RAJKOT SMART CITY DEVELOPMENT

# CITY DATA POLICY

RSCDL

# TABLE OF CONTENTS

# Message from the Desk of Municipal Comssioner

With the emergence of Smart Cities and data-driven innovation hubs across the country, the potential of data has become undisputed. Recognizing that timely and consistent access to data is an essential component of an open, transparent, collaborative and effective government, the city has decided to implement a City Data Policy (CDP) for the effective communication and coordination between citizens and government.

The main aim for designing this policy is to ensure data and information is utilised to its maximum potential. This policy will provide a framework for easier and effective data sharing among stakeholders and it is within the legal and legislative framework defined by the Rajkot city. It is expected to trigger a higher quotient of trust towards governments, increase collaboration and engagement with citizens, public and private entities, lead to innovation-driven problem-solving, and eventually a higher quality of life for citizens. For this to happen, it is of paramount importance to have a set of guidelines with necessary conditions built in to avoid misuse, while also generating awareness among various stakeholders and the steps and processes that need to be put in place to enhance impact of data.

For the effective and timely implementation of this policy, I urge the city and all the stakeholders to treat this policy document as a call for participation in our joint effort towards creating a robust open governance and innovation ecosystem for a more liveable urban future.

**Shri Udit Agarwal (I.A.S.)**
Municipal Commissioner
**Rajkot Municipal Corporation**

# 1. Need of Data Policy

City Government departments generate a large amount of data. The data is generated as part of daily activities of the municipal corporations and city governance departments. This raw data, if used appropriately, can help generate valuable information for Rajkot Municipal Corporation as well as entities external to Rajkot Municipal Corporation. This data can be used by these various stakeholders for economic, scientific and developmental purposes.

The need to facilitate sharing and utilisation of this large amount of data primarily points to the need of a structure defining the rules and regulations. Today most of the government information is stored in databases, and is not effectively utilised for public good. The current regime of data management does not enable open sharing of Government-owned data with other arms of the Government, nor does it expect proactive disclosure of shareable data available with data owners. Such regimes could lead to duplication of efforts and loss of efficiency in planning activities focused on city development. Hence, City Data Policy (CDP) of Rajkot Municipal Corporation aims to provide an enabling ecosystem and a platform for providing proactive and open access to the data generated through public investments and public revenue available with various departments of Rajkot Municipal Corporation/Rajkot Smart City Development Limited, other government departments etc.

This document is intended to put in place a formal data governance mechanism at Rajkot Municipal Corporation and is an attempt to further enhance the existing data initiatives of Rajkot Municipal Corporation by having in place a robust, complete and inclusive CDP. Having such a policy in place will make the city of Rajkot being identified as one of those cities that is data-driven and data-sufficient.

# 2. Scope of the Data Policy

This Policy document covers principle considerations concerning the use of data (access and utilisation of datasets including appropriate privacy management), and the principles governing the data sharing program for Rajkot, thereby defining the expectations for departmental participation and governance of the data program.

The policy is intended as a resource for city administrators such as Municipal Commissioner, Smart City CEO, and other officials such as City Data Officer, heads of various government departments, Data Coordinators, Data Champions and external agencies - parastatal, civic, private) interested in engaging with the data initiatives of the City.

This Policy will apply to all data and information created, generated, collected and archived by Rajkot Municipal Corporation / Rajkot Smart City Development Limited. This policy applies to any person/user, organisation, administrators, contractors, etc. who intends to access information or assets through any data portal of Rajkot Municipal Corporation / Rajkot Smart City Development Limited Specifically, the Data Policy applies to the following information assets of Rajkot Municipal Corporation / Rajkot Smart City Development Limited:

1. Data/information collected, captured, aggregated, processed and shared by Rajkot Municipal Corporation / Rajkot Smart City Development Limited
2. Citizens data / information
3. Personnel data/information relating to employees of Rajkot Municipal Corporation / Rajkot Smart City Development Limited

# 3. Data Policy Goals

Following are the Open Data Policy Goals for the City of Rajkot

- Increased Transpareney and Accountability thus fostering greater trust on government.
- Increase public participation in government data analysis and deliver solutions or ideas for betterment of city governance.
- Improved resource or asset visibility, social audit and open government.
- Better decision making thereby leading to more efficient and cost-effective solutions.
- Deepen open innovation, add co-creation.
- Foster data driven decisions by diverse players in urban economic ecosystem.
- Foster advanced research in academic and research institution.
- Helps cities develop new business models Empowers communities through sharing of data promotes.
- Development of emerging technologies like AI, ML and Blockchain Enhanced Goveinment to Government (G2G), Government to Business (G2B) Government to Academia (G2A) collaboration.

# 4. Data Management Principles

Implementing a data policy calls for a process that takes care of various aspects of data management. Having such a process in place will impart a sense of uniformity in the way in which data is generated/ created to the point that it is destroyed/refined/reused. This process will be applicable to all types of data at all levels, all categories and classifications.

## 1. Data Standards

Data standards are the rules that help keep the publishing and organization of open data orderly and efficient. Open Data sets should adhere to standards in order for it to be comparable, allow analysis, derive insights and interoperable to make it more inclusive.

Some data standards should be followed:

- Principles and Policy standards
- The Open Definition
- File formats for data
- Domain specific Policy standards
- Generic technical standards
- Domain specific technical standards

## 2. Data Categorization

Data will be categorized into two broad categories:

• **Personal Data:** Personal data is that data which is specific to a particular individual. It is the responsibility of the civic administration that they do not, in any case, publish personal identifiable data/information or parts of personal data/information on any of their Open Data Platforms or Datasets. It becomes the responsibility of the CDO to ensure that all personal data is anonymised before it is published.

• **Non-Personal Data:** Non-personal data is that data which cannot be identified or referenced to any individual. Anonymous data is also Non-personal data when all personal indicators and identifiers are eliminated for that particular data element.

## 3. Data Classification

Data classification is the process of organising data into categories for its most effective and efficient use. Further to data categorization, there is a need to classify the data basis its intended usage and stakeholders. There may be some data which can be open to the public,

while some other may be confidential and restricted. Such distinctions need to be appropriately defined to prevent misuse and maintain confidentiality. Rajkot Municipal Corporation will prepare the negative list of data which will be periodically reviewed. Further, all datasets will be tagged such as open data, shareable, or others.

Personal and non-personal data will be broadly classified into following levels:

| Class | Definition |
|---|---|
| Public/Shareable Data | Those data not covered under the scope of negative list and non-sensitive in nature. This data is available for public consumption and use |
| Internal Use | Information which could only be disclosed to RMC/RSCDL employees for manageing operations or delivery of public services on day to day bases. |
| Negative List | Non-shareable data as declared by the departments/organisations. |
| Restricted Data | Data which are accessible only through a prescribed process of registration and authorization by respective departments/ organisations. |
| Sensitive Data | Sensitive data as defined in various Acts and rules of the Government of India. |
| Protected Data | Data which needs to be proteced for e.g. Identify of citizens and disclosure/notification needs to be issued by RMC/RSCDL in case of any breach or loss of data. |

## 4. Types Access / Sharing of Data

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

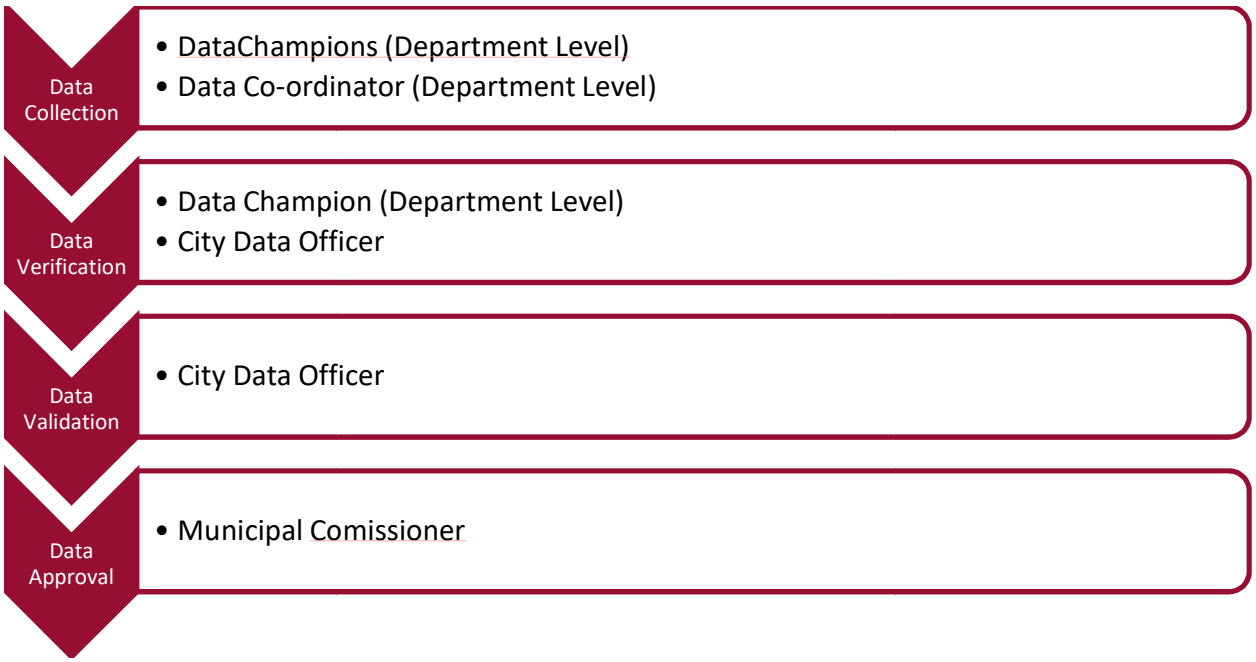
**Registered Access :** Data sets which are accessible only through a prescribed process of registration / authorization by respective departments /organizations will be available to tne recognized institutions /organizations / public users, through defined procedures.

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

## 5. Data flow/approval framework

Rajkot Municipal Corporation will set up enterprise processes to control the existing available data within the City administration. At every stage of data generation, the concerned stakeholders shall approve and authorise the data usage. Data flows may vary according to different scenarios, such as data being circulated between departments, uploading data on the open data portal, sharing data with third party, etc. Each dataset has a trustee accountable for data quality and security. Appropriate data flow and approval mechanisms should be in place for such situations. Also, all applications must be developed in compliance with the Government of India's India Enterprise Architecture (IndEA) framework.

The Approval framework for Rajkot City data is as follows:




## 6. Data Archival

Archiving is defined as secured storage of data/ documents, such that the same is rendered inaccessible by authorised users in the ordinary course of business, but which can be retrieved by an administrator designated by the HoD for the document in question. Based on certain aspects such as compliance with statutory and regulatory requirements, responses to inspections from regulator and availability of documents for any statutory assessments, the following archival schedule and mode of archival is defined.

| Data archival buckets | Archival period | Additional guideline |
|---|---|---|
| 1 | Permanent | - |
| 2 | 10 years | Deactivated records – archival period will be 10 years after retention period is over |
| 3 | 5 years | • Agreements/ Work Orders/ Collateral documents – 5 years after expiry of the document<br>• Employee Records – 5 years after retention period is over<br>• Legal Records – 5 years after the dispute is closed<br>• Audit reports- 5 years from the end of the fiscal period in which the audit or review was concluded |

Mode of Archival

| | |
|---|---|
| Electronic ( E) | Records to be archived in electronic form. These records may be stored on assigned servers or on tapes as per rules and guidelines issued by the technology department |
| Physical (P) | Record to be archived in physical form. The records may be archived in the premises or vendor premises. |
| Original Form (O) | Records to be retained in the original form in which they were created or used i.e. either electronic or physical. |

Data will be stored in the main database for 6 Months in a live state so that whenever a report needs to be generated, the data will be extracted from main database. Data older than 6 months will be archived. If report duration extends beyond 6 months, the data will be retrieved from archival to generate the report.

Please change the time periods as appropriate and decided by the city authority

## 7. Retention

Data Retention: CDO to undertake retention of data, i.e. the maintenance of documents to be accessed by an authorised user at a later stage. Retrieval schedule of the data will be as per the rules and regulations defined by the Government of India:

**1) Electronic:** Records to be retained in electronic form. These records may be stored on shared drives with access to only authorised individual/ group of individuals.

**2) Physical:** Record to be retained in physical form. These records may be kept in file cabinets or any other storage units assigned to each department with proper labelling so as to enable quick identification of the records.

**3) Original Form:** Records to be retained in the original form in which they were created or used i.e. either electronic or physical.

## 8. Data security and privacy

Data Security and privacy will be managed by RMC/ RSCDL under the preview of the Data Policy.

A) Physical Security:
   a. The premises will be physically secured, access control devices should t control devices should be available for accessing the premises entry and exit should be monitored.

B) Network Security:
   a. Appropriate firewalls, IPS, SSL devices etc. should be used to ensure Network secirotu. The solution should support encryption across network and between client and server mechanism for transferring data.

C) System Security :
   a. Adequate access control procedures should be followed to secure the entire IT system, physically and logically.

   b. The access controls procedures should cover all stages in the life-cycle of user access, from the initial registration of new users to the final de-registration of users who no longer require access to information systems and servicel.

   c. The system should have e factor authentication mechanism either through One Time Password (OTP) or soft token - based technologies for access control and user authentication.

D) Application Security.
   a. The solution should have appropriate authentication mechanisms.
   b. Application user authentication & authentication related tranbactions should be encrypted.
   c. Operating system should be hardened on which the application is installed.
   d. I web application firewall shall be deployed to secure the web layer.

E) Audit Trails & logs.
   a. Event logging should create an accurate record of user activity such as which users accessed which system, and for how long.
   b. The solution should log all types of events especially those related to security.

F) Data Protection
   a. The solution should support SSL encryption mechanism for transferring data across network.
   b. The data transferred across network should be encryptedtrsing (PKl) Public Key Infrastructure.
   c. Access to all system resources including data files, devices, processes and audit files should be provided to the intended users only.
   d. All mobile applications should be designed and developed in a way that It ensures security of the application and data on the device.
   e. Ensure to protect documents by assignins security parameters and criteria in ordfi io provide more effective proiection foi an electronig document in order to rnaintain Confidentiality, Authorization, Accountability, Integrity, Authentication and Nop-repudiation.

G) Session Management :
   a. The system should limit to only one session per user or process ID.
   b. The system should put a limit on the maximum time length.of anidle sessit,rt, whish should ensure that automatic session termination takes place after expiry of the specific time length.
   c. Mandatory password change after predefined time period.

H) Data Warehouse SecuritY :
   a. Users must not have access to the data warehouse prompt-of the.applicatior Access to the data warehouse prompt must be restricted only to the databar, administrator.
   b. “Super user" rights for the data warehouse - must only be given to the administrator and activities of these accounts must be properly logged.

I) Application Deployment.
   a. All unused ports should be blocked at server machines.
   b. The application server should be segregated from Internet zone through firerwall or other filtering

J) Information Security Governance.
   a. The employees working on the prolect should be be made aware of his or her responsibility with respect to Information Privacy and Information security
   b. Employees working on the project shall undergo security awareness training during iraining.

K) Compliance to Security Standards

a. Software/Hardware syrstem should be in compliance with ISO/IEC 27001:2015.

L) Security Information and Event Management System (SIBM)
   a. SIEM shouldbe available for Real-time analysis of securityile5ts generatedby applications and infrastructure.

M) Database Activity Monitoring (DAM)
   a. DAM should be available to monitor all database

# 5. Data Publishing Plan

**Develpoment:** Developed facilitating cost saving in terms of software and licenses and also provisioning community participation in terms of further development of product with modules of data visualization consumption, APIs to access datasets etc.

**Metadata:** Resources (Datasets/Apps) shall be published along with standard metadrtta along with conlolled vocabularies on government sectors, juridictions, dataset tpyes, access mode etc. Besides facilitating.easy access to datasets, this shall be ertremely useful in the future for federation/integration of data catalogs.

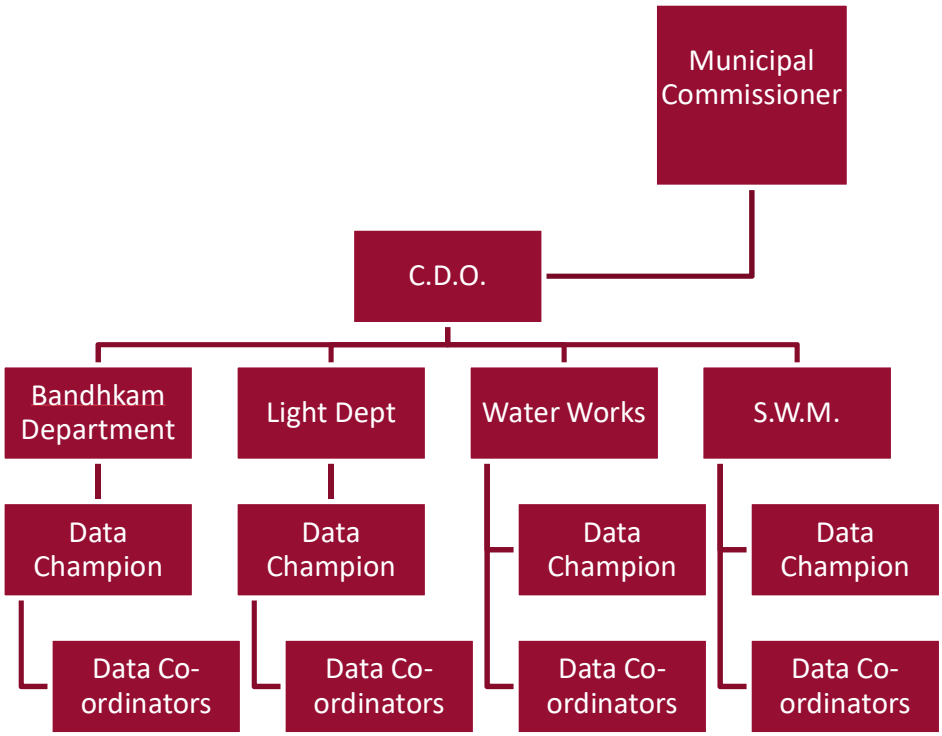
**Soeial Media Connect:** To support wider reach and dissemination of datasets anyone can share the information about any dataset published on the platform with his/her social media pages.

**Citizen Engagement:** The Platform has also a strong component of Citizen Engagement. Citizens can express their views well as rate the datasets w.r.t there aspects (euality, Accessibilityand Usbility) the scale of 5. They can also embed the Resourcer (nuturets/Apps) in their blogs or web sites. Facility to contact the Chief Datr Officers is also available-on the Platform.

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

# 6. Data Management and Team Structure

Rajkot Municipal Corporation has set up a data team for management of data at the city level and enabling coordination between various departments for making sure that the data is available as and when required. The quality of data available and its use depends largely on the efficacy of the team that is put in place. The data team structure for Rajkot Municipal Corporation is as follows:




Rajkot Municipal Corporation will also engage and secure buy-in from both internal and external stakeholders on key decisions. Further, this will help navigate through complicated hurdles (e.g. bureaucratic, political etc.) and to take prompt decisions and actions pertaining to collection, segregation and release of data.

# 7. Roles and Responsibilities

Key officials of the data team and their roles and key responsibilities are defined below:

## 1. City Data Officer (CDO)

The CDO will be the officer responsible for implementation of the DataSmart Cities Strategy at the city level. CDO will work with city leadership to assess and tap the potential of data and set up data culture across the organisation and outside the organisation. CDO will report directly to City Leadership and act as single point of contact to all internal and external stakeholders in the city. City leadership also needs to deploy dedicated skilled resources to drive the data initiative through CDO.

The key responsibilities of CDO are as follows:

- Ensure that the CDP evolves as per the needs of various stakeholders of the city and relevant upgrades to policy are carried out time to time accordingly.
- Coordinate with MDO (Mission Data Officer) to align with mission data strategy and priorities with respect to open government initiatives and policies.
- Organise regular meetings of the City Data Alliance (CDA).
- Coordinate with officers of various other government departments/agencies within the city for the effective implementation of CDP.
- CDO along with team of Data Champions/ Coordinators must assess the data requirements of various stakeholders in smart city ecosystem. External stakeholders may also need to be engaged to understand the data needs. CDO 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.
- Publish Data Catalogues and Datasets/Feeds on Open Government Data Portal and ensure that such datasets are updated at regular intervals as needed and create mechanisms for continuous feedback from citizens and stakeholders on type of datasets to be published.

- Assess all the operational IT Projects for identifying public datasets/feeds. Data Champions and Data Coordinators in respective departments must prepare integration plan with respective IT vendor/ integrator to ensure compliance as per CDP.

- Assess all proposed or under implementation projects to identify the datasets/feeds which could generate public datasets/feeds or may be useful for internal analysis. CDO must work with concerned System Integrator/vendor to ensure compliance of smart solutions with CDP.

- Assess all periodic and recurring MIS needs to identify the datasets/feeds which could be shared with other departments through data exchange. CDO will also assess third party funded reports related to city operations for e.g. City Mobility plan, Health Plan etc

## 2. Data Champions (DCs)

Data champions will be senior functionaries who would champion the implementation of the CDP in their respective departments/ organisations.

The key responsibilities of Data champions are as follows:

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

- Actively publish/enable publishing of datasets/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 CDP.

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

- Undertake activities to engage with their stakeholders and evolve their department's strategy on data in line with the deliberations

## 3. Data Coordinators

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

The key responsibilities of Data Coordinators are as follows:

- Aggregate the data demand from various channels.
- Sensitizing the department employees over the importance of data quality etc.
- Perform collection, interpretation and recording of data in accordance with CDP standards and CDO guidelines.
- Perform data validation and ensure data quality.
- Sort and organise the data; both hard copy and electronic versions.
- Transmit data report to Rajkot Municipal Corporation / Rajkot Smart City Development Limited or CDO via Internet.
- Update Rajkot Municipal Corporation / Rajkot Smart City Development Limited website or Rajkot Open Data Portal with latest data records.
- Assist department staff in data entry when required.
- Provide data management updates in all internal and external meetings as required.
- Analyse data for quality improvement purposes.
- Prepare data for reporting, meetings and presentations for the concerned department and Rajkot Municipal Corporation / Rajkot Smart City Development Limited at large.
- Ensure data management procedures comply with CDP.
- Provide statistical analysis and longitudinal analysis of data.
- Prepare and submit data required for audits

# 8. Standard Oerating Procedures

Standard Operating Procedures (SOPs) are sufficient guidelines designed to achieve consistency in specified situations by postulating a standard practice in performing those functions. These SOPs are designed with a view to enhance and standardise data operation and management.

## 1. SOP for data collection

Data collection is referred to as the method of collecting information in a systematic way. This is the first step towards data-driven decision making and evidence-based governance. Data collection provides both a baseline to measure and a target to improve. In order to get the primary datasets we have to collect data from different wards, zones which is under Rajkot Municipal Corporation, Rajkot Smart City Development Limited employee and Other Government departments. Data which is not going to be used for any kind of analysis or will not be used for any communication purpose should not be collected at all.

**If the request is received from external agency:**

1. If the request is received from external agency, it should be directed to the CDO.
2. Depending on the requested data or the data which need to be collected, the CDO shall direct the request to Data Champion (DC) of the concerned department.
3. The DC of the concerned departments checks for the requested data. If data is available with department, DC shall instruct the Data Coordinators or the concerned personnel to gather the data in requested format.
4. DC will take approval of data from their HOD.
5. If fresh data is needed to be captured /acquired, the Data Champion/Data Coordinators in consultation with CDO shall take appropriate action.

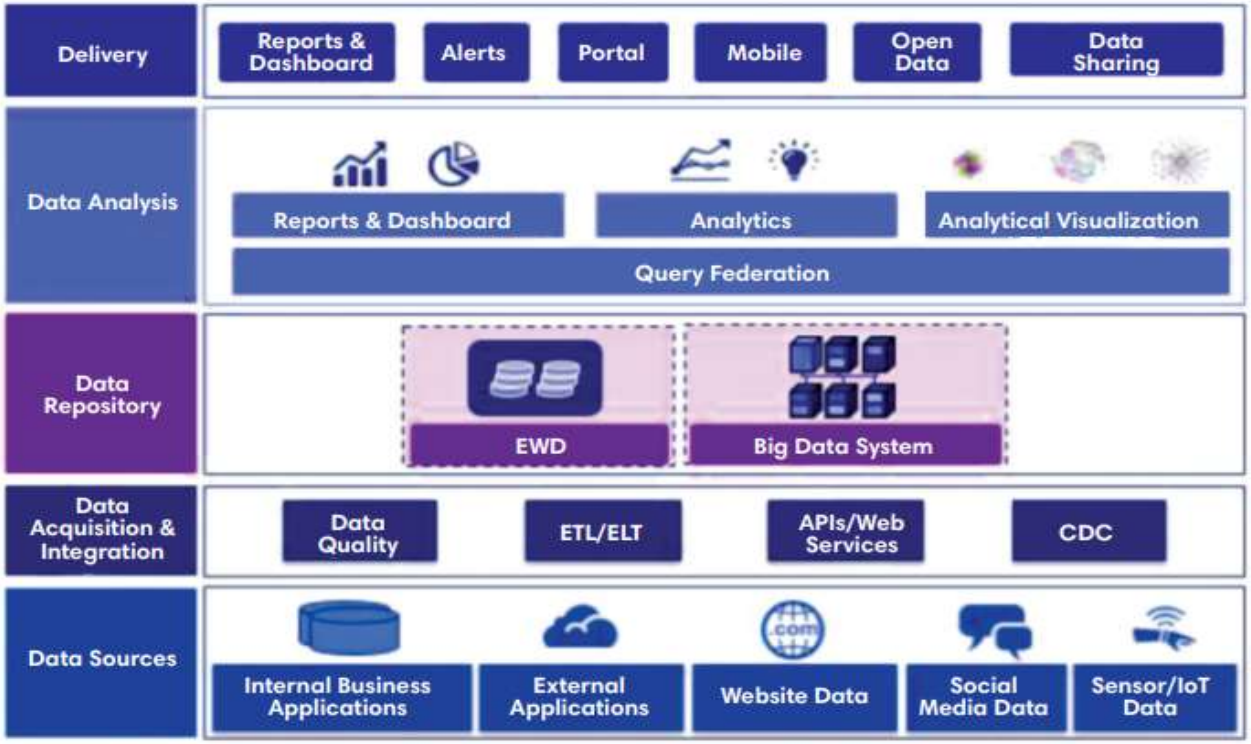
**If the request is received from internal departments:**

1. If the request is received from internal departments, it should be directed to the Data Champion for the concerned department.
2. The DC of the concerned departments checks for the requested data. If data is available with department, DC shall instruct the Data Coordinators or the concerned personnel to gather the data in requested format.
3. If fresh data is needed to be captured /acquired, the Data Champion/Data Coordinators in consultation with CDO shall take appropriate action.

## 2. SOP for electronic data collection

Additional factors need to be kept in mind when any form of data is collected, accessed, transferred or stored electronically. Cities ecosystem comprises of various technology solutions ranging from Sensors, IoT, SCADA, Electronic camera, GIS, payments system etc. which generate loads of structured data every second on different dimensions. Cities could also leverage various unstructured data emerging from different sources and third-party systems like social media, internet, websites, videos, images etc.

1. The data should be collected with consent of the end-user who may be a citizen or Rajkot Municipal Corporation / Rajkot Smart City Development Limited employee.
2. Special care needs to be taken for data privacy and security. The measures mentioned in above sections of the Policy shall be followed to maintain confidentiality and security of data.
3. For data collection and integration with the IoT systems / sensors, the following Data Architecture highlighted below shall be followed by Rajkot Municipal Corporation / Rajkot Smart City Development Limited



## 3. SOP for data processing and cleaning

Raw data may be old and inaccurate and can have an adverse impact on results. Data cleaning will be done to ensure that data is correct, consistent and useable by identifying any errors or corruptions in the data, correcting or deleting them, or manually processing them as needed.

1. While collecting the electronic data, the IT applications/ IT systems should be developed in such a way that under any circumstances these applications/systems should not accept any wrong data/null data.
2. If there is existing data, identify discrepancies which may come from different sources.
3. The collected data shall be properly processed and cleaned before performing any kind of analysis.
4. If needed, commercial software available in the market can be used with prior approvals from the concerned authorities.

Extract, Transform and Load (ETL) is the common methodology used for data integration and processing. 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 analysed, and stored (loaded) into a data warehouse or other system.

Specifically, Data Champion/ Data Coordinator will keep in mind the following points while collecting data from their respective department:

- Spaces in extra columns Compliance
- Ward-wise Compliance
- Blank Cells Compliance
- Standard format Compliance
- All NA Compliance
- Special Characters Compliance
- Split Sheets Compliance
- Datasets Compliance
- Data Completeness

## 4. SOP for quality assessment of datasets

Quality assessment of data is needed to ensure that the quality standard is met i.e. accuracy, free from any sort of legal issues, privacy of an individual is maintained and does not compromise with the National security. Data quality shall be assessed from the perspectives of adequacy, appropriateness, accuracy and reliability, authenticity, consistency and validity.

- Responsibility for quality check of data rests with the CDO while publishing data on the open data portal or sharing it with stakeholders upon request.
- Under some circumstances, a special committee under the leadership of Municipal Commissioner (Rajkot Municipal Corporation ), comprising of CDO, and Additional/Assistant Commissioner may be formed for data quality assessment.

## 5. SOP for data publishing

National Data Sharing and Access Policy (NDSAP) defines standards for publishing datasets and feeds on the open data portal. CDOs must ensure adherence towards defined standards and classification. This SOP describes the steps required for publishing data to the Smart Cities Open Data Portal:

**1. Understand the requirement:** Follow proper procedures to collect the relevant data to be uploaded on the open data portal. Understand the publishing options and the available datasets.

**2. Process the data:** Ensure data is in an appropriate format to be published on the open data portal. It does not contain any personalized information, is open, authenticated and free from defects.

**3. Prepare to deploy/publish data:** Follow procedures specific to the publishing option you have selected and work with the appropriate team (when necessary) to publish your data. NDSAP recommends that datasets should be published in an open format and should be machine readable. Data format can be chosen from the list recommended by NDSAP, highlighted in the Appendix.

**4. Publish metadata:** Follow established metadata procedures as per NDSAP and any other guidelines laid down by the city to publish metadata on the Portal and create linkages between data and metadata.

**5. Obtain approvals and finalise deployment:** Obtain the appropriate management approvals for your data based on your selected data publishing option, either from CDO or Municipal Commissioner.

## 6. SOP for engaging stakeholders

The concept recognizes the value of enhancing engagement among all four stakeholders of the quadruple-helix model—Government (Rajkot Municipal Corporation Departments, Rajkot Smart
City Development Limited and Other Departments of Governments), citizens, academia, and industry, along with improvements in the internal workflow and decision-making processes of city governments.

CDOs along with the team of Data Champions/Coordinators shall assess and document the data requirements of various stakeholders in the city ecosystem, along with frequency of consumption and level of granularity.

Key activities may include:

- Identifying stakeholders from various age groups and ethnicities and engaging them in city initiatives
- Organising workshops, hackathons/events to promote brainstorming over required datasets
- Decision making and consultation with data experts to zero down on the most important datasets required on the portal
- Data ideation with public forum to gain understanding of citizen/industry needs
- Formation of city data alliances

## 7. SOP for data collection, processing and analysis for on-field surveys

Field survey is defined as collection and gathering of information at the local level by conducting primary surveys. On-field surveys may be required in situation where data from digital sensors or existing datasets are inadequate. Surveys may be administered to fulfil a certain gap, with a well-defined problem for investigation. Data collected from on-field surveys can help Rajkot Municipal Corporation / Rajkot Smart City Development Limited get a snapshot of how things are at a specific time. The survey research may be descriptive, analytical or evaluative. Field surveys are cost-intensive hence they may be conducted after a thorough mapping of their objectives and outcomes.

Survey data processing consists of four important steps:

1. **Survey designing :** Survey is a research strategy and not a method. CDOs/concerned team should choose the most appropriate method based on purpose. CDO shall help the concerned team devise suitable surveys/questionnaires by clarifying the objective, determining sample and deciding upon the survey mode to finally create the questionnaire. A research method should not advise the questions, but other way around. Posing difficult to answer questions, in such case a simple rule or scale can be used to help respondents.

2. **Data collection :** Survey can employ a range of methods including questionnaires, interviews or even focus group discussions. Data entry in either format may happen manually or electronically. Going-forward, all the e-governance IT applications/systems shall be designed in such a way that manual processes get replaced by automated process without much intervention of humans. As most of process would be automated and handled by an e-mode, data will be available for further analysis.

3. **Data processing :** Before any analysis is possible, ensuring accuracy and quality of data is paramount. Survey form data is always prone to errors, omissions and other inconsistencies. This data inconsistency and incompleteness, if not edited and corrected on time, can complicate the analysis and may even result in wrong analysis. Data processing shall comprise of various steps necessary for preparing the data for analysis, including editing, data classification, removing redundancies, and preparation of tables. This is an important step when the survey instrument collects qualitative data, which needs to be then represented in a format for analysis.

4. **Data analysis :** Data analysis covers the final step of characterizing and interpreting research findings. In situations where the digital tools are employed for the survey, and the data can be processed easily. Data analysis will involve computation of certain indices or measures along with searching for patterns of relationship that exist among the data groups. The task of analysing quantitative data may be accomplished through statistics. Descriptive statistics is to be used for organising raw data obtained in the process of research, such as tabulation and classification of data. Inferential statistics, also known as sampling statistics, will be used for making inferences or conclusions from the data collected from a sample and drawing generalisations on the entire population.

## 8. SOP for data analysis

Data analysis or analytics is an encompassing and multi-dimensional field that uses mathematics, statistics, predictive modelling and machine-learning techniques to find meaningful patterns and knowledge in recorded data. This will help in using intelligent techniques to uncover actionable insights from the relevant data.

Rajkot Municipal Corporation / Rajkot Smart City Development Limited to establish analytical capabilities within the Data Team to accomplish data analysis on the cleaned and processed data.

- **Descriptive Analytics:** It helps in answering "What is happening?" For Example: Using past financial performance to predict a customer's likely financial performance. Descriptive anal)'tics can be useful in the sales cycle,

- **Diagostic 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 ehtire sales process, analyzing lead source, number of communications, types of commuriications, social media, documents, CRM data, etc. Properly tuned predictive analytics can be used to support sales, marketing, ot 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 clinicaliy obese, then add filters for factors like diabetes and LDL cholpterol levels to determine where to focus trealment. The same prescriptive model can be applied to almost any industry target group or prollem.

## 9. SOP for Data Monetization

SoP for Data Monetization A committee comprising of representatives from various departments of RMC/RSCDL, Traffic Police, city educational institutions and market experts would be formed to be CiW bata Committee. In the context of Right to Information act, this committee would decide the data which can be monetized, fix the price, and review it from time to time. The data can be differentially priced for academic research and commercial used.

RMC will create data monetizing platform, it will form revenue sources that would help the data ecosystem to become self-sustainable.

# 9. City Data Alliance (CDA)

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 CDP so as to inform and evolve the CDP effectively.

The responsibilities of the CDA will include:

A. To act as an advisory group to the city leadership on the CDP.
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 based on city data.
E. To support industry to design solutions using emerging technologies like Artificial Intelligence (AI), Machine Learning (ML) and Blockchain.
F. To assess and design use cases critical to the citizens of the city.
G. To generate awareness among various stakeholders towards open government initiatives.
H. To bring city's stakeholders on a 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 the city.
K. To deliver research papers using city data on civic problems in the city. l) To design and develop prototypes/ solutions on civic problems in the city.
L. To organise data-challenges on complex civic problems.
M. To organise hackathons and support shortlisted solutions at city level.
N. To set up scholarship for postgraduate and graduate students to work with the City Data Team along with the CDO.
O. To publish the progress report every month.
P. To prioritise the datasets/feeds for publishing on the open data platform.
Q. To sensitise ecosystem partners to share data for solving civic challenges.
R. To support, engage and encourage network/groups/members of data enthusiasts in the city.
S. To improve city capacity over data driven governance and policy formulation.
T. To support CDOs by extending resources (like interns, researchers, technology experts), funds (program sponsorship etc.) and technology (solutions etc.).
U. To share data available with the partners on data platforms to promote city data.

# 10. Stakeholders and Collaboration

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

- **Government Agencies :** Government agencies operating with dedicated administrative structure in city namely Traffic Police, City Police, Central/State Government Departments, Government Autonomous Bodies etc. (apart from City Administration).
- **Funding Agencies:** Funding agencies which regularly work with City Administration in different domains for e.g. World Bank, ADB, DFID, etc.
- **Industry:** Key flagship manufacturing/service industry promoters/players in the city/state.
- **Academia:** Representatives from leading universities/colleges/schools in the city.
- **Policy Advocacy groups and NGOs:** Policy advocacy groups and NGOs working in different domains/ areas like Slums, Health, Education, Environment, Participatory Governance, Mobility etc.
- **Start-ups and Incubators:** Representatives from start-ups and incubators in the city/state.
- **City Businesses:** Representatives from local small and medium business communities.
- **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.
- **Local Elected Representatives:** Local elected representatives to further the interest of citizens/ communities towards data driven policy governance and policy formulation.
- **Professional Representatives:** Representatives from various professional services like Doctors, CA, and Engineers etc.

# 11. Annextures

1. Open Government Data (OGD) Platform Policy
   Government of India has listed down policy for external users to access the data hosted on the Open Government Data Platform. This policy can be access at https://data.gov.in/policies. The same needs to be implemented and followed by Rajkot while implementing the policy document. Further, the Government Open Data Use Licence – India has been approved. This can be accessed at https://data. gov.in/sites/default/files/Government_Open_Data_Use_Licence_India.pdf

2. Standardization of Data Access as per National Data Sharing and Access Policy (NDSAP)
   NDSAP defines standards for publishing datasets and feeds. These standards need to be adhered to by Rajkot while sharing its datasets.

   Open Source Driven: Datasets are considered to be open by default unless classified as internal, sensitive, protected or restricted.

   Metadata: Datasets and apps must be published along with proper metadata. Besides facilitating easy access to datasets, using a common data taxonomy/structure shall be extremely useful in the future for federation/integration of data catalogues. Key Metadata elements for catalogues/resources include:

   Catalogues :
   • Title (Required): A unique name for the catalogue (group of resources) viz. Current Population Survey , Consumer Price Index , Variety-wise Daily Market Prices Data, City-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 (Optional): 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 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.

Resources (Datasets/Apps) :

• Category (Required): Choose from the drop-down options of whether it is a Dataset or an Application.

• Title (Required): A unique name of the resource viz. Consumer Price Index for etc.

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

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

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

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

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

• NDSAP Policy Compliance: This field is to indicate if this dataset is in conformity with the National Data Sharing and Access Policy of the Govt. of India.

**If Resource Category is Dataset**

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

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

**If Resource Category is App**

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

• Datasets Used: Datasets used for making this app.

• Language: Language used for app.

3. Data Formats: NDSAP recommends that datasets should be published in an open format, which can be accessed without the need for a software licence and should be machine readable. The data could be published in any of the following formats:

• CSV (Comma separated values)
• XLS (Spreadsheet - Excel)
• ODS (Open Document Formats for Spreadsheets)
• XML (Extensive Mark-up Language)
• RDF (Resources Description Framework)
• KML (Keyhole Mark-up Language used for Maps)

- GML (Geography Mark-up Language)
- RSS/ATOM (Fast changing data e.g. hourly/daily)

4. State level Data policy

The Rajkot City Data Policy is designed so as to apply to all shareable non-sensitive data available either in digital or analog forms and generated using public funds by various State departments and Subordinate offices/organisations/agencies. It is designed to promote data sharing and enable access to Government-owned data that could be used for planning and development.