

# Data Policy – Detailed Document

*Department of Registration and Stamps, Government of Maharashtra*

August 2021

## Table of Contents

<b>1. Executive Summary</b> .....	3
<b>2. Introduction</b> .....	5
<b>3. Purpose</b> .....	5
<b>4. Scope</b> .....	6
<b>5. Approach</b> .....	6
<b>6. Data Landscape</b> .....	7
<b>7. Principles</b> .....	10
<b>8. Data Management Framework</b> .....	13
<b>9. Policy Statements</b> .....	16
<b>9.1 Data Governance</b> .....	16
<b>9.2 Metadata Management</b> .....	25
<b>9.3 Data Quality</b> .....	27
<b>9.4 Data Integration</b> .....	29
<b>9.5 Data Security</b> .....	30
<b>9.6 Data Privacy</b> .....	31
<b>9.7 Data Storage</b> .....	32
<b>9.8 Data Usage and Sharing</b> .....	33
<b>10. Recommendations</b> .....	36

## 1. Executive Summary

Data is an essential resource for organizations. The success of an organization is affected by the quality of the data used within its business processes. Effective data management is the key to maximizing the quality of data, allowing the organization to deliver high quality services and increase revenue.

Considering this, the Department of Registration and Stamps would like to understand the context of setting up a Data Management Framework and related recommendations. The goal of this exercise is to first acknowledge that data is a key asset and then to improve both the data management functions and the data stored within the department's various datasets.

### Approach

Following is the approach which has undertaken to define the data policy for the department:

1. Assessment of department's current data landscape as per the key business processes, datasets and applications used
2. Understand current challenges (e.g., ad-hoc reporting, legacy record digitization, etc.) and future aspirations (e.g., monetization and innovation)
3. Define key principles and data management framework comprising of 8 key areas i.e., data governance, metadata management, data quality, data integration, data security, data privacy, data storage and data usage and sharing
4. Outline the policy statements across each domain which serve as standards to implement these domains and framework
5. Key recommendation to be undertaken by the department to implement the framework and address its challenges and achieve its aspirations

### Data Management Framework, Key Focus Areas & Recommendations

The following key focus areas have been identified by the department to be considered to address its challenges and aspirations. These areas have been taken as an input and compose the department's data management framework:

#### Accountability and Data Trust

- Challenge in establishing *accountability on data assets* to be addressed by setting up a robust *data governance* structure to align roles and responsibilities related to data management
- Lack of *data driven collaboration* across desks to be addressed by enhancing the *discovery and consistency* of data across desks and applications by managing the metadata to promote data use across desks
- Ensuring *trust in data* by improving *quality of data* stored across applications by incorporating controls across the data lifecycle

#### Collaboration

- Lack of robust *integration policies and standards* for data driven collaboration by sharing data within and across other departments, applications etc. by setting up a strategic *data integration* layer in the department
- Limited understanding of *classifying datasets* as per security and privacy practices and

regulations including corresponding implications on classified data. This is enabled by defining required classifications and assessing datasets used across desks and incorporating corresponding *data security and privacy* measures.

### **Use, share and monetize**

- Extensive use of *ad-hoc reports* using data from excel sheets to be automated by implementing a *data analytics platform* to automate reporting and analytics use cases for the management
- Explore options for *data monetization and open data sharing* to enhance *revenue generation* and collaboration opportunities respectively
- Assess key data retention and archival requirements as per data usage, sharing and regulatory requirements and implement them across the data lifecycle

### **Key Recommendations**

- Operationalize data management framework by assigning required roles and responsibilities
- Develop Strategy roadmap to implement Metadata design using appropriate technology solution
- Identify CDE (critical data elements) and enhance Data quality as per industry specific standards
- Develop centralized integration platform for data exchange with various internal & external parties
- Enhance data security & privacy by applying standard data classification areas to existing data sets and meet regulatory requirements for sensitive data elements
- Digitize historic data by using OCR (Optical character recognition) technology
- Optimum sharing recommendations using data analytics platform for reporting and analytics
- Data monetization strategy using data partnership plan to come up with flexible schemes
- Data sharing recommendation using open data platform for efficient collaboration, sharing and innovation

The policy statements contained in this document are framed considering the Department's vision and each desk is required to implement these and any implementation required as part of the same should be evaluated as per business needs.

The management teams of each desk are requested to acknowledge that their vision, leadership, and commitment will ultimately decide how effectively their department embraces the aim of these policies, and that this will determine whether they achieve effective management of the data given into their trust. The stewardship of government services is a significant and privileged responsibility. It is a responsibility that can be effectively realized when executives, staff and contractors are committed to data management best practice.

## 2. Introduction

Department of Registration & Stamps provides services for registration of documents, marriage to the citizens through various applications. iSarita is the main application used by the department for providing registration services. To support the iSarita there are multiple applications such as Public Data Entry (PDE), eStepin, eSearch, eRegistration etc. The department started its computerization journey in 2002 with the launch of Sarita (Stamp and Registration Information Technology Application) made by CDAC which was a standalone system. In 2012, a centralized application iSarita developed by NIC was launched. The application has been in use till date. eRegistration application was launched in 2014 for Leave & License agreement and later extended to CIDCO, MHADA, MIDC, Agreement for First Sale, SRA, etc.

Given the nature of services provided by the department, it is imperative for the department to ensure that the data is preserved and made available on demand to various stakeholders. The data from Sarita period (2002) onwards is currently available in digital (Index II) and scanned images format (Registered Documents). With the digitization of documents till 1985, this data is expected to increase further. One of the major challenges for the department is to ensure usability of the legacy data as most of the data is in the form of Scanned images.

Secondly a sizable number of documents are in Marathi language thus further hampering the usability of the same. This is a major roadblock while integrating the applications for other departments such as Income Tax, MCGM, Land records, etc. for data sharing. This is further complicated by lack of API integration policy.

Considering the sensitivity & criticality of the data, ensuring Data privacy and security is another major challenge for the department especially, in absence of any defined policy.

Therefore, it is necessary for the department to define a Data policy which will serve as the guiding light to manage the legacy data, facilitate data sharing with relevant stakeholders, ensure proper storage along with with sufficient measures to ensure data security & privacy.

## 3. Purpose

In its quest to be a smart and digitally empowered department and contribute to the larger Ease of Doing Business agenda of the government, the Department of Registration and Stamps aims to standardize its data practices. Data is a key enabler of the department's operations and its approach to efficient management and governance of data is based on a robust Data Management framework defined in this document and is focused on establishing ownership, stewardship, accountability, and custodianship of its data assets.

One of the key focus areas for the department is to harness the potential of its legacy and future data assets. The document is designed to lay down the key initiatives for data management domains which would lead to enhanced trust and transparency in the data assets of the department

The document is intended to facilitate the adoption of a department level data management framework to direct the department's desks to undertake specific interventions in areas of data management to unlock the potential of data driven opportunities.

The policy statements have been defined to lay down the requirements for the desks for holistic management of their respective data assets. The desks are required to understand the defined statements and effectively apply them to the data assets which are owned by them. The Data Management Function will ensure that the statements are comprehensible to the desks and clarify any ambiguities. The Data Management Function will assist the desks in the implementation of changes required for compliance to the statements.

The statements are intended to support the significant goals of:

- Enhancing the trust in legacy and future data assets across the department
- Maximizing the usage of data assets by ensuring integrity and availability of trust-worthy data across the department
- Unlocking significant monetization opportunities through the data assets
- Enabling data publishing in the open domain to promote innovation in the community
- Enhancing the reputation of the department amongst citizens, ministry, other government departments and vendors by efficiently managing the data across its lifecycle

## 4. Scope

The policy statements in this document will be applicable to:

- 4.1. All offices under the Department of Registration and Stamps
- 4.2. All government bodies working in conjunction or in addition to the department within the Maharashtra state and having the need for data exchange
- 4.3. Partners, vendors and implementation agencies working with the Department of Registration and Stamps

## 5. Approach

A holistic approach has been undertaken in defining the data policy for the department. The key activities and considerations are as follows:

- Understanding the current data landscape of the department outlined in section 6. The landscape considered the key business processes of the department (Figure 1), the data sets required across the lifecycle of the process (Figure 2) and the applications used to manage the datasets across the process lifecycle (Figure 3)
- The current challenges with respect to the data sets such as use of ad-hoc reports for executives, inability to digitize legacy records, etc. were taken into consideration
- Aspirations for monetizing data sets and addressing needs of agencies and other departments were observed

- Based on the understanding of the current state of operations and aspirations, 9 key data management principles were identified and mapped to the required data management domains to address the needs of the department
- Based on the identified domains, a holistic data management framework has been defined for adoption.
- Each of the domains have been further explored as per the needs and best practices to lay down detailed policy guidelines for adoption

## 6. Data Landscape

The Citizen begins his registration journey with the e-Search application for identifying and verifying the transaction history for a particular property. Post this the citizens complete the Public Data entry (PDE) and visit the Sub Registrar Office (SRO) for completing the Registration process. At the SRO then using the iSarita application completed the registration. The iSarita application interacts with the Land Records application to check the details of 7/12/Property Card with details mentioned in the document. Party verification is done using PAN No and AADHAR verification. Post the completion of registration, the data is auto pushed to Department of Land Records and Municipal Corporation of Greater Mumbai for completion of mutation respectively

The figure 1 below shows the various steps involved in the major processes such as Registrations (Physical/ Marriage/ e-Registration), e-Filing, Refund at Department of Registration & Stamps. The figure also shows the activities performed by citizens/ Department officers as well as the interaction with various department and other department applications.

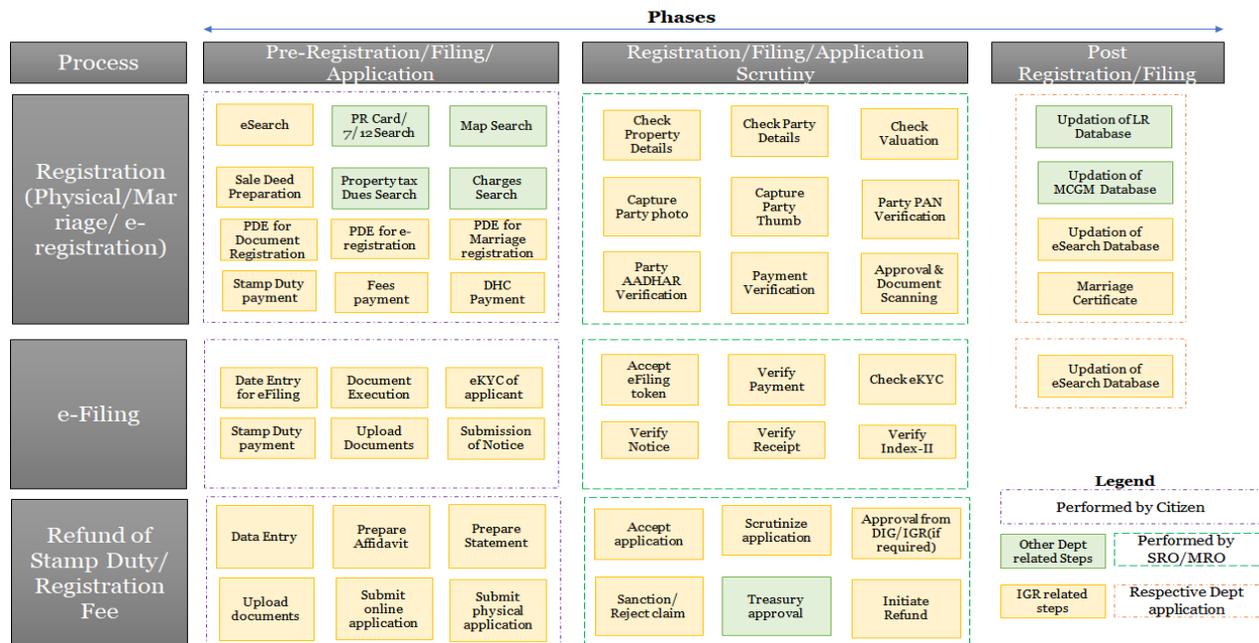


Fig 1: Major Process Flows

Currently the data captured by the department is stored in following two databases:

1. MSSQL
2. PostgreSQL

Sarita period data (2002 to 2012) is stored in MSSQL while iSarita period data (2012 – Present) is in PostgreSQL. Currently NIC is working to convert the data from MSSQL to PostgreSQL.

Department has around 170+ TB of data most of which is in the form of Scanned images. The datasets captured in major activities across the processes are shown in Figure 2. The figure also shows the various stakeholders involved in interacting with the datasets.

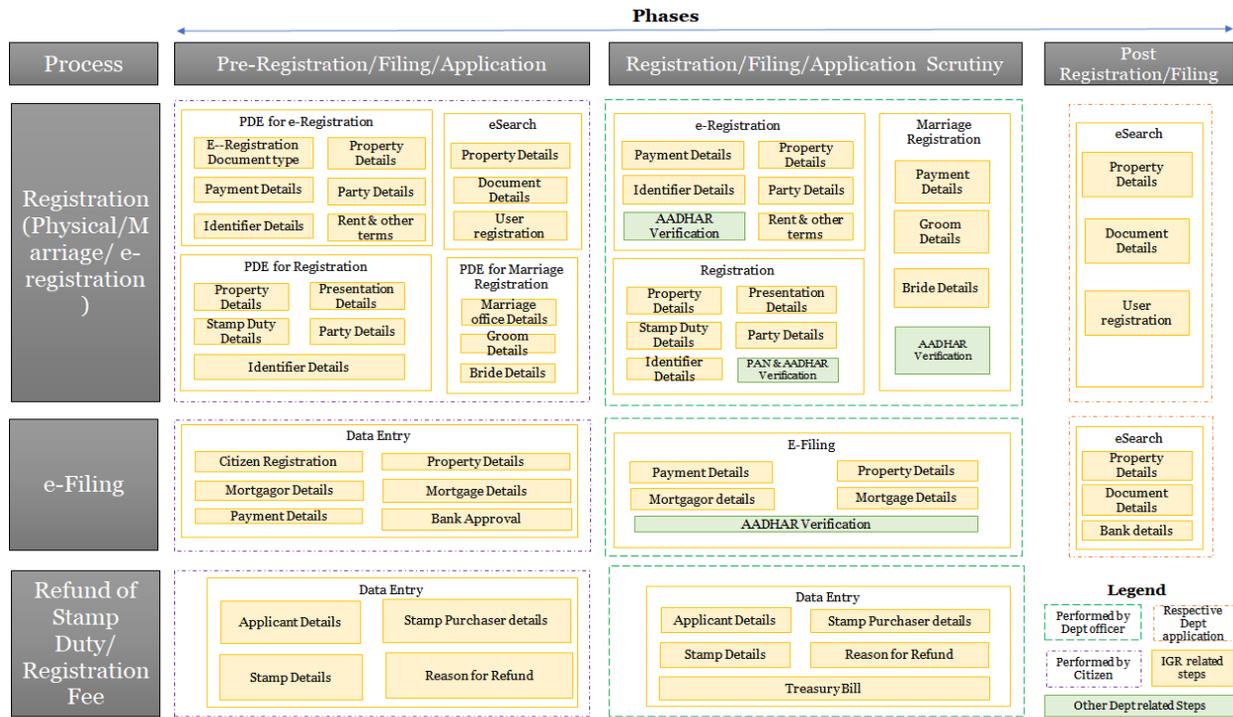


Fig 2: Datasets captured

Currently the department has more than 60 applications which provide various services to citizens. Some of the applications interact with applications from other departments such as Income Tax, UIDAI, Land records, MCGM, GRAS etc for exchange of data during registrations. The Data Flow between various applications of the department is shown in Figure 3.

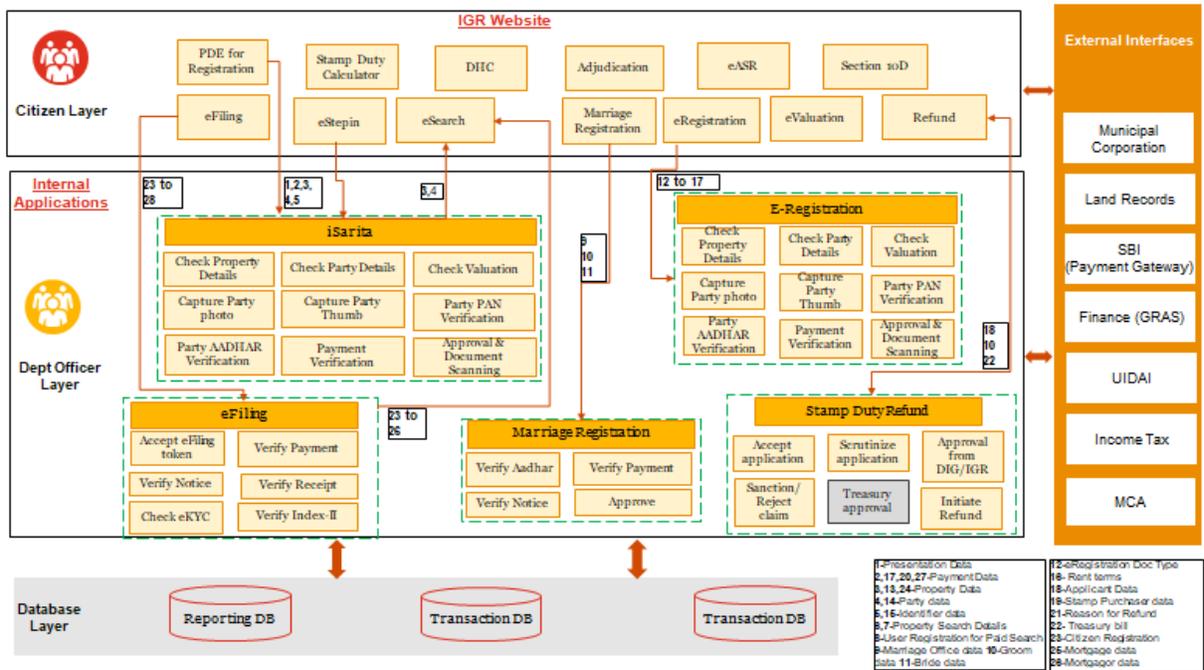


Fig 3: Data Exchange in applications

Currently, the department shares the data based on the requests received from various government departments. The details are given below:

#	Department	Data fields shared	Scope of data	Mode of sharing
1	CID	<ol style="list-style-type: none"> <li>Name of Lessor and Lessee</li> <li>Permanent address of Lessor and Lessee</li> <li>Age of Lessor and Lessee</li> <li>Gender of Lessor and Lessee</li> <li>Details of Property under consideration (for Rent)</li> </ol>	Leave and License registrations across the State	API
2	Pune Police Commissioner	<ol style="list-style-type: none"> <li>Name of Lessor and Lessee</li> <li>Permanent address of Lessor and Lessee</li> <li>Age of Lessor and Lessee</li> <li>Gender of Lessor and Lessee</li> <li>Details of Property under consideration (for Rent)</li> </ol>	Leave and License registrations in Pune City	Access to dedicated portal
3	MCGM	<ol style="list-style-type: none"> <li>Name of Buyer and Seller</li> <li>Permanent address of Buyer and Seller</li> <li>Age of Buyer and Seller</li> <li>Gender of Buyer and Seller</li> </ol>	Property transactions in Mumbai	API

		5. Details of Property under consideration		
4	Land Records	<ol style="list-style-type: none"> <li>1. Name of Buyer and Seller</li> <li>2. Age of Buyer and Seller</li> <li>3. Gender of Buyer and Seller</li> <li>4. Details (7/12 &amp; 8A) of Property under consideration</li> </ol>	All transactions in the State involving transfer of land ownership	API
5	Income Tax	<ol style="list-style-type: none"> <li>1. Name of Buyer and Seller</li> <li>2. Permanent address of Buyer and Seller</li> <li>3. Age of Buyer and Seller</li> <li>4. Gender of Buyer and Seller</li> <li>5. PAN of Buyer and Seller</li> <li>6. Details of Property under consideration</li> <li>7. Consideration amount</li> <li>8. Stamp Duty Paid</li> <li>9. Registration fee</li> </ol>	All transactions in the State with consideration amount greater than Rs. 5 lakhs	Manual
6	Anti-Corruption Bureau	<ol style="list-style-type: none"> <li>1. PAN of Buyer and Seller</li> <li>2. Transactions</li> </ol>	All transactions in the state	Manual

## 7. Principles

The department has laid down the following set of 9 principles which will anchor the development of a holistic data management framework. The principles and the framework will guide the development of specific policy statements across the data management domains. The principles will further help stakeholders in the department to imbibe new ways of working with data to enhance operations and engage in a seamless manner with citizens, other government organizations, third parties and employees.

Principle	Definition	Mapped Data Management Domain(s)
Asset	The department shall treat data as an asset to realize its potential for use and monetization	<ul style="list-style-type: none"> <li>• Data Governance</li> <li>• Data Usage and Sharing</li> </ul>
Owned	All data must have owner(s) and data management responsibilities must be assigned to an employee who is accountable for the data as a part of their responsibilities. Ownership shall be governed through the formation and socialization of the required data councils.	<ul style="list-style-type: none"> <li>• Data Governance</li> </ul>

Described	<p>All data must be defined, described and consistent. Policy, processes, and tools must be in place to support the appropriate level of description of all data used and managed by the department to ensure the same. This information enables:</p> <ul style="list-style-type: none"> <li>● Standardization and consistency of data assets</li> <li>● Increased quality and trust in data</li> <li>● Maximum discoverability and reuse of data</li> <li>● Wider use of data by internal functions, vendors and other government departments</li> <li>● Machine based understanding of what the data stands for, and therefore the development of intelligent solutions such as OCR based scanning agents, image recognition solutions, etc. to streamline operations</li> </ul>	<ul style="list-style-type: none"> <li>● Data Governance</li> <li>● Metadata Management</li> </ul>
Trusted	<p>Trustworthy data is a key to unlocking the full potential of a truly data driven organization. Enhanced trust on the department’s data assets will enable the following:</p> <ul style="list-style-type: none"> <li>● Creation of a trusted and high-quality data ecosystem to improve decision making</li> <li>● Enhanced data sharing amongst government departments (IGR, Land records, Municipal Bodies, RTO, Income Tax, ACB, etc.), citizens, vendors and private entities</li> <li>● Collaboration and transparency amongst the desks (e.g., Registration desk, Stamp desk, Accounts desk, Refund desk, Legal desk, etc.) of the department regarding data assets</li> </ul>	<ul style="list-style-type: none"> <li>● Metadata Management</li> <li>● Data Quality</li> <li>● Data Security and Privacy</li> <li>● Data Usage and Sharing</li> </ul>
Transparent	<p>The department shall be transparent regarding the management of data across its lifecycle from collection to disposition. E.g., Personal data (Name, Address, etc.) collected on iSarita needs to be used and processed as per the consent of the data principal.</p>	<ul style="list-style-type: none"> <li>● Metadata Management</li> <li>● Data Security and Privacy</li> <li>● Data Storage</li> <li>● Data Usage and Sharing</li> </ul>
Access	<p>The principle of access emphasizes on:</p> <ul style="list-style-type: none"> <li>● Classifying data into appropriate categories to identify sensitive data. For Example: Citizen Biometric Data captured</li> </ul>	<ul style="list-style-type: none"> <li>● Data Security &amp; Privacy</li> <li>● Data Storage</li> </ul>

	<p>during registration</p> <ul style="list-style-type: none"> <li>• Maintaining adequate access for individuals across the department to eliminate the risk of exposing sensitive data. e.g., JDR Pune City will have access to data of Pune City only.</li> <li>• Periodic audits to review and maintain access</li> <li>• Consideration of protecting the privacy of information relating to citizens, third parties, property related data etc.</li> <li>• Any data captured from other parties should be collected by consent of citizens, vendors and employees and the consent preferences should be maintained by the department e.g., UIDAI number for eKYC.</li> </ul> <p>Department must ensure that data and information systems are stored/hosted in the Militarized / Demilitarized zones are secure, robust, and resilient.</p> <p>Maintain data lifecycle with emphasis on the duration of data storage, archival &amp; disaster recovery.</p>	
Ethical Data Use	The department should build ethical practices to use data for business purposes to ensure fairness & transparency in delivery of data driven services.	<ul style="list-style-type: none"> <li>• Data Governance</li> <li>• Data Usage and Sharing</li> </ul>
Collaboration	<p>Collaboration is not limited to integration between various systems &amp; applications.</p> <p>It also implies that:</p> <ul style="list-style-type: none"> <li>• While designing the integration data services such as APIs/web-services, the department should strive to keep flexible designs so it could be used with future requirements. Department will have to consider the wide variety of data which will capture various formats</li> <li>• Shared data is reasonable and adheres to consent of citizens, vendors &amp; third parties</li> <li>• Departments should also review the data they are capturing and if it fits the existing purpose</li> </ul>	<ul style="list-style-type: none"> <li>• Data Governance</li> <li>• Data Integration</li> <li>• Data Usage and Sharing</li> </ul>
Culture	The department shall invest in upskilling its employees to adopt new ways of working with data to ensure the emerging trends of data management are adopted while using data.	<ul style="list-style-type: none"> <li>• Data Governance</li> <li>• Data Usage and Sharing</li> </ul>

## 8. Data Management Framework

This section outlines the approach to defining the data management framework, the vision and mission for data management and the Data Management Framework adopted at the Department of Registration and Stamps to achieve the department's data related aspirations.

### Approach

We have undertaken a phased approach towards defining the data management framework for the department as per the inputs from the data landscape assessment outlined in section 6, data related aspirations of the department and emerging trends in the industry. Based on the same, 8 key areas of focus were identified and mapped to respective data management domains in the framework. These focus areas are:

- Establishing **ownership and accountability** across the data lifecycle for each desk's data to enhance trust in data through a robust **data governance** strategy.
- On establishing governance, it is imperative to ensure current and future data sets available in the department are **discoverable, defined as per their context** for business and technical stakeholders and standardized through a **metadata management** program.
- As the data is discovered and standardized through department wide definitions, improvement of **trust in data** is ensured by availability of **high-quality data** from source systems. This is enabled through implementation of **data quality** standards to ensure data is fit for purpose.
- With improved confidence and trust in data assets, the desks shall start **sharing data** across the departments and with external parties to explore opportunities for monetization and innovation where **data integration platform** plays a major role to optimize efforts and costs.
- While implementing the integration platform, it is imperative to consider the **security and privacy implications** across the data landscape to comply with laws and regulations. This is catered through the **data security and privacy** domain.
- Exploring storage options for variety and volumes of data is a key to managing the data assets and is addressed through the **data storage** domain.
- In addition to the adoption of policies across domains, a **strategic data platform** will augment existing efforts for use and sharing of data for **reporting, monetization, and innovation**. Collaboration through the open data platform and setting up of the strategic data platform are addressed in the **data usage and sharing** domain.

Following is the vision and mission for data management aligned with the overall vision and mission of the department.

**Vision for Data Management:** To create a trusted and accountable data ecosystem within the department to enhance citizen service delivery with regards to document registration and revenue collection.

**Mission for Data Management:** To embed a culture of data literacy and collaboration and create an environment where the department is able to use data ethically to create innovative solutions for document registration and revenue generation using the right set of people, policy, process and technology related interventions.

### **Data Management Framework**

As per the outlined approach, the department has adopted the framework outlined in Figure 4 to achieve its vision and mission for data management. The framework comprises 2 key and 3 supporting layers comprising 8 data management domains. The layers of the framework are:

- **Data governance:** The data governance layer overlays the different domains of data management and outlines the accountabilities, committees, roles, and responsibilities of key stakeholders to drive adoption of data management policies, processes, and technologies across the department for ethical and optimal use of data.
- **Data management:** The data management layer comprises the key data management domains which need to be enabled for successful implementation of data driven use cases for the department. The specific interventions required to enable these domains are addressed in the section ‘Policy statements’. The domains addressed are:
  - *Metadata management:* The domain focuses on effective access to high quality integrated metadata across the documents and applications of the department outlined in the Data Landscape section. It is supported by the use of an automated solution acting as reference to the department’s metadata.
  - *Data Quality:* The domain focuses on establishing practices to deliver trustworthy and fit for purpose data to the department by setting up rules across the data lifecycle to govern it.
  - *Data integration:* Data integration involves the collection of data from different sources and consists of solutions fostering a harmonious internal and external communication between the internal and external applications of the department. The domain is a key factor in enabling data sharing and collaboration.
  - *Data Security:* Data Security addresses the requirements of processes, people, and technology required for classification, secured access, authorization to ensure security of data.
  - *Data Privacy:* Addresses the requirements of controls to be established for protecting personal data used by the department
  - *Data storage:* The domain addresses the capability of the department to store multiple data types such as Spatial data, citizen demography, maps, valuation data, etc. in a scalable, reliable and cost-effective manner across cloud and on-premise options.
  - *Data Usage and Sharing:* The domain highlights the controls required for external and internal sharing of data. In addition, it highlights the critical data driven use cases to be undertaken by the department to generate insights from its existing data.
- **Impact:** Built on a tried and tested integrated framework it will help departments and respective desks to improve the overall performance and productivity of the department with less turnaround time for data processing and increasing user trust in data.

- **Operational Support:** To drive the data management initiatives, it is important to make sure it is managed by PMO activities where there is a dedicated team / people to monitor the ongoing progress, align with the business plan and responsible for communications with the end users for user adoption, training etc.
- **Enablers:** The enablers outline the requirements for the right set of skilled people, processes, and technologies across the data management domains to ensure the data program is operational in the long term and addresses the current challenges as well as the emerging needs of the department in the future.

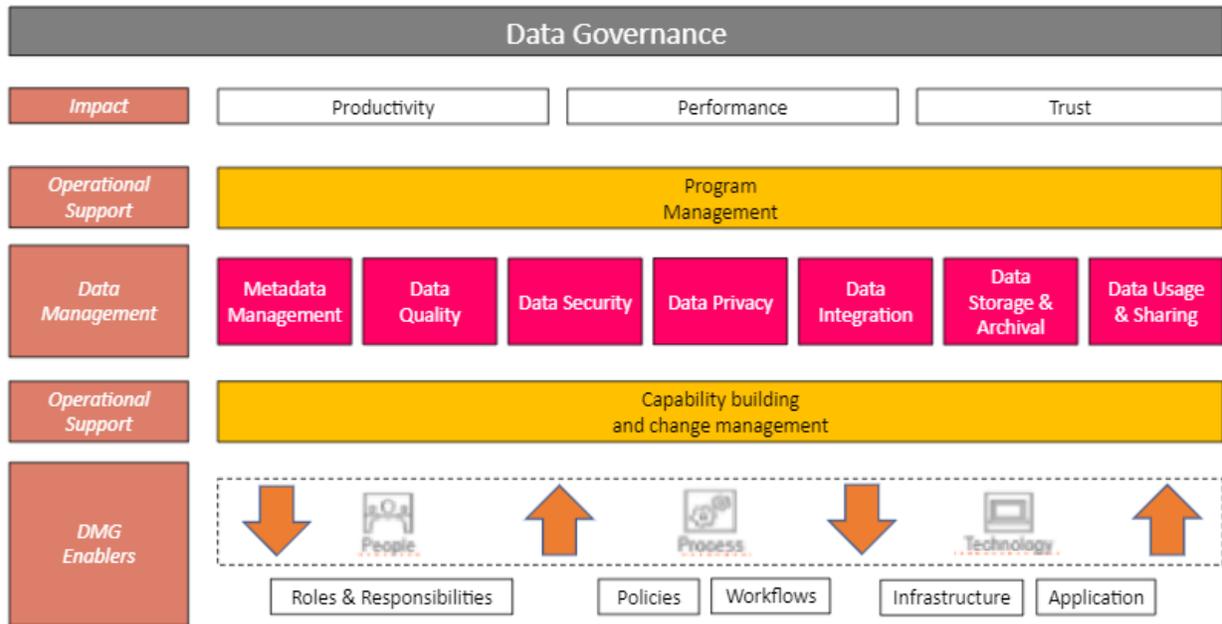


Fig 4: Data Management Framework

## 9. Policy Statements

The section lays down the data policies for the department across the domains of the data management framework outlined in section 6.

### 9.1 Data Governance

The domain outlines the roles and responsibilities, committees, performance management KPIs and the artifacts required to drive the implementation of the data management framework.

#### 9.1.1 Committees

- The department’s Data Management Function shall be a 3 layered team as outlined below:

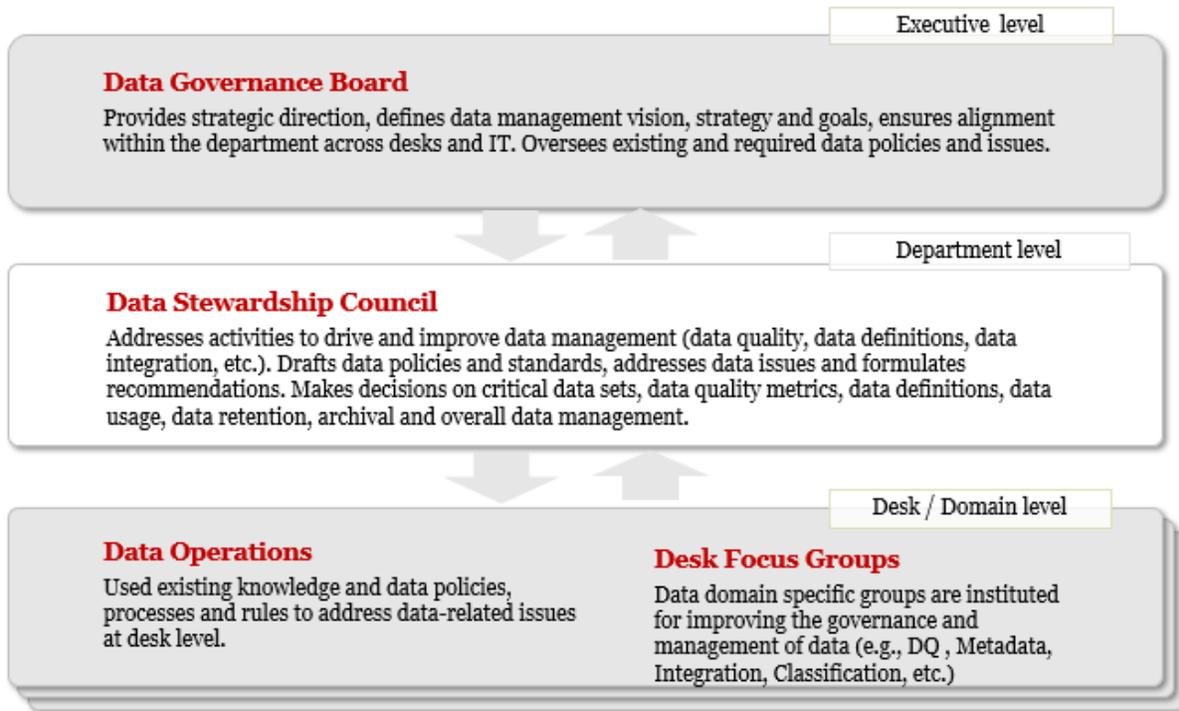
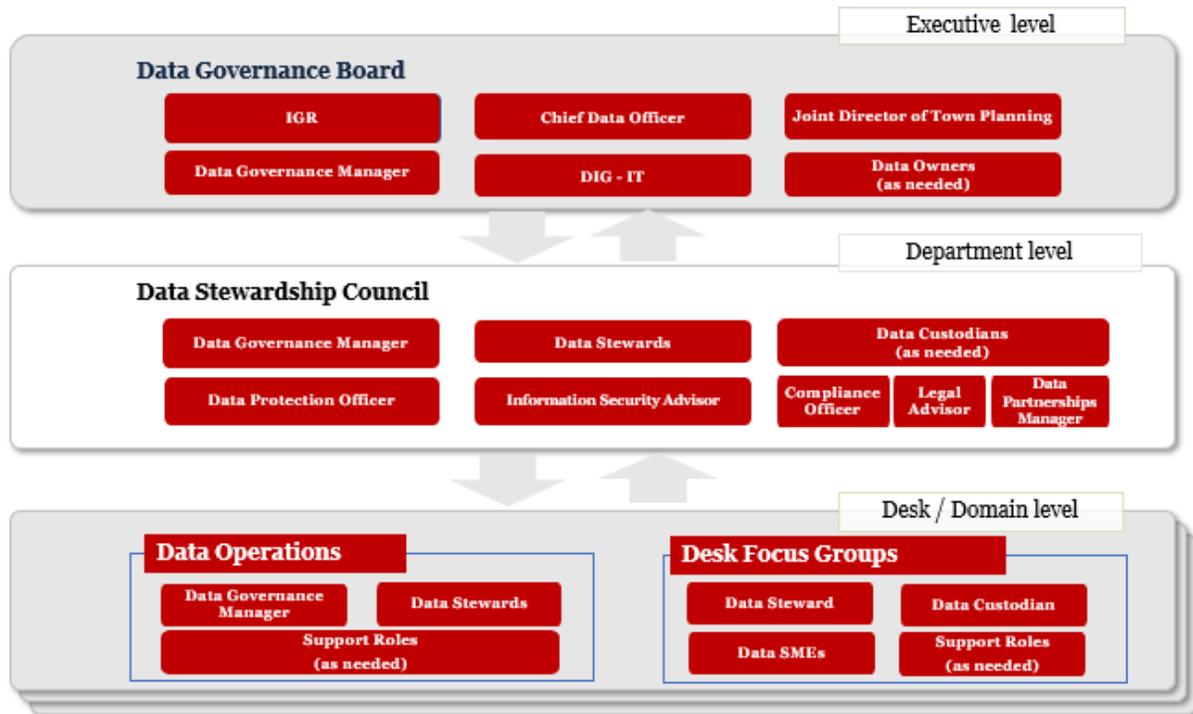


Fig 5: Data Management Function

- Following are the roles to be represented in the committees highlighted above:



Fi. 6: Data Management Committees and Composition

### 9.1.2. Roles and Responsibilities

The Department of Registration and Stamps will need the team for establishing the accountabilities and responsibilities for strategizing and implementing its data management framework. The responsibilities and skills required for delivering these roles are as follows:

#### a) Chief Data Officer

Responsibilities	Role Type: Core
<ul style="list-style-type: none"> <li>• <b>Define data management strategy, framework, and methods</b> by which the department acquires, analyses, governs and manages data related to Property Registration, Marriage Registration etc. in digital and offline modes.</li> <li>• <b>Consult with process/data owners and data stewards</b> for the desks in implementing data policies and framework.</li> <li>• Identify global technology trends in managing processes related to registration and <b>provide insights to the steering/leadership committee.</b></li> <li>• Analyze inputs on existing data issues and <b>challenges</b> (e.g., generating insights from scanned documents) <b>and provide guidance for resolution.</b></li> <li>• Liaise with the Compliance Officer to understand and address non-compliance to data policy.</li> <li>• <b>Reduce costs and redundancies</b> that result from multiple isolated data and technology programs across various desks.</li> </ul>	

### Skills and Qualifications Required

- Strong leadership capabilities with effective stakeholder management and capacity to defend strategic decisions with respect to data management and governance.
- Extensive experience of strategizing and implementing Data Management and Governance related projects.
- Expert domain knowledge of Government Land Departments/ Urban Planning Department/ Real Estate Industry.
- 15+ years of experience working in data management and governance areas
- B.E./B. Tech. in Computer Science/Information Technology/Electronics and Communications with an MBA preferably
- Professionals preferably certified in Data Management Frameworks such as DAMA, DCAM, etc.

### b) Data Governance Manager

#### Responsibilities

Role Type: Core

- **Define and implement data management strategy and framework** across the department.
- **Assess, design, and implement improvements in policies and processes** related to data management domains like data quality, metadata management, data storage, etc. in collaboration with the Data Owners and Data Stewards of desks.
- **Establish and maintain clearly defined data roles and responsibilities** at desk levels.
- **Oversee and collaborate on defining and designing data governance artifacts** (e.g., Issue register), SLAs for data exchanges, data retention, escalation timelines, data quality, etc.
- Collaborate with the Data Partnerships Manager to **explore opportunities for sharing the department's data** publicly for innovation.
- **Conduct training and awareness sessions** for adoption of changes related to data management.

#### Skills and Qualifications Required

- Past experience as project manager as well as experience of handling Data Management/Governance projects
- Effective communication, stakeholder management and interpersonal skills
- Expert knowledge of Government Land Departments/ Urban Planning Department/ Real Estate Industry.
- Proven success in developing and executing strategies, continuous improvement, analytics & advanced technologies to improve core governance values.
- Sound knowledge on Data Governance tools and technologies.
- 10+ years of experience working in data management and governance areas
- B.E./B. Tech. in Computer Science/Information Technology/Electronics and Communications with an MBA preferably
- Professionals preferably certified in Data Management Frameworks such as DAMA, DCAM, etc. and Project Management (e.g., Prince2, PMP, etc.)

### c) Data Owner

Responsibilities	Role Type: Core
<ul style="list-style-type: none"> <li>Accountable for <b>strategizing the management of their desk specific domain data</b> and consulting the Data Governance Manager on a need basis.</li> <li>Responsible for ensuring the <b>implementation of data policies and processes for their data domain</b>.</li> <li>Responsible for <b>optimizing the storage and use of data</b> and generating new opportunities for <b>generating insights out of the data</b>.</li> <li><b>Approves</b> domain/desk specific data projects, data requirements and data definitions /business terms.</li> <li><b>Authorizes access to data</b> for the domain (based on access controls established)</li> <li><b>Approves datasets or insights</b> from the data domain to be published as <b>Open data</b> using methods such as OpenAPIs, File dumps, etc.</li> </ul>	
Skills and Qualifications Required	
<ul style="list-style-type: none"> <li>Strong communication skills to liaise with department’s stakeholders for implementation of data policies and processes.</li> <li>Understanding of the business processes of the department and desks and their implications on data.</li> <li>15+ years of experience working in government departments preferably in smart cities, urban development, registrations, etc.</li> </ul>	

### d) Data Steward

Responsibilities	Role Type: Core
<ul style="list-style-type: none"> <li>Responsible for <b>gathering data requirements</b> from desks, improving data quality, definitions of business metadata, data classification and exploring avenues for generating insights from data.</li> <li><b>Defining data policies</b> in collaboration with the Data Governance Manager and implementing them for the desk.</li> <li>Ensure that corrective <b>actions to address data quality deviations</b> for key desk data are defined and implemented.</li> <li><b>Drive and coordinate data-related changes</b> for the desk.</li> <li><b>Reporting</b> data-related requests and issues <b>to the Data Stewardship Committee</b>.</li> <li>Collaborate with Data Custodians to <b>ensure that requirements are properly implemented across all applications</b> hosting the desk’s data.</li> <li><b>Design and maintain dashboards and performance indicators</b> to monitor progresses towards target state for the desk’s data.</li> </ul>	
Skills and Qualifications Required	
<ul style="list-style-type: none"> <li>Strong communication skills to liaise with business and technical stakeholders for implementation of data policies and processes.</li> <li>Understanding of the business processes of the department and desks.</li> <li>Experience of working with the department and understanding of functional and process requirements and implications on related applications and data sets and defining related</li> </ul>	

- policies for data management.
- 10+ years of experience working in data management and governance areas
  - B.E./B. Tech. in Computer Science/Information Technology/Electronics and Communications
  - Professionals preferably certified in Data Management Frameworks such as DAMA, DCAM, etc. and Project Management (e.g., Prince2, PMP, etc.)

### e) Data Custodian

Responsibilities	Role Type: Support
<ul style="list-style-type: none"> <li>• Develop an <b>understanding of how data is stored, processed, and transmitted</b> within the department and externally.</li> <li>• <b>Implement applicable data policies</b> (e.g., data retention periods, technical metadata maintenance, etc.) for respective applications as per the guidance from data stewards.</li> <li>• <b>Explore cost effective data storage options</b> for applications to address growing needs of the departments to digitize legacy records.</li> <li>• Responsible for <b>adherence to data integration SLAs for data exchange</b> with other applications.</li> <li>• Responsible for <b>conducting root cause analysis for application-level issues</b>, and remediating data related escalations related to respective applications.</li> </ul>	
Skills and Qualifications Required	
<ul style="list-style-type: none"> <li>• Strong communication skills to liaise with business and technical stakeholders for implementation of data policies and processes.</li> <li>• Understanding of the business processes of the department and desks.</li> <li>• Experience of working in complex technical environments and leading technical experts across software development lifecycles preferably for government entities.</li> <li>• 10+ years of experience working in data management and governance areas</li> <li>• B.E./B. Tech. in Computer Science/Information Technology/Electronics and Communications</li> </ul>	

### f) Data Partnerships Manager

Responsibilities	Role Type: Support
<ul style="list-style-type: none"> <li>• <b>Define data monetization use cases</b> for the department in collaboration with the data stewards and owners.</li> <li>• Explore the Indian data marketplace to <b>identify opportunities to establish data sharing partnerships</b> for the department to monetize data.</li> <li>• <b>Define pricing models</b> for data monetization.</li> <li>• Consult the data owners and data stewards on <b>potential open data initiatives and data monetization opportunities</b>.</li> <li>• Collaborate with the data owners and data stewards to <b>formulate and agree on the pricing models for data monetization</b>.</li> <li>• <b>Define and monitor the KPIs</b> for data partnerships and monetization.</li> <li>• Liaise with the National, State and City Level Open data platform teams to <b>publish</b></li> </ul>	

**datasets from the department.**

### **Skills and Qualifications Required**

- Experience with government agencies, authorities, regulators and private organizations leading data partnership teams.
- Successful track record of identifying and implementing data monetization opportunities by influencing executive level stakeholders.
- Understanding of implications of the regulations on data partnerships and monetization opportunities.
- Proven experience in managing complex negotiations with key stakeholders.
- 10+ years of experience working in partnerships and strategic alliances preferably in government sectors
- B.E./B. Tech. in Computer Science/Information Technology/Electronics and Communications with an MBA preferably

### **g) Data Protection Officer (DPO)**

#### **Responsibilities**

**Role Type: Support**

- Oversee **applicability of data privacy laws and delivery of department's data privacy initiatives** (e.g., consent management, rights management, personal data classification, data sharing, etc.) and provide **guidance on best practices to the Data Management Function.**
- **Track changes to privacy regulations** and advise the desks on the same.
- **Conduct Data Protection Impact Assessments** periodically to identify gaps in data privacy policies and practices.
- **Lead dispute settlements in courts or outside courts** with parties who have filed legal suits related to data privacy against the department.
- Report any non-compliance to Privacy Regulations and **identify the root cause to address the gaps.**
- **Act as decision making authority** for various privacy templates and SLAs.
- **Address grievances and requests of the data principals** as per the guidelines of the law.
- **Organize training sessions** for staff on data privacy practices.

#### **Skills and Qualifications Required**

- Strong leadership capabilities with effective stakeholder management and capacity to defend strategic decisions with respect to data privacy
- Extensive experience of leading Data Privacy related projects for compliance to regulations like GDPR, CCPA, POPI, etc. as a Data Protection Officer in the past
- Knowledge of Indian PDP Bill, 2019 preferred
- 15+ years of experience working in data privacy and regulations
- Candidates with Degree/Diploma/Certification in Cyber Law to be preferred

**h) Information Security Advisor**

Responsibilities	Role Type: Support
<ul style="list-style-type: none"> <li>• Understand best practices followed and <b>design and maintain policies, methodologies, guidelines for data security.</b></li> <li>• Define the processes to be followed to <b>ensure adherence to security requirements, mode of tracking, templates, escalation matrices,</b> etc.</li> <li>• <b>Advise the Data Management function</b> on data security issues and resolution mechanisms.</li> <li>• <b>Consult the data owners and data stewards</b> in data classification exercises.</li> <li>• <b>Escalate, track and document issues / requests</b> related to data security and responsible for resolving and remediating data security escalations in collaboration with required people.</li> </ul>	
Skills and Qualifications Required	
<ul style="list-style-type: none"> <li>• Individual with sound knowledge of data security and cyber security</li> <li>• Experience in designing data security strategy and security architecture</li> <li>• Expert in tools or systems which provide security control</li> <li>• Strong communication skills to liaise with stakeholders for implementation of security protocols.</li> <li>• 10+ years of experience working in data security domain preferably in government sectors</li> <li>• B.E./B. Tech. in Computer Science/Information Technology/Electronics and Communications</li> </ul>	

**i) Compliance Officer**

Responsibilities	Role Type: Support
<ul style="list-style-type: none"> <li>• <b>Conduct periodic data audits</b> against the defined data policies and document non-compliances across desks.</li> <li>• <b>Calculate Compliance Audit Score</b> and share it with the Data Function.</li> <li>• <b>Maintain the department's risk register</b> with data related risks.</li> </ul>	
Skills and Qualifications Required	
<ul style="list-style-type: none"> <li>• Experience conducting IT / data audits in the government sector</li> <li>• Strong understanding of IT and data landscape including policies and processes for data management and governance</li> <li>• 10+ years of experience conducting IT audits preferably in government sectors</li> <li>• B.E./B. Tech. in Computer Science/Information Technology/Electronics and Communications with an MBA preferably</li> <li>• Certifications such as CISA/CDCAP preferred</li> </ul>	

**j) Legal Advisor**

Responsibilities	Role Type: Support
<ul style="list-style-type: none"> <li>● <b>Assess the applicability of changes to laws/bills for data related implications</b> on the department and advise on key actions to be undertaken. Some of these laws/bills are:               <ul style="list-style-type: none"> <li>○ The Registration Act, 1908</li> <li>○ The Maharashtra Stamps Act, 1958</li> <li>○ The Special Marriage Act, 1954</li> <li>○ The Indian Personal Data Protection Bill, 2019</li> </ul> </li> </ul>	
Skills and Qualifications Required	
<ul style="list-style-type: none"> <li>● Extensive experience and understanding of Indian Laws and their implications on the department's data.</li> <li>● Experience in managing regulatory compliance asks preferably in IT/Data Security/Data Privacy space</li> <li>● In-depth knowledge of department's business processes and a part of department's legal desk</li> <li>● Experience in liaising with advocates for legal disputes</li> <li>● 15+ years of experience as legal advisor in government sectors</li> <li>● B.A. LLB/ LLB graduate preferably with an LLM</li> </ul>	

The key roles identified for the interim council are as following:

#	Role	Officer & Designation	Involvement	Hire / Align
1	Executive Chair	Shri. Shravan Hardikar (IAS), Inspector General of Registration & Controller of Stamps	Part-time	Align
2	Chief Data Officer	Shri. Suhas Mapari, DIG-IT	Part-time	Align
3	Data Governance Manager	To be hired	Full-time	Hire
4	Data Protection Officer & Legal Advisor	Shri. Kale, Desk Office 14	Part-time	Align

5	Data Owners	Shri. Bharat Garud, Desk Office 4 (Registration) Shri. Sanjeev Deshmukh, Desk Office 5 (Stamps) Shri. Khomane, Desk Office 6 (Audit) Shri. Sanjeev Deshmukh, Desk Office 12 (Refund) Shri. Abhijeet Ketkar, Desk Office 15 (Valuation)	Part-time	Align
6	Data SME	Shri. Santosh Hingane, JDR (DIG - Pune Office)	Part-time	Align
7	Data Stewards	5 data stewards to be hired and aligned to Desk Offices for Registration, Stamps, Audit, Refund and Valuation	Full-time	Hire

### 9.1.3. Performance Management

- The department shall track and monitor the following performance management KPIs to gather statistics on its data management program:
  - Assignment of the Data Governance roles across the department
  - Number of periodic Data Management Function Meetings completed
  - Development / revision of the department’s Data Policy
  - Data related trainings and awareness sessions completed
  - Number of participants in training and awareness Sessions
  - Number of non-compliances against the Data Policy during compliance audits
  - Reduction in Turnaround Time (TAT) to resolve data related issues reported to the Data Management Function
  - Number of change requests resolved and closed
- Improvement programs must be undertaken in collaboration with the desks and IT teams to ensure the above KPIs are achieved
- The department shall engage the stakeholders (re-aligned internally or hired/contracted externally) identified in the roles mentioned in 7.1.2. on a comprehensive change management program comprising of trainings and other upskilling initiatives to enable them deliver on their responsibilities related to data management and governance.

### 9.1.4. Artifacts

- The department shall maintain a ‘Data Governance Approval Register’ which will serve as a record of all decisions with justifications made by the CDO for the program
- The department shall maintain a ‘Data Management Issue Tracker’ to capture record of all issues, their resolution methodology and status related to data management

- The artifacts shall be version controlled by the Data Governance Manager

## **9.2 Metadata Management**

### **9.2.1. Planning**

- The department shall create a metadata management plan and roadmap to create and manage a data catalog for its data assets across the desks and applications. The catalog will help in improving the consistency, trust, literacy, and usage of data assets across the department.
- The department shall prioritize data sources to be included in the data catalog, along with the definition of their business and technical metadata. E.g., iSARITA and eRegistration being a major application needs to be catalogued in the first phase.

### **9.2.2. Attributes and Content**

- Data stewards and data custodians should define the required metadata values for each data element for their desks and applications respectively and including the following:
  - Criticality of the data elements to the business
  - The business, technical and operational metadata for the identified critical data elements (CDE)
- The data stewards will collaborate with the data SMEs in their desks to collect / update / delete the business metadata of the business terms owned by the desk:
  - Data owner, data steward, data custodian
  - Definition of the business term
  - Criticality
  - Associated policies for access, use and retention
  - Security and privacy classifications
  - Data quality business rules
  - Data quality metrics (E.g., # of null values, # of invalid records, etc.)
- The data custodians will be responsible to collect/update/delete the technical and operational metadata of the identified CDEs:
  - Application name
  - Database name
  - Table name
  - Column name
  - Data type
  - Primary key
  - Foreign key
  - Load Statistics
  - Change Log
- The desks, in collaboration with the Data Management Function, will establish data ownership, stewardship and custodianship to maintain the metadata for respective data assets.
- The data stewards and custodians will ensure that business, technical and operational metadata are version controlled using specific attributes to identify the version that they were captured against.

### 9.2.3. Solution Features

- The Data Management Function will adopt a data catalog solution that meets the following key requirements:
  - Data inventory: Registry of datasets by its owners considering the criticality of the datasets, number of users of the dataset and the complexity of the dataset
  - Business glossary: Registry of business terms for data objects, attributes, relationships, and values with contextual meanings. E.g., a data object such as 'party name' should have a single definition and consistently be used across the department.
  - Data dictionary: Registry of business, technical and operational metadata with the rules and semantic relations amongst them
  - Storage: All the business, technical and operational metadata derived from business glossary and data dictionary registered in the data catalog should be stored in a secure environment
  - Metadata repository: To store and provide access to the harvested metadata
  - Browsing portal - To provide controlled access to metadata to enterprise users with capabilities of search to browse data assets
  - Workflow management - To implement processes for lifecycle management of metadata
  - Version control - To capture versioning of metadata for audit purposes
  - Data lineage - To capture semantic and technical source to target maps of data elements and capability to track and visualize it for users
  - Impact Analysis - To identify the impact of change to the data across upstream and downstream data operations
  - Metadata Exploration - To enable users search, browse and retrieve metadata from the data dictionary seamlessly
  - Automated metadata capture - To connect to applications using a wide range of connectors to harvest technical metadata automatically

### 9.2.4. Processes

- A standard metadata creation process shall be defined and adopted for capturing metadata of new data assets as per the data governance operating model.
- A standard issue management process shall be defined and adopted for registering / escalating / resolving / remediating issues related to metadata as per the data governance operating model.
- A standard metadata updation process shall be defined and adopted for capturing changes related to metadata as per the data governance operating model.

### 9.2.5. Performance Management

- The department shall establish key performance indicators (KPIs) to measure quality of its metadata. KPIs shall include, at minimum, the following:
  - Completeness (degree to which business glossaries and data dictionaries are completed)
  - Accuracy (degree to which definitions and descriptions align to desk's context)

- Consistency (degree to which definitions of metadata are consistent across the department).
- The data stewards will collaborate with Data Management Function and data custodians to define additional metrics, as needed, to track the quality of metadata.
- Data stewards will be responsible for developing and publishing the reports based on the metrics to the data owners.

## **9.3 Data Quality**

### **9.3.1. Critical Data Elements**

- All the data elements used by the department will be assigned a Data Criticality Level by respective data stewards for their desk's data assets using the following evaluation criteria:
  - Highly Critical Data: Catastrophic to severe impact if data is unavailable or not fit for purpose where (a) activities need to resume as soon as possible; (b) High quality data is not maintained. E.g., Party details, Property details
  - Critical Data: Important impact if data is unavailable or not fit for purpose where (a) business activities can be postponed for no more than a few days; (b) Alternate approach needs to be considered to limit the impact on the operations. E.g., Payment details
  - Medium Critical Data: Moderate impact if data is unavailable or not fit for purpose where business activities can be postponed for a longer period. E.g., Stamp details in refund process
  - Non-Critical Data: Low or no material impact if (a) data is unavailable; (b) data quality is not met. E.g. Document details
- Critical Data Element (CDEs) are data elements with a Data Criticality Level of Highly Critical Data / Critical Data

### **9.3.2. Data Quality Dimensions**

- The data stewards will define data quality rules based on desk specific operations in consultation with the data custodians. Rules are developed using a data quality business rule / dimension framework. Rules can be created for 1 or more data quality dimensions. These rules will be across the following data quality dimensions:
  - Completeness - Extent to which required data must be populated.
  - Uniqueness - Uniqueness means nothing will be recorded more than once based upon how that thing is identified; uniqueness is a measure of unwanted duplicated records within or across systems for a particular field, or data set. E.g., Party Details, Property Details, etc.
  - Validity - Describing the degree to which data conforms to acceptable business content. This can include: Format, Pattern or Data type. E.g., Postal code should comprise a 6-character string following XXXXXX, with X representing a number.
  - Accuracy - Describing the acceptable margin of error against reality to support the intended purpose(s) of the data.
  - Consistency - Describing the degree of consistency between different data points and conformity to reference data. E.g., If the party title is "Mr." the gender must be "male"

- Reasonableness: Describing the degree to which data conforms to expected reasonable ranges for certain values and dates. E.g., The age of a citizen must be between X and Y
- Timeliness - Describing the acceptable latency between data capture and data usage.
- For each of these dimensions, the desk data steward will:
  - Assess the impact of data quality and validate whether the data is fit-for-purpose for the associated business processes of the desk
  - Determine if there would be a business benefit as the quality of data increases and accordingly decide the course of action for the data element
  - Define the thresholds for data quality rules against each dimension
- The critical data elements will be profiled to determine the metrics against data quality dimensions such as:
  - Completeness
  - Validity
  - Uniqueness
  - Accuracy
  - Timeliness
  - Consistency
  - Reasonableness

### **9.3.3. Processes**

- The data governance manager shall define the following processes for data quality management:
  - Critical Data Identification process
  - Data Quality Issue Logging process
  - Data Quality Remediation process

### **9.3.4. Performance Management**

- The following data quality metrics should be consolidated by the data governance management and published to the data owners and data stewards periodically in the form of dashboards. These metrics are:
    - # of CDEs identified
    - # of CDEs having business rules
    - # of CDEs having thresholds established
    - # of DQ issues identified, triaged and resolved on time
    - % of DQ issues not addressed per SLA expectations
- The dashboards for at least 24 months should be conserved to measure progress
- Data quality business rules and metrics will be reviewed and updated at least annually or on an adhoc basis as needed by data stewards to enhance the quality of data assets.

## 9.4 Data Integration

### 9.4.1. Data Integration Layer

- The department shall design and implement a central integration layer to connect internal and external applications (e.g., iSARITA, e-Search, GRAS, DMA, Land Records etc.) to enable the exchange of data sets and documents required for its business processes. The layer should ensure/facilitate:
  - Data exchange – mapping and physically moving data from one application to another
  - Data exchange quality - delivery of datasets as per the data quality policies in section 7.3
  - Detect data delivery failure
  - Access auditing – logging users, services and requests
  - Performance monitoring – monitoring data volumes, frequency and availability
  - Security controls – ensuring controlled access to data

A service bus can be considered as an option for implementing the layer.

- The department shall consider and incrementally plan to migrate its internal and external data exchanges into and out of its applications through layer. The data stewards in collaboration with the data custodians shall create a plan to achieve the same. The high priority applications to be migrated are iSarita, eRegistration, etc.
- The department shall plan and avoid use of point-to-point integration methods to exchange data within and outside the department.
- The department shall prefer using the following methods for data exchange:
  - File based data exchange – transferring a data file to a central physical location, where it may be processed, validated, and transformed before being collected by the receiving application
  - Message based data exchange – exchanging information through formatted messages via a service bus, typically in a publisher/subscriber model or broadcast model
  - Database to database data exchange – typically used with ETL/ELT systems, data may be passed through an intermediary database for transformation and validation before routing to its final destination
- The department shall consider using the following formats for data exchange:
  - Value separated data formats – such as CSV and Tab-delimited files
  - XML and JSON data formats

XML and JSON data formats shall be the preferred mechanism for data transfer between Entities.

### 9.4.2. Service Level Agreement

- The department shall create binding service-level agreements on data shared and exchanged internally and externally. The agreement shall comprise of the following:
  - Data purpose: the overall intent for why the data should be shared and exchanged

- Data frequency: agreed duration and how often the data will be shared and exchanged
- Data quality (See section 7.3)
- Data volume: the amount of data each party commits to sending and receiving
- Availability of service: planned uptime, or service availability windows
- Variety of data: the structure of the dataset, including data model and definitions
- Change control process: the mechanism of informing data consumers of changes to the underlying data sets or data formats
- Exception escalation path: the mechanism for investigating data errors, service outages, and exceptions to the SLA
- SLA monitoring frequency: the frequency at which the service level shall be measured

The Data Management Function shall triage and address any issues when the agreement is not adhered by the parties.

## 9.5 Data Security

### 9.5.1 Data Classification

Information Security Advisor is responsible for clear management and ownership of data security to comply with internal and regulatory requirements (e.g.: safeguarding and controlling the sharing of customer sensitive information with other parties)

- Scope of Data:
  - Desks are collecting sensitive information from customers e.g., Aadhar Card, Pan card which is being used for processing application
  - To protect data, the first step is Data Classification and identifying datasets
- Data Classification Categories:
  - Classification levels assigned to identified datasets and artifacts based on the impact assessment of the potential damage.
  - The table illustrates the data classification and impact level for each supported by descriptions and rationale for the impact

#### *Data Classification Definitions*

#	Data Classification	Impact Level	Description
1	Public / Open	None / Insignificant	<p>Any data that is intended to be disclosed to the public. Such information is defined as having no local, national or international restrictions on access and usage.</p> <p>Annual Statement of Rates (ASR) which is current published by the department on yearly basis is an example for open / public data</p> <p>Data on No of Documents registered, and revenue earned across state</p>

2	Confidential	Low	Data that must be afforded limited confidentiality protection due to its use in the day-to-day operations. Disclosure of such data could have limited adverse impact on the functioning or reputation of the department. Such data relates to the internal functioning of the desk services and will not have general relevance and applicability to external parties or marketing websites
3	Sensitive	Medium	Data that requires robust protection due to its critical support to decision-making within the department. Data that could disclose deals, sensitive pricing data or vulnerabilities exploitable by those with malicious intent.
4	Secret	High	Data that requires substantial and multilevel protection due to its highly sensitive nature. Disclosure of such data could have a serious and sustained impact on an individual's security, compromise regulations etc.

- Ownership:
  - Data owners should be assigned from a security perspective with defined roles and responsibilities to control the users accessing the data
  - Departments should develop a RACI Matrix detailing access rights for individuals
- The department shall define access rights to classified data and provision access to user groups accordingly.

### 9.6 Data Privacy

- The department shall conduct an assessment to determine the compliance gaps in its personal data management practices. This shall be followed by a strategy to remediate the identified gaps.
- The DPO in collaboration with the data stewards and custodians shall undertake a classification exercise to see if the data is sensitive or critical. Following are the classifications as per the Bill:

#	Data Classification	Description
1	Personally Identifiable Information	Any data which can be used to uniquely identify a person
2	Sensitive Personal Information (SPI)	This classification comprises of the following types of data: <ul style="list-style-type: none"> <li>● Financial data</li> <li>● Health data</li> <li>● Official identifier</li> <li>● Sex life</li> <li>● Sexual orientation</li> <li>● Biometric data</li> <li>● Genetic data</li> </ul>

		<ul style="list-style-type: none"> <li>● Transgender status</li> <li>● Intersex status</li> <li>● Caste or tribe</li> <li>● Religious or political belief or affiliation</li> </ul>
3	Critical Personal Information (CPI)	To be notified by the government in the future

- Data classified as SPI can be stored and processed outside India, but a copy of the information needs to be stored in India
- Data classified as CPI needs to be stored and processed in India only
- The DPO shall be responsible to define a ‘Privacy by Design’ policy and get it published on its official channels.
- The department shall establish a lawful purpose for collection, processing, storage, sharing and retention of personal data.
- The department shall collect minimal personal data required for its business operations.
- The department shall retain personal data only for the time period it is required for.
- The department shall institute a breach reporting procedure for personal data related breaches and notify impacted citizens on any such occurrence as soon as feasible.
- The department shall de-identify or anonymize personal data while using it for business processes unless and until the data is required as-is for the process to be completed.
- The department shall develop a digital solution to empower citizens with the following rights on their personal data:
  - Confirmation of processing of their personal data
  - Access to the current state of the personal data
  - Correction of the any discrepancies in the personal data
  - Erasure of any personal data records
  - Portability of personal data records to other departments or organizations
- The department shall maintain records of all the processing activities it undertakes on personal data.
- The department shall enable citizens access their personal data in its current format and provide confirmation on the processing activities undertaken on personal data.

## 9.7 Data Storage

### 9.7.1. Options for Storage

- The Data Management Function shall conduct a baseline assessment of the existing data storage options and infrastructure used by the applications.
- The data stewards and custodians shall define the future storage requirements for their respective business processes and applications.
- The department shall prefer using cloud storage options (e.g., the Government Community Cloud) for its existing and future applications to harness the operational benefits of cloud.
- The department shall explore blockchain based solutions for data storage
- The department shall establish a data backup and disaster recovery plan for its applications to ensure business continuity

### 9.7.2. Data Migration

- The data stewards and data custodians shall create a strategic data migration plan to consolidate the databases used by their desk's applications. E.g., the databases for iSARITA i.e., MSSQL and PostGRESQL shall be consolidated and migrated to PostGRESQL to enhance performance of the application and improve the user experience.

### 9.7.3. Data Retention and Archival

- The data stewards shall define the retention and archival periods of data owned by their desks based on its usage, type, classification, value and regulatory requirements. E.g., The data related to registered documents; Party details needs to be retained for perpetuity.

The data stewards shall define the disposal and destruction rules based on the data type and classification for their data assets.

## 9.8 Data Usage and Sharing

### 9.8.1. Data Analytics Platform

- The department shall develop a central data analytics platform to cater to its analytics and reporting requirements and improve the usage of its data assets.
- The platform shall capture data from all the internal applications of the department and required external sources of the department and aggregate it to generate automated dashboards for the department and desks. This shall lead to improving the efficiency of decision making and reduction of manual reporting scenarios prevalent in the department currently. E.g., Department would be able to track the no of documents registered and time required to complete the registration by SRO and thus able to identify top performing SRO
- The dashboards, reports and analytics use cases shall be discussed, agreed, and prioritized with the end users by the respective desk level data stewards in collaboration with the members of the analytics platform team. The initial set of reports to be implemented are:

#	Report Type	Report description
1	Article-wise report	<ul style="list-style-type: none"> <li>• Report about article wise documents registered, stamp duty and registration fees paid</li> <li>• Report on document registration and revenue under Top 10 articles</li> </ul>
2	Monthly collection report	<ul style="list-style-type: none"> <li>• Monthly document registration and revenue collection report</li> <li>• Report about the month-on-month revenue collection against the target revenue of the department</li> <li>• Report on month-on-month document registration by the the department</li> <li>• Report about division wise collection and document registration</li> </ul>
3	Office wise report	<ul style="list-style-type: none"> <li>• Periodic Report about the office-wise documents registered, stamp duty paid, registration fees and DHC collected</li> <li>• Office wise reports on document registrations with higher value (SFT reports)</li> </ul>
4	Comparative report	<ul style="list-style-type: none"> <li>• Article wise comparative about document registration and revenue collection (year on year)</li> </ul>

		<ul style="list-style-type: none"> <li>Year-on-year trend of total document registration and revenue collection</li> </ul>
5	Payment mode report	<ul style="list-style-type: none"> <li>Report on collection of revenue through various payment modes</li> <li>Office wise or Article wise report on DHC collection</li> </ul>
6	ASR (Annual Statement of Rates) reports	<ul style="list-style-type: none"> <li>Analytic Reports about the trend of property sale transaction values with reference to ASR in each zone to arrive at new ASR</li> </ul>
7	Income Tax Department Reports	<ul style="list-style-type: none"> <li>Reports submitted to IT department against enquiries</li> </ul>
8	Ad-Hoc Reports	<ul style="list-style-type: none"> <li>Analytic reports about document registration trend, revenue collection, modes of registration, etc.</li> <li>Report documents registered using PDE and eStep-In</li> <li>Analytic report on office-wise average time for document registration</li> </ul>

- The department shall prefer a cloud native architecture for developing the platform considering its benefits over on-premises implementations.
- The data governance manager shall define the following processes to manage the platform:
  - Analytics request management process
  - Change management process
  - Issue management process
- The data governance manager shall maintain an ‘Analytics Register’ comprising all the requested and deployed use cases on the platform.
- The department shall undertake a change management plan to train its staff on using the platform.

### 9.8.2. Data Monetization

- The data partnerships manager in collaboration with the Chief Data Officer shall develop a data monetization plan to generate revenue out of its data assets.
- The desk data owners and stewards in collaboration with the data partnerships shall develop use cases to monetize their respective data assets. One such use case is monetization of historical property sale transactions data set from e-Search with real estate aggregators such as 99acres, magicbricks, etc. on a consumption-based model.
- The data partnerships manager shall liaise with the other departments, agencies and private entities to agree on the pricing schemes for sharing data assets with them. These schemes can be based on the following models:
  - Subscription model
  - Consumption-based model
  - One-time fee model
- The data partnerships manager will maintain a ‘Data Monetization Register’ to capture the identified and implemented data monetization use cases and the revenue generated from them.

### 9.8.3. Open Data

- The department shall create a plan to share some of its data assets on open data platforms (e.g., data.gov.in) to enable collaboration, sharing and innovation using the data assets.

This shall enable agencies and other departments liaising with the department for data assets to leverage the platform and access data as and when needed. In addition, this will also enable use of the department's data assets for developing data-based solutions for economic benefits. Some of the data assets can be stamp duty collection, property valuation data set, etc. Department can also consider publishing on the Digital India website (digitalindia.gov.in) for presence on the government initiative

- The data partnerships manager will liaise with the desk data owners and desk data stewards to identify data assets which can be published on open data platforms.
- The data partnerships manager will ensure that the data assets being published by the department are as per the NDSAP Open data license.
- The data partnerships manager, in collaboration with the desk data steward, shall ensure that the metadata for open data assets are documented and published with the data assets for enhancing its discovery and usability.
- The data governance manager, in collaboration with the data partnerships manager will define and implement the following processes for managing the department's open data assets:
  - Open data identification and classification
  - Open data compliance validation
  - Open data maintenance
  - Open data performance management
- The department shall use the following formats defined and approved by National Data Sharing and Accessibility Policy, 2012 to publish its open data assets:
  - 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)
- The department shall monitor the following KPIs to measure the progress of its open data plan and gather statistics on the published data assets:
  - # of downloads per published data asset
  - # of identified and prioritized data assets
  - # of identified data assets that have been published
  - # of updates performed on published data assets

#### **9.8.4. Data Sharing Guidelines**

- The department shall, wherever legally and ethically correct, promote sharing of its data assets internally and externally. This shall be undertaken using technology enablers highlighted in section 7.4 of this document as far as feasible.
- The department shall establish on its official website a channel to manage the submission and the reception of data sharing requests. These requests should be channeled to the Data Management Function.

- The data partnerships manager in collaboration with the data governance manager shall define and implement a data sharing process to cater to the internal and external data sharing requirements of the desks.
- The data governance manager in collaboration with the data partnerships manager shall define and follow an internal data sharing agreement template that shall be used when data is shared between information systems within the department.
- The data governance manager in collaboration with the data partnerships manager shall define and follow a third-party data sharing agreement template that shall be used when data is shared outside the department. The Data Sharing Agreement shall include, at minimum, the following:
  - Purpose of the data sharing
  - Information about requesting and sharing entities
  - Lawful basis for sharing
  - Sharing details (date, duration, volume, mode, etc.)
  - Liability provisions
  - The agreement shall be signed by the respective desk data owner and the appropriate representative from the receiver prior to sharing data
- The data partnerships manager shall maintain a data exchange directory to register all its existing data sharing and exchange internally and externally, including information on which sources are being requested, by whom, how often and for what purpose. The directory shall also record a list of requests for data sharing not approved by the department's Data Management

## 10. Recommendations

In this section we have summarized the initiatives captured in the various areas of Data Management Framework.

In order to overcome existing challenges and to be able to efficiently manage the data you could consider below initiatives:

- **Roles and Responsibilities:** Identify people for designated tasks with appropriate skill sets for each of the core areas
- **Metadata Strategy & Technology** - Define your data and implement using efficient technology. For the department to have consistency and trust in data, a Metadata Management solution is recommended which will cover the organization's requirements
- **Data Quality Upliftment** - Improve your data, a key aspect for data management is governing the quality of the data and establishing standards for future data collection
- **Data Integration Platform** - Integrate your data, a centralized integration layer (e.g., Service bus) will facilitate efficient and effective exchange of data between internal and external parties, application, departments etc.
- **Data Security & Privacy** - Define data classification for existing datasets across Department's processes. Implement best practices to identify and protect Personal Data
- **Data Analytics Platform** - Recommendation to develop a central platform for reporting and analytical requirements. Some of these use cases are monthly registration and revenue dashboards generated from eRegistration, iSARITA and GRAS, ad-hoc historical trend analysis requested by the department, etc.

- **Data Monetization** - With a vision of generating revenue from data, the Department can consider data partnerships, for e.g. Establishing contractual agreements on historical property sale data with estate aggregators like 99acres, magicbricks etc.
- **Open Data Platform** - There are initiatives by the government of India to develop an open data platform ([digitalindia.gov.in](http://digitalindia.gov.in)). Department can consider publishing datasets here to enable collaboration, sharing & innovation.
- **Data Sharing & Retention** - Abiding with regulations and using appropriate technology for sharing and documentation archival.