Guidance Note

For Center of Excellence on AI for **Sustainable Cities**

Table of Contents

Problem Statements in Urban India

Approaching AI solutions for Cities

The traditional methods employed by municipal agencies are slower in adapting to the demands of rapidly expanding urban environments. For municipal functionaries, manual interventions are time intensive and can be vulnerable to errors. In this context AI-platforms capable of seamlessly integrating with municipal operations, can be truly transformational.

It may be appreciated that the problem statements discussed here are reflective of the current scenarios in cities. The proposals therefore are encouraged to articulate future scenarios that more deeply address the root-cause of the problems articulated here, and even the ones that are not. The consortiums can consider the following principles while articulating the solutions:

- Urban problems are networked and hierarchical, with underlying negative and positive feedback loops that AI models are well placed to resolve.
- Proposal are encouraged to address effective combinations of problem statements expressed in this document.
- The AI innovation landscape is demonstrating aggregation of multiple capabilities such as visual, text, and audio among others. Proposals are encouraged to leverage these developments to craft solutions that stretch the problem scenarios discussed below.
- Cities are adaptive institutions, but with steeper learning curves. Enhancing the adaptability of the functionaries for the below problem statements, in itself can be a key value driver for AI.

1. Urban Governance Challenges

The operational environments of cities are complex, with sector-specific challenges. Organizationally, the capabilities to adapt to emerging challenges is weak, and resources are drawn into reactive management of crises. Due to the interplay of these organizational and operational factors, there is immense potential for AI capabilities to improve the functioning and effectiveness of ULBs. Some specific aspects of the problem, to be addressed are as below:

1.1 **Document processing workflows for municipal functions:** Processing for citizen documentation such as registries and licensing involves analysis of handwritten documents of variable quality and resolution. To address this, pre-processing work is needed on the handwritten documents to enhance data quality. There is critical need for systems that can extract information from these documents. Following handwritten text recognition, the AI interventions should correct words and associate text to complete sentences. After text extraction, the AI should highlight key features within the document, such as location, names, and summary information and automate database changes.

1.2 **Improving the precision of inspections:** AI systems should enhance the precision of inspections for establishments in the regulated sector, which may include restaurants, hospitals, cowsheds, catering businesses, and distributors. As an example, the systems can refine the process of selecting establishments for inspection by crawling consumer reviews available on online review platforms.

1.3 **Safety and Security in Urban Areas:** Urban areas are often labeled as the "crime hot spots" in the country. The safety and security of the urban dwellers remain a matter of concern for law enforcement agencies and the administration. AI-enabled applications can proactively predict and prevent crime and provide timely aid to victims.

1.4 **Technical analysis & processing of bids, quotes, contracts, and invoices:** AI systems should facilitate the comparison of procurement documents, offers, call for proposals etc. with their existing in-house IT components, including APIs, applications, and more. The system should examine the textual content within these documents and cross-reference it with metadata associated with the components. Pre-processing steps may be implemented for APIs, to extract descriptive information from the metadata files. The system may also introduce capabilities to compare technical diagrams with the metadata of their IT components, further enhancing evaluation of technical parameters in the procurement process. Further, drawing insights from decisions, opinions, recommendations & reports by oversight authorities such as the Central & State Authorities, the AI system should act as a comprehensive assistive resource for cities. The AI system should enhance the detection of mandatory information, identify contract clause themes, and, where relevant, spot unfair terms.

1.5 **Shorter response time to citizen grievances & services:** Citizen services often have long processing times. Additionally, many applications are filled in incorrectly, prolonging processing times. AI systems can reduce the processing time by screening the applications for errors and presenting the functionaries with information on decisions in previous and similar cases. The system can incorporate decisions from previous, comparable cases and thus support the functionaries in their assessment and decision.

1.6 **Creating collaborative human-AI cyber-infrastructure:** The AI applications should be adaptable to the user challenges of city workforce; hence multi-lingual capabilities are important given language barriers of the field staff. Further, a lot of urban management is enabled by a highly mobile and distributed ground force. Thus, the AI applications need to be mobile capable both on Android and iOS platforms, with scalable computing requirements. The wide variety of sensors available on mobile devices can be further leveraged. Therefore, diagnostic and predictive applications of AI can be a force multiplier for the on-ground workers. A co-pilot type of scenario may be envisioned.

1.7 **Fragmented technology stacks:** City technology systems are deeply fragmented across frameworks, solutions, vendors, codebases, field device networks and access levels. AI solutions that can bring down the fragmentation in technology shall be highly valuable in increasing efficiency in city governance.

1.8 **Data quality management & synthetic data augmentation:** Cities data resources are highly fragmented with data spread across formats, frequencies, languages, and quality. AI systems that enable standardization, missing data augmentation, improving data quality, and data translation will unlock significant productivity gains in cities.

1.9 **Sandbox for urban R&D:** Currently piloting and prototyping of new urban solutions is a long-drawn process. An AI enabled sandbox that provides resources for urban R&D can introduce significant efficiency and unlock better returns on capital investment.

1.10 **Furthering inclusivity:** Currently, there is a perceived barrier to inclusivity with urban interventions designed normatively in climate resilience, urban mobility, solid and liquid waste management, water governance, and infrastructure. AI interventions can revolutionize access to public services by introducing voice controlled and gesture driven interactions.

Indicative data sources:

1. **Municipal Data:** Existing datasets on births/death, taxation, data related to land use, permits, and zoning among others.
2. **Geospatial Data**: maps, satellite imagery, and terrain models for accurate representation of the physical environment.
3. **IoT Sensors**: real-time information on traffic flow, air quality, energy consumption, and environmental conditions.
4. **Government Records**: Demographic data, land use, permits, zoning, and urban planning and development.
5. **Social Media and Crowdsourced Data**: real-time insights into events, opinions, and sentiments within the urban environment.
6. **Mobile Apps**: location-based services and user-generated content

2. Urban Digital Twins for city planning

A modern city is a complex system of systems that need to ensure high livability whilst operating in dynamic and uncertain environment where changes occur along multiple dimensions in seemingly independent manner. Therefore, for city leaders it is critical to come up with adaptive responses in an ever-shrinking window of opportunity. The specific aspects of the problem are as below:

2.1 Develop city models amenable to simulating scenarios: City leaders need simulate-able representation city scenarios amenable to automated analysis through what-if and if-what simulations. Currently, such models are cognitively and informationally prohibitive.

2.2 Reflect agent interactions: model the city as a collection of autonomous actors interacting with each other to achieve their stated objectives,

2.3 Downstream effects with spatial and temporal dimensions: The systems should dynamically simulate second & third order positive and negative externalities. This may entail real-time modelling of spatial and temporal aspects of the city such as traffic flow, energy consumption, energy efficiency, air quality, and carbon emissions.

Indicative data sources:

1. **Municipal Data:** Existing datasets on births/death, taxation, data related to land use, permits, and zoning among others.
2. **Geospatial Data**: GIS data, satellite imagery, and aerial photography for accurate mapping and terrain modelling.
3. **IOT Sensors**: Data from sensors deployed throughout the city, including traffic sensors, air quality monitors, and smart meters for energy consumption.
4. **Social Media and Crowdsourced Data**: Data generated by citizens through social media and crowdsourcing platforms to capture real-time events and opinions.
5. **Mobile Apps**: Data from mobile applications that provide location-based services and collect user-generated content.
6. **Mobility Data**: Data from public transportation systems, including bus and subway routes, vehicle tracking, and ridership information, Data from ride-sharing services, bike-sharing programs, and transportation network companies to analyze mobility patterns.
7. **Environmental Monitoring**: Data from environmental sensors measuring air quality, temperature, humidity, and pollution levels. Weather forecasts and historical weather data to incorporate weather conditions into the digital twin.
8. **Utility Data**: Data from utility providers on water, electricity, and gas consumption.
9. **Disaster & emergency services data**: Data from emergency services, including response times, incident reports, and disaster planning.
10. **Waste Management Data**: Data on waste collection schedules and routes to optimize waste management processes.

3. Urban mobility

Urban mobility in Indian cities is marred by unplanned urban growth, traffic congestion, growing pollution, and inadequate public transport systems. This has resulted in inefficient transportation, lower productivity, negatively impacting the quality of life and environment. The existing transportation planning and management systems have fallen short against the rapid urbanization and escalating traffic congestion levels. Some specific aspects of the problems to be addressed are as below.

3.1 Managing Traffic Congestion: Predicting and addressing traffic jams through route optimization and real-time demand estimation & management.
3.2 Last-Mile Connectivity: Bridging the gap between transport hubs and final destinations.
3.3 Multi-Modal Transport: Integrating various transport modes for seamless connectivity and promoting ride-sharing and carpooling.
3.4 Freight Transport Management: Efficiently planning and scheduling freight transport routes in urban areas during peak hours.
3.5 Emergency/Green Corridors: Designing effective routes for emergency services and considering environmental impact.

3.6 Pedestrian and Cycling Infrastructure: Improving sidewalks, crossings, cycling lanes, and dedicated parking facilities.
3.7 Road Quality Maintenance: Ensuring smooth mobility by upgrading roads, flyovers, and prioritizing maintenance tasks.
3.8 Smart Street Lighting: Managing streetlights for safer movement and energy efficiency.
3.9 Weather-Responsive Infrastructure: Adapting infrastructure to weather conditions for safe mobility and air quality monitoring.

Indicative data sources:

1. **Mobility & user data:** Traffic data (historical and real-time), Real-time service updates, Last-mile connectivity and multi-modal integration data, location data of users
2. **Infrastructure and Safety Data:** Pedestrian safety and cycling infrastructure data, Parking management and road maintenance data, Street lighting management data
3. **Enforcement and Emergency Data:** Maintenance scheduling and traffic signal optimization data, Traffic violation detection and tracking data, Emergency/green corridor route data.
4. **Environmental and Resource Data:** Weather-responsive infrastructure and air quality monitoring data

4. Climate & disaster resilience

Indian cities with a variety of geographical features and climatic conditions, are highly vulnerable to different kinds of climate changes, posing a significant threat to its densely populated areas. There is a need to address this for predicting, managing, and mitigating climate change effects and associated risks. Addressing climate resilience requires proactive measures to assess risks, build adaptive capacity, and implement strategies that protect and bolster urban systems, ensuring the well-being and resilience of communities in the face of climate change. Some specific aspects of the of the problems to be addressed are as below.

4.1 Early Warning Systems: Developing early warning systems for natural disasters and extreme weather to provide timely alerts to at-risk communities, enhancing preparedness and response to events such as flooding and heatwaves.
4.2 Climate Risk Assessment: Assessing vulnerability, predicting climate risks and developing adaptation strategies for extreme weather events such as flooding and heatwaves.
4.3 Resource Management: Optimizing the use of resources such as energy, water, and waste management systems by recommending efficient strategies. There is a need to improve energy planning by assessing renewable energy potential (including electric vehicle charging and solar) and managing the grid effectively to enhance energy efficiency and sustainability. The AI systems should further help assess the “green” & “strategic” aspects of procurement.
4.4 Urban Planning and Resilient Infrastructure: Optimizing infrastructure plans while considering climate change impacts and simulating scenarios for decision-making.
4.5 Air Quality Monitoring and Pollution Control: Monitoring and predicting pollution levels by analyzing air quality sensor data, meteorological conditions, and emission sources.

Indicative data sources:

1. **Satellite Imagery and Sensor Data:** Advanced satellite imagery and sensor data used to monitor and predict weather patterns.
2. **Utilities' Data:** Detailed information on energy, water, and waste consumption patterns for efficient resource management.
3. **Built Infrastructure Data:** Comprehensive data on existing buildings and structures for urban planning and development.
4. **Grid Management Data:** Data related to the operation and management of power grids, ensuring efficient energy distribution.
5. **Meteorological Data:** Comprehensive weather data including temperature, humidity, wind speed, and precipitation forecasts.
6. **Emission Source Data:** Data identifying sources and levels of various emissions, crucial for environmental monitoring

5. Water management

Effective management of water resources in India is crucial to ensure the availability of safe and sustainable water supply, while minimizing environmental impacts. Challenges such as aging infrastructure, population growth, water scarcity, and pollution require the development and implementation of an integrated management system that we need an optimal and efficient water management mechanism that is data-driven AI based solution to tackle the challenge in the urban context. Some specific aspects of the of the problems to be addressed are as below.

5.1 Water Consumption & Network Management: Efficiently manage water consumption in a city, including optimizing water resource allocation, pump operations, reducing non-revenue water (NRW), predictive maintenance, distribution network pressure adjustment to enhance resource efficiency and reduce energy consumption.

5.2 Water Quality Monitoring: Effectively monitor water quality and its surrounding circumstances using historical data.

5.3 Leakage Reduction: through collated and coordinated information using data from pressure sensors, acoustic sensors, flow meters, and other relevant sources for leakage detection.

5.4 Real-time Decision Support: Real-time insights and recommendations to water management authorities for timely intervention and protection of water quality, water body rejuvenation, in the context of pollution detection.

5.5 DPR, Project Design, Financial Management and Optimization: Leverage Generative AI to resolve challenges in investment planning, cost optimization, and financial feasibility assessment of water management strategies.

Indicative data sources:

1. **Historical Water Consumption and Demand Data:** Records of past water usage and demand patterns, essential for predicting future needs.

2. **Water Quality & Pollution Data:** Detailed information on water purity, contaminants, and pollution levels for health and safety monitoring.
3. **Data from Pump Operations and Sensors:** Operational data from pumps, pressure sensors, acoustic sensors, and flow meters for system efficiency.
4. **Real-Time Water Supply Data:** Up-to-the-minute information on water availability and distribution, crucial for managing supply systems.
5. **Historical Financial Data for Water Management:** Financial records related to water management activities and infrastructure investments for budget planning and analysis.
6. **Grievance & Community Feedback Data:** Public input for inclusive and responsive water system management.

6. Solid waste management

With its rapidly increasing urban population, India faces a significant challenge in solid & liquid waste management. There is a need to develop AI systems capable of enabling better preparedness in assessing, tracking, effort optimization and disposal of waste generated in the cities. AI systems for efficient waste management in Indian cities should address waste generation, collection, separation, diversion, and disposal /recycling processes. Some specific aspects of the of the problems to be addressed are as below.

6.1 Waste Collection, Segregation and Route Optimization: Address inefficient waste collection schedules and routes leading to high fuel consumption, increased costs, and greenhouse gas emissions. Further, inefficient waste sorting and segregation processes lead to reduced recycling rates and contamination of recyclable materials. Current systems also lack accurate forecasting of future waste generation levels, hindering the optimization of waste management infrastructure and capacity planning.

6.2 Illegal Dumping Detection: Address difficulty in monitoring urban areas and identifying instances of illegal dumping in real-time, resulting in delayed response and environmental harm.

6.3 Early Detection of Contamination: Remedy insufficient monitoring and early detection of contamination in waste, posing environmental and health risks.

6.4 Public Awareness and Grievance Management: Resolve lack of effective grievance management and community engagement in waste management decision-making

Indicative data sources:

1. **Historical Waste Generation Data:** Records of past waste production, key for predicting future waste management needs.
2. **Demographic Trends Data:** Information on population changes, age distribution, and urban migration for city planning.
3. **Real-Time Sensor Inputs:** Instantaneous data from various sensors, crucial for dynamic urban management and response.
4. **Mobility Data:** Information on vehicle route plan & metrics such as fuel & load, road usage patterns, congestion levels, and traffic flow for efficient transportation planning.

5. **Visual Characteristics Data for Waste Sorting:** Data aiding in the identification and segregation of waste types for recycling and disposal.
6. **Weather Patterns Data:** Information on local weather trends, essential for infrastructure planning and emergency response.
7. **Data for Monitoring Environmental Hazards and Diseases:** Information critical for identifying and mitigating environmental risks and health concerns.
8. **Data for Vehicle Maintenance and Resource Allocation:** Insights for optimizing vehicle upkeep and effective resource distribution.
9. **Public Input and Crowd-Sourced Requirements Data:** Community feedback and data sourced from the public for inclusive urban development.

7. Liquid waste management

India's rapidly expanding urban areas face acute challenges in liquid waste management, particularly sewage and septage. The current infrastructure struggles with the escalating volumes, leading to water pollution and health hazards. Advanced AI systems are needed to improve the efficacy of sewage and septage management, encompassing aspects like collection, treatment, and recycling. Some specific aspects of the of the problems to be addressed are as below.

7.1 Optimized Sewage Collection and Treatment: Enhance efficiency in sewage system operations, reducing operational costs and environmental impact.
7.2 Real-time Monitoring of Septage Management: Implement systems for immediate detection and response to septage disposal issues.
7.3 Accurate Prediction of Sewage Flow: Utilize data to forecast sewage volumes, aiding in effective infrastructure planning.
7.4 Contamination Detection: Improve methods for early detection of pollutants in sewage to safeguard public health.
7.5 Community Engagement: Foster public awareness and participation in sustainable sewage and septage practices.

Indicative data sources:

1. **Sewage Flow Data:** Historical and current data on sewage volumes for predictive management.
2. **Urban Demographic and Infrastructure Data:** Information on population and existing sewage infrastructure for strategic planning.
3. **Sensor Data in Sewage Systems:** Real-time data from sensors within the sewage network for operational monitoring.
4. **Public Health and Environmental Data:** Information linking sewage management to health and environmental outcomes.
5. **Grievance & Community Feedback Data:** Public input for inclusive and responsive sewage system management.

8. Built Infrastructure

Urban built infrastructure includes buildings, structures, and public spaces, green spaces, river front, transportation systems, etc. The built environment is the most important driver of

livability in urban areas. The interventions on sustainability and resilience help in improving the quality of life of the citizens and strengthen the foundations of economic vibrancy. Some specific aspects of the of the problems to be addressed are as below.

8.1 Demand management, cost optimization of maintenance & scheduled downtime of the city assets

8.2 Supporting interpretation of complex urban information across multiple datapoints such as satellite imagery, aerial photography, street view images, and social media posts.

8.3 Associations between urban features such as space/street network design, landscape features, cultural ecosystem services and outcomes such walkability, public open space vitality, transit user experience, recreation activities, and health outcomes.

8.4 Community assessment, planning, disaster risk management.

Indicative data sources:

1. **Urban Asset Management Data:** Comprehensive data on city assets including buildings, transportation systems, and public spaces for effective demand management, maintenance scheduling, and cost optimization.
2. **Satellite Imagery Analysis Data:** High-resolution satellite data for monitoring urban sprawl, land use changes, and environmental impacts in urban areas.
3. **Aerial Photography Interpretation Data:** Detailed aerial photographs providing insights into urban topography, infrastructure development, and planning.
4. **Street View Image Data:** Ground-level imagery capturing the real-time condition of streets, buildings, and public spaces for urban analysis
5. **Urban Design and Outcome Analysis Data:** Data correlating urban design features like space/street network, landscape elements, and cultural services with outcomes such as walkability, public space vitality, transit experiences, recreation, and health.
6. **Environmental and Public Health Data:** Data linking the built environment with environmental conditions and public health outcomes, essential for sustainable urban planning and citizen well-being.

Annexure 1: Data platforms, infrastructure & resources currently available in Smart Cities

The Mission has created a rich data ecosystem which can be leveraged for training AI models The data is available through multiple platforms, spread across:

1. Open Data Portal

The Smart Cities Mission Data Portal, accessible at smartcities.data.gov.in, is designed to empower cities through data by making the right data available to the right people at the right time. The portal serves as an open data platform where all 100 Smart Cities can share their datasets. It has classified its catalogue into 25 sectors and provided 75 templates for standardization to be used by the cities.

2. AMPLIFI Portal

Assessment and Monitoring Platform for Livable, Inclusive, and Future-ready Urban India (AMPLIFI) 2.0 represents a significant advancement in data-driven urban policymaking. The platform serves as an aggregator of data from myriad sources, facilitates in-depth analyses across data points and offers downloadable access to all the datasets. This data has been collected through multiple assessment frameworks of Smart Cities Mission including Ease of Living Index, Municipal Performance Index, Data Maturity Assessment Framework, Climate Smart Cities Assessment Framework and Urban Outcomes Framework 2022. The data disseminated on AMPLIFI 2.0 aims to enable research, analysis, and evidence-based policymaking for Government, Citizens, Academia, and Industry. It presents the performance of 126 Indian cities across 28 indicators under five climate centric themes: Urban Planning, Green Cover, and Biodiversity; Energy and Green Buildings; Mobility and Air Quality; Water Management; and Waste Management.

3. India Urban Data Exchange (IUDX)

City operational data from Integrated command and control centers (ICCCs) with 40+ Cities & 10+ Sectors

The MoHUA has developed a data exchange platform **vis. India Urban Data Exchange (IUDX)** which is a completely open source, based on an underlying framework of open APIs, data models, and the security privacy and accounting mechanisms that facilitates, easy and efficient exchange of data among disparate data silos. IUDX provides a way for accessing data in a unified, common format, allowing for sharing of data among multiple public and private stakeholders, as well as opening data for third party developers to create innovative new applications and citizen services. A number of datasets across various sector of the cities can be accessed on IUDX in a secured manner. Additionally, many applications have been developed using such datasets in the cities. IUDX has onboarded a diverse set of datasets from

systems/solutions deployed across Smart cities such as IoT Sensors (AQM, Flood, Rainfall, etc.), Intelligent Traffic Management, Intelligent Transport Management, Solid Waste Management, Smart Parking, Water Scada, Surveillance/RLVD/ANPR Cameras etc.

IUDX platform facilitates secure and consent-based data access in a unified, common format, thus facilitating data sharing and monetization among entities, fostering a data marketplace for innovative applications. The datasets already onboarded on IUDX can be accessed via IUDX Catalogue (https://catalogue.cos.iudx.org.in/).

S. No	Dataset (Categories)	Dataset Description/Type of Data Resources
1	Environmental info	Real-time data from Air Quality Monitoring sensors, Rain sensors
2	Solid Waste Management	Pick up locations, SWM vehicle live location, Route info, Weight of waste, Employee attendance, Fuel disbursement, Bin levels & details
3	Citizen grievances	Grievance for civic issues like Cleanliness, Water etc.
4	Bus Transit	Bus Routes, Trips, Stops, Unit info, Live location, Fare, ETA, Occupancy, Automatic Fare collection
5	E Bike Sharing	e-bike Stations, Real-time availability, Fare
6	Metro Rail	Stations, Routes, Schedules, Fare
7	Smart Parking	Parking lots, Occupancy, Real-time availability, Fare
8	POIs or Smart element locations	Locations of WiFi, PA, ECB, VMD, Digital Kiosks, Smart Poles, ANPR/RLVD Cameras, CCTV Cameras etc.
9	Adaptive Traffic	Traffic junctions, Corridors, Traffic density
10	Ambulance	Live location, On duty status (Tentative frequency: 5 to 10 seconds)
11	Flood data	Water level from Flood sensors
12	Metrological info	Rainfall (history, forecast)
13	Safety index	Safety index of streets, places
14	Video sample/feeds	Surveillance Camera/RLVD/ANPR/Any other Camera feeds

15	Revenue collection	Property Tax, Professional tax, Trade license, Vehicle registration, Facility booking revenue
16	Streetlights	Locations, Energy consumption
17	Traffic Violations	Location and Type of violation, Payment status
18	Water Distribution	Tank capacity, Supply/day, Water level, Pressure, Quality
19	Road information	Road quality, Road Assets, Road signages, Road Network
20	GIS	City assets on maps
21	Manhole Sensors	Level of wastewater/slur, Locations of sensors
22	Mosquito Density	Density/Count of Mosquitos
23	EV Charging N/w	Real-time occupancy, Fare, Locations, etc.
24	Smart Toilet	Usage count, Revenue, etc.
25	Solar Scada	Generation, Consumption, Location of plants, etc.

*****