Smart City
MISSION TRANSFORM-NATION

Ministry of Housing and Urban Affairs
Government of India

स्वच्छ भारत
एक कदम स्वच्छता की ओर

DataSmart Cities: Empowering Cities through Data

CONFIDENTIAL

Smart Cities Mission

Ministry of Housing and Urban Affairs (MoHUA)

Government of India (GoI)

Table of Contents

Setting the context 3

Definitions 6

Chapter I: Contours of DataSmart Cities 7

1. Building blocks 10
2. Data Strategy 11

2.1 Open Government Initiative, Government of India 12

2.2 Open Data Platform (data.gov.in): 12

2.3 Government Open Data License-India: 13

2.4 DataSmart Cities: An evolving Data Strategy for Smart Cities 13

2.5 Reference Policies and Guidelines 14

2.6 Implementing DataSmart Cities 14

2.7 Data Governance 16

2.7.1 Mission Data Officer (MDO) 16

2.7.2 City Data Officer (CDO) 17

2.7.3 Data Champions and Coordinators 20

2.7.4 Data Champions (DC): 21

2.7.5 Data Coordinators: 21

2.7.6 City Data Alliance (CDA) 21

2.7.7 Smart Cities Data Network (SCDN) 24

2.7.8 City Data Policy 24

2.7.9 City Technology Projects and Data Platform 27

2.7.10 Data Categorization and Classification 27

Chapter II: Why Data is important for Smart Cities 31

1. Benefits and Case studies 34
2. Data Architecture @ Smart City 38
3. Life-Cycle of Data: Explained 39
4. Maturity Model for Data driven Governance 42
5. Data Sharing Models 43

Chapter III: City Data Maturity Assessment 46

How smart is City about data? 46

CONFIDENTIAL

What is DataSmart Readiness Index? 46
Annexure I: India Urban Data Exchange (IUDX) 51
Why IUDX is relevant for DataSmart Cities? 52
Annexure II: Reference City Data Policies 61

CONFIDENTIAL

Setting the context

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.

While the value of data as highlighted above is accepted and understood by all stakeholders, not many success stories of data driven governance are evidenced from our cities. The importance of using data to resolve critical problems faced by cities cannot be exaggerated. In fact, smart cities are all about getting the right data to the right people at the right time to solve relevant use case scenarios. However, certain fundamental challenges have acted as barriers in the adoption of data driven approaches for problem resolution.

1) *Lack of a "Culture of data":* One of the key challenges has been, what we could call the lack of a "culture of data". Despite the availability of a large amount of very useful data with different agencies, not much of it is used to draw insights and create actionable intelligence for city governance. Collaboration around data is missing amongst different stakeholders and the power of data as a potential economic resource is not harnessed appropriately. Data governance is a issue which has not been fully addressed by the cities. City governments tend to resolve complex problems faced by their cities by delivering solutions through a fragmented, over-simplified and piecemeal approach. This can be seen in almost every sphere of work, whether it be urban planning, mobility or water management to name a few. Un-accounted cross-play between different spheres, mechanical adherence to procedures, large number of assumptions, inability to diagnose real issues are some of the ways in which this over-simplification manifests itself, thereby leading to delivery of suboptimal solutions.

2) *Lack of a City Data Policy:* Even if the city government and its stakeholders realize the value of data and want to unlock the power of this valuable resource, there is often a lack of clarity on data policy which restricts them from doing so. A data policy is essential to understand the contours of data sharing, privacy, security and ownership in the context of the city. Certain types of data (e.g. an individual's tax payments) are clearly private and should not be shared. On the other hand, certain

CONFIDENTIAL

types of data (e.g. air quality sensor readings) are for unrestricted public consumption. Vast amounts of data are in the "grey zone" where clear policies are required that balance privacy, legal and public benefit considerations. Data policy is also needed to define the contours of collaboration between various governmental/ non-governmental entities on sharing and access of data. Data policy should also spell out the key actors within important city organizations who would act as data champions and data contributors. The data policy should lay out the roadmap of the city in terms of milestones in the adoption of open data, data exchange platforms. Data policy should answer critical questions regarding data ownership and safety. Lack of a clear data policy prevents cities from adopting data driven decision making as critical issues highlighted above remain unanswered.

3) *Lack of City Data Alliance:* Government data alone is not sufficient to solve city problems. Data is also collected by businesses, communities, institutes, universities in silos. Thus, it is critical for any city to assess the data needs of various stakeholders. All-important entities, both government and non-government, which generate and hold data crucial to better planning and functioning of the city, need to engage together to understand , create and promote data driven solutions for the issues faced by the city. Non- governmental organizations and other entities who work on issues regarding data privacy and security, also need to work in close engagement with other entities. However, such an engagement between different stakeholders in the city does not exist and therefore it acts as a barrier in the city's efforts to resolve issues through data driven approaches. The need for a data alliance in the city cannot be exaggerated.

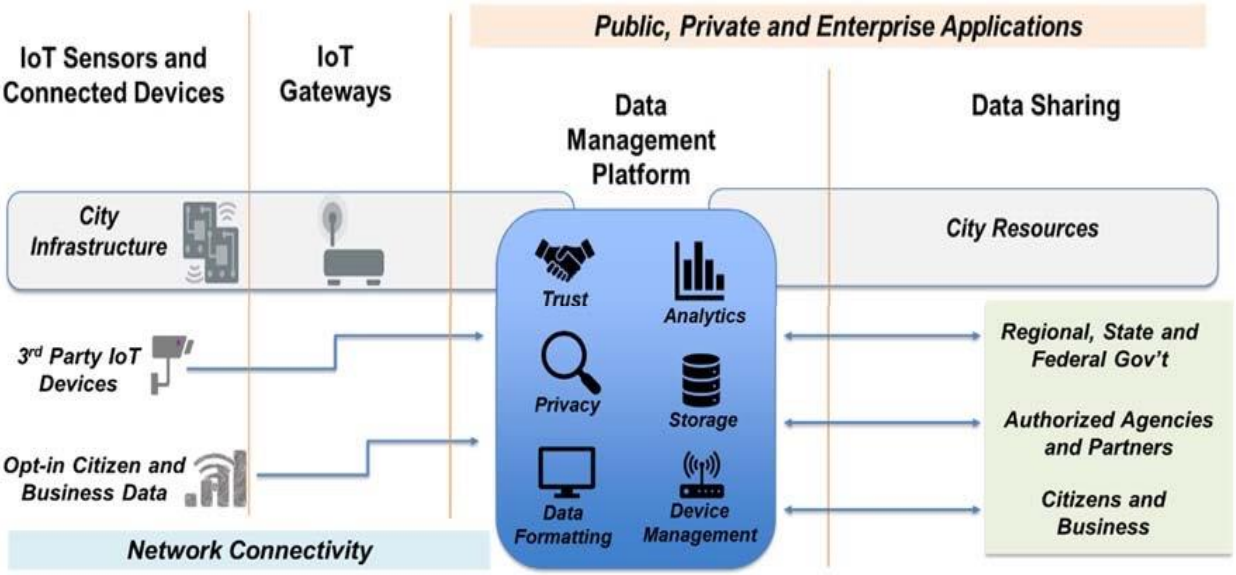
4) *Lack of appropriate Data platforms:* As mentioned earlier, a lot of data exists in silos in different departments. Even if the culture and policy considerations are overcome, there is often a technical incompatibility between these silos, which prevents the effective sharing and exploitation of data 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.

5) *Lack of awareness around use of data:* Even if cities adopt policies and create open data or data exchange platforms, lack of awareness around the use of data prohibits its widespread use. Engagement with civic organizations, communities and other relevant stakeholders is important to create information, education and awareness around the use of data. A culture of open dialogue would help understand use cases for the potential use of data to create solutions. Moreover, even amongst businesses there seems to be a lack of understanding on the economic benefits of better coordination around data. It's important to understand that "data" is not the goal, but the means to create solutions needed by the city and hence it is important for cities to 'broker' the right engagement amongst stakeholders. Lack of legal support or clarity over privacy matters impacts the open government initiatives like open data. Thus, there is evident need for addressing these barriers through dialogue on technology, processes and concerns of the community.

CONFIDENTIAL

Unless all stakeholders in the city come together and resolve to create, nurture and further the spirit of data driven empowerment, collaboration and governance, not much headway can be made in this direction. The question is how do we make this happen?

Smart Cities Mission aims to address these barriers through 'DataSmart Cities', an evolving policy framework on data for smart cities, which aims to be a catalyst for adoption of data-centric governance for the entire smart cities eco-system comprising of people, processes and technology.



Source: ATIS (https://www.atis.org/scde/)

As mentioned above, Data can be the foundation on which several use cases can be addressed by the cities.

CONFID

Definitions

Terms	Definition
Data	*Facts and statistics collected together for reference or analysis.*
Data Set	*A collection of related sets of information that is composed of separate elements but can be manipulated as a unit by a computer*
MDH	*Mission Data Hub*
Meta Data	*Data about data*
MDO	*Mission Data Office*
Open Data License	*The Open Database License (ODbL) is a copyleft ("share alike") license agreement intended to allow users to freely share, modify, and use a database while maintaining this same freedom for others.*
SCDA	*Smart City Data Alliance*
Open Standards	*An open standard is a standard that is publicly available and has various rights to use associated with it, and may also have various properties of how it was designed (e.g. open process).*
SCDN	*Smart Cities Data Network*
Open Government	*Open government is the governing doctrine which holds that citizens have the right to access the documents and proceedings of the government to allow for effective public oversight.*
OGD	*Open Government Data*
Negative List	*List of Prohibitive data sets/feeds*
NDSAP	*National Data Sharing and Access Policy*
SPV	*Special Purpose Vehicle*
Data Agencies	*Agencies which are consumers and suppliers of public data.*

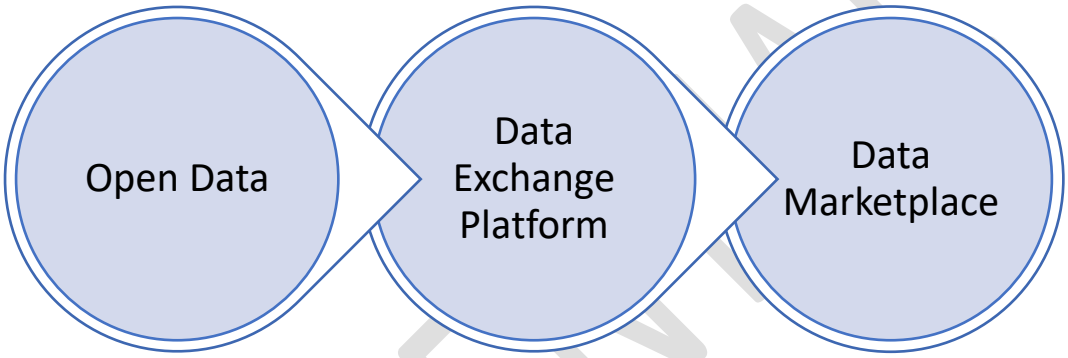
CONFIDENTIAL

CONFIDENTIAL

Chapter I: Contours of DataSmart Cities

Contours of DataSmart Cities

This section provides a roadmap on data for smart cities. The roadmap below depicts a hierarchy of data platforms as the work on data gains maturity in cities:



- DataSmart Cities begin with sharing data on open data platforms.
- Open Data platforms allow Smart Cities to publish data in consumable and machine readable format.
- Such publishing of data empowers data driven decision making.

- Data Exchange platforms allows sharing of dynamic data between multiple stakeholders.
- Such platforms create better situational awareness about cities.
- Promotes Open Innovation and co-creation at city level.

- Data marketplace allows buyers and sellers to transacts through a trusted platform.
- Can help create long term financial models for smart city solutions
- Can facilitate economic growth by helping businesses take more informed decisions.

The chart below highlights the purpose and outcomes of the platforms mentioned above:

MODEL	PURPOSE	OUTCOME
OPEN DATA PLATFORM	• To create a better understanding of cities • To empower communities through data • To provides insights for data driven decision making • To strengthen research around cities • Provides free and open static and dynamic data sets covering historical data in 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

CONFIDENTIAL

		• Better decision making thereby leading to more efficient and cost-effective solutions. • Deepen open innovation and co-creation.
DATA EXCHANGE PLATFORM	• To allow stake holders 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 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 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 secure platform • One to one relationship among publisher and consumer • Compliance of legal framework around data • Data Exchange and Payment through data portal • Central and Decentralized architecture.	• Evolution of marketplace would help strengthen the urban economic ecosystem • Enhanced G2G, G2B, G2A and B2B collaboration • Helps cities develop new business models • Empowers communities through sharing of data • Promotes development of emerging technologies like AI, ML and Blockchain

CONFIDENTIAL

1. Building blocks

	Open Data	Data Exchange	Data Marketplace
Policy	• NDSAP • City Data Strategy • Open Data Policy • Terms of Use	• NDSAP • City Data Strategy • Data Sharing and Access Policy • Terms of Use	• NDSAP • Data Procurement Policy • Data sharing and access policy • Terms of Use
Technology	• Open Data Platform	Data Exchange Platform	Data Marketplace Platform
Processes	• Data Sets Identification and Publishing	• Data Feeds Identification and Sharing	• Data Procurement Processes

Standards	• APIs
	• Standard Data Collection Formats
	• File Formats Standards
	• Publishing Standards
		• APIs
		• Data Ontology and Taxonomy
		  • Standard Data Collection Formats
		  • File Formats Standards
			Will evolve over time

Key Actors	• City Mayor • City Municipal Commissioner • City Data Officer	• City Mayor • City Municipal Commissioner • City Data Officer	• City Mayor • City Municipal Commissioner • City Data Officer

CONFIDENTIAL

	• Data Champions • Data Coordinators • City Data Alliance	• Data Champions • Data Coordinators • City Data Alliance • Data Producers and Consumers	• Data Champions • Data Coordinators • City Data Alliance • Data Buyers and Sellers Network

2. Data Strategy

Smart Cities Mission

India is urbanizing exponentially with the expected influx of over 200 million more people upto 2030. This rapid urbanization presents a great opportunity for India to leapfrog stages of development, and at the same time poses challenges in addressing chronic urban problems. The rate of development of urban infrastructure in India has lagged population growth resulting in demand and supply gaps across sectors such as water, waste management, energy, mobility, the built environment, education, healthcare and safety.

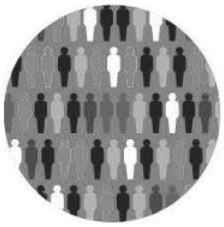
Recognizing these challenges and opportunities, MoHUA, Government of India (GoI), launched the Smart Cities Mission (SCM) in June 2015. Since the launch of the mission, 100 smart cities have been identified and cities have started executing civic infrastructure and ICT projects as per the mission guidelines. 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.

CONFIDENTIAL

Infographics:

Total Winning Proposals

100



Total Urban Population Impacted

99,630,069

Total size of Projects

INR 2,05,018 Cr

Source: http://smartcities.gov.in

2.1 Open Government Initiative, Government of India

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). For more details on NDSAP, please visit the following link: https://data.gov.in/

2.2 Open Data Platform (data.gov.in):

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.

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.

CONFIDENTIAL

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.

2.3 Government Open Data License-India:

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. For more details on Open Data License, visit the following link: https://data.gov.in/

2.4 DataSmart Cities: An evolving Data Strategy for Smart Cities

SCM intends to implement DataSmart Cities for all 100 Smart Cities through their urban local bodies (ULB) in order to leverage the potential of data for solving complex urban problems impacting their citizens.

Cities like Pune and Surat have already started publishing data sets through City Data Portals (http://opendata.pmc.gov.in) and (https://surat.data.gov.in/). Smart Cities Mission intends to unlock civic data for all 100 cities as per NDSAP. Smart Cities Mission will drive and support implementation of this strategy in consonance with NDSAP .

The implementation of this initiative in Smart Cities will lead to following benefits:

a) Data driven governance and policy formulation: Data will help City administrators in making better decisions and policies for the city. Data empowers city officials, citizens and communities and helps promote evidence based decision making.

b) Promotion of Data Sharing and Exchange: Smart Cities are about doing more with less and thus it requires better situational awareness at any given point of time for effective decision making to respond in real time. Unfortunately data is locked up in different 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.

c) Promotion of Multi-disciplinary research on Civic Issues: Research requires local data which could unlock research on civic issues like transport, traffic, solid waste etc. Multi-Disciplinary researchers may provide different perspectives or solutions on civic issues to City Administration.

CONFIDENTIAL

d) Co-Creation and Open Innovation: City governments will be able to work with entrepreneurs, Industry, and Academia to promote co-creation and open innovation to design cost-effective and contextual solutions to address civic issues by empowering them with required civic data.

e) Emergence of innovative Technologies: Data is fuel for development of solutions based on emerging technologies like AI, ML, Blockchain etc. Data platforms will help cities become data ready to kick-start innovation in emerging technologies.

f) Enhance Transparency and accountability: Smart Cities intend to deliver reliable services to their citizens through various Smart Solutions. DataSmart Cities initiative 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 city government and citizens.

g) Enhance Civic Engagement: Data sharing will lead to enhanced citizens and community engagement over various civic issues. This will enhance partnerships and collaborations between institutions, communities and stakeholders in different ways. Enhanced civic engagement enables delivery of sustainable outcomes.

2.5 Reference Policies and Guidelines

National Data Sharing and Access Policy:

The National Data Sharing and Access Policy will act as a reference point for ***DataSmart Cities***. 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.

For more details on NDSAP, please refer: https://data.gov.in/

2.6 Implementing DataSmart Cities

DataSmart Cities will be implemented in consonance with NDSAP Guidelines. Following institutional structure will support its implementation.

i. Mission Director, SCM, will nominate a **Mission Data Officer (MDO)** for SCM. The MDO will be an officer not below the rank of a Director in MoHUA. The MDO will be the officer responsible for implementation of the strategy at the national level. The MDO will engage with all the 100

CONFIDENTIAL

smart cities, NDSAP Nodal Ministry (MEITY), NIC and other relevant agencies and organisations to achieve the 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. The MDO will prepare and continuously review a negative list in line with NDSAP and other government policies in consultation with SCDOs.

ii. A **Mission Data Hub (MDH)** will be constituted within the Smart Cities Mission office. The MDH cell will function as Project Management Unit (PMU) to support the MDO in implementing DataSmart Cities strategy. The cell would be constituted of relevant experts in legal frameworks, data science, data analytics, communications and other relevant fields. The cell would act as 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, create data analytics capabilities within the mission. The MDH Cell will provide coordination, implementation, monitoring and hand holding support to Smart Cities. The MDH cell will review progress with smart cities every month based on defined KPIs, chalk out plans, share ideas, brainstorm new use cases, enable peer-to-peer learning and build and share solutions around data sharing, privacy, exchange, data driven governance amongst various stakeholders of the smart cities and release a status report on the implementation of the strategy every quarter. In short, the MDH cell would be the backbone for implementation of strategy at the national level.

iii. Municipal Commissioners of Smart cities will designate a senior official of the rank of a Chief Technology Officer/Chief Information Officer/Department Head IT or above as the **City Data officer (CDO).** All smart cities will designate their CDOs within 15 days from the announcement of this strategy. The CDO will be the officer responsible for implementation of this data strategy at the city level. The MDO at the national level, and the CDOs in the cities will continuously identify key data sets and high value data feeds to be published on the OGD Portal. High value data sets are those data sets, which create large scale empowerment of communities and enable transformative data driven governance and co-creation possible in areas which are most critical to the city's needs.

iv. **Smart Cities Data Network (SCDN)** will be a network of selected CDOs from 100 smart cities, along with representatives from other Ministries of Government (both State and Central), industry associations, research organizations, academic institutions and legal firms amongst others. The network will be termed as. MDO will be the convener of the SCDN. The SCDN will meet at least once every quarter, physically or virtually. SCDN will act as an advisory body for DataSmart Cities strategy of SCM. SCDN will help define policy contours, bring in expertise in understanding the data landscape in other parts of the world, identify best practices in Indian cities, help in interpreting the legal framework around data. SCDN, thus will help the Mission in its effort to use data as a tool for empowerment of the society.

v. **City Data Alliance (CDA)** comprising of various key stakeholders including city government, other key actors in policy making, various government department and agencies, representatives of leading academic and research institutions in the city, community organizations, entrepreneurs

CONFIDENTIAL

and advocacy groups should be created in the cities. The Municipal Commissioners of smart cities will facilitate the creation of active CDAs in their cities. The CDO will be the convener of the CDA.

Data-Smart Cities Strategy is applicable to all 100 Smart Cities under Smart Cities Mission of Ministry of Housing and Urban Affairs.

2.7 Data Governance

2.7.1 Mission Data Officer (MDO)

The Key responsibilities of MDO are as follows:

a) MDO will manage MDH cell at SCM Office under MoHUA.

b) MDO will work with Smart City Data Officers (SCDOs) of 100 Smart Cities.

c) MDO will provide coordination, implementation, monitoring and hand holding support to all SCDOs and will work as the convener for the Smart Cities Data Network (SCDN). *Refer section 2.1 for more details on SCDN.*

d) MDO will prepare the Mission Data Strategy and Plan (MDSP).

e) MDO will, along with the SCDOs, identify and suggest the list of data sets/feeds to be published by Smart Cities on Data Platform. The list of data sets/ feeds so prepared should be reviewed every month by the MDO in consultation with CDOs.

f) MDO will engage with CDOs, NDSAP Nodal agency (MEITY), NIC and other relevant agencies, to achieve the Mission goals outlined under Mission Data Strategy.

g) MDO will monitor the Key Performance Indicators as set by MDH cell and will prepare regular dashboard and report to assess progress as per Mission Data Strategy and Plan.

h) MDO will engage with the Smart Cities Stakeholders to identify the data needs and demand on regular basis.

i) MDO will guide SCDOs in ensuring security and privacy aspect as per NDSAP and other prevailing laws.

j) MDO will review the data quality of data available under SCM on Data portal.

k) MDO will publish white paper every quarter to bring forth the overall progress in implementation of the strategy.

CONFIDENTIAL

l) MDO will formulate data taxonomy for data sets/feeds. The SCDOs will refer the data taxonomy to publish data sets/feeds.

m) MDO will conduct data challenges, maturity benchmarking of cities on continuous basis with the help of MDH.

2.7.2 City Data Officer (CDO)

The Municipal Commissioners of Smart cities will designate a senior official of the rank of a Chief Technology Officer/Chief Information Officer/Department Head IT or above as the City Data officer (CDO). The CDO will act as custodian and driver of City Data plan (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.

The key responsibilities of CDO is as follows:

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 first City Data Policy shall be prepared by the Smart Cities within two months from announcement of this strategy. The policy should be created post engagement with relevant stakeholders. The CDA 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 to align with mission data strategy and priorities with respect to Open government initiatives and policies.

c) Organise regular meetings of CDA.

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: CDOs will publish data Catalogues and Data Sets/Feeds 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.

Tips: City Leadership needs to provide full administrative and logistic support to CDO. This only can make possible a city wide change in data culture. Leadership commitment and communication will be a crucial aspect in achieving goals highlighted under the strategy Thus, it is advised that City leadership active interest in appointing suitable CDO and providing full support to him/her to successfully implement the City Data Strategy. City Data officer will report directly to City

CONFIDENTIAL

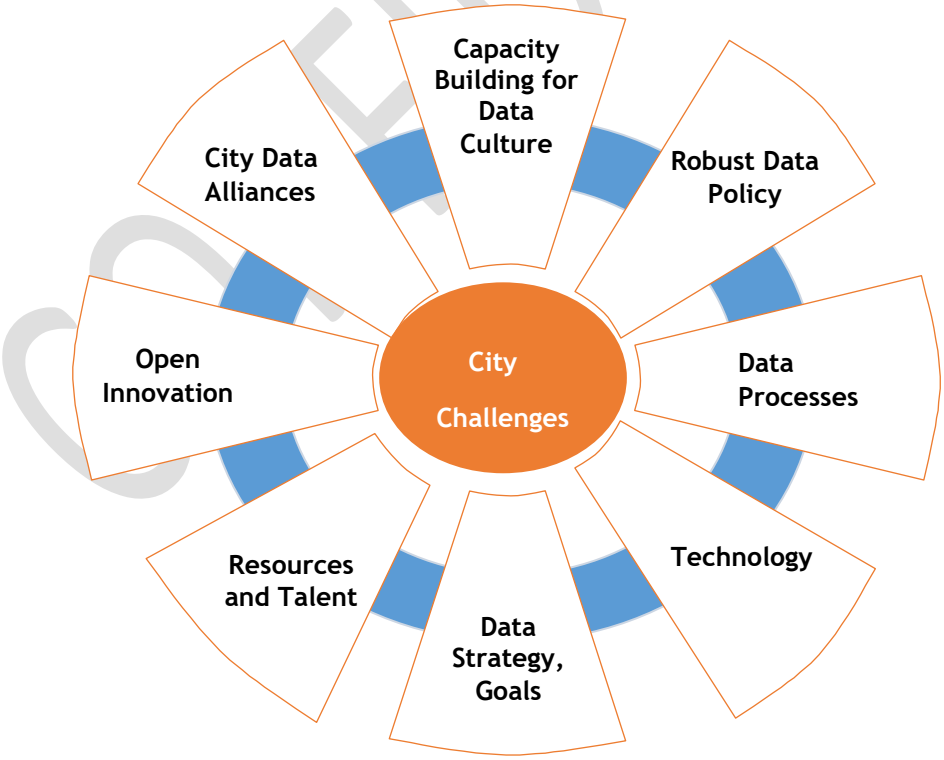
Leadership and act as single point of contact to all internal and external stakeholders in the city. Leadership need to also deploy dedicated skilled resources to drive the data initiative through City Data Officer. Integrating right talent with required business knowledge and technical skill set is key requirement for driving the city data office. City Leadership need to assign dedicated budget to on-board resources for City Data Officer. Apart from hiring resources, dedicated funding should be committed for capacity building, change management, setting up open data platform and knowledge sharing.

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.

Key Building Blocks of Data Culture

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.



CONFIDENTIAL

Key Building Blocks are explained below:

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.

a) 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.

b) 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 organisational structure chain.

c) 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 analysed form. The analysed data should be accessible on multiple devices and other desired channels.

d) 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.

CONFIDENTIAL

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 focussed data collection and may need to be trained not only on the tools but also on the importance of collecting data properly.

e) 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.

f) 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.

g) 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.

2.7.3 Data Champions and Coordinators

The CDO cannot work in isolation. Active participation from data agencies will be key to successful data collaboration within the city. Data champions will be senior functionaries, not below the rank of a Head of Department or equivalent, who would *champion* the implementation of the CDP in their respective departments/ organizations. They would be the flag bearers of the policy in their departments/ organizations and would work to align their teams to imbibe the principles of data driven decision making in their day to day functioning. They would also make their respective teams aligned to the value of collaborative work on data, as siloed approach to data ownership and use will not foster the development of integrated approaches to resolution of potential use cases within the context of each city.

CONFIDENTIAL

2.7.4 Data Champions (DC):

a) Data Champions (DCs) in respective departments/government agencies shall identify the data sets/feeds, derived information, intelligence or data challenge with respect to day to day operations of the department.

b) DCs will 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 (DC0s) within the department to streamline processes of data reporting, collection and analysis etc. Data Champions 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.

Tips: DCs needs to act as trainers and lead the team of data coordinators at the department level. It is critical for appropriate senior functionaries to be designated as DCs in each department by the respective city heads. The DC will the nodal point for implementation of the CDP within the department and will function to supervise the team of data coordinators on day to day basis. DC will be first touch point of CDO in different city organizations and must undertake continuous capacity building programs for their DCOs and other staff.

2.7.5 Data Coordinators:

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

b) Data Coordinators will also aggregate the data demand from various channels.

c) Data Coordinators will also be responsible for sensitizing the department employees over the importance of data quality etc.

2.7.6 City Data Alliance (CDA)

A CDA for a city is envisaged to be network of government departments, agencies, private sector companies, community organizations, city policy makers, domain & legal experts, research, academic institutions, incubators, entrepreneurs, etc, within the city who come together voluntarily as a collective to diagnose city problems which need resolution, act as an advocacy group for the formulation of the city data policy (CDP) which defines the collective approach of the city on issues related to data. The CDA will provide a collaborative framework to create and define use cases to solve critical city problems through the use of data, catalyse 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

CONFIDENTIAL

and evolve the CDP effectively. Thus, a CDA is not merely an advocacy organization or a think tank, but it is a network of organizations who not only understand the value of data towards improving quality of life, employability and sustainability in the city, but also be the chief driver of adoption of data driven governance and co-creation within the community. 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.

Roles of City Data Alliance (CDA):

The key objectives for setting up City Data Alliance (SCDA) are as follows:

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:

CONFIDENTIAL

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 SCDO 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.

Need for 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 for e.g. World Bank, ADB, DFID, AFB etc.

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.

CONFIDENTIAL

2.7.7 Smart Cities Data Network (SCDN)

The Smart Cities Data Network (SCDN) is conceived as a national forum that will provide guidance related to the implementation of this strategy to the office of the MDO and CDOs of all smart cities. The SCDN will be a national forum which will have as its members representatives of CDAs from few Smart Cities chosen on rotation basis from year to year, some CDOs from the cities, Municipal Commissioners from a few cities, representatives from various Government departments/ agencies from both the Central/ State levels, academia and research organizations, industry associations, experts and other key members from the Government/ private sector actively working in fields related to data governance/privacy/management etc.

SCDN will act as an advisory group to the office of the MDO and will contribute actively to evolve the Mission Data Strategy. The recommendations of SCDN would become inputs for policy formulation around data at the National level. The CDA in each city will the alliance at the city level translating national policies and guidelines to a locally meaningful manner, and providing city level feedback and requirements to the SCDN through CDO.

SCDN Activities and Responsibilities

a) Advise the MDO on Data Policy guidelines, and recommendations in accordance with NDSAP and other relevant laws in this regard.

b) Advice on capacity building around data in Smart Cities.

c) Extend advisory support in evaluating the functioning of City Data Alliances (CDAs) and recommend suitable course of action as per observations, so as to strengthen the working of CDAs in each of the Smart Cities.

d) Provide advisory support to the MDO over various relevant policies/ laws/ rules and regulations about data sharing/ privacy/ management etc.

e) Provide advice on organization of key events like hackathons, conferences, etc.

f) Advice on global best practices on data for cities.

Tips: *Set up desk or online mechanism to understand the data needs of various Smart Cities Stakeholders. Engage the stakeholders and address the concern related to data accessibility, data availability, data quality etc. Also set up discussion forums etc. to gauge the community sentiment or feedback on data initiatives.*

2.7.8 City Data Policy

CDO will formulate the CDP with milestones in consultation with CDA and MDO. The policy so created will be taken up for approval by the Municipal Commissioner. The policy should among other things, include the following:

CONFIDENTIAL

a) Stakeholders: Identify relevant stakeholders and respective data needs. For Ex: Employees will require real time data which could enhance their situational awareness on the ground.

b) Identify and create plan to provision Data Sets and Feed for use cases as identified to be of benefit to the city.

c) Current IT Systems: Assess City IT Infrastructure and identify key IT systems (Project management System, Water and Sewerage Management System) which could provide public data sets/feeds.

d) Proposed Smart Solutions/Projects: Assess smart solutions or projects under implementation as per Smart City Proposal and identify data sets/feeds which could be published through Data Platform. Work with SI or implementation partner to provision to integrate the solutions with Data Platform.

e) Identify Data Sets from reports or plan published by City government and agencies through public funds. Also identify different data sets/feeds which would be published by the city from time to time at regular intervals.

f) Road Map with Milestones: Identify short-term, medium-term and long-term goals along with timeline for publishing datasets and frequency to update it.

Tips: *City Data Policy is the first significant step in the direction to provide conceptual clarity over accessing and sharing protocols over city data. Data Policy must address concerns like data classification, categorization, archival, security, privacy, ownership etc. Absence of City Data Policy also act as barrier towards setting up data economy at the city level. City Data Policy also provides clarity around ownership of data, legal framework, terms of use etc.*

City Administration need to set up enterprise processes to leverage the existing available data with City administration. City Leaders and officials must establish processes to allow data to flow between departments and users seamlessly internally without any user interventions. Idea is to unlock the data available in-siloes with different stakeholders and make it available seamlessly.

Institutionalizing Data Culture in Cities: Key Considerations for City Data Policy

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

a) 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.

CONFIDENTIAL

b) 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.

c) 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.

d) 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.

Solving Urban Challenges

Hackathon for Urban Challenges: 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 program to solve its problems through structured challenge process. First step is to identify the correct problem statement and provide support data to the participants. City Data Officer also need to set up multi -disciplinary evaluation committee and mentors to assist participants during solution design stage. Shortlisted solutions could be extended support in next stage to test the solution on ground.

CONFIDENTIAL

2.7.9 City Technology Projects and Data Platform

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. For ex: Ease of Living or Swach Sarvekshan etc. SCDO 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.

2.7.10 Data Categorization and Classification

Data will be categorized into two broad categories:

a) Personal Data: Personal data means data consisting of information which is related to a living 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 City on Data Platform under any data sets. Data sets must be anonymized before publishing.

b) 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

CONFIDENTIAL

rendered anonymous in such a manner that the data subject is not or no longer identifiable. In other word, anonymization means excluding any personal identifiers from data sets.

Personal and Non Personal Data will be classified into following category:

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

2.7.11 Data Sets and Feeds Guidelines:

Data sets/feeds must be prioritized over following criteria:

a) Data sets which promotes co-creation and open innovation.

b) Data/Information which are frequently requested by stakeholders.

c) Data/Information which leads to academic research on civic issues.

d) Data/Information which helps in improving operational efficiency of City administration.

e) Data/Information which leads to enhanced transparency and accountability.

f) Data/Information which leads to citizen convenience.

g) Data/information which promotes multidisciplinary research using insights or analysis.

h) Data/Information which leads to institutionalization of a culture of data driven governance.

i) Data/Information which leads to better situational awareness.

j) Data/Information which leads to mandatory disclosures.

CONFIDENTIAL

k) Data/Information which forms the basis of reports/research funded by public funds by city administration.

l) Data/Information which improves inter-department coordination.

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:

- 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.
- Description (Required): Provide a detailed description of the catalogue e.g., an abstract determining the nature and purpose of the catalogue.
- 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.
- 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.
- Sector & Sub-Sector (Required): Choose the sectors(s)/sub-sector(s) those most closely apply(ies) to your catalogue.
- 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.

d) 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.

e) Machine Readable: Data Sets and Feeds should be machine readable.

f) Formats: As per NDSAP following data formats should be published:

a. CSV (Comma separated values)

b. XLS (Spread sheet - Excel)

CONFIDENTIAL

c. ODS (Open Document Formats for Spreadsheets)

d. XML (Extensive Markup Language)

e. RDF (Resources Description Framework)

f. KML (Keyhole Markup Language used for Maps)

g. GML (Geography Markup Language)

h. RSS/ATOM (Fast changing data e.g. hourly/daily)

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

h) **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.

i) **Archiving:** CDO will define and set up process for archiving process. 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.

j) **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. For more details on Open Data License, please refer: https://data.gov.in/

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

Terms of Use: Smart City will publish Terms of Use to restrict the misuse of data and indemnify the city administration in case of any misuse by end user. Refer Annexure III for standard terms of use for data platform.

CONFIDENTIAL

Chapter II: Why Data is important for Smart Cities

Introduction: Why is Data Important for Smart Cities?

Cities are engines of growth for the economy of every nation, including India. Nearly 31% of India's current population lives in urban areas and contributes 63% of India's GDP (Census 2011). With increasing urbanization, urban areas are expected to house 40% of India's population and contribute 75% of India's GDP by 2030. This requires comprehensive development of physical, institutional, social and economic infrastructure.

The magnitude of challenges facing Indian cities, coupled with their complexity and diversity make it immensely important to foster innovation in products, services, infrastructure and processes to create community-centric, cost-effective and sustainable solutions. To meet these objectives, it is important to build the right partnerships and networks, create enabling environments for stakeholder engagement, and put in place an ecosystem which breeds innovation. Government as producer of services and communities as consumers is the business-as-usual way of thinking about governance. What is required today is governance in the quadruple-helix framework wherein communities and Governments co-create solutions with partnership of Industry and academia. Together they comprise the innovation eco-system. Harnessing of strengths of Indian innovation ecosystem would lead to promoting and nurturing such innovations in Indian cities.

Data is the life-blood of the modern city. Evidence based approaches to city planning, identification of real needs for deployment of infrastructure, coordination between multiple stakeholders depend on the availability of good quality data. Data led approaches are central to create mitigation strategies for climate change, for monitoring energy use in cities and taking stock of crime and promoting well-being of citizens. Without access to credible data, no KPIs can be monitored and evaluation of decisions made becomes extremely difficult. With the explosion of personal computing devices, availability of faster communication networks, reduced cost of storage, the amount of data being produced from city operations and various forms of communication has grown multifold. With deployment of IoT devices, sensors and other methods to 'sense' the city, the sources of real-time data are increasing every day. Also, with ever evolving technologies, data collection, aggregation and processing is becoming easier and affordable in real time with each passing day.

CONFIDENTIAL

If leveraged effectively, data can make considerable difference to quality of life of its citizens and enable city administration to deliver more with less resources. Thus, it is extremely critical for Smart Cities to get its data strategy right as high-quality data will lead to high quality of information and intelligence which is a must for effective decision making on service delivery and policy formulation.

Illustrative scenarios are as follows:

Area	Scenario	Data Centric Decision Making
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?

CONFIDENTIAL

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

CONFIDENTIAL

Impact of data centric decision making in various stages are as follows:

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 Design	• Deliver Project Objectives as desired • Proper utilization of existing resources • Leads to convergence • High Project Success Rate
Project Implementation	• No Scope, Time and Cost Creep • No additional burden on city finance • Low Project Failure rate
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

1. Benefits and Case studies

Government cannot solve all its problems alone and neither can businesses. Thus, Open Government is the way forward. It leads to following benefits:

a) Increased Transparency and Accountability

What	The "Our City" portal makes it easy for citizens in Moscow, Russia, to submit requests or complaints about city services – for example, about street cleaning, garbage collection, or potholes – and receive updates when the issue is addressed. The portal also shows a map of where citizens have submitted requests and complaints.
Why	Citizens often do not know what services the city provides, or how to request them. Administrators may not be aware of problems unless citizens report them, and may not receive due credit for addressing them.
How	Citizens can submit complaints via the online portal or a mobile app. In each case, the citizen briefly explains the problem and, optionally, attaches photos. Then, the case is automatically assigned to a subcontractor or city authority that has strict rules regarding time frames for resolution. Finally, citizens receive a personalized response when the problem is fixed, typically with photos of the result.
Potential	As well as improving the efficiency of infrastructure and services, the data generated by such platforms could help to prevent problems by informing better on-time maintenance or management decisions

b) Responsive Administration

CONFIDENTIAL

What	CityScore is an online dashboard showing how Boston's city government is performing against its targets in 24 areas – from sign placement to emergency call response, from garbage collection to pothole repairs. Daily activity updates make performance and progress transparent to the public and city administrators. A single, combined number summarizes how the administration is performing overall.
Why	The metrics measured by CityScore are a gauge of how well the city government is serving its citizens. Tracking performance against targets enables problem areas to be quickly identified and remedied, and offers citizens the opportunity to hold administrators to account. It will historically have been much more challenging to collect and review the necessary data manually.
How	Some data is automatically collected by sensors, such as the percentage of buses arriving on time. In other cases, city workers use their mobile devices to record when they have completed a task. Then, this data is automatically aggregated and divided by the target figure to generate a daily, weekly or quarterly score: above 1 means the city is exceeding its targets, below 1 means it is falling short.
Potential	An open-source version of the CityScore platform is available on GitHub for other cities to adopt and adapt, offering the potential to engage citizens and improve services around the world.

c) Leads to Open Innovation and Co-creation

What	Dublin and Melbourne are among cities leading the way in making their data publicly available for businesses and citizens to use, as well as providing tools for visualization and analysis. Melbourne decided to make available data it has collected for many years on areas such as land use and pedestrian movements; Dublin's website Dublinked is a partnership of local authorities and the National University of Ireland.
Why	City governments use data to inform the making of policies and decisions. Opening up that data to the general public is intended to drive innovation and collaboration in the development of new urban solutions. Dublinked, for example, has led to the formation of startups and launching of new software applications.

CONFIDENTIAL

How	The data comes from various sources: in Melbourne, examples include revaluations of property for taxation purposes and remote digital counting of pedestrian movements. Dublinked has over 250 data sets covering areas from planning and land use to recreation and amenities; public and private companies can also share their data on the website, alongside the data provided by local authorities.
Potential	As more cities emulate Melbourne and Dublin by creating an open data culture, more citizens and startups are expected to find ways to use the data to improve city services and create economic and social value.

d) Smarter and Efficient Operations

What	Several cities are pioneering methods to engage citizens in planning and budgeting. Melbourne, for example, has twice used "people's panels" to guide decisions where recommendations included increasing city taxes to fund climate change programmes. Quito's platform Mi Ciudad, which shares data about the city – including services, infrastructure and budget allocation – is being developed to enable more citizen participation; currently, a percentage of the municipal budget is allocated based on decisions by citizens at neighbourhood assemblies. "Youth Lead the Change" - a participatory budgeting programme in Boston founded in 2013, was the first of its kind in the United States; over 2,000 youth aged 12 to 25 decide together how to spend one million dollars of the city's capital budget, developing and voting on project ideas and playing a role in their implementation.
Why	When city administrations are making the inevitable trade-offs involved in spending limited budgets, it is challenging for them to truly represent the views of all citizens. Melbourne's "people's panels", in particular, exemplify the use of new technology to engage citizens in prioritizing the relative importance of issues, rather than merely listing issues they think are important.
How	In Melbourne, expressions of interest were sought from members of the community and representative groups of around 24 people were selected and provided with extensive data. Quito's platform enables citizens to download data and an app is being developed to enable citizens to participate in the budget planning process. Boston's programme invites public submission of ideas, Which youth develop into concrete proposals.
Potential	Participatory budgeting can engage communities to take more interest in the running of their city – in the case of Boston, specifically engaging youth. When city governments are confident that their decisions reflect what citizens actually want, they are more likely to take decisions that are courageous and meaningful.

CONFIDENTIAL

e) Engaged Communities and Citizens

What	Treepedia will allow city dwellers to view the location and size of trees within their communities and to submit input to tag, track and advocate for more trees in their cities. It is designed to teach urbanites a greater appreciation of the green spaces around them and reinforce the role of the green canopy in responding to climate change.
Why	The green canopy is an important part of urban life. Trees help mitigate extreme temperatures, provide a natural respite from traffic, noise and congestion, and improve the quality of life for those living in urban environments. However, the average citizen is often removed from understanding the individual features of their unique environmental habitats. If citizens are better engaged, they can play a more active role in campaigning on behalf of the landscapes they dwell in.
How	Treepedia utilizes Google Street View panoramas to create a Green View Index, which enables scientists and city dwellers to monitor urban tree coverage and take action to improve it. The addition of a crowdsourcing and citizen-interactive platform will enable citizens to add information on the location of trees. MIT Senseable City Lab developed Treepedia in partnership with the World Economic Forum's Global Future Council on Cities and Urbanization and its Global Shapers Community.
Potential	The project has the potential to be scaled to additional cities where Google Imagery is available, and utilize satellite data where it is not. Ultimately, by inspiring citizens to seek the planting of more trees in their cities, it can help address liveability and mitigate climate change.

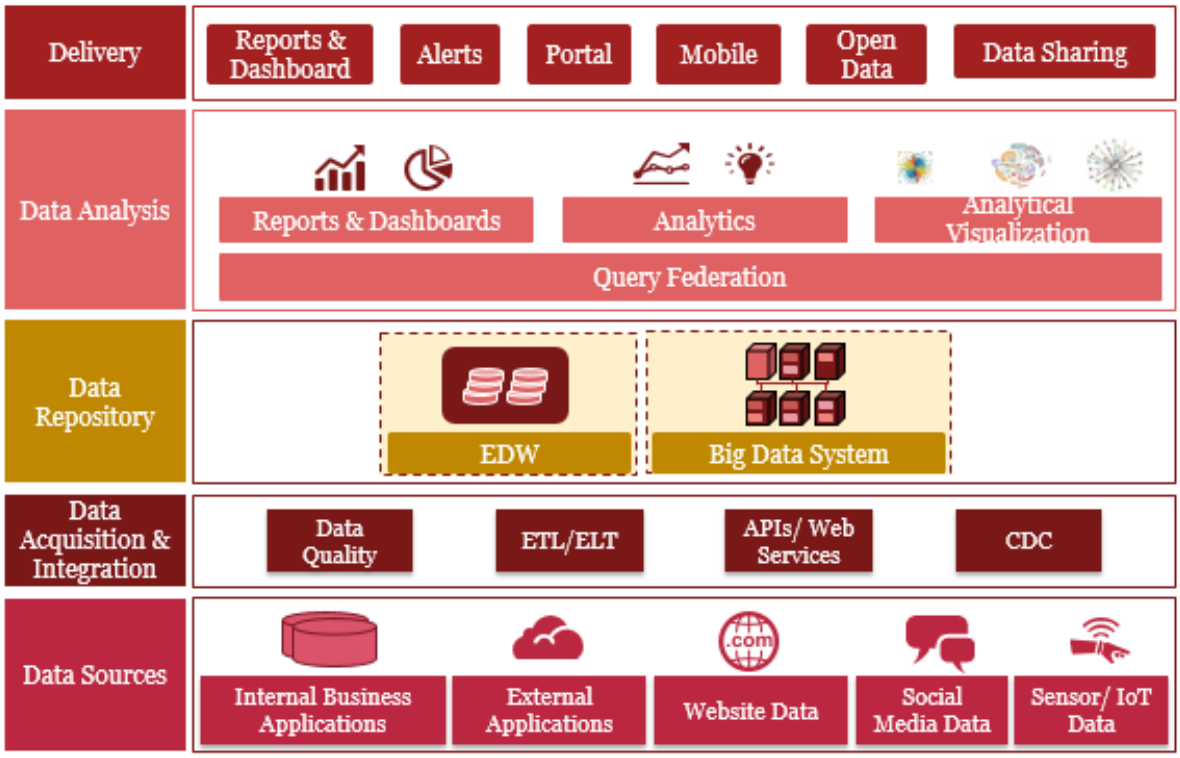
f) Informed Decision making

What	Cities are developing online platforms to make it easier for citizens to find out about planning issues and land use: zoning regulations, local development plans, heritage sites, flood risks, building permits, planning applications and so on. Examples include Myplan.ie, which covers all land in Ireland, and Buildingeye.com, which has been launched in San Francisco, Palo Alto and Pittsburgh.
Why	Ireland's 31 separate planning authorities – including city and county councils, town councils and borough councils – have together adopted over 400 statutory plans, in a variety of formats that make it difficult for citizens and local authorities alike to get an effective overview of factors that could affect their decisions. In the US, Buildingeye.com makes it easier for neighbours and community groups to find out what is happening in their area, reducing the hundreds of phone call inquiries that city workers process every day.
How	The platforms combine data such as street maps, aerial photography, historical maps, architectural heritage, areas of special protection, nature reserves, floodplains, population census and education services. Myplan.ie took five years to develop, in a collaborative approach involving local authorities, government departments and

CONFIDENTIAL

	agencies, and independent experts, which incorporated information into Ordnance Survey Ireland's web mapping service. Users of Buildingeye.com can browse a map and click icons representing proposed projects, colour-coded to reflect the planning status and see associated drawings and documents.
Potential	Such platforms can radically improve the efficiency and effectiveness of local authority plan making, enabling the planning service to perform better with fewer resources. For individuals and organizations – such as property developers, architects, surveyors, lawyers and planning consultants – they can lead to significant efficiency gains from no longer having to piece together information from multiple sources.

2. Data Architecture @ Smart City



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.

  Refer Data Collection in next section i.e. Lifecycle of Data.

- 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.

  Refer Data Integration in next section i.e. Lifecycle of Data.

- 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.

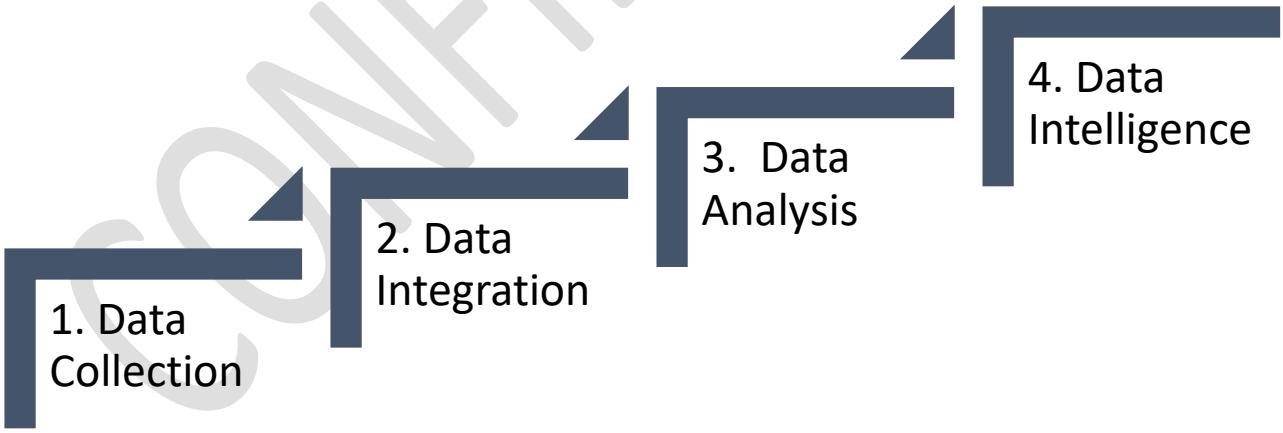
- Data Analysis Layer: Comprises of engines to process the structured and unstructured data on various dimensions for various purpose. Data Analysis Layer engine helps user to derive information, intelligence and knowledge out of processed data stored in Enterprise Data Repository for analysis or decision-making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

  Refer Data Analysis in next section i.e. Lifecycle of Data.

- 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.

CONFIDENTIAL

3. Life-Cycle of Data: Explained



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."[1]
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.

a) 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.

b) 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

[1] Source : http://www.atis.org/smart-cities-data-sharing/smart-cities-data-sharing.pdf

CONFIDENTIAL

- 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.[2]

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?"

[2] https://www.sas.com/en_us/insights/data-management/what-is-etl.html

CONFIDENTIAL

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.[3]

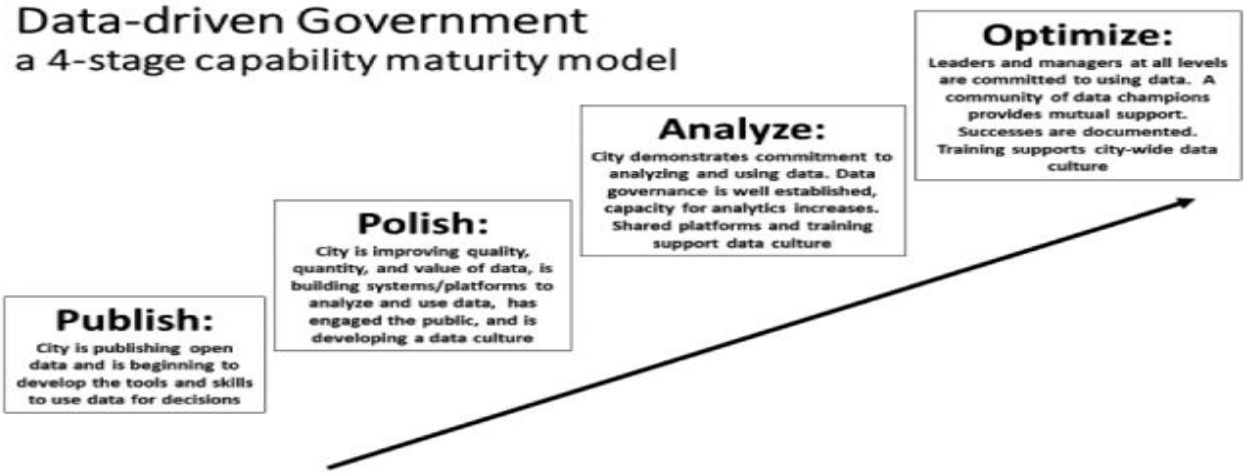
4. Maturity Model for Data driven Governance

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:

CONFIDENTIAL

[3] http://www.ingrammicroadvisor.com/data-center/four-types-of-big-data-analytics-and-examples-of-their-use



Source: https://datasmart.ash.harvard.edu/news/article/analytics-excellence-roadmap-866

The City of Chicago in 2015 developed a mathematical model to predict which of its 15,000+ restaurants and food establishments were most likely to cause foodborne illness. A predictive analytics model enabled restaurant inspectors to prioritize their visits according to risk, visiting first the places most likely to make customers sick. This project used data from across city government including from the Police Department, Department of Public Health, Department of Business Affairs and Consumer Protection, and 311, as well as external data such as weather and social media. The results were impressive – a 25% improvement in operational efficiency and the ability to find critical violations of the health code seven days faster.

https://datasmart.ash.harvard.edu/news/article/analytics-excellence-roadmap-866

5. Data Sharing Models

City Government cannot solve all problems facing the city in silos, neither can businesses and citizens can do it independently. Data sharing between all stakeholders is the only way which can catalyze collaborative problem solving. In the long run this is the way forward towards building foundation of data economy. Various ways of collaboration around data sharing to drive open innovation are as follows:

1. API: API stands for Application Programming Interface. An API is a software intermediary that allows two applications to talk to each other. In other words, an API is the messenger that delivers your request to the provider that you're requesting it from and then delivers the response back to you.[4] **Example:** The Monetary Authority of Singapore's APIs provide financial institutions and application service providers more opportunities to serve their customers better. Users can also automate the extraction of data and illustrate trends easily and quickly with these APIs. - See more at: https://www.smartnation.sg/resources/open-data#sthash.zh4h6zdc.dpuf

2. Data Pooling: Data is provided by various data producers from across the globe. The concerned takes this data and organizes it on its website by category. Site visitors can select a category to see the history of data available, compare two set of data side by side, or view data by source.

   Example: Safaricom, one of Kenya's leading mobile companies, shared a year of anonymized phone data with Harvard researchers to map how migration patterns contributed to the spread of malaria in Kenya. By combining Safaricom's data on call locations with national infectious disease data, researchers were able to estimate and map routes that contributed to the spread of the disease. - See more at: https://medium.com/internet-monitor-2014-data-and-privacy/mapping-the-next-frontier-of-open-data-corporate-data-sharing-73b2143878d2

   Intelligence Product: Any team of researchers can use existing data to create predictive models and machine learning techniques to guide data-driven services. Examples: The municipality of Fredensborg, in Denmark, has produced a GHG inventory since 2008, using data delivered directly by the national government in the Danish Carbon Emission Inventory Portal. Johan Vedel, Head of Climate & Energy Planning in Fredensborg, points out that having an inventory has enabled the municipal government to set realistic and achievable goals, take action wherever it makes the most sense within the municipality, and measure the effect of those actions. Since 2008, Fredensborg has registered a 38 percent reduction in CO2 emissions. See more at: https://www.opengovpartnership.org/stories/open-data-key-unlocking-cities-potential-on-climate

3. Prizes and Challenges: Public data will be used for the challenge and available through some data platform, Participants will also be encouraged to include data from their own organizations and add this data to the Data Registry.

   Example: City Hall will soon announce a series of City Data Challenges. These will be vitally important in helping us deliver on the promise of open data for the capital. Their aim will be to use data analytics to uncover new insight and ultimately new approaches to service delivery or meeting city challenges, or to use data in the creation of web-based apps which genuinely improve the lives of

[4] **Source:** https://www.google.co.in/search

CONFIDENTIAL

Londoners. How, for example, could aggregated crowd data be used to help London Government make better decisions about the city. - See more at: https://data.london.gov.uk/data-challenges/

4. Research Partnerships: Data shared by data holders with researchers through a secured protocol to avoid any data leakage or misuse..

   Example: The University of Chicago and Chapin Hall worked with the City of Chicago to build on analytics work. They analyzed 311 data, predicting when and where graffiti, potholes, and other problems are likely to occur. Working with Chapin Hall at the University of Chicago, fellows built analytics models that can be used to prevent problems and deliver city services more proactively. - See more at: https://dssg.uchicago.edu/project/predictive-analytics-for-smarter-city-services/

5. Trusted Intermediary: It is type of data which has been taken from trusted sources. For **Example:** Google Waze shares crowdsourced user data with local governments around the world in order to improve traffic management[5]. See more at: http://datacollaboratives.org/explorer.html

CONFIDENTIAL

[5] **Source: http://datacollaboratives.org/explorer.html**

Chapter III: City Data Maturity Assessment

How smart is City about data?

In order to build a sustainable 'Data Culture' through the DataSmart Cities initiative, it is important to understand the DataSmart readiness for each city.

DataSmart Readiness Index (DRI) has been designed by the Mission to help smart cities assess their readiness regarding the use of data for evidence based decision making

What is DataSmart Readiness Index?

Monitoring Parameter	Key Performance Indicator	Response
Data Strategy and Data Culture	City Data Strategy formulated and published?	Y/N
	Goals and milestones outlined under City Data Strategy	Y/N
	Data Sets/Feeds prioritization approach formulated for G2G, G2B data sets/Feeds?	Y/N
	Budget allocated for Software and Infrastructure to support Open Data Initiatives?	Y/N
	Budget allocated for Change Management and Capacity Building?	Y/N
	Budget allocated for Information, Education, and Communication (IEC)? (Y/N)	Y/N
Data Assessment (*Departments:* Water, Waste water,	Current assessment done for data available with Municipal Corporation covering following dimensions: a) Data Purpose a. Reporting b. Administrative c. Planning	Y/N

CONFIDENTIAL

Solid waste management, Property tax, Building permission , Market and License fees, Education, Health Roads, estates departments)	a) Data Collection a. Type of Collection i. Sample ii. 100% iii. Random b. Mode of Collection i. Manual ii. Electronic iii. Field Collection	Y/N
	b) Frequency of collection a. Real Time b. Hourly c. Daily d. Weekly e. Monthly	Y/N
	c) Level of Details /parameters: a. Macro b. Micro	Y/N
	d) Is it open/available on paper/electronically	Y/N
	e) Budget Estimates are prepared using data set	Y/N
	f) Is there any penalty applicable if data is collected and furnished by city administration	Y/N
	g) Is it mandatory to publish the data if asked under Right to Information Act (RTI)	Y/N
Data Policy	Has the City Data Policy been drafted and approved? (Y/N)	Y/N
	Does the City Data Policy have following components?	
	Data Classification	Y/N
	Data Categorization	Y/N
	Data Archival and Retention	Y/N
	Data Security	Y/N
	Data Privacy	Y/N
	Terms of Use	Y/N
Processes	Is there any formal enterprise process/SoP for Data Collection?	Y/N

CONFIDENTIAL

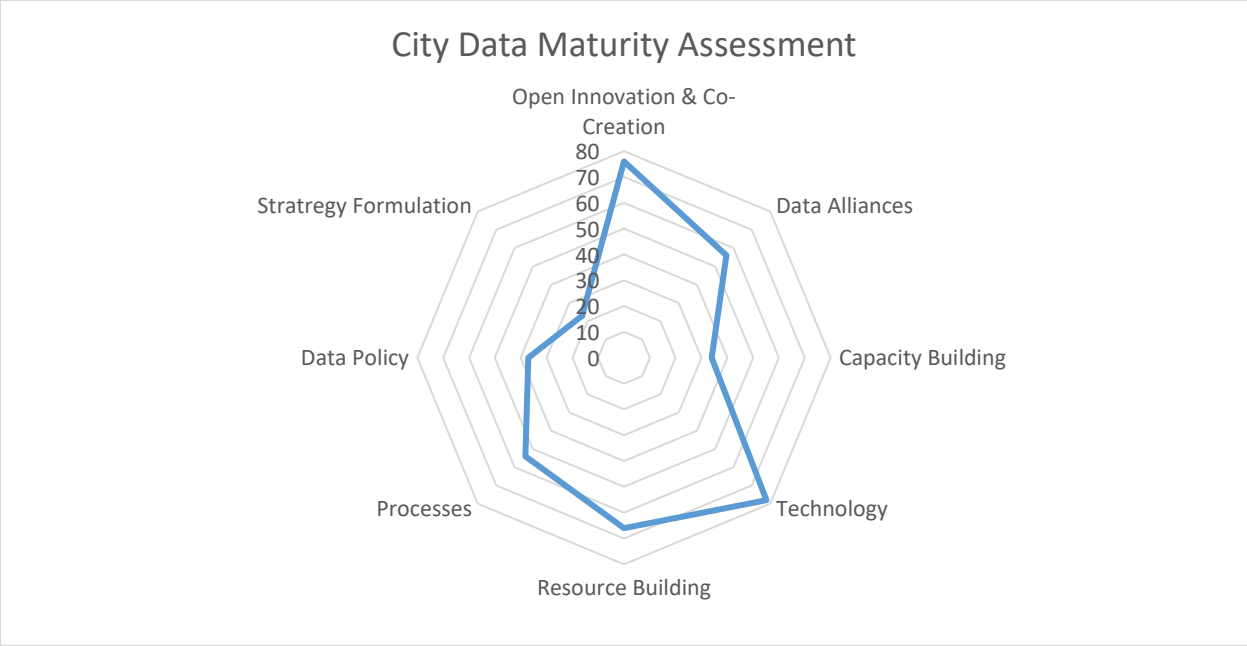
Are the processes defined on an enterprise level?	Is there any formal process/SoP for data processing and cleaning?	Y/N
	Is there any formal process/SoP for data for quality assessment of data sets?	Y/N
	Is there any formal enterprise process/SoP for Data publishing as per Open Data Norms?	Y/N
	Is there any formal process/SoP for engaging stakeholders to assess the data needs?	Y/N
	Is there any process/SoP for data collection, processing and analysis for on field Survey?	Y/N
	Do the Processes defined include provisions for data analysis?	Y/N
Technology	Technology solution deployed to integrate Sensors and field devices to capture the data at the source?	Y/N
	IT Infrastructure (Portal/Platform/APIs) deployed to support Open data initiatives to host data sets?	Y/N
	Automated tools deployed for Extraction, Transform and Load (ETL) capability?	Y/N
	Data Analytics Tools deployed to enable data analysis and generate visualization?	Y/N
	Dashboard and Scorecard platform deployed to communicate the data insights to stakeholders?	Y/N
Capacity Building	Personnel to be trained Identified?	Y/N
	Training Material and Trainers Identified?	Y/N
	Data Champions and Data Coordinators appointed and trained across departments?	Y/N
Data Alliances	Potential Alliance Partners On-boarded?	
	Industry Private Players/Startups and Incubators	Y/N
	Universities and colleges	Y/N
	Innovators/Think Tanks	Y/N
	Citizens and Communities	Y/N

CONFIDENTIAL

	Other Government Departments? (Police, Fire, Water, Sanitation, Urban Development, etc.)	Y/N
Resource Building	City Data officer appointed by Smart City Leadership	Y/N
	Have the following key resources been hired?	Y/N
	Data Champion/Expert (Y/N)	Y/N
	Data Coordinator (Y/N)	Y/N
	Data Architect	Y/N
	Data Engineer	Y/N
Open Innovation & Co-Creation	Has city conducted Innovation Hackathon Events?	Y/N
	Has city conducted Co-creation Challenges?	Y/N
	Has city conducted Data Challenges for Academia and Students?	Y/N
	How many Research Papers written on City Issues are using the City Data?	Y/N

CONFIDENTIAL

Outcome:



Toolkit can be downloaded here : << Link here>>

CONFIDENTIAL

Annexures

Annexure I: India Urban Data Exchange (IUDX)

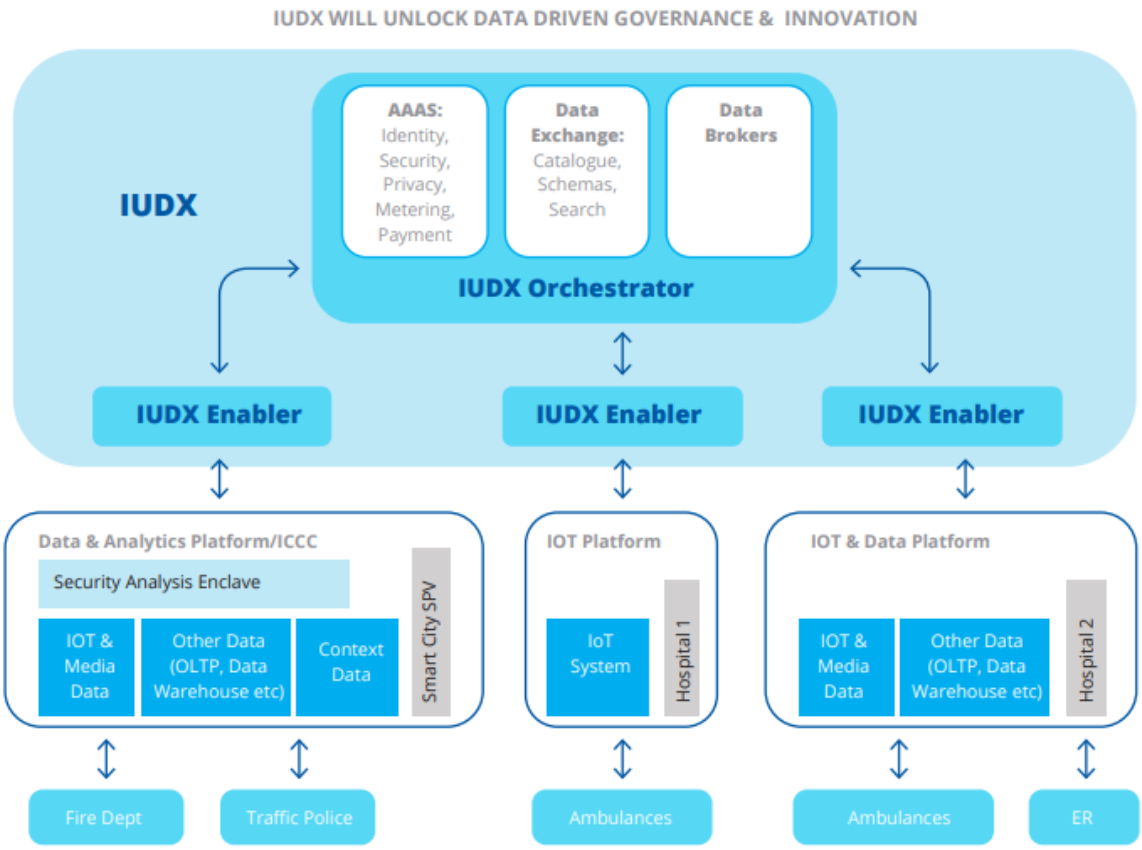
Cities generate a large amount of data daily through a variety of stakeholders. This data can be harnessed to unlock innovation, catalyze new business models and deliver data-driven governance and policy-making across urban India. At the same time, the use of data needs to be balanced in a manner that 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.

CONFIDENTIAL



Why IUDX is relevant for DataSmart 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.

An illustration of various urban services and Integrated Command and Control Centre (ICCC) in a smart city is shown below:

Figure 1

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.

CONFIDENTIAL

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.
3) 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.
4) 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.

CONFIDENTIAL

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.

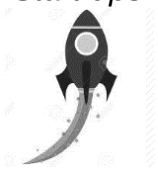
Academia and Research benefits:

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.

Stakeholder specific IUDX Use Cases

A sample set of use cases benefiting various stakeholders in the smart cities eco-systems, by use of a standardized data exchange framework, are also highlighted as follows:

Stakeholder	***Scenario***	***Solution***
Smart City SPV	Smart City SPV CEO would like to fully leverage the investment made under various ICT projects especially ICCC to solve urban challenges like Water Supply, Health, garbage, Emergency Response, Street Light, Environment, Traffic, etc. Data feeds like traffic camera, CCTV feeds, city bus transport data etc. are under control of different layer of administration like traffic, police and transport deptt. etc. and remain in various silos. Consequently, the City ICCC is not fully operational and utilized. As a CEO he needs a solution to leverage the potential of ICT solutions and take it to the next level to engage the communities in the city to make his City truly smart which works for its citizens.	By integrating the IUDX platform with the existing ICCC platform, it is possible to share data from different applications using standardized APIs. Two systems can easily start sharing and consuming data thereby addressing the challenges posed by different SI/Vendors. SI successfully integrated IUDX and configured various use cases to leverage the potential of ICCC platform in solving civic issues that needed data from more than one source. The city SPV was able to build a standard operating environment among various line departments by cross- leveraging their respective strengths and authorities. The IUDX platform also allowed SPV to integrate variety of sensors, devices from different suppliers without investing in a different application every time while buying a particular field device, sensor etc. This helped avoid vendor lock-in as standard based data sharing was enabled. IUDX also provides various mechanism that ensure data control, access authorization and secure exchange with seamless sharing of data, the city was able to build monetizable data sets to help sustain their operations.

System Integrators	SoftInc. Pvt. Ltd (a System Integrator-SI) is working with City Public transport company to set up smart devices on buses to monitor and track the real time location of buses. Client has asked SI to provide an analytic applocation to identify un-viable routes. SI estimates license cost to onboard the licensed analytics solution but cost estimate is on higher side. Client instructed SI to work on an open source solution to reduce the cost. SI is worried about the efforts and its cost implication.	Open Smart Cities Consortium India (OSCI) platform has set up a Marketplace to allow startups to distribute innovative new applications. Many creative analytics applications are operational and are available on OSCI marketplace. SI identifies a suitable application and start-up partner on the OSCI Marketplace that can address the client need,. SI works with the start-up to integrate the start-up application with its solution to meet the client requirements. SI saved lot of effort by leveraging a working application available on OSCI marketplace.
OEMs	OEM has launched environment sensors to check the Air Quality Levels, which could only work with its proprietary application code. OEM is finding it challenging to provide affordable solutions to Smart Cities.	OEM enables its product integration using IUDX framework allowing the developer community to design innovative solutions using its product for Smart Cities. A local start-up designed an innovative algorithm on the IUDX platform to predict the air quality in the city using environment sensors. OEM integrates the algorithm with its product line and offers a high-quality and affordable solution to Smart Cities. OEM also started conducting workshops with Smart Cities and the start-up community under OSCI ecosystem to educate them on its product's features. As a result, the OEM was able to align with many other start-up's to provide customized solutions.
Start-ups 	Data Inc. , a start-up, has deep expertise in Artificial Intelligence (AI). Data Inc. has designed a predictive algorithm using AI to predict the next swine-flu outbreak in the city. Data Inc. needs a lot of city data and lab data to perfect its algorithm. The challenge is the availability of quality data e.g. lab test data of patients visiting private labs.	On the IUDX platform, Data Inc. got access to masked public and private labs historical data of Swine flu cases, which preserved the privacy of individuals. Data Inc. also got access to spatial and demographic data of the city. Data Inc. used these disparate data sets and perfected its algorithm for predicting Swine Flu outbreak in the city. Data Inc. deployed the prototype and showcased it to City Healthcare officials showing a risk heat map. City Officials were amazed that algorithm was able to predict the outbreak with 95% accuracy.
Citizens		Citizens have experienced improved quality of life as Smart City SPV is able to leverage the data generated through various technology projects deployed across the city. Smart City is

		working with various start-ups to deploy innovative open source solutions to address the urban challenges.
Academic Researchers		Sumit is doing his PhD in Mathematics from a reputed university. He conceptualized the mathematical model which could potentially reduce traffic congestion. But he needs real time traffic data to design working prototype before he could showcase the product to the industry. He shared his limitation with his Professor guide who suggested him to check availability of data on IUDX platform of the city. Sumit was able to access the daily traffic data at no cost as he registered himself as a University Researcher. He was able to design his working prototype.
Universities		The University has set up Data Science Center under Department of Statistics. They have also designed the certificate course for teaching data science using various statistical techniques. University Data Science cell is approached by Smart City -City Data Officer to design solutions around City Urban Challenges. Data Science cell deployed 3 students to work over select areas with City Data Officer. University Students using the data from IUDX platform designed an algorithm which could work to predict the outbreak of communicable diseases in the city. City Data Officer using developer deployed the open source code on OSCI platform which is now being used by 10 other cities.
Developer Community		City Data Officer has set up a developer community group to invite like-minded developers to work on open source code for city operations. Developer community contributed to the development of solutions and in turn developers got hands on experience on designing and working over Smart City solutions. Also developers who have experience on different components like devices, interface, security etc. contributed to add best features to make the product rich in features. CDO received overwhelming response from developer community.

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

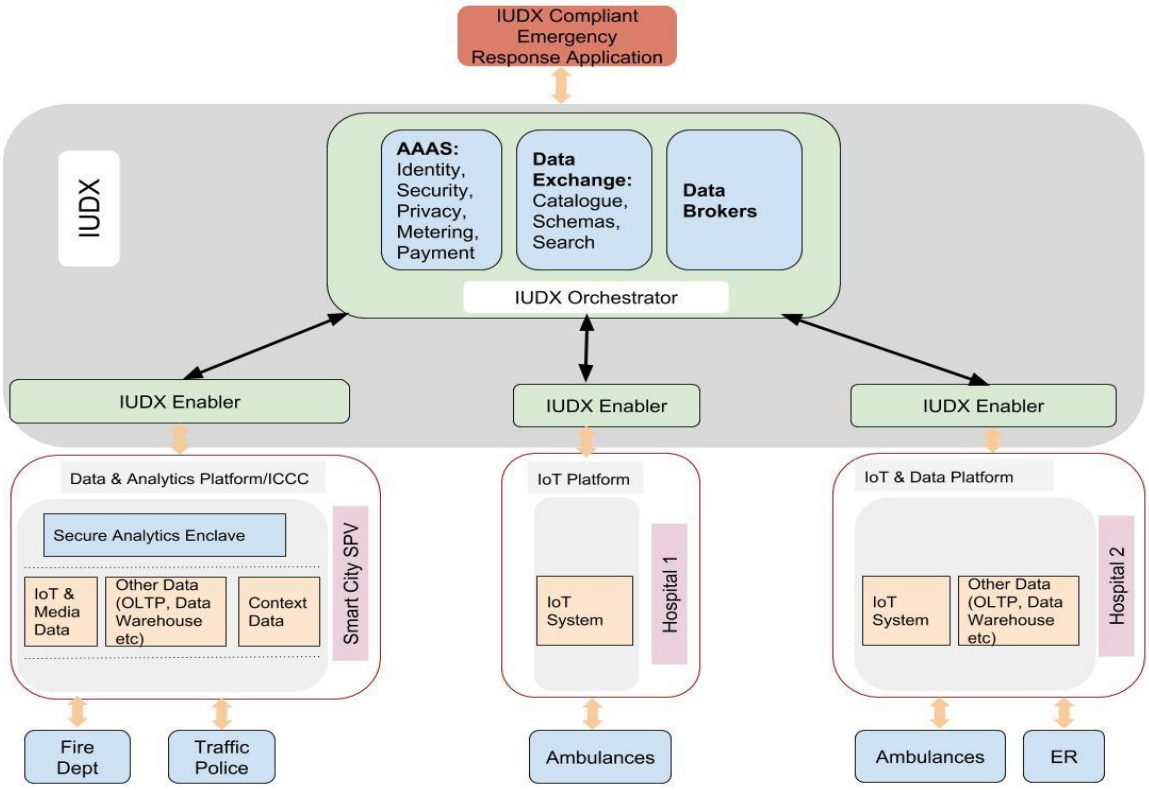
standardized publisher APIs along with common data schemas, will enable vendor neutrality for IoT devices.

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 data[6].
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.

Use Case 1: Emergency Response (ER)



[6] Context data digitally represents the existing physical, organizational and other facets of a city. Examples include GIS information, utility grids, road networks, organizational structure, building and other structures etc.

Figure 6: Illustration of how IUDX supports use case 1.

Consider the use case of a new emergency response application which coordinates across three systems: Fire department, traffic police and a bunch of hospitals. Typically, the hospital systems will be separate from a smart city deployment. However, with IUDX enablement of each platform, it will become very easy to coordinate across these different systems. When a new emergency response application is registered with IUDX, it gets permission to access the fire alarm data from the fire department, the hospital ER as well as ambulance data from the hospital system and be allowed to post requests for green corridor to the traffic police. Once the application is live, it interacts with the city data platform and the hospitals' data platforms to thus effectively coordinate the emergency response. With IUDX the same application **can now run at all IUDX enabled cities**.

Use Case 2 :Streetlight Management

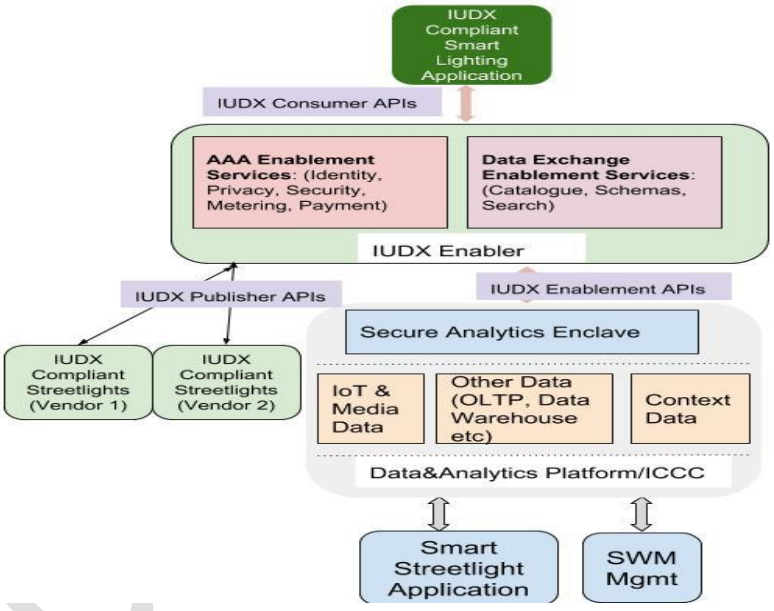


Figure 7: Enabling vendor agnostic deployment of smart streetlight solutions

Consider a smart city platform that hosts a smart streetlight application. With IUDX enablement of the platform, vendors can agree on common schemas for exchanging data between the smart streetlights and the city's platform. This will allow new streetlight devices from multiple different vendors, that are IUDX compliant, to work with the city's platform. Similarly, a new streetlight management application that is IUDX compliant, can

manage all the streetlights (including the non-IUDX compliant ones), in a more efficient way, for example, by combining analytics from traffic junction cameras to detect pedestrian activities. The traffic analytics can run in the secure analytics enclave, to ensure the privacy of video data. **More importantly, all these devices and applications will be able to run seamlessly across all IUDX compliant cities**.

CONFIDENTIAL

Annexure II: Reference City Data Policies

City	URL
San Francisco	https://datasf.org/resources/
Chicago	https://www.cityofchicago.org/city/en/narr/foia/data_disclaimer.html
Seattle	https://data.seattle.gov/
New York	https://opendata.cityofnewyork.us/wp-content/uploads/2018/04/Open-Data-Policy_TSM_v1.4_FINAL_04.02.18-1.pdf
Data for London - A City Data Strategy	https://data.london.gov.uk/dataset/data-for-london-a-city-data-strategy
Quick Guide to City Data	**https://futurecities.catapult.org.uk/project/quick-start-guide-to-city-data/**
Centre for Cities	https://www.centreforcities.org/data/
San Diego	https://www.sandiegodata.org/reports/municipal-open-data-policies/

CONFIDENTIAL

CONFIDENTIAL