CITY DATA POLICY

For

Faridabad City

Contents

1. Introduction……4
2. Need for the policy……5
3. Scope……6
3.1 Open Data Policy Goals……7
3.2 Ownership of Data……7
4. Data Management……8
4.1 Data Standards……8
4.2 Data Categorizations & Classifications……9
4.3 Data Archival and Retention……11
4.4 Data Security & Privacy……12
4.5 Data management & Team Structure……13
4.6 Data flow & Approval……15
4.7 Data Set Provisioning and Identifying Dataset……16
5 Stakeholder & Collaboration……18
6 Roles & Responsibilities……20
6.1 City Data Officer (CDO)……20
6.2 City Data Champions……22
6.3 City Data Coordinators……23
6.4 City Data Alliances……25
7 Data Policy Budget Allocation……27
8 Open Data Policy Implementation……27
9 Standard Operating Procedures……29
9.1 SoP for Data & Electronic Data Collection……30
9.2 SoP for Data Processing and Cleaning……32
9.3 SoP for Data Quality Assessment & Dataset……33
9.4 SoP for Data Publishing as per Open Data Norms……34
9.5 SoP for Engaging Stakeholders to Assess Data Needs……35
9.6 SoP for Data Collection, Processing & Analysis as per field Survey……36
10 Annexures……37

Abbreviations and/or Definitions

Terms	Description
MoHUA	*Ministry of Housing and Urban Affairs*
ULB	*Urban Local Bodies*
MCF	*Municipal Corporation Faridabad*
FSCL	*Faridabad Smart City Limited*
SPV	*Special Purpose Vehicle*
CDP	*City Data Policy*
CDO	*Chief Data Officer*
MDO	*Mission Data officer*
DC	*Data Champion*
OGP	*Open Government Data*
API	*Application Programming Interface*
AI	*Artificial Intelligence*
ML	*Machine Learning*
RWA	*Resident Welfare Associations*
G2G	*Government-to-Government*
G2B	*Government-to-Business*
G2C	*Government-to-Citizen*
C2G	*Citizen-to-Government*

1. Introduction

City Data Policy is the first significant step in the direction to provide conceptual clarity over accessing and sharing protocols over city data. City Data Policy also provides clarity around ownership of data, legal framework, terms of use etc. As an extension of this, Municipal Corporation Faridabad (MCF) and Faridabad Smart City Limited (FSCL) has decided to design a city data policy document in line with National Data Sharing and Accessibility Policy (NDSAP), Government of India.

Efficient sharing of data among data owners and inter-and-intra governmental agencies along with data standards and interoperable systems is the need of the hour. A data policy is essential to understand the contours of data sharing, privacy, security and ownership in the context of the city. Data policy is also needed to define the contours of collaboration between various governmental/ non-governmental entities on sharing and access of data.

Policy document will guide the City Administration to set up enterprise processes to leverage the existing available data with City administration. It is intended as a resource for city administrators such as Municipal Commissioner, Smart City CEO, other officials such as City Data office, Heads of various government departments, data coordinators and data champions and external agencies interested in partnering with the data initiatives for the City. It will help to establish processes to allow data to flow between departments and users seamlessly internally without any user interventions.

2. Need for the Policy

City Government departments generate a large amount of data. This data is generated as a part of daily activities of the municipal corporations, smart city and city governance department.

In parallel to it, there is also large chunk of data being generated, on daily basis, by other than government agencies like RWAs, market associations, Shopping malls, etc.

The need to facilitate sharing and utilization of these large amount of data primarily points to the need to a structure defining rules and regulations.

The current regime of data management does not enable open sharing of Government owned data with other arms of the government nor does it expect proactive disclosure of sharable data available with data owners. Such regimes could lead to duplication of efforts and loss of efficiency of planning of activities focused on development.

Here, City Data Policy of Faridabad Municipal Corporation & Faridabad Smart City Limited aims to provide an enabling provision and platform for providing proactive and open access to the data generated through public fund available with various government departments along with data generated by various agencies other than government departments.

3. Scope of the Policy

This Policy will apply to all data and information created, generated, collected and archived by Faridabad Smart City Limited (FSCL)/Municipal Corporation Faridabad (MCF)/other government departments of Faridabad.

This policy also applies to all data and information created, generated, collected and archived by other than government agencies of Faridabad who willing to share data for the public interest.

The purpose of this policy is to set guidelines for incorporating a data framework and implementing the same into existing and future systems and procedures.

This policy will also aid in determining which data sets should be made public and shareable, how to make data sets public and shareable and how to maintain and modify the existing published data sets.

With the help of published data sets, the city will enable the public to:

- Assist in distinguishing and proposing the viable and proficient solutions for government
- Promote innovative methodologies/activities for social advancement
- Create economic/financial advancement across the city

3.1. Open Data Policy Goals

The major goals of this Open Data Policy are:

- Data Availability and Accessibility to citizens within defined Legal and Legislative Framework
- Increased Transparency and Accountability thus fostering greater trust on Government
- Improve public understanding of City Operations and other Information concerning their Communities
- Foster data driven decisions by diverse players
- Empower communities through sharing of data
- Empower City employees to be progressively effective, better coordinated internally, and recognize chances to better serve the public
- Enhancing Government to Government (G2G), Government to Business (G2B), Government to Citizen (G2C) and Citizen to Government (C2G) collaborations.
- Encourage the development of innovative technology solutions that improve quality of life.

3.2 Ownership of the Data

- Data will remain property of agency/department/ ministry/ entity which generates and collect it.
- Once open access data is declared by the data owning agency/department then only it can be shared or published.
- It can be shared through Disclaimer.

4. Data Management

Implementing a data policy calls for a process that takes care of the Data Lifecycle 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.

The data management with its various components viz. data categorization and classification, data retention and archival, data security and privacy are based on different parameters of concerned stakeholders/departments/associations. And these component standards and lifecycle is decided by all related stakeholder department/associations in collaboration with City Data Cell and City Data Alliance. This can also be further revised based on requirement and necessities arises in future.

4.1 Data Standards

Data Standards are rules that help keep the sharing, publishing and organization of data orderly and efficiently. Data set should adhere to standards in order for it to be comparable, allow analysis, drive insights and interoperable to make it more inclusive. It becomes imperative to have in place metadata so that available data can be effectively used. This helps achieve uniformity in the way different data features are inferred. The responsibility of adherence to data standards should lie on the shoulders of the City Data Officer (CDO). Some of the data standards are

- Principles and policy standards
- The Open Definition
- File Formats for data
- Domain specific policy standards
- Generic technical standards
- Domain specific technical standards, etc

Some of the important features of metadata are:

- **Standardized name of the data element:** A commonly known and accepted name by which the data element is called across the organization. In case there are different nomenclatures for the same element, the same should be formally documented.
- **Standardized definition of the data element:** Synonymous to the standard name the data element should also have a known and formally accepted definition that will be used to define the element across the organization.
- **Rules for Data Quality:** The data should meet the norms of quality which may include but are not limited to, Accuracy, precision or resolution, completeness, format, consistency, range, pattern etc.

4.2 Data Categorization & Classification

4.2 (a) Data Categorization

Data can be categorized in the below two categories;

- **Personal Data:** Personal data is that data which is specific to a particular living individual. Even incomplete portions of information when combined gives or leads to the identification of an individual are also termed as personal data. It hence becomes enormously commanding to 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 Data Sets. It becomes the responsibility of the CDO to ensure that all personal data is anonymized before it is published.
- **Non-Personal Data:** Non-personal data is that data which cannot be identified or referenced with any living individual. Anonymous data is also Non-personal data when all personal indicators and identifiers are eliminated for that particular data element

4.2 (b) Data Classification

Personal and Non-Personal data can be further classified into the following:

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

4.3 Data Retention & Archival Policy

4.3 (a) Data Retention Policy

Retention is defined as the maintenance of documents which can be accessed by an authorized user as and when required.

Mode of Retention:

Electronic (E)	Records to be retained in electronic form. These records may be stored on shared drives with access to only authorized individual/ group of individuals.
Physical (P)	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 labeling so as to enable quick identification of the records.
Original Form (O)	Records to be retained in the original form in which they were created or used i.e. either electronic or physical.

4.3 (b) Data Archival Policy

Archiving is the process of moving data that is no longer actively used to a separate storage device for long-term retention. Archive data consists of older data that is still important to the organization and may be needed for future reference, as well as data that must be retained for regulatory compliance. Archiving is defined as secured storage of data/ documents, such that the same is rendered inaccessible by authorized users, but which can be retrieved by an administrator designated authority.

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.

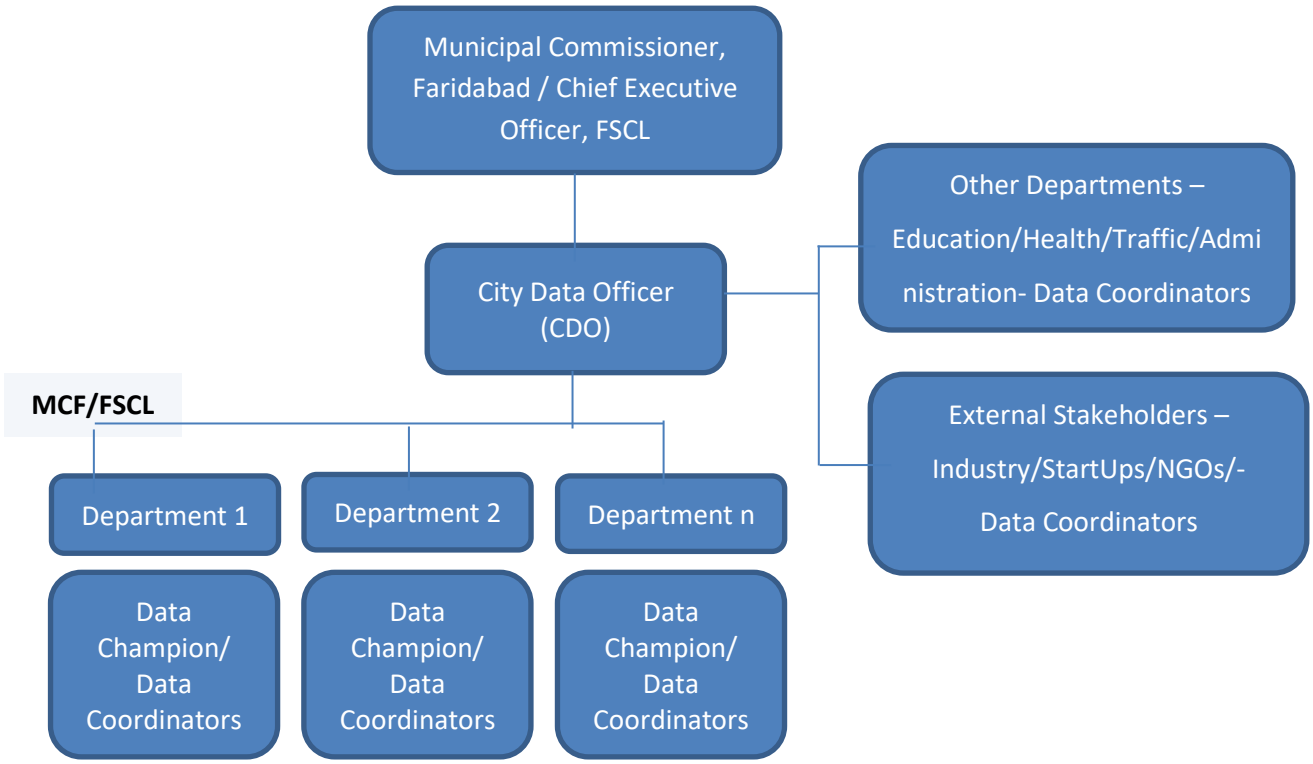
4.4 Data Security and Privacy

Managing security including privacy of data is crucial for building and maintaining trust between end users/all stake holders and is the major element of the city data policy.

- Data collection, sharing and analysis must be ring-fenced by a privacy first approach to guarantee protections for residents and users.
- Clear-cut identification of restricted data, sensitive data, negative list and sharable data by the concerned stakeholders so as to apply specific security & privacy to that category of data.
- Need to establish or comply with existing standards and certifications for data privacy and security.
- Data dissemination should be only to authenticated and authorized stakeholders (both internal and external) through data fiduciaries.
- It is the responsibility of City Data personnel/s to build up a data secured ecosystem. The organization should set down clear standards on the Do's and Don'ts for the entrance and utilization of Data and Information. Accentuation/emphasis ought to be laid on for what reason the information can be utilized and for what reason it can't be what reason it can't be
- For management of Privacy and Security of Data, it is recommended that all data access must be through some secured channels & encryption process by using any of the security features like Application Programming Interface (API)/Web Services calls to ensure appropriate security controls.
- Except for open data, it is recommended that direct access to data be prohibited.

4.5. Data Management & Team Structure

A data team has to be setup 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 Data management is as follows:



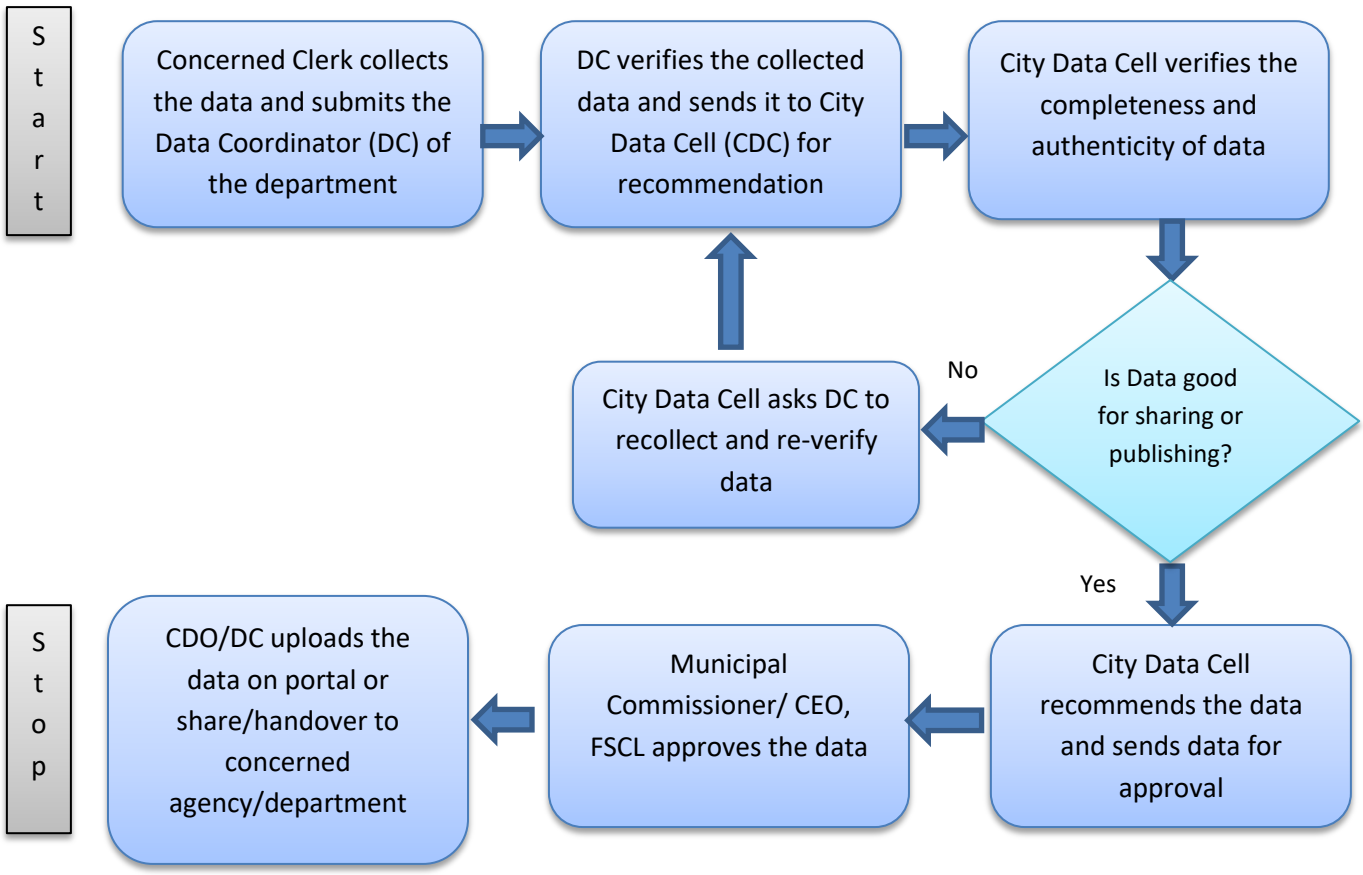
City Data Cell should be formed to engage participation and secure buy-in from both internal and external stakeholders on key decisions. Further, sufficiently empowering this cell can help navigate through complicated hurdles (e.g. bureaucratic, political etc.) and to take prompt decisions and actions pertaining to collection, segregation and release of data.

Though Open Data Policies may be written for a specific city or ULB under consideration, yet there is always going to be the need of close collaboration and integration with State Level and National Level entities.

The various entities involved at various levels that adhere to set protocols and guidelines as prescribed by GoI.

- National Level
  - Mission Data Officer (MDO)
  - Data Analytics and Management Unit (DAM Unit)
  - Smart City Data Network
- State Level
  - Officer In charge of Urban Data Work in coordination with State leave Data/Statistics department of Municipal Administration and existing municipal data alliance
  - State level data analytics and Management Unit
  - State Level Data Alliance
- City Level
  - City Data Officer
  - City Data Cell
  - City Data Alliance

4.6 Data Flow/Approval frameworks



Note:

- Exchange/sharing of information/data across different departments in SPVs, Corporation and State agencies are in accordance with framework decided by the concerned stakeholders in purview of the NDSAP & City Data Policy guideline.
- Exchange/sharing of information/data from private/external agencies/associations are in accordance with common framework decided by the SPV, Corporation and the external agencies/associations in purview of the NDSAP & City Data Policy guideline.

4.7 Data Set Provisioning and Identifying Data Set

Identifying important dataset that will be published on Open Data portal would have direct and indirect impact on the citizen (in terms of city infrastructure and services) and other stakeholders. Faridabad City has envisioned having open data policy in place to critically identify these important datasets.

Hence it becomes necessarily important to exhaustively identify the data sets that would have varying level of impact on citizens and other stakeholders of the city. While deciding on selecting the data sets the below factors can be considered for.

- Demand/Need of Public Data Sets
    - Identify the need and demand for the Data Set under consideration.
    - particular datasets and the priority with which it is required
    - Engage the important stakeholders demanding this Data Set.
    - Segregate the so identified data as Restricted, Non-Restricted, Sensitive, Personal etc.
- Social Impact
    - Datasets which potentially could create a positive social impact should be included
    - Such positive impacts can include generation of jobs & employment, boosting education and literacy, etc.
    - Datasets with negative influence should be avoided from being published.
- Data Market Potential
    - Publishing certain datasets can encourage innovation; encourage entrepreneurship eventually resulting in generation of economic value.
    - Such data sets that could boost startups and entrepreneurship, innovation and change should be prioritized for publishing.

- Legal and Compliance requirement
    - Datasets that have a legal binding to be published, for instance under the ambit of RTI, should be identified and made public.
    - There can be legal and compliance requirements for mandatory disclosure of a certain dataset to the public.
- Evading Conflicts/Opposition
    - Some Data sets may have the potential to expose unattended civic issues and problems and hence bring to fore the people/dept. responsible for these issues.
    - Utmost care should be taken to ensure that publishing such data sets does not hamper the continuity of the open data initiative.

5. Stakeholders and Collaboration

In order to successfully implement the Open Data Initiative it will have to identify the various stakeholders within MCF/FSCL and external to it. After having identified them it is equally important to collaborate effectively with these stakeholders. In initiatives like these it is of utmost importance to successfully establish Government to Government (G2G), Government to Business (G2B), Government to Citizen (G2C) and Citizen to Government (C2G) collaborations.

- **Government agencies**
  - Government agencies operating with dedicated administrative structure in city namely Traffic police, City Police, Central/State Government departments, Government Autonomous Bodies
- **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.
- **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:**
    - Representatives from local elected representatives to further the interest of citizens/communities towards data driven policy governance and policy formulation.
- **Professional Representatives :**
    - Representatives from various professional bodies like Doctors, CA, and Engineers etc.
- **City Businesses**
    - Representatives from local Small and medium business communities

6. Roles and Responsibilities

6.1. City Data Officer (CDO)

The CDO will act as custodian and driver of City Data Policy (CDP) and a flag bearer of open government initiative. 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 report directly to city leadership and act as single point of contact to all internal and external stakeholders in the city. CDO will be supported by a pool of Data Coordinators. The CDO will work with all the departments including those which are outside the jurisdiction of the smart city & Municipal Corporation. For example, Traffic, Police, power utilities, telecom networks, gas distribution, etc. Following are the roles and responsibilities of CDO:

- The CDOs will create a City Data Policy (CDP) for Faridabad Smart City which will be reviewed periodically to keep it contextual to the need of the times.
- Coordinate with MDO to align with mission data strategy and priorities with respect to Open government initiatives and policies. City Data Officer along with a team of data champions/ coordinators must assess the data requirements of various stakeholders in the smart city ecosystem.
- Coordinate with MDO to align with mission data strategy and priorities with respect to Open government initiatives and policies.
- Coordinate with officers of various other government departments/agencies within the city for the effective implementation of City Data Policy.
- Organize regular meetings of City Data Alliance (CDA).
- City Data officer will engage with the external stakeholders to understand the data requirements of the city.
- City Data Officer must identify cross-cutting data sets and publish it on a platform based on data needs of various stakeholders in a routine manner.

- City Data Officer must design a program to solve its problems through a structured challenge process like hackathon, events 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
- Publish Data Catalogues and Data Sets/Feeds on OGD/FSCL/MCF portal: CDOs will publish data Catalogues and Data Sets/Feeds on OGD/FSCL/MCF defined platforms/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.

6.2 Data Champions

Data champions (DCs) will be senior functionaries, not below the rank of a Head of Department or equivalent, who would champion the implementation of the City Data Policy at the department level. DCs need to act as trainers and lead the team of data coordinators at the department level. DC will be first touch point of CDO and must undertake continuous capacity building programs for their staff. Following are the roles and responsibilities of Data Champions:

- Shall identify the data sets/feeds, derived information, intelligence or data challenge with respect to day to day operations of the department.
- 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.
- 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.
- DCs will undertake activities to engage with their stakeholders and evolve their department's strategy on data in line with the deliberations.

6.3. Data Coordinators

Data Coordinators are designated for each department of FSCL/MCF as the main point of contact and accountability for open data in their department and are sometimes the individual in charge of individual database, data sets or information systems. Data Coordinators will assist the Data Champions at the department level as reporting staff. This Data Coordinator will act as data management experts who will coordinate, gather, collect and sanitize the data for Open Data Portal development. Overall data coordinator is the person who knows how to get access to data. For the successful roll out and implementation of Open Data Policy, set of roles and responsibilities to be defined for Data Coordinators. To achieve the same, following shall be the responsibilities of appointed data coordinators:

- Aggregate the data demand from various channels
- Serving as a key point of accountability for timelines and questions about datasets
- Implementing privacy, metadata and other standards & practices
- Sensitizing the department employees over the importance of data quality etc.
- Support to Data Champion for implementation of city data policy at department level
- Understand the public demand to identify high value data sets so as to make them available in a way that maximizes the number of users and portal usage
- Assist the data champion and city data officer for continuous updation and improvement in open data policy
- Provide the guidance and training to department users regarding open data portal and its usage
- Engage with other departments to prototype and execute data-related projects
- Maintain data integrity and security and ensure the proper sorting and organizing of datasets

- Should prioritize and classify the data as per the policy standard. e.g. Data to be categorized as Personal or Non Personal Data, further it should be classified in the defined category like Secured, Protected, Restricted data
- Collaborate with other public sector organizations/departments to define best practices and lessons learned for similar projects
- Data updation and maintenance on periodic basis based on the timelines/granularity defined by FSCL/MCF
- Assist with other key open data initiatives led by the MCF/FSCL, MoUHA or any other government bodies in coordination with CDO and Data Champion
- Inventorying department data sets for the Managing accuracy, quality and completeness of the data
- Assisting departments and Data Champion with analysis of city data sets
- Determining what datasets are appropriate for public disclosure
- Managing access to and use of the data, including documentation in coordination with CDO and Data Champion

6.4 City Data Alliance (CDA)

Role

The CDA will provide a collaborative framework to create and define use cases to solve critical city problems through the use of data, catalyze the right set of collaborations and networks to make available such data and undertake continuous dialogue between various stakeholders in the city around the City Data Policy so as to inform and evolve the CDP effectively. The alliance will undertake education and awareness about data in the community, understand and address concerns on data privacy and security, build use cases for city problems, create data collaborations between various government and private agencies for solving relevant use cases and continuously evolve the culture of data in the city's context.

Responsibilities

- To act as an advisory group to the city leadership on the City Data Policy.
- To assess the data needs of various Smart City stakeholders.
- To promote data driven governance and policy formulation.
- To design and implement solutions and analysis using city data.
- To support industry to design solutions using emerging technologies like AI, ML and Blockchain.
- To assess and design use cases critical to the citizens of the respective cities.
- To generate awareness in various stakeholders towards open government initiatives.
- To bring Smart Cities stakeholders on common platform to influence the city data priorities.
- To facilitate data for co-creation and collaboration over civic issues
- To provide critical feedback to the city over the quality and relevance of data provided by Smart City.

- To deliver 4 Research paper annually using City Data on Civic Problems in Smart City
- To design and develop two prototype/ solutions annually on Civic Problems in Smart City
- To organize a data-challenge every half yearly on complex civic problems
- To organize a Hackathon annually and support shortlisted solutions at city level
- To set up scholarship for postgraduate and graduate interns to work with Office of CDO.
- To publish the progress report every month
- Prioritize the Data Sets/Feeds for publishing on Data Platform: To sensitize ecosystem partners to share the data for leveraging data for solving civic challenges
- To support, engage and encourage network/groups/members of data enthusiasts in Smart City
- To improve city capacity over data driven governance and policy formulation
- To support CDOs by extending resources (like interns, researchers, technology experts), funds (program sponsorship etc.) and technology (solutions etc.)
- To share data available with partners on Data Platform to promote city data.

7. Data Policy Budget Allocations

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

8. Open Data Policy – Implementation Plan

Once Open Data Policy is approved and formalized, the next step for FSCL/MCF would be to implement the Policy while adhering to the norms and guidelines as stated in this policy document. Implementing an Open Data policy may not be a simple process as it involves the participation of various stakeholders like Civic Administrators, Political Champions, external departments etc. Apart from having a Data Policy in place it also becomes equally important to have an implementation plan and roadmap in place to successfully implement the open data initiative.

Referential Implementation plan while implementing its Open Data Policy are as

1. Initiation
    a. Identify Stakeholders
    b. Analyze existing structure, readiness in terms of People, Process and Platform
    c. As-Is-assessment of current Data Platforms, Data Process & Data Practices
2. Planning
    a. Define and Formalize Open Data Policy.
    b. Define Scope in terms of Data Sets
    c. Define team structure with Roles & Responsibilities
    d. Plan Quality & Risk Management

3. Execution
    a. Publish and Implement Open Data Policy
    b. Data Set Finalization
    c. Stakeholder Engagement
    d. Design Open Data Platform
    e. Publish Data Set on Open Data platforms
4. Continuous Improvement
    a. Maintenance and Support
    b. Dataset review and revalidation
    c. Assess performance of Accountable Stakeholders
    d. Refine and Redefine the Open Data Policy

9. Standard Operating Procedure

Standard Operating Procedures (SOPs) are succinct formal documents designed to achieve consistency in specified functions by specifying standard practice in performing those functions.

In order to have responsibilities defined, RACI Matrix defined. It is a powerful tool to assist in the identification of roles and assigning of cross-functional responsibilities.

RACI Definitions:

- Responsibility = person or role responsible for ensuring that the activity is completed
- Accountable = person or role responsible for actually doing or completing the activity
- Consulted = person or role whose subject matter expertise is required in order to complete the activity
- Informed = person or role that needs to be kept informed of the status of activity completion

9.1 SoP for Data & Electronic Data Collection

9.1(a) for manual collection

The purpose of this SoP is to define the Data collection processes in the collection and management of open data held or on behalf of Faridabad City and the responsibilities of individuals involved. It also aims to describe good practice in data collection and management techniques. This SOP should be used when any form of data is collected, accessed, transferred or stored.

	Stakeholders				
Open Data Activities	**CEO, FSCL/Municipal Commissioner**	**City Data Officer**	**Data Champion**	**Data Coordinators**	**Other Representatives (Alliances)/Dept.)**
Identification of Data Sets (after receiving request directed to data coordinators of concerned department)	I	A	C	R	I/C
Defining standard format or dataset format for collection/gathering of data	I	A	C	R	I/C
Data Collection	I	R	C	A	C
Data Sanitization	I	A	C	R	I/C
Data Categorization and Classification	I	A	R	R	I/C
Data Standardization	I	A	R	R	I/C
Data Approval	I	R	C	I	I
Publishing the data set	I/C	A	I	R	I

9.1(b) for electronic collection

The purpose of this SoP is to define the Electronic Data collection processes in the collection and management of open data held or on behalf of Faridabad City and the responsibilities of individuals involved. It also aims to describe good practice in data collection and management techniques electronically. This SOP should be used when any form of data is collected, accessed, transferred or stored.

	Stakeholders				
Open Data Activities	**CEO, FSCL/Municipal Commissioner**	**City Data Officer**	**Data Champion**	**Data Coordinators**	**Other Representatives (Alliances)/Dept.)**
Identification of Data Sets	I	A	C	R	I/C
Defining standard format or dataset format for collection/gathering of data	I	A	C	R	I/C
Data Collection	I	R	C	A	C
Data Sanitization	I	A	C	R	I/C
Data Categorization and Classification	I	A	R	R	I/C
Data Standardization	I	A	R	R	I/C
Data Approval	I	R	C	I	I
Publishing the data set	I/C	A	I	R	I

9.2 SoP for Data Processing & Cleaning

The purpose of this SoP is to define the Data processing and cleaning while doing the collection and management of open data held or on behalf of Faridabad City and the responsibilities of individuals involved. It also aims to describe good practice in processing of data and cleaning techniques. This SOP should be used when any form of data is collected, accessed, transferred or stored.

	Stakeholders				
Open Data Activities	**CEO, FSCL/Municipal Commissioner**	**City Data Officer**	**Data Champion**	**Data Coordinators**	**Other Representatives (Alliances)/Dept**
Data Identification and Collection	I	R	R/C	A	C
Defining standard format or dataset format for collection/gathering of data	I	A	C	R	I
Data Sanitization	I	A	C	R	I/C
Data Categorization and Classification	I	A	C	R	I/C
Analysis and Accuracy Validation	I	A	C	R	I/C
Periodic gathering and collection of dataset as per defined frequency and granularity	I	R	R/C	A	I/C
Security Check with all parameters involving removal of personalized data	I	A	C	R	I/C

9.3 SoP for Data for Quality Assessment and Data Sets

The purpose of this SoP is to define the processes for the quality assessment and data sets of open data held or on behalf of Faridabad City and the responsibilities of individuals involved. It also aims to describe good practice in quality assessment and Datasets management. This SOP should be used when any form of data is collected, accessed, transferred or stored.

	Stakeholders				
Open Data Activities	**CEO, FSCL/Municipal Commissioner**	**City Data Officer**	**Data Champion**	**Data Coordinators**	**Other Representatives (Alliances)/Dept.**
Data collection as per standard defined	I	R	R/C	A	I/C
Data Sanitization	I	A	C	R	I/C
Security Check involving removal of personal data	I	A	C	R	I/C
Units/Currency other terms standardization	I	A	C	R	I/C
Periodic gathering and collection of dataset as per defined frequency and granularity	I	R	R/C	A	I/C

9.4 SoP for Data Publishing as per Open Data Norms

The purpose of this SoP is to define the data publishing processes as per the Open Data Norms of open data held or on behalf of Faridabad City and the responsibilities of individuals involved. It also aims to describe good practice in data publishing and management techniques as per Open Data Norms. This SOP should be used when any form of data is collected, accessed, transferred or stored.

	Stakeholders				
Open Data Activities	**CEO, FSCL/Municipal Commissioner**	**City Data Officer**	**Data Champion**	**Data Coordinators**	**Other Representatives (Alliances)/ Dept**
Data Categorization and Classification	I	A	C	R	I
Data Standardization	I	A	C	R	I
Analyzing the Data	I	A	C	R	I
Consulting and Identifying new data set	I	R	R/C	A	I/C
Defining new problem areas to tackle upcoming problems Constantly enhancing the quality of data sets	I	A	C	R	C
Keeping concern stakeholders updated about the new guidelines and open data norms	I	R	C	R	I/C

9.5 SoP for Engaging Stakeholders to Assess the Data Needs

The purpose of this SoP is to define the processes for engaging stakeholders to assess the data needs of open data held or on behalf of Faridabad City and the responsibilities of individuals involved. It also aims to describe good practice in engaging the stakeholders to assess the data needs. This SOP should be used when any form of data is collected, accessed, transferred or stored.

	Stakeholders				
Open Data Activities	**CEO, FSCL/Municipal Commissioner**	**City Data Officer**	**Data Champion**	**Data Coordinators**	**Other Representatives (Alliances)/Dept**
Organizing workshops, hackathons/events and trainings	I/C	A	C	R	I/C
Decision making and consultation with data experts	I	A	I	R	I/C
Data ideation with public forum	I/C	A	C	I	I/C
Formation of city data alliances	I/C	A	R	I/R	I/C
Identifying stakeholders from various age groups and ethnicities and engaging them in city initiatives	I	A	R	I/R	I/C
Brainstorming and role modeling sessions with concerned stakeholders and subordinates	I	A	R	I/R	I/C

9.6 SoP for Data Collection, Processing and analysis for on field survey

The purpose of this SoP is to define the Data collection, processing and analysis processes for on field survey of open data held or on behalf of Faridabad City and the responsibilities of individuals involved. It also aims to describe good practice in data collection, processing and analysis techniques for on field survey. This SOP should be used when any form of data is collected, accessed, transferred or stored.

	Stakeholders				
Open Data Activities	**CEO, FSCL/Municipal Commissioner**	**City Data Officer**	**Data Champion**	**Data Coordinators**	**Other Representatives (Alliances)/Dept**
Electronic device usage for identification and collection of Data	I	C	C	R	A/C
Quality assessment of data	I	C	C	R	I/C
Defining standard format or data set format for collection/gathering of data	I	I	C	R	C
Data Sanitization	I	I	C	R	I/C
Data Classification and standardization	I	I	C	R	I/C
Measuring and reviewing the accuracy of data	I	I	C	R	I/C
Identification of new data sets using electronic devices	I	C	R/C	R	I/C
Periodic gathering and collection of datasets as per defined frequency and granularity	I	I	R/C	A	I/C

Annexures

1. **Standardization of Data Access as per National Data Sharing and Access Policy (NDSAP)**

   National Data Sharing and Access Policy define standards for publishing data sets and feeds. These standards need to be adhered to by Faridabad City while sharing its data sets.

   - **Open by Default:** Datasets are considered to be open by default unless classified as internal, sensitive, protected or restricted.
   - **Meta Data:** Datasets 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.
   - **Data Catalogue:** As per NDSAP metadata elements for data sets or feeds are defined as follows:
     - **Title (Required):** A unique name for the catalog (a 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 catalog 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 catalog. Example: rainfall, weather, monthly statistics.
     - **Group Name (Optional):** This is an optional field to provide a Group Name to multiple catalogues to show that they may be presented as a group or a set.

- **Sector and Subsector (Required):** Choose the sector (s)/subsector (s) that 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/Application) cater to namely entire country, state, district, city etc.

- **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.
- **Machine readable:** Data sets and feeds should be machine readable
- **Formats:** As per NDSAP following data formats should be published:
  - CSV (Comma separated values)
  - XLS (Spreadsheet - Excel)
  - ODS (Open Document Formats for Spreadsheets)
  - XML (Extensive Markup 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)

References

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