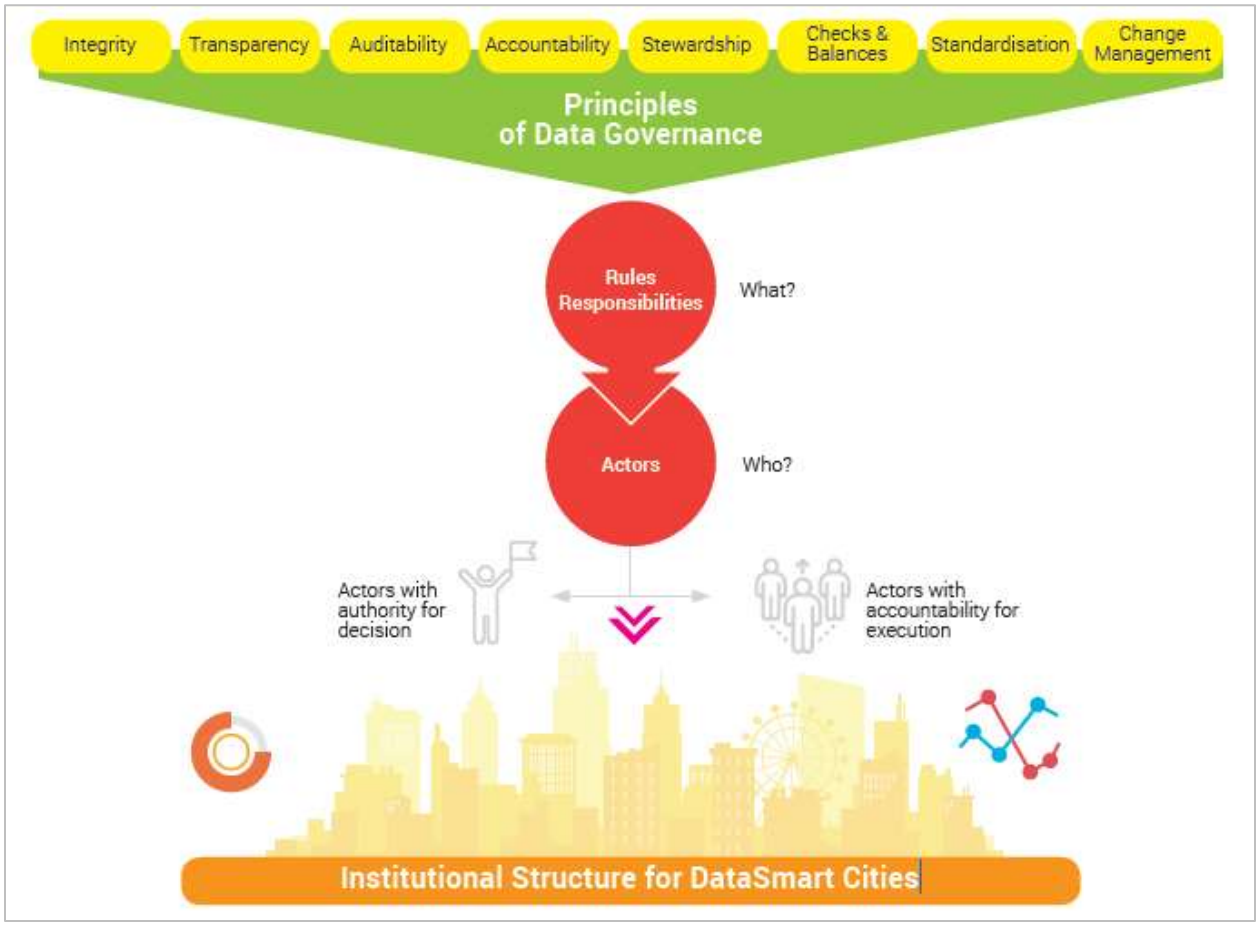


Figure 5: Institutional Governance Structure Scenario

7.2 Institutional structures at city level

The following table provides a quick view of the actors and their core responsibilities at the city level.

What (Responsibilities)	Who (Actors)
A city level Analytics and Management Unit (City Data Cell) integrated with existing Integrated Command and Control Centers (ICCC) and/or any other existing data platforms	A pool of staff with skills in data science, ICT and domain knowledge (Data Champions and Data Coordinators)
Implementation of data strategy at the city level which includes: - Creation of City Data Cell - Identification of key data sets and publish high-value data sets on the OGD Portal - Convening stakeholders and create networks/ alliances - Ensuring compliance with NDSAP, privacy compliance and required disclosures. - Prioritize data security by ensuring the Confidentiality, Integrity and Availability of data through risk management processes and best practices - Creating systems of resilience and recovery with respect to data	City Data officer (CDO) - a senior official of the rank of a Chief Technology Officer/Chief Information Officer/Department Head IT or above. CDOs will be supported by a pool of Data Coordinators. The CDO will work with all the departments including those which are outside the jurisdiction of the municipal corporation. For example, power utilities, telecom networks, gas distribution, etc.
Commitment and enhanced participation from non-State stakeholders (i.e., citizens, industry, academia) on data-sharing, co-learning and co-creation, creating an open Data Culture, bridging capacity-building needs, creating awareness about importance of Data Culture.	A City Data Alliance comprising of various key stakeholders including city Governments, other key actors in policy making, various Government department and agencies, representatives of leading academic and research institutions in the city, community organizations, entrepreneurs and advocacy groups

Figure 6: Overview of City Level Institutional Governance Structure

7.2.1 Detailed City level roles and responsibilities

I. City Data Officer (CDO)

Role

The CDO will act as custodian and driver of City Data Policy (CDP) and a flag bearer of open government initiative in respective city. CDO's major responsibility is to put data to its right use i.e. for generating insights, using data for effective service delivery or infrastructure delivery, improving civic operations by making real time decision making etc.

City data officer will work with city leadership to assess and tap the potential of data and set up data culture across the organization and outside the organization.

City Data officer will report directly to City Leadership and act as single point of contact to all internal and external stakeholders in the city. Leadership need to also deploy dedicated skilled resources to drive the data initiative through CDO.

Core objective of setting up City Data Office is to focus on setting up data driven governance culture across organization. It is implied that city leaders will be providing the required leadership support to drive the data driven decision making through seamless data collection, processing and analysis across all departments/government agencies.

Responsibilities

a. The CDOs will create a City Data Policy (CDP) for their respective smart cities which will be reviewed every month to keep it contextual to the need of the times. The policy should be created post engagement with relevant stakeholders. The SCDA would act as advisory body for the review of CDP from time to time. It will be responsibility of Municipal Commissioner to ensure that the policy evolves as per the needs of various stakeholders of the city and relevant upgrades to policy are carried out time to time accordingly.

b. Coordinate with MDO (Mission Data Officer) to align with mission data strategy and priorities with respect to Open government initiatives and policies.

c. Organize regular meetings of Smart City Data Alliance (SCDA).

d. Coordinate with officers of various other government departments/agencies within the city for the effective implementation of City Data Policy.

e. Publish Data Catalogues and Data Sets/Feeds on OGD portal on OGD Portal and will ensure that such data sets are updated at regular time intervals as needed and create mechanisms for continuous feedback from citizens and stakeholders on type of data sets to be published. The CDO will be responsible for publishing of such data sets/ feeds as mandated as part of Mission Data Strategy.

IMPORTANT FUNCTIONAL FACTORS

a. **Current IT Projects:** CDO will assess all the operational IT Projects for identifying public data sets/feeds. Data Champions and Data coordinators in respective department must prepare integration plan with respective IT vendor/integrator to ensure compliance as per CDP.

b. **Smart Solutions:** CDO will assess all proposed or under implementation project to identify the data sets/feeds which could generate public data sets/feeds. Data Officer must work with concerned System Integrator/vendor to prepare plan to ensure compliance of smart solutions with CDP.

c. **Data Reports and Plan:** CDO will assess all periodic and recurring MIS needs to identify the data sets/feeds which could be shared to other department through data exchange. CDO will also assess third party funded reports related to city operations for e.g. City Mobility plan, Health Plan etc.

d. **City Data Platform:** CDO will set up a City data page on ***data.gov.in*** in order to publish the data sets on Data Platform. CDO will publish the data sets and data feeds as per City Data Policy. CDO will remove outdated data sets/feeds. CDO will also integrate the data portal with City Website Homepage. CDO would be responsible for publishing of mandatory data sets as communicated to the cities by the MDO.

II. Data Champions

Role

Data champions (DCs) will be senior functionaries, not below the rank of a Head of Department or equivalent, who would champion the implementation of the City Data Policy in their respective departments/ organizations.

DCs needs to act as trainers and lead the team of data coordinators at the department level. DC will be first touch point of CDO in different city organizations and must undertake continuous capacity building programs for their CDOs and other staff.

Responsibilities

a. Shall identify the data sets/feeds, derived information, intelligence or data challenge with respect to day to day operations of the department.

b. Actively publish/ enable to publish data sets/feeds identified as relevant to the resolution of critical use cases for the city. They will work closely with the CDO for active implementation of the City Data Policy.

c. DCs will be assisted by the Data Coordinators within the department to streamline processes of data reporting, collection and analysis etc. DCs will be responsible for data quality.

d. DCs will undertake activities to engage with their stakeholders and evolve their department's strategy on data in line with the deliberations.

III. Data Coordinators

Role

Data Coordinators will assist DCs at the department/government agency level as reporting staff.

Responsibilities

a. Aggregate the data demand from various channels.
b. Sensitizing the department employees over the importance of data quality etc.

IV. City Data Cell

Role

The main activities City Data Cell would be to manage the OGD Platform, provide Technical Advice to the departments, handhold for dataset contribution as well as capacity building of Data Contributors and City Data Officer.

Responsibilities

1. Management of OGD Platform

OGD Platform would be managed and hosted at the << >> adhering to the Guidelines of the Government and Data security policies. The architecture would be scalable and of high availability.

2. Technical Advice

Departments would be provided with technical advice with respect to preparation of datasets, contribution of datasets, explanation of metadata and the entire workflow of data publishing, feedback management etc.

3. Capacity Building

City Data Cell will be responsible for conducting trainings and provide references and help on the Awareness and Sensitization of the policy, contributing datasets to the OGD Platform, provide advisory on conversion of data to digital format to Data Contributors, City Data Officer and other Members who are the actors under the Policy.

V. Smart City Data Alliance (SCDA)

Role

The SCDA will provide a collaborative framework to create and define use cases to solve critical city problems through the use of data, catalyze the right set of collaborations and networks to make available such data and undertake continuous dialogue between various stakeholders in the city around the City Data Policy so as to inform and evolve the CDP effectively.

The alliance will undertake education and awareness about data in the community, understand and address concerns on data privacy and security, build use cases for city problems, create data collaborations between various government and private agencies for solving relevant use cases and continuously evolve the culture of data in the city's context.

Responsibilities

a. To act as an advisory group to the city leadership on the City Data Policy.
b. To assess the data needs of various Smart City stakeholders.
c. To promote data driven governance and policy formulation.
d. To design and implement solutions and analysis using city data.
e. To support industry to design solutions using emerging technologies like AI, ML and Blockchain.
f. To assess and design use cases critical to the citizens of the respective cities.
g. To generate awareness in various stakeholders towards open government initiatives.
h. To bring Smart Cities stakeholders on common platform to influence the city data priorities.
i. To facilitate data for co-creation and collaboration over civic issues
j. To provide critical feedback to the city over the quality and relevance of data provided by Smart City.
k. To deliver 4 Research paper annually using City Data on Civic Problems in Smart City
l. To design and develop two prototype/ solutions annually on Civic Problems in Smart City
m. To organize a data-challenge every half yearly on complex civic problems
n. To organize a Hackathon annually and support shortlisted solutions at city level
o. To set up scholarship for postgraduate and graduate interns to work with Office of CDO.
p. To publish the progress report every month
q. Prioritize the Data Sets/Feeds for publishing on Data Platform:

r. To sensitize ecosystem partners to share the data for leveraging data for solving civic challenges

s. To support, engage and encourage network/groups/members of data enthusiasts in Smart City

t. To improve city capacity over data driven governance and policy formulation

u. To support CDOs by extending resources (like interns, researchers, technology experts), funds (program sponsorship etc.) and technology (solutions etc.)

v. To share data available with partners on Data Platform to promote city data.

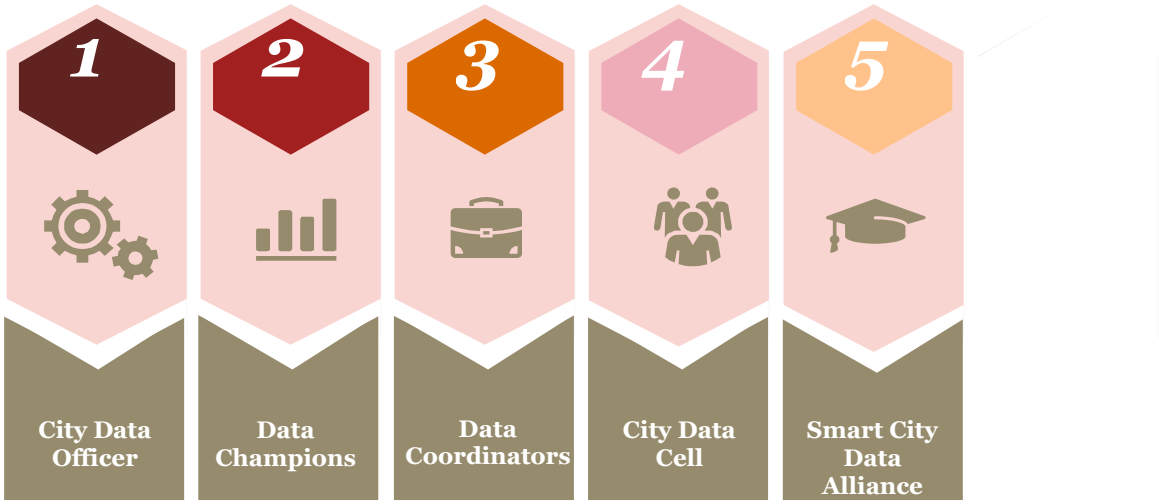


Figure 6: City Level Institutional Governance Structure

7.3 City Stakeholders to come together in CDA:

Following stakeholders should come together to set up City Data Alliance to assess, strategize, plan , implement and review the City Data Policy :

a. **Government Agencies:**
Government Agencies operating with dedicated administrative structure (apart from City Administration) in Smart City namely Traffic Police, City Police, Central/State Government Departments, Government Autonomous Bodies etc.

b. **Funding Agencies:**
Funding Agencies which regularly works with city administration in different domain.

c. **Industry:**
Key flagship manufacturing/service Industry promoters/players in the Smart city/state.

d. **Academia:**
Representatives from leading Universities/Colleges/Schools in the Smart city.

e. **Policy Advocacy Groups and NGOs:**
Policy Advocacy groups and NGOs working in different domains/areas like Slums, Health, Education, Environment, Participatory Governance, Mobility etc.

f. **Start-ups and Incubators:**
Representatives from startups and incubators in the Smart City/State.

g. **City Businesses:** Representatives from local Small and medium business communities

h. **Citizens and Communities:**
Representatives from Communities and citizen interest groups to further the interest of citizens/communities towards data driven policy governance and service delivery.

i. **Local Elected Representatives:**
Representatives from local elected representatives to further the interest of citizens/communities towards data driven policy governance and policy formulation.

j. Professional Representatives:
Representatives from various professional bodies like Doctors, CA, and Engineers etc.

7.4 Publishing & Management of Resources (Datasets/Apps)

Contribution of datasets/apps is by login into a simple web based Dataset Management System through https://----------. Resources to be contributed under Catalogs are processed through a predefined workflow, ensuring compliance with the policies. City Data Officer nominated by the City Leadership is authorized to publish datasets in open format on OGD Platform.

City Data Officer has the facility to create any number of Contributors for contributing Datasets/Apps for their Department. Once the Contributor is created by the City Data Officer, a mail is sent to the mail id of the contributor. The Contributor then can login and contribute datasets along with its metadata for further approval. However, the responsibility on the relevancy and quality of datasets published on the OGD Platform rests with City Data Officer.

Every Department should prepare a Negative list. The datasets, which are confidential in nature and would compromise to the county's security if made public, are put into this list. The datasets which contain personal information are also included in this list.

7.4.1 Registration for Creating Datasets/Apps

City Data Officer (CDO)

City Data Officer will get his login credentials on his email-id mentioned in the nomination letter. He can use the credentials to login into the web based database management system at https://----------.

Data Contributors (DCs)

The CDO can register any number of Data Contributors into the portal.

Login Credentials for the DCs will be sent on their registered email ids.

Catalogs and Resources (Datasets/Apps) are contributed by the Data Contributors by logging in to the web-based interface available at https://----------

7.4.2 Contribution, Approval and Publishing Process for Catalogs and Resources

1. City Data Officer (CDO) can Log In into the OGD Platform using his/her EMAIL ID provided in the CDO nomination letter.

2. Nominate Data Contributors. They can be Directors/Jt. Directors or Heads of respective divisions/units. They will coordinate, identify, prepare and release datasets of their division/unit on OGD Platform.
3. Data Contributors will try to prepare list of datasets which can be contributed. Prepare and contribute the metadata in predefined format for the datasets.

4. Catalogs/Resources along with the metadata contributed by the Data Contributor pass through a predefined workflow to the CDO; who in turn ensures that it is in compliance with the City Data Policy and other State level and National Level Data Policy. Catalogs/Resources are published only after approval of CDO.

5. CDO can edit the catalog/resources or can send it back to the Data Contributor for review/modification or pushes those to City Data Cell for publishing on the OGD Platform.

6. The City Data Cell pushes the catalogs/resources from staging area to the production area and publishes on OGD Platform.

7. Ministry/Department should organize a workshop on OGD Platform for Chief Data Officer, Contributors and technical team for 2 hours that would help the stakeholders to understand the process better. City Data Cell would impart requisite training during the session about how to proceed and optimally make use of the platform for uploading Catalogs/Resources.

8. Detailed step by step procedures for resources (dataset/apps) contribution and help materials are available at City Open Data Platform https://----------.

7.4.4 Consumption of Datasets from e-Gov Applications

Metadata discovery is a way to fetch datasets and metadata information from different eGov applications. Desired eGov applications have to define the source and type of tool to be used to fetch dataset and metadata from their servers. The https://---------- tool would crawl at source at specified time to fetch metadata/datasets.

DATA SOURCE FORMATS FOR METADATA DISCOVERY

Programming languages or Scripts or SQL queries can be used to write some programs to fetch the data from the relational database or some other sources to generate the desired metadata and dataset information. The https://---------- tool can fetch the data catalog information with the help of the programming language. The language passes pre-defined metadata information to the data import tool to create the data catalogs automatically without any human interface. Datasets and related metadata may be created from a variety of different sources including:

- Relational databases
- CSV files or Spreadsheets
- XML files
- Web Services or data in JSON format

Relational Databases
All most all the eGov applications use a relational database to store their reports and MIS data. Stored procedures or Database Triggers can be written to periodically export the data in to a dataset in CSV or XML format which can be consumed by available data import tool to create the dataset catalog.

CSV Files or Spreadsheets
CSV files or Spreadsheets are the most efficient and common source of data import tool. eGov applications can specify the location from which data import tool has to fetch the dataset and metadata. Metadata and dataset can be put in a zip file with predefined naming convention. With the help of the available tool designated location can be traversed, the metadata and dataset can be extracted in CSV files or Spreadsheets to convert it into data catalogs.

XML files
Structured XML files are one of the good sources for the data import. eGov applications can pass the metadata information and the details about the metadata in a predefined XML file. Data import tool can traverse the designated location of the XML at predefined time to fetch the data and convert it into searchable human readable data catalogs.

Web Services or Data in JSON Format
REST and SOAP based web service over HTTP can be used as a stream of data or services. Using REST based web services provides more flexibility to share data over XML as well as JSON format. eGov application has to submit the definition of the fields to be exposed by the web services which need to be mapped with the metadata information to create the catalog. At the given period of time with the help of metadata mapping the tool will import the data from the source apps and will convert to data catalogs.

7.4.5 View & Respond to Queries on Published Datasets

Citizens can browse, search, filter, sort and access the datasets on the OGD Platform. Citizens also have the option to send their queries and feedbacks about the published datasets. This feedback would be available on the dash board of the CDO to take further necessary action.

7.4.6 Respond to Suggestions for new Datasets

The OGD Platform has a strong Citizen Engagement feature built in. While browsing through the catalogue of datasets, if one is not able to find the dataset which is of interest to him then he can request for the same through suggestions module. Suggestions already made for particular datasets are displayed and one can also endorse the same.

The suggested list i.e. the requirement for new datasets is sent to the respective department's Data Champions. This would facilitate the CDO to prioritize his release of datasets on the platform. They are expected to send in their response on the same.

7.4.7 Review Analytics & Plan

The Dash Board of the CDO metrics would be available for the datasets contributed by all the contributors of that City. Feedback related to datasets would also be available along with the suggestions from citizen with respect to the requirement of new datasets. This feature would facilitate him to watch the analytics and accordingly plan his course of action.

7.4.8 DOs for Data Contribution and Approval

- Identify and prioritize the release of datasets; categorize the type of access granted for them and publish as many high value datasets as possible.
- Contribute datasets which are in the Open List and do not fall under the Negative List.
- Ensure that the quality standards are met i.e. accuracy, free from any sort of legal issues, privacy of an individual is maintained and does not compromise with the National security.
- Ensure that the datasets being published through a workflow process are in compliance with NDSAP. Details on original source of the dataset and methodology of the data collection should be provided in metadata.
- Prepare and contribute the metadata in predefined format for the Catalogs and Resources (Datasets/Apps). The key metadata elements are Title, Description, Category, Sector/Sub-Sector, Dataset Jurisdiction, Keywords, Access Method, Reference URLs, Data Group Name, Frequency, Granularity of Data and Policy Compliance. All the metadata elements must be filled with utmost quality and ease of use.
- Pricing of data, if any, would be decided by the data owners as per the government policies.

- Ensure that data being contributed to the OGD Platform are in machine readable or in specified open data format only. The advisable formats are:

  - CSV (Comma separated Values)
  - XLS (spread sheet- Excel)
  - ODS (Open Document Formats for Spreadsheets)
  - XML (Extensive Markup Language)
  - RDF (Resources Description Framework)
  - KML (Keyhole Markup Language used for Maps)
  - GML (Geography Markup Language)
  - RSS/ATOM (Fast changing data e.g. hourly/daily)

- Ensure that the data being uploaded on the OGD Platform is as complete as possible, reflecting the entirety of what is recorded about a particular subject and is de-normalized. The datasets also should be optimized by adding redundant data or by grouping data before uploading.

- Priority should be given to data whose utility is time sensitive. Real time information updates would maximize the utility the public can obtain from this information.

- Replace any Not Available, Not Reported or missing values in the data with 'NA'.

- Metadata that defines and explains the raw data should be included as well, along with formulas and explanations for how derived data was calculated.

- Keywords must be defined in data catalog to make it search friendly.

- Provide link to the reference documents (if any) or website for detailed information and explanation on the method of calculation or source of data.

- Read the process manual and available resources at https://---------- for any kind of help or references

- Prioritize the release of datasets and take relevant action on the basis of feedbacks and suggestions received on the OGD Platform from citizen's pertaining to the Department.

- Ensure the correctness of login details on the OGD Platform by sending a mail to City Data Cell official email id << >> , in case of any change.

7.4.9 DON'Ts for Data Contribution and Approval

- Don't contribute datasets which fall under the negative List e.g. the datasets which are confidential in nature and are in the interest of the country's security.

- Don't impose 'Terms of Service', attribution requirements, restrictions on dissemination and so on, which act as barriers to public use of data.

- Don't impose cost on the public for access of datasets, as imposing fees for access skews the pool of who is willing (or able) to access information.
- Don't publish hand written note, as it is very difficult for machines to process. Scanning text via Optical Character Recognition (OCR) results in many matching and formatting errors. Information shared in the widely used PDF format is very difficult for machines to parse. Hence, the data in these formats should be avoided.
- Data in non-Unicode formats should be avoided.
- Don't contribute datasets with any special characters (e.g. @, %, $, &, etc.) or missing values.
- Don't provide any explanation, including the method of calculation or source of data in data file to be attached in the web form.

7.4.10 Metadata Elements for Catalogs/Resources and their Description

A. Catalog

Title (Required)
A unique name for the catalog (group of resources) viz. Current Population Survey <Year>, Consumer Price Index <Year>, Variety-wise Daily Market Prices Data, State-wise Construction of Deep Tubewells over the years, etc.

Description (Required)
Provide a detailed description of the catalog e.g., an abstract determining the nature and purpose of the catalog.

Keywords (Required)
It is a list of terms, separated by commas, describing and indicating at the content of the catalog. Example: rainfall, weather, monthly statistics.

Group Name
This is an optional field to provide a Group Name to multiple catalogs in order to show that they may be presented as a group or a set.

Sector & Sub-Sector (Required)
Choose the sectors(s)/sub-sector(s) those most closely apply(ies) to your catalog.

Asset Jurisdiction (Required)
This is a required field to identify the exact location or area to which the catalog and resources (dataset/apps) caters to viz. entire country, state/province, district, city, etc.

B. Resources (Datasets/Apps)

Category (Required)

Choose from the drop down options. Is it a Dataset or an Application.

Title (Required)
A unique name of the resource viz. Consumer Price Index for <Month/Year> etc.

Access Method (Required)
This could be "Upload a Dataset" or "Single Click Link to Dataset".

Reference URLs
This may include description to the study design, instrumentation, implementation, limitations, and appropriate use of the dataset or tool. In the case of multiple documents or URLs, please delimit with commas or enter in separate lines.

*** If Resource Category is Dataset**

Frequency (Required): It mentions the time interval over which the dataset is published on the OGD Platform on a regular interval (one-time, annual, hourly, etc.).

Granularity of Data: It mentions the time interval over which the data inside the dataset is collected/ updated on a regular basis (one-time, annual, hourly, etc.).

Access Type
It mentions the type of access viz. Open, Priced, Registered Access or Restricted Access (G2G).

*** If Resource Category is App**

App Type (Required): It mentions the type of App being contributed viz. Web App, Web Service, Mobile App, Web Map Service, RSS, APIs etc.

Datasets Used: Datasets used for making this app.

Language: Language used for app.

Date Released: It mentions the release date of the Dataset/App.

Note: It mentions the anymore information the contributor/controller wishes to provide to the data consumer or about the resource.

City Data Policy Compliance
This field is to indicate if this dataset is in conformity with the City Data Compliance.

8. Driving Data Collaboration within States and at National Level

State Level-

City will continue to empower themselves with better data culture across the vertically and horizontally integrated structures. Relevant departments and agencies in the respective States will play an important role to achieve this integration. Two aspects of integration where the State agencies will play a direct role are mentioned below:

Sectoral data integration in verticals - Respective State-level departments and agencies related to one particular sector can share data in a common platform. These will gradually need to follow standard data protocols suggested by the national policies and offer cities an integrated platform for accessing sectoral datasets managed by specific departments (such as department of commerce, water resources etc.).

Horizontal integration of data through State data agencies - This will largely happen through State level agencies responsible for the collation of data from multiple sectors and thus offer a centralized data repository. The departments for municipal administration under the Urban Development Department (UDD) and the department of statistics in respective States can play a pivotal role in this regard through their MIS platforms. It is imperative that City will work in coordination with these State-level data entities to realize the full potential of the data hubs and make the desired shift towards a complete data lifecycle, i.e., from data collection to data intelligence. Development of a State level Urban Observatory would add value in this context.

National Level-

A Mission Data Officer (MDO) will be nominated for SCM by the Mission Director, SCM. The MDO will be an officer of the rank of Director or above in MoHUA. The MDO will be the officer responsible for implementation of the Strategy at national level. The MDO will engage with all 100 smart cities, NDSAP Nodal Ministry (MEITY), NIC and other relevant agencies and organizations to achieve outcomes outlined under the strategy. The MDO will continuously identify key data sets, and high-value data feeds to be published on the OGD Portal.

A **Data Analytics and Management Unit (DAM Unit)** will be constituted within the Smart Cities Mission office to act as a key support structure to the MDO to evolve the strategy over time, create capacity building within the smart cities ecosystem, coordinate with different stakeholders, advise on legal frameworks, and create data analytics capabilities within the mission.

A Smart Cities Data Network (SCDN) will be a network including inter alia, selected CDOs from 100 smart cities, along with representatives from other Ministries of Government (both State and center), industry associations, research organizations, academic institutions and legal firms. MDO will be the convener of SCDN.

City will work in collaboration with the all of the above national level bodies to define policy contours, bring in expertise in understanding the data landscape in other parts of the world, identify best practices in Indian cities, and help in interpreting the legal framework around data to use data as a tool for empowerment of the society.

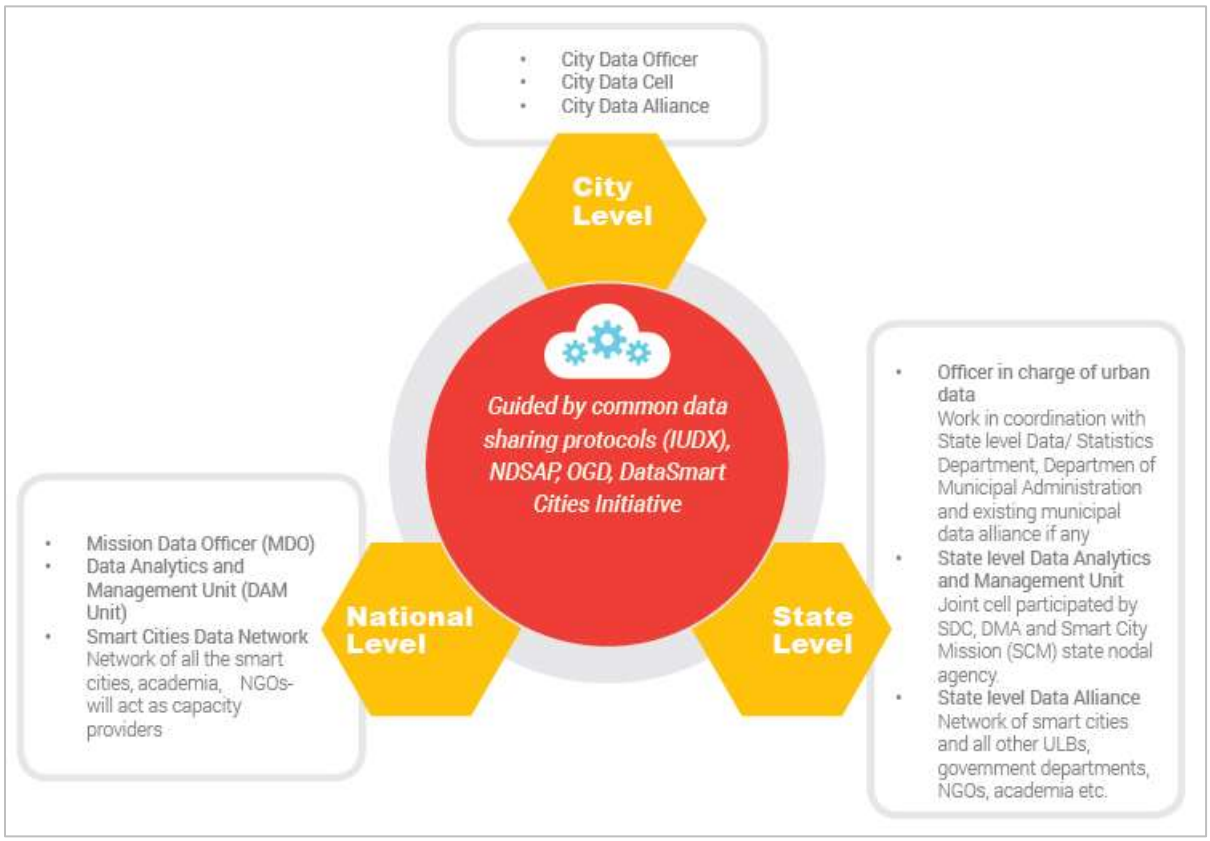


Figure 7: Data Collaboration at different levels

Bhopal endorsing NDSAP (National Data Sharing and Accessibility Policy)

NDSAP Overview

The Union Government through the Ministry of Science and Technology has formulated the National Data Sharing and Accessibility Policy (NDSAP), while Ministry of Electronics & Information Technology (MeitY) is the nodal Ministry to implement the policy.

The Department of Science and Technology (DST) under Ministry of Science and Technology has formulated the NDSAP through close collaboration with other line ministries and MeitY by creating "data.gov.in" through National Informatics Centre (NIC). The NDSAP had identified MeitY as the nodal Ministry for the implementation of the policy through NIC, while the DST continues to be the nodal department on policy matters. The policy was notified by DST (Govt. of India Gazette dated 17th March 2012).

The National Data Sharing and Accessibility Policy (NDSAP) is designed so as to apply to all sharable non-sensitive data available either in digital or analogue forms generated using public funds by various Ministries/Departments /Subordinate offices/Organizations/ Agencies of Government as well as States. The NDSAP is designed to promote data sharing and enable access to GoI owned data for national planning, development and awareness.

The objective of this policy is to facilitate access to GoI owned shareable data (along with its usage information) in machine readable form through a wide area network all over the country in a periodically updatable manner, within the framework of various related policies, acts and rules of GoI, thereby permitting a wider accessibility and usage by public.

Different types of datasets generated both in geospatial and non-spatial form by different ministries/departments are supposed to be classified as shareable data and non-shareable data. Data management encompasses the systems and processes that ensure data integrity, data storage and security, including metadata, data security and access registers. The principles on which data sharing and accessibility need to be based include: Openness, Flexibility, Transparency, Quality, Security and Machine-readable.

NDSAP once implemented would lead to:

1. Opening up of the information out of the Government System

2. Making available of the Accurate, Reliable and Unbiased information
3. Providing single data OGD Platform for the country for data sharing
4. Establishment of a platform to promote innovation in government applications
5. Enhancing government Transparency, Accountability and Public Engagement
6. Effective utilization of Government data by providing meaningful visual representations
7. Enabling development of Innovative Applications around datasets or mash-up from multiple datasets hence giving different perspectives to government data

DataSmart Cities will be implemented in consonance with NDSAP Guidelines.

Bhopal is also working in conjunction with the NDSAP Policy.

National Data Sharing and Access Policy defined standards

National Data Sharing and Access Policy defines standards for publishing data sets and feeds. SCDOs must ensure adherence towards defined standards and classification:

a. **Open by Default:** Data sets are considered to be open by default unless classified as internal, sensitive, protected or restricted.

b. **Meta Data:** Data sets and feeds must be published with proper metadata. Information about the datasets being published using common data taxonomy/structure is needed as it helps in providing easy access through Data Platform.

c. **Data Catalogue:** As per NDSAP metadata elements for data sets or feeds is defined as follows*:*

d. **Title (Required):** A unique name for the catalogue (group of resources) viz. Current Population Survey, Consumer Price Index, Variety-wise Daily Market Prices Data, State-wise Construction of Deep Tube wells over the years, etc.

e. **Description (Required):** Provide a detailed description of the catalogue e.g., an abstract determining the nature and purpose of the catalogue.

f. **Keywords (Required):** It is a list of terms, separated by commas, describing and indicating at the content of the catalogue. Example: rainfall, weather, monthly statistics.

g. **Group Name:** This is an optional field to provide a Group Name to multiple catalogues in order to show that they may be presented as a group or a set.

h. **Sector & Sub-Sector (Required):** Choose the sectors(s)/sub-sector(s) those most closely apply(ies) to your catalogue.

i. **Asset Jurisdiction (Required):** This is a required field to identify the exact location or area to which the Catalogue and Resources (dataset/apps) caters to viz. entire country, state/province, district, city, etc.

j. **Open data:** Data Sets and feeds should be published in formats specified under NDSAP i.e. Open format. Data should be provided in freely available formats which can be accessed without the need for a software license.

k. **Machine Readable:** Data Sets and Feeds should be machine readable.

l. **Formats:** As per NDSAP following data formats should be published:

- CSV (Comma separated values)
- XLS (Spread sheet - Excel)
- ODS (Open Document Formats for Spreadsheets)
- XML (Extensive Markup Language)
- RDF (Resources Description Framework)
- KML (Keyhole Markup Language used for Maps)
- GML (Geography Markup Language)
- RSS/ATOM (Fast changing data e.g. hourly/daily)

m. **Maintenance of Data Sets/ Feeds:** CDO will ensure that published data sets and feeds are up to date and relevant.

n. **Support:** CDO will provide required technical and non-technical support over the queries/inputs/suggestion received from users through email, portal or through social media platforms like Facebook, Twitter.

o. **Archiving:** CDO will define and set up process for archiving data. Every data set/feeds catalogue must contain archiving information. Data Sets published over open data portal will be retained as per retention policy. For specific file type (geo-spatial files), recent copy must be made available to users through Data Platform.

p. **Ownership:** All data sets/ feeds remain property of publisher i.e. CDO. The CDO will endorse Government Open Data License to ensure that published data is not misused or misinterpreted by its users.

The CDO is responsible for ensuring that data sets /feeds published are regularly updated as per defined frequency. CDO is responsible for managing the necessary compliance as per NDSAP. CDO is also responsible ensuring proper privacy compliance and required disclosures

Bhopal endorsing MPDSAP (Madhya Pradesh Data Sharing and Accessibility Policy)

The City Data Policy is to be read in conjunction with the **MPDSAP (Madhya Pradesh Data Sharing and Accessibility Policy).**

Objective of the Policy

The objective of MPDSAP Policy is to facilitate the availability and access to data and information in both human readable and machine readable forms through an electronic network in a proactive and periodically updatable manner, within the framework of various related Policies, Acts and Rules of the Government of Madhya Pradesh and Government of India.

Scope of the Policy

MPDSAP (Madhya Pradesh Data Sharing and Accessibility Policy) is designed so as to apply to sharable non-sensitive data available either in digital or analog forms and generated using public funds by various State Departments/ Subordinate Offices/ Organizations/ Agencies. It is designed to promote sharing and enable access to Government owned data that could be used for planning and development.

Implementation of the Policy

1. National data portal data.gov.in being operated by NIC will be used for wider dissemination of data under this policy as well as through the Government of Madhya Pradesh portal.

2. The Department of Information Technology will be responsible for Implementation of this policy in close collaboration with the State Unit of NIC.

3. Each Department will make resources available under their own budget allocations for implementation of this policy.
4. The Nodal Department will issue the detailed implementation guidelines similar to the National Implementation Guidelines issued by DeitY, GoI.

5. All the department will provide atleast three high value datasets within six months after notification of the policy. The department will endeavor to upload all existing data in one year's time and hereafter as and when data is available.

6. The standards of NDSAP will be followed to ensure interoperability. The state nodal department will develop their state portal on the lines of the national portal developed by NIC.

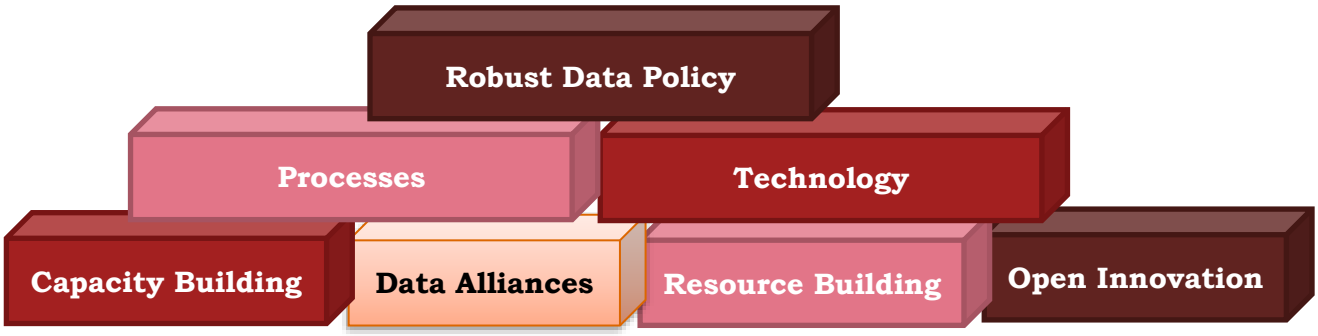
7. An oversight state committee will be constituted for facilitating the implementation of the policy and its provisions thereof under the chairmanship of the chief secretary and convener of the head of the nodal department in the state.

8. Government department will encourage community participation for development of application using the data made available under this policy with a view to empower the citizens.

Driving the Data Culture Transformation

In order to create a self-sustaining, usable data management ecosystem various elements will be required to come together so true value may be derived from the data being captured.

1. Building Blocks of Data Culture



Robust Data Policy

The aim should be to build a Safe, Secure, Easily Accessible environment for data collection, collation and analysis. The data governance policy should serve as clear documented set of guidelines for ensuring the proper management of managing digital information across all stages of the data management lifecycle.

Processes

Processes need be formulated in a way so as to collect data in an efficient manner. Data collected should be accurate, clean and be supported by layered reviewing capability basis importance of data and available resources in the organizational structure chain.

Technology

The technology should be future ready to manage the pre-empted/strategized data needs.

Open Source and Proprietary technologies that provide tremendous data capture and analysis capabilities are readily available. These should be chosen bearing in mind factors like data usage goals (data indexing, depth of analytics required), storage and indexing needs, technical expertise available and financial appetite.

In most cases a lot of data is already available with various government bodies that could of use to the Smart City. The chosen tools and corresponding processes should also be able to cater to the data demands from various entities.

It should be noted that ease of data collection should be the top priority at data collection points since the users collecting the data may not be technological sound in some cases.

Data should be secure and accessible according to the user rights both in raw or analyzed form. The analyzed data should be accessible on multiple devices and other desired channels.

Capacity Building

Post identification of tools and formulation of processes, users at all stages will require training on how to collect data and using the analytics tools.

After the technology intervention and processes have been identified, focus on capacity building, technical training will happen.

At times the resources deployed at the last mile may not understand the value of meticulous and focused data collection and may need to be trained not only on the tools but also on the importance of collecting data properly.

Data Alliances

While sustainable data influx and can be ensured using above building blocks, it is importance not to ignore already present data at various other government and private sources. Such data can help serve as an important supplement to the existing data sources and on channels as desired.

This data might not be captured in the same format and will need cleaning and normalization before analysis can be done.

Resource Building

In order to handle the complexities of the data, it is important that the right resources be deployed at critical points in the data life cycle. These resources can serve to guide the teams at these critical points in the data life cycle so as to avoid spill overs, incorrect data entry, cleansing techniques etc.

These data scientists/data champions will be instrumental in generating useful insights and deliver value to the senior stakeholders.

Open Innovation

Innovation is an important element in the success of any technological initiative. In order to drive innovation Hackathons/Innovation competitions could be held. Better techniques tailored to the Smart City needs could be used or useful data could be shared with participants to drive new insights that may have been previously missed.

Strategy Formulation and Mind-set Transformation

The first step towards building a sustainable data culture is to identify current data goals, to strategize/pre-empt any upcoming data priorities. What are the final outcomes that are required from the data? What sort of value needs to be added from the analysis of said data? Who all will be using the data? How will the data be accessed?

The only way a sustainable 'Data-Culture' can be created if the mind-set of users from officers to last mile data collectors is in sync with the importance of collecting the data and the value that it adds by providing varied actionable insights.

Although the chosen technological platforms may be conducive to allow easy collection and management of data, re-iterating the importance of messages would go a long way in creating a data driven culture.

2. Institutionalizing Data Culture in Cities

Fundamental objectives in setting up data culture in smart cities is to make better decisions.
Following key initiatives are required to institutionalize the decision making culture in Smart City:

Assess Data Requirements

City Data Officer along with team of data champions/ coordinators must assess the data requirements of various stakeholders in smart city ecosystem. External stakeholders needs to be engaged to understand the data needs. City Data Officer must engage various internal stakeholder at Operational, Tactical and Strategic level to assess the data need to make decisions. Data needs and frequency of consumption needs to be outlined for internal stakeholders.

Define Use Cases

Department specific use cases needs to be outlined for decision making or policy formulation keeping in mind the need of stakeholders to enable data driven governance. Data driven governance has two aspects:

a. **Data for Decision making:** Decision making requires right data in right format at right time. Use cases must be outlined for internal stakeholders in each department. For example, In order to respond to citizen complaint on-ground staff need data about nature of problem, complexity of complaint, severity of complaint, photos and exact location to respond to the user complaint At the same time tactical level i.e. middle management need aggregated data to measure the performance of its ground staff to gauge health of operations through indicators like number of complaints resolved, re-opened etc. At strategic level, leadership can use city wide data to assess the trends, citizen satisfaction ratings, patterns to allocate resources across various departments based on data.

b. **Data for Policy Formulation:** For policy formulation city administration requires data from various sources i.e. primary and secondary research. City Data office leverage the existing data available in different formats in different systems or may use crowdsource data or third party data to design policy. For Ex : To design City Parking Policy, city administration must have data about number of registered vehicles in city,

expected growth in next 5 years, mobility trends and patterns of congestion in different areas etc. These data sets could be collected through primary research and also through secondary research if available over the web.

Stimulate Data Demand

Data demand need to be stimulated by city leadership in different stages of Project/policy conceptualization, design, implementation and its monitoring. City Leadership may outline the Key Performance Indicators and scorecard for each line of service to assess the performance of its resources i.e. manpower, capital, assets to stimulate the data demand across the enterprise.

Publishing Cross Cutting Data Sets

City Data Officer must identify cross cutting data sets and publish it on platform based on data needs of various stakeholders in routine manner. Cross cutting data sets at the city levels could be leveraged by any department for its own purpose. For Example: location of schools, rates of property in different areas on map etc.

Hackathon / Data Challenges etc. for Urban Innovation

Government alone cannot solve all its problems. It needs to bring academia and industry together to solve its urban complex problem through co-creation and open innovation. City Data Officer must design a program to solve its problems through a structured challenge process.

3. Impact of Data centric-decision making

Stage	Impact of Data Centric Decision Making
City Planning	✓ Equitable Access to community resources and Infrastructure ✓ Affordable Infrastructure and resources ✓ Desired Utilization of Infrastructure
Policy Formulation	✓ Improves Quality of Life ✓ Equitable distribution of resources ✓ Public Policy that works for its citizens
Project Implementation	✓ No Scope, Time and Cost Creep ✓ No additional burden on city finance ✓ Low Project Failure rate

Stage	Impact of Data Centric Decision Making
Service Delivery	✓ Meets Service Level Expectations ✓ High Scalability and Interoperability ✓ High Customer Satisfaction Index ✓ High Morale of Employees ✓ Proper management of resources
Project Operations and Sustenance	✓ Delivers outcomes in-line with Project Design

Illustrative scenarios

City Planning	Should we set up Public Hospitals in the city? Or should we launch an insurance policy to offer coverage in private hospitals instead of investing in infrastructure?	▪ Do we have data of who accesses public hospitals? Where do they stay in the city? How much they earn annually? ▪ What is the ratio of patients to doctor across the city? ▪ How much do people pay towards medical expenses? ▪ What is the total cost of setting up and managing hospitals? How much time it will take to set up a new hospital? ▪ Do we have Cost benefit analysis of investments in setting up and managing infrastructure versus spending on offering medical insurance plan?
Policy Formulation	Should we increase parking fees to discourage private vehicles during peak time?	▪ What is the trend and projected rate of growth of two wheelers and four wheelers in the city? ▪ How many citizens uses public transport today and what is YoY rate of growth? Will it be sufficient to cater the growing need? ▪ In which areas of city and what time of the day citizens faces parking issues? ▪ What % of vehicles are from outside the city? ▪ Do we have mobility analysis of the city which could help us to know from

		where traffic emanates and converge in the city? ▪ What is the price point which could discourage the vehicle owner to bring private vehicle on road and use public transport?
Project Design	Should we set up city wide Wi-Fi or City fiber to provide city wide connectivity?	▪ Projected growth of connectivity demand from household and businesses in the city? ▪ Projected growth of internet mobile users in the city? ▪ Area wise bandwidth consumption trend and projected growth? ▪ What is length of fiber network in city today? ▪ Cost benefit analysis of setting up Wi-Fi and City-wide fiber across the city?
Project Implementation	What is the status of project implementation?	▪ Do we have Cost Creep? If Yes, by what % of total project cost? ▪ Do we have Time Creep? If Yes, what is estimated impact on project cost?
Service Delivery	What is satisfaction level of citizens over garbage management services?	▪ Area wise trend analysis of garbage complaints ▪ Satisfaction rating analysis over garbage complaints ▪ Communicable diseases hotspots across cities and correlation with poor satisfaction rating
Project Operations and Sustenance	How can we predict next water leakage situation?	▪ Area wise water quality complaints trends assessment ▪ Area wise Low-Pressure complaints trends assessment ▪ Ageing analysis of pipelines across the city ▪ Maintenance Schedule trend analysis of past years

4. Data Architecture @ Smart City

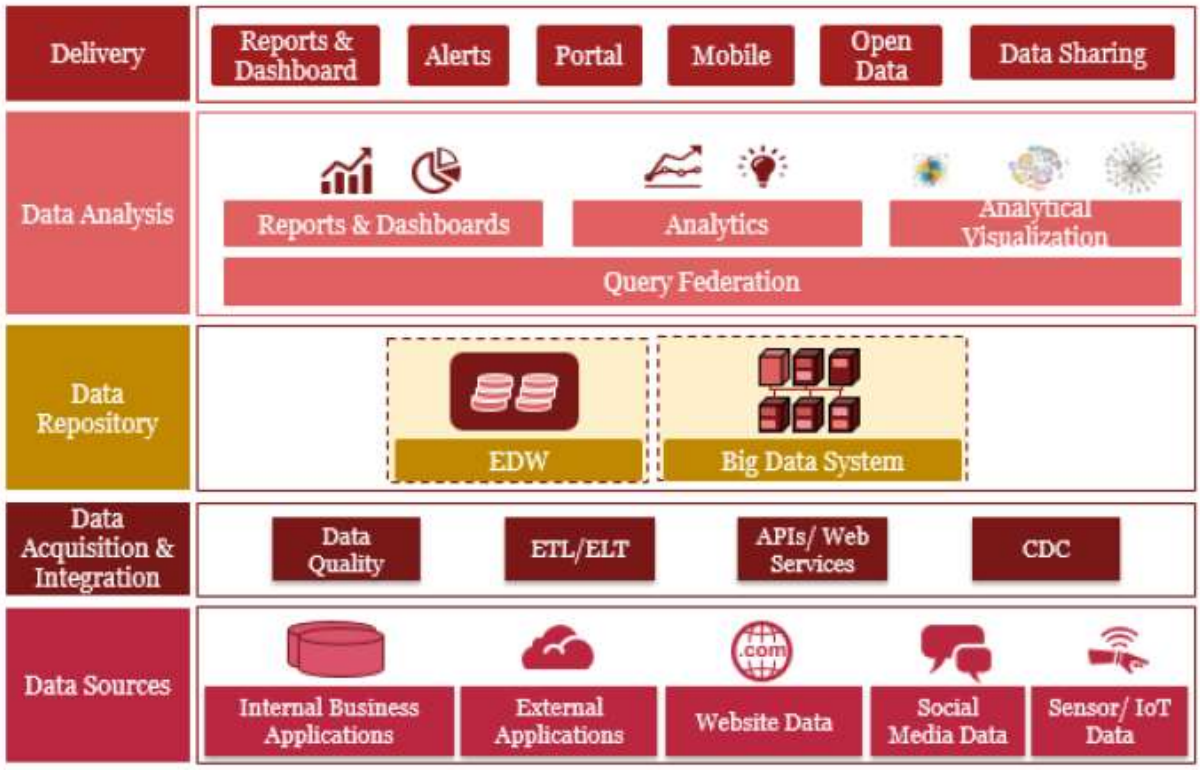


Figure 8: Smart City Data Architecture

The Smart Cities reference data architecture is explained below:

Data Sources

Smart Cities ecosystem comprises of various technology solutions ranging from Sensors, IoT, SCADA, Electronic camera, GIS, payments system etc. which generates loads of structured data every second on different dimensions. Smart-Cities ecosystem could leverage various unstructured data emerging from different sources and third-party systems like social media, internet, websites, videos, images etc.

Data Collection and Acquisition

Data Collection and Acquisition Layer acts an interface between data sources and Intelligent Platform (Command and Control Platform, Analytics Platform) which comprises of engines and solutions to customize and prepare the data collected through various sources and formats i.e. structured and unstructured for analysis purpose.

Enterprise Data Repository

Processed Structured and Unstructured Data is stored in Enterprise Data Warehouse from where it could be used by various users and applications for decision making.

Data Analysis Layer

Comprises of engines to process the structured and unstructured data on various dimensions for various purpose. Data Analysis Layer engine helps user to derive information, intelligence and knowledge out of processed data stored in Enterprise Data Repository for analysis or decision-making purpose.

Service Delivery Layer

Service Delivery layer comprises of various applications and systems which could be used to deliver information, intelligence and knowledge to end user.

Life-Cycle of Data

Figure 9: Life Cycle of Data

Data Collection

Data collection is becoming easier and affordable in real time with each passing day due to advancement in technologies.

a. **Field Data:** Smart Cities are deploying various IoT sensors, actuators, devices, cameras and solutions to capture the data from the field directly. Gartner estimates "that by 2020, there will be 9.7 billion connected things in Smart Cities, and 81% of those things will come from smart home and smart commercial building sensors."

b. **Operations Data:** Various system are deployed under various departments to manage city core operations like Water Supply, Surveillance, Traffic Electricity, Street Lights, Water and Sewerage Treatment, Health, Education, Fire Department, Disaster Response and License permits etc. These systems generate various data points in different formats which provides critical information and intelligence to officers to manage critical services and Infrastructure of the city but in silos.

c. **Third Party Platform and Mobile Apps:** Citizens also avail services from private businesses like radio taxi, food delivery, hospitals and labs etc. which could provide various insights and patterns which could be instrumental in policy formulation and city planning.

d. **Internet:** Various platforms engages citizens and communities which captures general sentiments which could provide insights mood or opinion of citizens and communities towards specific issue.

These data sources and systems could generate structured and unstructured data.

- **Structured Data:** Data which is generated by systems or humans and could be handled using existing or predefined models. Structured data could be stored in relational data bases and analyzed using basic search algorithms. E.g.: Location data, User transactions, Sensor data etc.

- **Un-Structured Data:** Data which is generated by systems or humans and cannot be handled using existing or predefined models. Un-structured data cannot be stored in relational data bases and is difficult to analyze using basic search algorithms. E.g.: Images, Video files, Audio Files etc.

Data Integration

Data from different systems using different technologies comes in different size, shape and format. In order to derive meaningful information from structured and unstructured data it is required to make data compatible for consumption. Various data types and formats generated from various smart cities systems which are as follows:

a. Field Sensors and Devices:
- Hierarchical files (JSON, XML, YAML, etc.)
- Real-time stream
- Objects
- Videos
- Images
- Locational Data

b. Operations Systems
- Relational Data Structures
- Blocks of raw data
- Flat text files
- Documents (.xls, .pdf, .ppt, etc.)
- Log files
- Financial Data
- Location Data

c. Internet and Social Media Content:
- Blogs
- Video
- Music
- Sentiments
- Images

Extract, Transform and Load (ETL) is the common methodology used for data integration. It is a three-step process which used for data integration to blend data from multiple sources. It's often used to build a data warehouse. During this process, data is taken (extracted) from a source system, converted (transformed) into a format that can be analyzed, and stored (loaded) into a data warehouse or other system.

Data Analysis

Analytics is an encompassing and multidimensional field that uses mathematics, statistics, predictive modeling and machine-learning techniques to find meaningful patterns and knowledge in recorded data. Applying intelligent techniques to uncover insight from the relevant data.
For examples:

a. Slice and dice to drill down the data till lowest entity

b. Trend analysis and pattern identification on time series (days, weeks, months, quarter or seasonal etc.)

c. Trend analysis and patter identification using various dimensions: Cost, Budget, domain specific parameters etc.

d. Comparison between various parameters in different geographies etc.

e. Visualization to view the trends and patterns for decision making. Converting the data into a more comprehensible and user-friendly format.

Data Intelligence

It is final stage of the journey from being informed to actionable insights and finally to actions using following techniques:

Descriptive Analytics:
It helps in answering "What is happening?"
For Example: Using past financial performance to predict a customer's likely financial performance. Descriptive analytics can be useful in the sales cycle, for example, to categorize customers by their likely product preferences and sales cycle.

Diagnostic Analytics:
It helps in answering "Why did it happen?"
For Example: For a social media marketing campaign, you can use descriptive analytics to assess the number of posts, mentions, followers, fans, page views, reviews, pins, etc. There can be thousands of online mentions that can be distilled into a single view to see what worked in your past campaigns and what didn't.

Predictive Analytics:
It helps in answering "What is likely to happen?"

For Example: Some companies are using predictive analytics for sales lead scoring. Some companies have gone one step further use predictive analytics for the entire sales process, analyzing lead source, number of communications, types of communications, social media, documents, CRM data, etc. Properly tuned predictive analytics can be used to support sales, marketing, or for other types of complex forecasts.

Prescriptive Analytics:
It helps in answering "What should I do about it?"
For Example: In the health care industry, you can better manage the patient population by using prescriptive analytics to measure the number of patients who are clinically obese, then add filters for factors like diabetes and LDL cholesterol levels to determine where to focus treatment. The same prescriptive model can be applied to almost any industry target group or problem.

5. Maturity Model for Data driven Governance

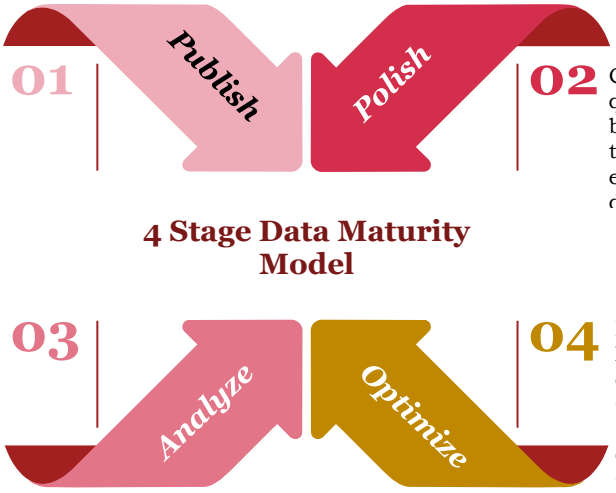
Smart Cities is all about getting right data at right time for effective decision making and policy formulation. Smart Cities are setting up various sensors, actuators, devices, systems and ICT solutions which could assist in making real time decision making to enable them to ensure proper utilization of assets and resources and to do more with less resources.

Smart Cities through integrated systems could leverage data available in silos with different stakeholders to derive information and intelligence for decision making and policy formulation. Advanced Analytics tools also helps in deriving intelligence using past data trends from various systems and is able to predict outcomes based on different scenarios.

Thus, it is critical to tap the data in various size and shape available in various formats to generate better situational awareness at any given point in time.

Broadly, there are 4 levels of maturity in data led governance frameworks:

City is publishing open data and is beginning to develop the tools and skills to use data for decisions



City id improving quality, quantity and value of data, is building systems/platforms to analyze and use data, has emerged the public and is developing a data culture

4 Stage Data Maturity Model

City demonstrates commitment to analyzing and using data. Data Governance is well established, capacity for analytics increased. Shared platform and training support for data culture

Leaders and Managers at all levels are committed to using data. A community of Data Champions provide mutual support. Successes are documented. City wide Data Culture nourishes.

Annexures

ANNEXURE 1

City Data Maturity Assessment

City Data Maturity Assessment is a framework for a deep self-evaluation for city governments to help them assess their readiness in becoming DataSmart Cities. The objective is to help cities emerge as 'Digital Leaders' in a paradigm of data-driven governance.

A strive towards data-driven governance brings two important aspects to focus, i.e., data as a process; and data for achieving outcomes. The first refers to a robust data governance mechanism along with the supporting institutional structures. The second refers to the availability of data itself, and its usage in decision-making.

The intent of DMAF is to provide a comprehensive yet pragmatic set of indicators to help cities assess their preparedness in both these aspects. DMAF based assessment focuses on "People, Process and Platform" approach in all Smart Cities as outlined in the DataSmart Cities Strategy. It seeks to promote a spirit of competitive benchmarking amongst the 100 identified Smart Cities that will enable them to assess themselves at varying degrees of data maturity with respect to a standardized framework covering aspects of enabling policies, governance structures, data management, capacity building, and stakeholder engagement.

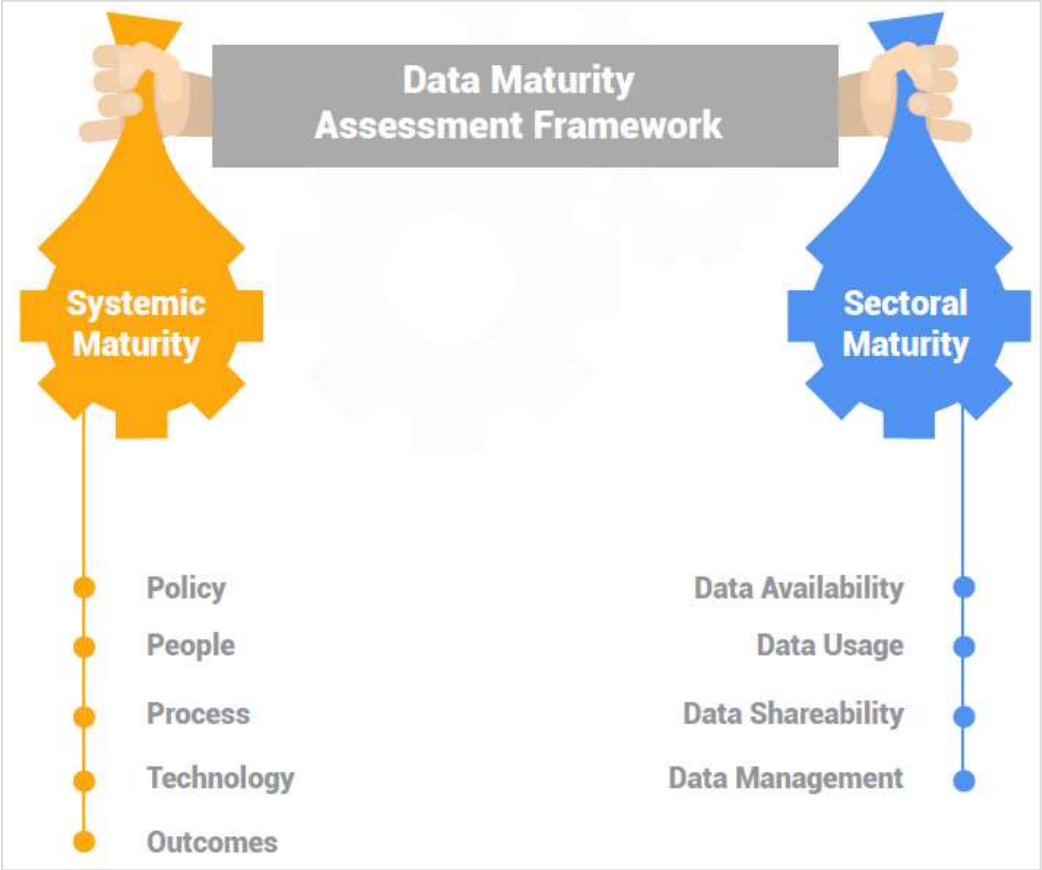
As an intermediate outcome, it is envisaged that DMAF will help cities to be better prepared in periodic self-evaluation through important key urban performance indicator sets proposed by the Government of India (such as 'Ease of Living Index' and 'Municipal Performance Index').

The Data Maturity Assessment Framework and associated evaluation will be carried out in the cities on an evolving basis at regular frequency, to allow cities to increase their understanding of data governance principles and institutionalize associated practices, leading to effective decision making and citizen centric delivery of services. This will also play an enabling role in furthering innovation, collaboration, co-creation and research.

Cities will be able to utilize the challenge to identify successful approaches and key learnings in a structured manner and connect with each other to replicate relevant approaches. Over time, it will engender a robust, datadriven innovation ecosystem that will help cities solve key urban challenges in a context-sensitive manner.

Key Pillars

DMAF comprises two key pillars - vis. **Systemic Maturity and Sectoral Maturity**. The Assessment Framework may include elements from both pillars depending on the maturity levels of cities and the stage of the challenge. Over time, the weightage of each pillar in the scoring methodology and certification process will shift in order to motivate and incentivize cities to first build foundational systems and then focus on achieving depth in key urban sectors as they implement the DataSmart Cities Strategy.



1. Systemic Maturity

Systemic Maturity measures the ability of cities to implement the DataSmart Cities strategy from the perspectives of people, processes, technology, policies, and outcomes at the city level. This pillar is the foundational cornerstone of a city's ability to ensure effective data governance, enhanced usage of data in decision-making processes, and drive cities towards better interdepartmental, interagency and systemic collaboration.

Policy: This component assesses the existence of robust policy mechanisms in the city around data governance, empowerment, protection, collaboration and innovation. It also includes the presence of necessary budgetary allocations to operationalize the policy.

People: This component assesses the presence of empowered city officials with the capacity to guide the development of city data policies, manage data governance, drive interdepartmental and inter-agency data exchange and to build city data alliances

Processes: This component assesses the effectiveness of the city's processes around data collection, usage, management, security, privacy, empowerment, collaboration, and innovation

Technology: This component assesses the quality and robustness of the city's information and communications technology infrastructure including digital platforms, sensors, IoT devices, data exchanges, big data and artificial intelligence.

Outcomes: This component assesses the quality of outcomes around data driven governance, ease of living, ease of doing business, collaboration and innovation in the city.

2. Sectoral Maturity

Sectoral Maturity measures the ability of cities to harness the power of data by focusing on availability, usage, sharing and control management of data in key urban sectors. This pillar recognizes that while data is the underlying language with which cities can identify, analyse and solve urban challenges, solutions are sector specific and hence the effectiveness with which problems are solved is a function of the maturity of data systems, governance, resources and collaborations in each sector.

Availability: This component measures the availability of real-time, reliable, systemic data in each sector that is geospatially enriched and integrated with key functions and processes.

Usage: This component measures the effective usage of data to drive decision making, improve service delivery, manage departmental functioning, foster inter- agency cooperation and ecosystem collaboration.

Shareability: This component measures the existence of data sharing processes, anonymization and machine readability of data sets and publication of open data in keeping with the City Data Policy and/or NDSAP guidelines.

Management: This component measures the existence of effective structures, systems and processes to manage data access and controls over departmental and sectoral data sets, robust backup and retention policies and continuity plans to deal with data loss and / or systems failure.

Guiding Principles

In line with the spirit of DMAF as expressed through its intent, following three principles guide the design of assessment methodology and indicators.

A. Context Relevant Assessment

It is important to ensure that DMAF is context relevant to the current state of Smart City evolution across India. It should support their efforts to consolidate existing efforts and successes while planting the seeds of their future growth trajectories. Thus, each cycle of assessment will need to be calibrated to provide the right balance between effective and relevant evaluation and setting aspirational targets for future development of the city's data culture.

B. Engendering Self Reflection

Assessments will only be of value when they will be able to provide cities the tools and inputs required for their growth. Therefore, assessments under the framework will be carefully calibrated to enable city officials realize the existing gaps in their data ecosystem, and to make effective decisions aimed towards closing these gaps in a sustainable manner.

C. Supporting Effective Planning for DataSmart Cities Strategy Implementation

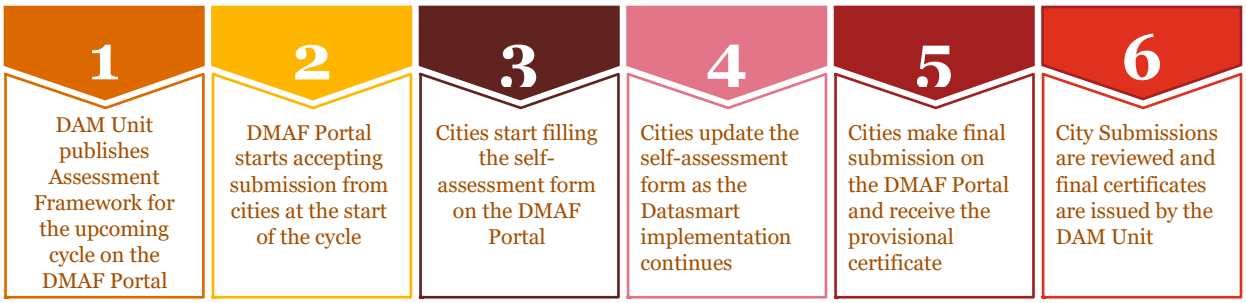
It is important to ensure that as the DataSmart Cities Strategy unfolds and evolves over time, cities are equipped to implement it effectively in their context. Thus, assessments under DMAF will always be geared towards enabling the necessary planning and resourcing efforts for cities to implement the strategy in a holistic manner to derive its intended benefits.

Responsibilities of the DAM (Data Analytics and Management Unit, GoI)

The DAM Unit will coordinate the execution of DataSmart Cities Challenge from the Ministry's office. It will provide cities the following areas of support during the process:

- Host capacity building sessions like Webinars with Experts, facilitating peer-to-peer learning to share best practices and approaches in implementing the DataSmart Cities Strategy.
- Issue guiding documents in the form of draft model policies, frameworks, guidelines, recommendations and use cases to support cities in the implementation of the DataSmart Cities Strategy in keeping with the focus areas of the current assessment cycle.
- Support cities with suggested strategies, approaches and access to experts and partners that are matched to their existing certification levels.

Assessment Process



The DAM (Data Analytics and Management Unit, GoI) Unit will publish the Assessment Framework for each cycle ahead of the commencement of the assessment cycle on the DMAF portal. At the beginning of the assessment cycle, the portal will start accepting submissions from cities to participate in the assessment cycle.

Each city will be able to participate and share necessary data, information templates, supporting documents as indicated on the assessment framework portal. As the city continues its implementation of the DataSmart Cities strategy, it may edit its submission multiple times during the challenge period until it is ready to make the final submission or the final day of the challenge period. At this point, the information submitted will be finalized and cannot be modified. On submission, cities receive a provisional certification from the DMAF portal.

In cases where the self-assessment is not supported by requisite documentation, the assessment scores may be adjusted in accordance with the verification protocol. The final certifications are issued by the DAM Unit after review.

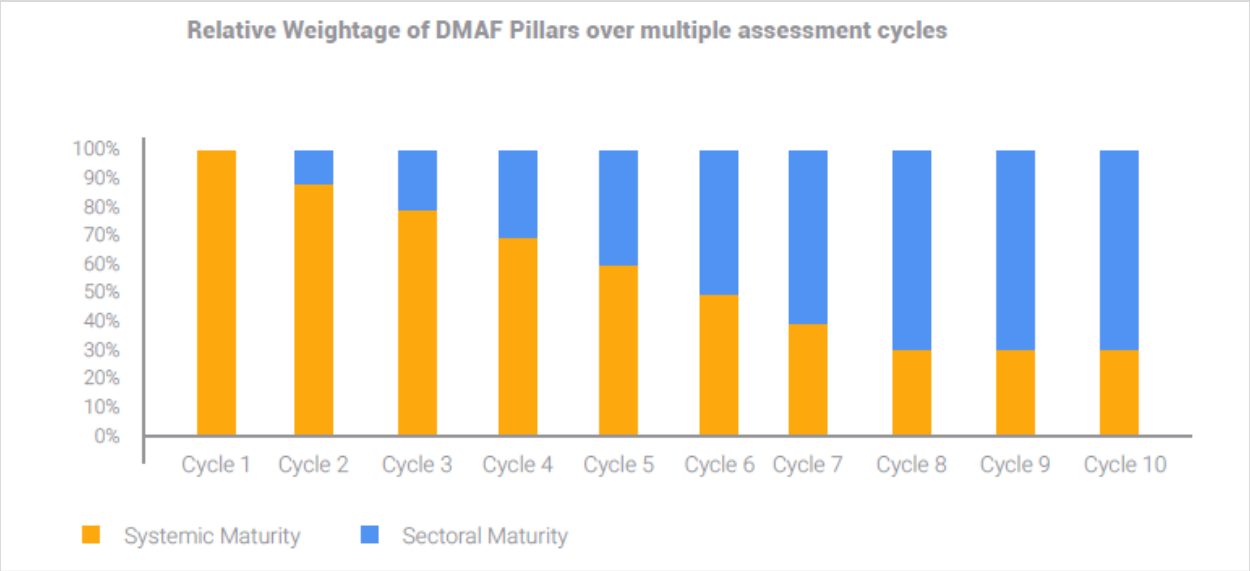
Each cycle of assessment will build on the previous cycle's framework, with a focus on shifting the goalposts for cities to evolve towards greater data maturity levels. This will enable cities and the Mission to adopt a tiered approach in implementing DMAF by collectively pursuing higher levels of achievements in each cycle.

It is important to note that while cities may be at different levels at a point in time, and may make progress at different paces, the DMAF assessment is primarily intended to be a self-evaluation tool for cities. The assessments and certifications therefore are measures to help cities calibrate their data policy, strategy and their implementation of DataSmart Cities initiatives.

Certification Mechanism

DMAF will enable city managers to assess their maturity on use of data for empowering their cities. The assessment results will be in the form of certification levels – starting from 'Initiator', 'Explorer', 'Enabled', and 'Connected' in increasing order, as described in the assessment framework for the assessment period.

Certification levels will be defined before the beginning of each assessment cycle and will be an evolution of the previous cycle's assessment framework. Thus, a city that is assessed at a connected level in the first cycle will have to continue to push forward with its efforts in subsequent cycles in order to retain its original certification level.



In the first cycle of the Challenge, focus of the framework will entirely be on the Systemic Maturity pillar, with subsequent cycles introducing the Sectoral Maturity pillar. As cities begin to grow on systemic maturity and come at par with each other, in subsequent cycles, the weightage assigned to Systemic Maturity pillar will be reduced in comparison to the preceding cycle, gradually reaching a minimum weight not below 30%.

Simultaneously, the Sectoral Maturity pillar will become more detailed and span across more sectors in each subsequent cycle, thus ensuring a holistic implementation of the DataSmart Cities strategy. This approach will ensure that cities are constantly motivated and incentivized to innovate and strengthen their adoption of the strategy to drive improved outcomes for all stakeholders.

Assessment Results

DMAF will enable city managers to assess their maturity on use of data for empowering their cities. The assessment results will be in the form of certification levels- starting from 'Initiator', 'Explorer', 'Enabled', and 'Connected' in increasing order, as described in the framework for the assessment period.

Cities will be invited by the DAM Unit to share their learnings with other cities through workshops, and video conferences and other online/offline platforms, so that other cities may benefit from their experiences.

Expectations from Cities

Through DMAF, cities will benefit by:

- Performing honest self-evaluations in the spirit of successfully adapting and implementing the DataSmart Cities strategy in their context.

- Work with other peer cities and the DAM Unit with regard to the assessment framework and help evolve it over time.

- Focusing on building preparedness and capabilities to implement the DataSmart Cities strategy in a sustainable manner.

ANNEXURE 2

India Urban Data Exchange (IUDX)

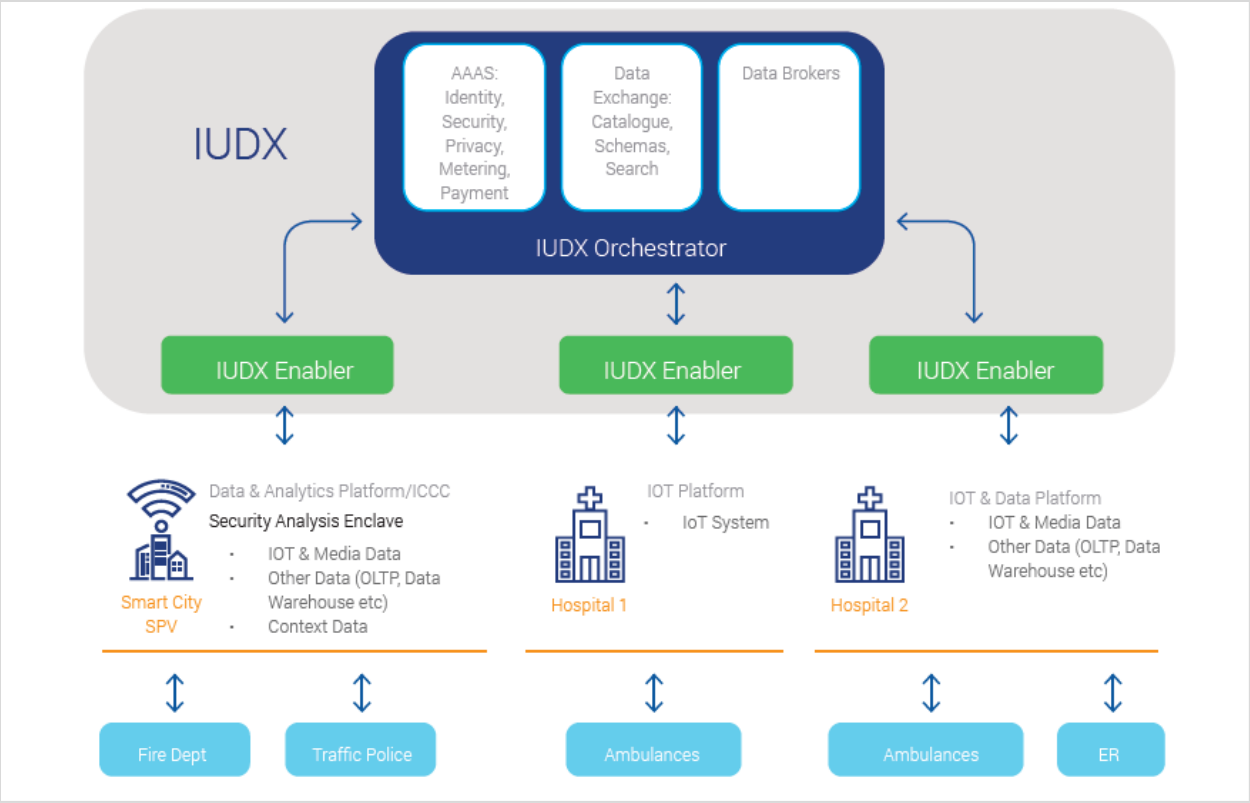
Cities generate a large amount of data daily through a variety of stakeholders. This data can be harnessed to unlock innovation, catalyze new business models and deliver data-driven governance and policy-making across urban India. At the same time, the use of data needs to be balanced in a manner that empowers citizens and protects their privacy. To ensure that various stakeholders are able to leverage data effectively for a variety of requirements across urban India, it is important to create a unified urban data exchange that enables consented and anonymized data access through virtual data rooms. Creating data exchange solutions will ensure streamlined access and drive better data-driven decisions across the country.

India Urban Data Exchange (IUDX) will be an open sources software platform that will facilitate secure, authenticated and managed exchange of data amongst various data platforms, 3rd party authenticated & authorized applications and other data sources, data producers and consumers, both within a city to begin with and scaled up across cities eventually at a national level, in a uniform & seamless way.

The platform will provide full control to the data owners as to what data to expose and to whom. Built-in accounting mechanisms will enable connect with payment gateways which will form the foundations for a data marketplace. The whole platform will be developer friendly, via definitions of open APIs (application program interfaces) and data schema templates (formats for interpreting data), so that a whole new application ecosystem gets created.

The Initial focus will be to enable data exchange between various city departments, between governments & citizens and governments & private sector within a city. Going forward, the initiative will scale up to data sharing between various cities & their stakeholders on a national level data sharing platform. It will directly address the issues that inhibit sharing & extraction of maximum value from the City's data.

IUDX WILL UNLOCK DATA DRIVEN GOVERNANCE & INNOVATION



IUDX Platform Functional Architecture

How IUDX is relevant to Smart Cities

Smart City projects have been conceptualized through a comprehensive process of citizen engagement and hence are very contextual, relevant to the city's needs and are being implemented by cities through their respective Smart City Special Purpose Vehicles (SPVs). The resulting outputs/outcomes are therefore quite different for each city, with each city investing in applications, infrastructure, and services that meet hopes and aspirations of their citizens and stakeholders. Each city chooses system integrators (SIs), hardware and software vendors (OEMs), and application developers that best meet its needs through a transparent bidding process.

While the first level implementations have just begun and are starting to show positive results on ground, they are largely independent, created in vertical silos, with no standardization of software components or their interfaces or the underlying data models. Data created by a specific application is usually available only to that application and cannot be leveraged more broadly. This limits the ability to gain broader insights from the enormous data generated within each project & department, for use between various stakeholders within the city and across different cities. The use cases in the subsequent sections demonstrate this through specific examples. This is a clear case where each city may optimize around its immediate objectives, but it may fail to achieve its goals of maximizing citizen benefits with their investments.

These challenges could be addressed through the establishment of a software platform, (hereinafter referred as **IUDX),** which will be completely open source, based on an underlying framework of open APIs, data models and security, privacy & accounting mechanisms that will facilitate, easy and efficient exchange of data among various stakeholders of Smart Cities by interconnecting disparate urban data silos, and enabling co-creation & innovation.

The cities are not expected to lose any autonomy in the prioritization of projects or in choice of vendors or System Integrators. Instead the cities will collaborate with them. The decisions on projects would be with the cities as it is today. However, they would benefit from leveraging the IUDX in their implementations or soliciting for proprietary platforms to be compliant to IUDX because this would make their projects more cost efficient and would reduce the time of deployment considerably.

This will enable efficient utilization of data & leveraging of investments, while creating more opportunities for different market players, & will spur innovation in smart cities ecosystem. Additionally, using common software components, such as IUDX, for certain key middleware functions which are similar across all projects, will also result in reduced costs & improved implementation speed.

It is expected that the entire Indian Smart City ecosystem will gain considerably from the initiative. Some of the major benefits for the different constituencies are outlined below:

Citizen and Community benefits:

1) Better and more innovative applications/solutions to help the citizens from industry, developers/entrepreneurs, start-ups and communities having variety of creative minds.

2) Rapid dissemination of learning and enabling high quality successful applications to be quickly available to all the Indian citizens.

3) Start-ups and other innovative application developers can quickly introduce new services without costly coordination with government agencies or SI.

City benefits:

1) Reduced development cost and faster development times. This is based on:

   a. The ability to reuse the code from successful smart city implementations.
   b. The ability to modularize implementation.
   c. The ability to source high quality solutions through marketplace.

2) No Vendor Lock-in. The standardized and open platform enables much greater flexibility and choice of vendors based on the evolving needs of the city.

4) New revenue opportunities from data. The data economy enabled by the IUDX platform will allow cities to monetize their data more effectively. It will also allow cities to broker third party data and benefit from this brokerage.

5) Unleashing innovation from entrepreneurs and community. This is based on the market place for plug-and-play applications that IUDX will enable.

Industry/Start-ups/Entrepreneurs benefits:

1) Improved ability to find skills and rapidly ramp up projects. This is through standardized & proven APIs and data models for external access and for internal interfaces between components.

2) Reduced cost of development. Open source code provides a solid license-free foundation to build upon, with standard interfaces simplifying implementation.

3) Ability to innovate. Developers can focus on innovation and differentiated value rather than build basic software.

4) SIs benefit from the reduction in heterogeneity. Because of IUDX standardized interfaces, they do not have to deal with variety of different platforms, each with specific capabilities and limitations.

5) Start-ups and other entrepreneurs can readily create innovative new applications through standardized APIs and Software Development Kits (SDKs) provided by IUDX. The involvement in a City's IUDX implementation can benefit the startup with business use cases and access to developer community. The start-ups will have opportunity to provide value added data services like labelling, curation, provenance, ontology creation, data model creation etc. and training and consulting services to city officials

6) The data economy enabled by the IUDX platform will allow third party data sources to monetize their data more effectively.

Academia and Research benefits:

1) Better engagement with key stakeholders through OSCI. Academia and researchers community would be able to analyze and provide insights to policy makers.

2) Improved access to data for research. It would promote multi-disciplinary research and analysis by making data available from various data sources.

We envision these services to seamlessly inter-work not only across the various city departments, but also across cities at a national level. IUDX compliant applications will be able to use consumer APIs to pull data from any of the underlying data platforms and using the publisher APIs to push data to any of the applications behind the individual platforms. Standardized APIs and data schema templates, will enable an IUDX compliant application to work in a city without needing any modification. Additionally, the standardized publisher APIs along with common data schemas, will enable vendor neutrality for IoT devices.

In summary, the IUDX platform will consist of:

a) Definitions of the APIs for the consumer, producer and enablement interfaces.

b) Definitions of the schema templates (where needed) for both IoT data as well as context data6.

c) Open source libraries that implement the APIs and other exchange related functions to help the application and device developer community to easily work with IUDX.

d) Open source implementation of the enablement and orchestration services.

In IUDX, There shall be clear definition of **data ownership and sharing mechanism**, under the control of the data owner. The platform itself is only a vehicle for the data owners to more effectively manage and share the data as per their discretion/policies.

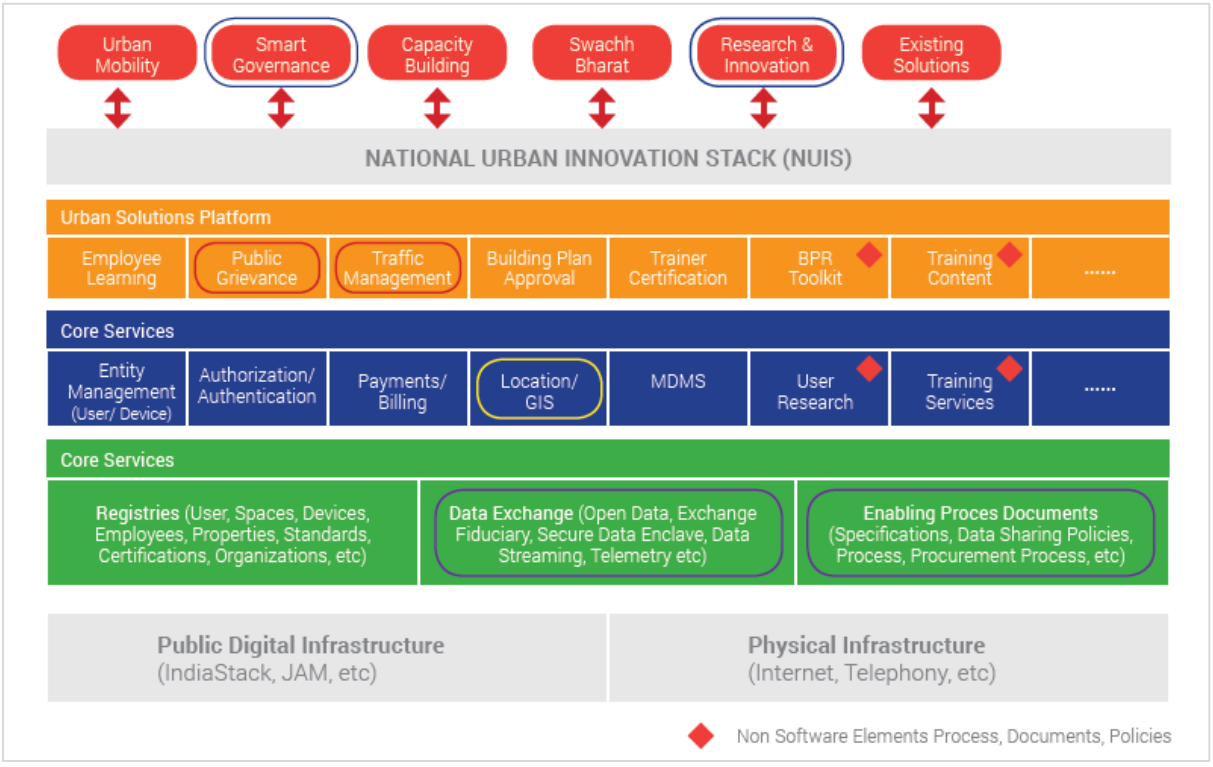
ANNEXURE 3

National Urban Innovation Stack (NUIS)

Strategy Platform can be defined as a set of Digital Infrastructure components needed for the management, analysis and use of data for a data-driven governance.

Smart Cities Mission, MoHUA and National Institute of Urban Affairs (NIUA) have jointly released a strategy paper on National Urban Innovation Stack (NUIS). This document lays down the architectural blueprint for the entire stack as shown in the figure below. NUIS stack provides a good framework of
where a DataSmart platform fits in the entire eco-system.

As shown in the figure below, the NUIS stack has three layers comprising Core Data Infrastructure, Core Service and the Urban Solution Platform. The DataSmart Cities Strategy will have an influence on all these layers. The red boxes around the examples in the NUIS stack demonstrate the overlap between the DataSmart Cities Strategy with the innovation stack. This is only representational but provides a good idea of where the intersection of data platform lies.

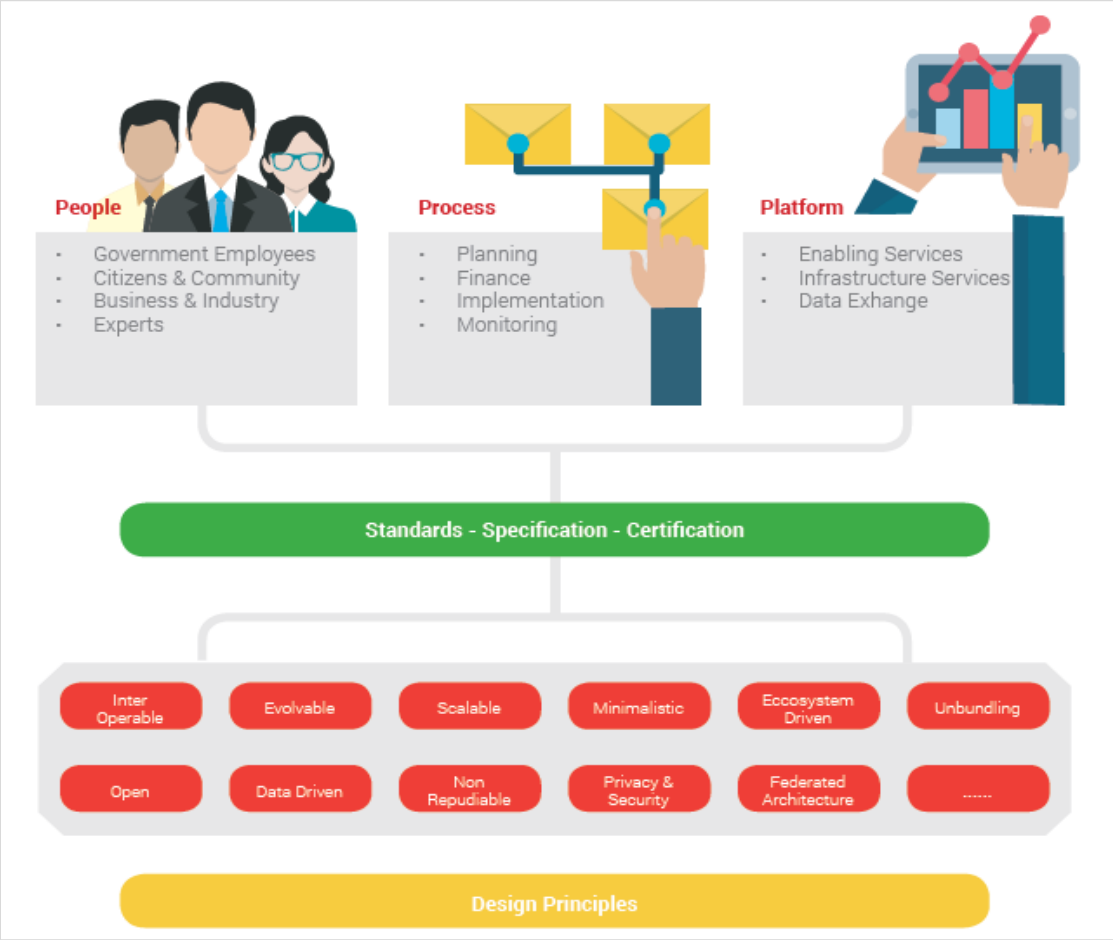


NUIS Architecture

The data platform is of prime importance to this Strategy. It defines the data access principles, the potential sourcing model and the publishing model for the city data.

Design Principles

Much of the design philosophy laid down by the NUIS stack is applicable to the DataSmart Cities Strategy Platform as well.



NUIS Stack design Principles

Just as has been indicated in the NUIS stack design principles, inter-operability, scalability, open yet secure are the key tenets of a Data Platform.

Non-reputability and quality are two implicit requirements without which the data loses much of its relevance.

ANNEXURE 4

National Data Sharing and Accessibility Policy (NDSAP) – 2012

1. Preamble

1.1 Asset and value potentials of data are widely recognized at all levels. Data collected or developed through public investments, when made publicly available and maintained over time, their potential value could be more fully realized. There has been an increasing demand by the community, that such data collected with the deployment of public funds should be made more readily available to all, for enabling rational debate, better decision making and use in meeting civil society needs. Principle 10 of the United Nations Declaration on Environment and Development (Rio de janeiro, June 1992), stated

".....each individual shall have appropriate access to information concerning the environment that is held by public authorities ..... and the opportunity to participate in the decision-making process. States shall facilitate and encourage public awareness and participation by making information widely available."
Section 4(2) of the Right to Information Act, 2005 reads

"It shall be a constant endeavor of every public authority to take steps in accordance with the requirements of clause (b) of sub-section (1) to provide as much information suomotu to the public at regular intervals through various means of communication, including internet, so that the public have minimum resort to the use of this Act to obtain information"

1.2 The principles on which data sharing and accessibility need to be based include: Openness, Flexibility, Transparency, Legal Conformity, Protection of Intellectual Property, Formal Responsibility, Professionalism, Standards, *Interoperability, Quality, Security, Efficiency, Accountability, Sustainability and Privacy*.

1.3 A large quantum of data generated using public funds by various organizations and institutions in the country remains inaccessible to civil society, although most of such data may be non-sensitive in nature and could be used by public for scientific, economic and developmental purposes. The National Data Sharing and Accessibility Policy (NDSAP) is

designed so as to apply to all sharable non-sensitive data available either in digital or analog forms but generated using public funds by various Ministries / Departments /Subordinate offices / organizations / agencies of Government of India. The NDSAP policy is designed to promote data sharing and enable access to Government of India owned data for national planning and development.

2. Definitions

2.1 Data –Data means a representation of information, numerical compilations and observations, documents, facts, maps, images, charts, tables and figures, concepts in digital and/or analog form.

2.2 Data Archive – A place where machine-readable data are acquired, manipulated, documented, and distributed to others for further analysis and consumption.

2.3 Data Generation– Initial generation/collection of data or subsequent addition of data to the same specification.

2.4 Dataset – A named collection of logically related features including processed data or information.

2.3 Geospatial Data – All data which is geographically referenced

2.4 Information – Processed data

2.5 Metadata – The information that describes the data source and the time, place, and conditions under which the data were created. Metadata informs the user of who, when, what, where, why, and how data were generated. Metadata allows the data to be traced to a known origin and know quality.

2.6 Negative List – Non-sharable data as declared by the departments/ organizations

2.5 Restricted Data –Data which are accessible only through a prescribed process of registration and authorization by respective departments / organizations.

2.6 Sensitive Data– Sensitive data as defined in various Acts and rules of the Government of India.

2.7 Sharable Data – Those data not covered under the scope of negative list and non-sensitive in nature

2.8 Standards – Any application that embeds data handling functions (e.g., data collection, management, transfer, integration, publication, etc.) and operates on data in a manner that complies with data format and data syntax specifications produced and maintained by open, standards bodies.

3. Need for the Policy

Evidence-based Planning of socio-economic development processes rely on quality data. There is a general need to facilitate sharing and utilization of the large amount of data generated and residing among the entities of the Government of India. This would call for a

policy to leverage these data assets which are disparate. The current regime of data management does not enable open sharing of Government owned data with other arms of the government nor does it expect proactive disclosure of sharable data available with data owners. Such regimes could lead to duplication of efforts and loss of efficiency of planning of activities focused on national development. Efficient sharing of data among data owners and inter and intra governmental agencies and with public calls for data standards and interoperable systems. Hence, National Data Sharing and Access Policy aims to provide an enabling provision and platform for providing proactive and open access to the data generated through public funds available with various departments / organizations of Government of India.

4. Objectives

The objective of this policy is to facilitate the access to Government of India owned shareable data and information in both human readable and machine readable forms through a network all over the country in a proactive and periodically updatable manner, within the framework of various related policies, Acts and rules of Government of India, thereby permitting a wider accessibility and use of public data and information.

5. Scope of this Policy

The National Data Sharing and Accessibility Policy will apply to all data and information created, generated, collected and archived using public funds provided by Government of India directly or through authorized agencies by various Ministries / Departments /Organizations / Agencies and Autonomous bodies.

6. Benefits of the data sharing policy

6.1 Maximizing Use – Ready access to government owned data will enable more extensive use of a valuable public resource for the benefit of the community.

6.2 Avoiding Duplication – By sharing data the need for separate bodies to collect the same data will be avoided resulting in significant cost savings in data collection.

6.3 Maximized integration – By adopting common standards for the collection and transfer of data, integration of individual datasets may be feasible.

6.4 Ownership information – The identification of owners for the principal datasets provide information to users to identify those responsible for implementation of prioritized data collection programs and development of data standards.

6.5 Better decision-making – Data and information facilitates making important decisions without incurring repetitive costs. Ready access to existing valuable data is essential for many decision making tasks such as protecting the environment, development planning, managing assets, improving living conditions, national security and controlling disasters.

6.6 Equity of access – Amore open data transfer policy ensures better access to all bonafide users.

7. Data Classification

Different types of datasets generated both in geospatial and non-spatial form by different ministries /departments are to be classified as shareable data and non-shareable data. The types of data produced by a statistical system consist of derived statistics like national accounts statistics, indicators like price index, data bases from census and surveys. The geospatial data however, consists primarily of satellite data, maps, etc. In such a system, it becomes important to maintain standards in respect of metadata, data layout and data access policy. All departments / ministries will prepare the negative list within one year of the notification of the policy, which will be periodically reviewed by the oversight committee.

8. Types of Access

8.1 Open Access – Access to data generated from public funding should be easy, timely, user-friendly and web-based without any process of registration / authorization.

8.2 Registered Access – Datasets which are accessible only through a prescribed process of registration / authorization by respective departments / organizations will be available to the recognized institutions / organizations / public users, through defined procedures.

8.3 Restricted Access – Data declared as restricted, by Government of India policies, will be accessible only through and under authorization.

9. Technology for Sharing and Access

A state-of-the-art data warehouse and data archive with online analytical processing (OLAP) capabilities, which includes providing, a multi-dimensional and subject oriented view of the database needs to be created. This integrated repository of data portals of various ministries / departments as a part of https://data.gov.in, will hold data and this repository over a period of time will also encompass data generated by various State Governments and UTs. The main features of the data warehouse need to include:

a) User friendly interface

b) Dynamic / pull down menus

c) Search based Report

d) Secured web access

e) Bulletin board

f) Complete Metadata

g) Parametric and Dynamic report in exportable format

10. Legal framework

Data will remain the property of the agency/department/ ministry/ entity which collected them and reside in their IT enabled facility for sharing and providing access. Access to data under this policy will not be in violation of any Acts and rules of the Government of India in

force. Legal framework of this policy will be aligned with various Acts and rules covering the data.

11. Pricing

Pricing of data, if any, would be decided by the data owners and as per the government policies. All Ministries / Departments will upload the pricing policy of the data under registered and restricted access within three months of the notification of the policy. A broad set of parameters would be standardized and provided as guidelines for the use of data owners.

12. Implementation

a) The Department of Science & Technology serving the nodal functions of coordination and monitoring of policy through close collaboration with all Central Ministries and the Department of Information Technology by creating https://data.gov.in through National Informatics Centre (NIC).

b) All sharable data will be made available on 'as-is where-is' basis.

c) Detailed implementation guidelines including the technology and standards for data and metadata would be brought out by Department of Information Technology, Government of India.

d) All the data users who are accessing/using the data shall acknowledge the ministry/department in all forms of publications.

e) All Ministries/Departments will upload at least 5 high value datasets on https://data.gov.in within three months of the notification of the policy.

f) Uploading of all remaining datasets should be completed within one year

e) Thereafter, all datasets are to be uploaded regularly every quarter.

f) https://data.gov.in will have the metadata and data itself and will be accessed from the portals of the departments/ministries.

g) The metadata in standardized formats is to be ported on https://data.gov.in which enables data discovery and access through departmental portals. All metadata will follow standards and will minimally contain adequate information on proper citation, access, contact information, and discovery. Complete information including methods, structure, semantics, and quality control/assurance is expected for most datasets.

h) Government will design and position a suitable budgetary incentive system for data owners for increasing open access to the sharable data.

g) An oversight committee will be constituted for facilitating the implementation of the policy and its provisions thereof

h) Department of Information Technology will constitute a coordination committee for implementation.

13. Budget Provisions

The implementation of National Data Sharing and Access Policy is expected to entail expenditures for both data owners and data managers for analog to digital conversion, data refinement, data storage, quality up-gradation etc. Budgetary provisions and appropriate support for data management for each department / organization by Government of India would be necessary.

14. Conclusion

14.1 While policies provide official mandate, facilitation of optimum accessibility and usability of data by the implementers pre-suppose a trajectory of proper organization of data, with access services and analysis tools that provide the researchers and stakeholders with added value. For data to be reused, it needs to be adequately described and linked to services that disseminate the data to other researchers and stakeholders. The current methods of storing data are as diverse as the disciplines that generate it. It is necessary to develop institutional repositories, data centers on domain and national levels that all methods of storing and sharing have to exist within the specific infrastructure to enable all users to access and use it.

14.2 National Data Sharing and Access Policy aims at the promotion of a technology-based culture of data management as well as data sharing and access. It opens up, proactively, information on available data, which could be shared with civil society for developmental purposes, their price details if any, and methods for gaining access to registered and restricted use. The policy has limited its scope to data owned by the agencies, departments/ Ministries and entities under the Government of India and forms a statement of the Government of India of its commitment to transparency and efficiency in governance. Department of Science & Technology will continue the process of evolving the policy further, keeping in tune with technological advancements and the National requirements and enrolling the State Governments.

ANNEXURE 5

Madhya Pradesh Data Sharing and Accessibility Policy (MPDSAP)

1. Preamble

Government Of India {GoI} issued the National Data Sharing and Accessibility Policy {NDSAP} vide Gazette of India Notification No. 11, Part 1, Sec 1 on 17, March 2012 {Annexure 1}. It mandates all Central government departments to make data and information created, generated, collected and archived using public funds directly or through authorized agencies of various organs of GoI, available for planning and development. Madhya Pradesh Data Sharing and Accessibility Policy is to be read in conjunction with NDSAP.

2. Objectives

The objective of this policy is to facilitate the availability and access to data and information in both human readable and machine readable forms through a electronic network in a proactive and periodically updatable manner, within the framework of various related Policies. Acts and Rules of the Government of Madhya Pradesh and Government of India.

3. Scope

The Madhya Pradesh Data Sharing and Accessibility Policy (MPDSAP) ins designed so as to apply to all sharable non-sensitive data available either in digital or analog forms and generated using public funds by various State department/Subordinate offices/ organizations/agencies, It is designed to promote sharing and enable access to Government owned data that could be used for planning and development.

4. Implementation

a) National data portal data.gov.in being operated by NIC will be used for wider dissemination of data under this policy as well as through the Government of Madhya Pradesh portal.

b) The Department of Information Technology implementation of this policy in close collaboration with the State unit of NIC.

c) Each department will make resources available under their own budget allocations for implementation of this policy.

d) The nodal department will issue the detailed implementation guidelines similar to the National implementation guidelines issued by DeitY, GOI.

e) All the Department will provide at least 3 high value datasets within six month after notification of this policy, The department will endeavor to upload all existing data in one year's time and hereafter as and when data is available.

f) The standards of NDSAP will be followed to ensure interoperability. The State nodal department will develop their state portal on the lines of the national portal developed by NIC.

g) An oversight state committee will be constituted for facilitating the implementation of the policy and its provisions thereof under the chairmanship of the chief secretary and convener of the head of the nodal department in the state.

h) Government department will encourage community participation for development of application using the date data made available under this policy with a view to empower the citizens.

ANNEXURE 6

Emerging Technologies and the impact on City Data

Emerging technology needs to be a core part of every city's strategy. While each city's strategy for how to best exploit them will vary, these technologies will have the most significant global impact across industries. The evolution of these technologies and the benefits derived from them depend, to a large extent, on the availability of relevant data and their usage.

Technology	Description	Sample Use Cases
 *Blockchain*	Distributed electronic ledger that uses software algorithms to record and confirm transactions with reliability and anonymity. The record of events is shared between many parties, and information once entered cannot be altered, as the downstream chain reinforces upstream transactions.	• Identity management • Voting • Peer to peer transactions • Supply chain management • Smart contracting • Provenance / traceability • Asset registration/ownership • Trade finance • Record management
 *Drones*	Air- or water-based devices and vehicles, for example, Unmanned Aerial Vehicles (UAV), that fly or move without an onboard human pilot. Drones can operate autonomously (via onboard computers) on a predefined flight plan or be controlled remotely.	• Insurance claim validation • Precision farming • Infrastructure inspections • Railway safety • Cargo delivery • Construction site management • Forestry management • Facility inspection (wind, turbine, oil rig, etc.)
	A network of objects – devices, vehicles, etc. – embedded with sensors, software, network connectivity and compute capability that can	• Inventory and material tracking • Real-time asset monitoring

 *Internet of Things*	collect and exchange data over the internet. IoT enables devices to be connected and remotely monitored or controlled. The term IoT has come to represent any device that is now "connected" and accessible via a network connection. The industrial IoT is a subset of IoT and refers to its use in manufacturing and industrial sectors.	• Connected operational intelligence • Customer self-service • Usage and performance benchmarking • Data integration and analytics • Connected service parts management • Remote service • Real-time market insights • Flexible billing and pricing models
 *Robots*	Electro-mechanical machines or virtual agents that automate, augment or assist human activities, autonomously or according to a set of instructions – often a computer program.	• Manufacturing • Hazardous industries • Hotels and tourism • Service industry • Automation of predictable tasks • Data management
 *Virtual and Augmented Reality (VR & AR)*	A computer-generated simulation of a three-dimensional image or a complete environment, within a defined and contained space that viewers can interact with in realistic ways. VR is intended to be an immersive experience, sometimes with the addition of information or visuals to the physical world, via graphics and audio overlay, to improve the user experience for a task or a product.	• Immersive journalism • Virtual Workplaces • Virtual Showrooms • Manufacturing/product design • Architecture & construction • Education & training • Big data management • Entertainment • Healthcare • Merchandising
 *Artificial Intelligence*	Software algorithms that can perform tasks that normally require human intelligence, such as visual perception, speech recognition, decision-making, and language translation. AI is an "umbrella" concept that is made up of numerous subfields, such as machine learning, which focuses on the development of programs that can teach themselves to learn, understand, reason, plan, and act (i.e., become more intelligent) when exposed to new data in the right quantities.	• Managing personal finances • Trading systems • Real-time fraud and risk management • Automated virtual assistants • Underwriting loans and insurance • Customer support, transactions, and helpdesks • Data analysis and advanced analytics

Data will form a basis for the development of these technologies in years to come. While some of these technologies like AI and robots will require the availability of relevant and structured data, other technologies such as blockchain and drones will create data that will loop back into the data ecosystem thereby enhancing the Data Culture of cities. This system of data creation and usage which will form an integral part of DataSmart Cities.

References

1. Ministry of Housing and Urban Affairs Publication under Smart Cities Mission – *"DataSmart Cities – Empowering Cities Through Data"*

2. Ministry of Housing and Urban Affairs Publication under Smart Cities Mission – *"Data Maturity Assessment Framework – Design and Methodology"*

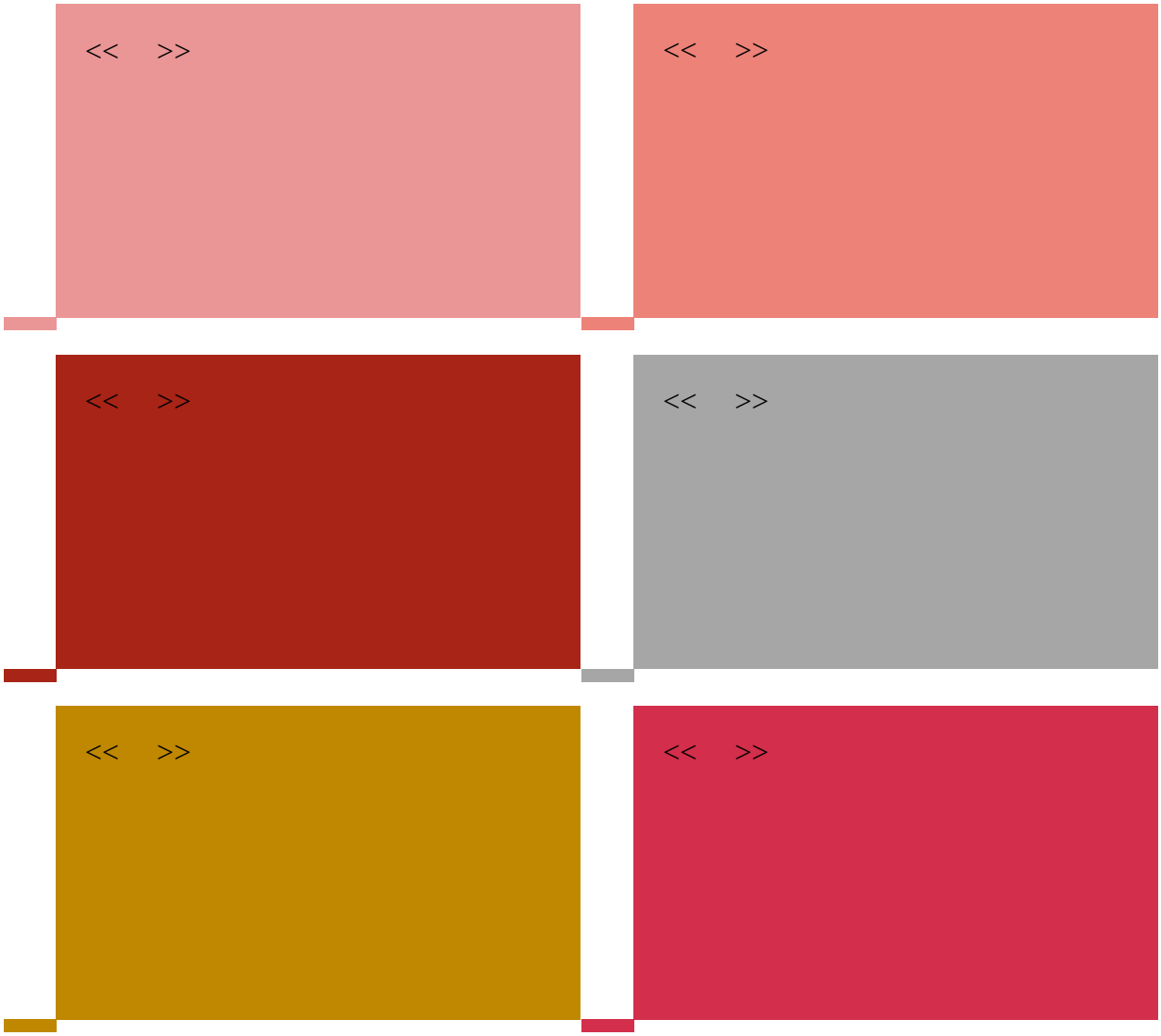
3. *National Data Sharing and Accessibility Policy (NDSAP Policy)*, Open Government Data Division, National Informatics Centre
https://data.gov.in

4. *NDSAP Implementation Guidelines*, Open Government Data Division, National Informatics Centre
https://data.gov.in

5. *Madhya Pradesh Data Sharing and Accessibility Policy (MPDSAP)*
http://www.mpsedc.com

6. Pune Municipal Corporation
https://opendata.punecorporation.org

Core Team




Thank You....