



सत्यमेव जयते

Government of Rajasthan

**Chief Minister's Rajasthan Economic Transformation
Advisory Council (CMRETAC)**

DEPARTMENT OF PLANNING, RAJASTHAN

**POLICY STUDY ON USING DATA
FOR BETTER POLICY FORMULATION AND
EVIDENCE-BASED DECISION MAKING**



2023



indicc
ASSOCIATES

Technical Support Organisation
CMRETAC



सत्यमेव जयते

**Chief Minister, Rajasthan
& Chairman, Economic
Transformation Advisory Council**



CHAIRMAN'S MESSAGE

Every state is important in the scheme of national development. We cannot assure the progress of India without the progress of the states. The Constitution binds us in a federal polity where every order of the government (Union, State and Local) has an important role to play. State governments are certainly closer to the people and hence bear an enormous responsibility towards ensuring effective delivery of goods and services.

In this endeavor, they have a direct, indirect, and enabling role to play. Rajasthan is committed towards that goal and has been at the forefront of many reforms since long. Our sincere and unceasing efforts, during the pandemic and otherwise, have been recognized widely.

The Bhilwara COVID-19 containment model has been recognized as a replicable model globally. Ours was also one of the first states in India which came up with a comprehensive strategy for economic revival in the wake of the pandemic. Besides taking a plethora of immediate steps to extend social and economic relief to the people during COVID-19, Rajasthan has also introduced several transformative measures in the recent past to boost the economy of the state. MSME Facilitation Act, 2019; Food Processing Policy, 2019; Tourism Policy, 2020; Mukyamantri Chiranjeevi Swasthya Bima Yojana, 2021; Handicraft Policy, 2022; Rajasthan Investment Promotion Scheme, 2022; Rajasthan Right to Health Care Act, 2022; Indira Gandhi Urban Employment Guarantee Scheme, 2022 and Vision 2030 are some of the path breaking initiatives undertaken by the government.

We also started a practice of 'thematic' annual budgets for converging our efforts and energy on most pressing issues and have ensured that governance is truly decentralized. Our campaigns on 'Prashasan Shehron Ke Sang Abhiyan' and 'Prashasan Gaon Ke Sang Abhiyan' are examples of that spirit.

While our efforts are incessant, structural slowdown and unexpected shocks like the pandemic can derail the economy. This calls for continuous preparedness on our part.

Creation of Chief Minister's Rajasthan Economic Transformation Advisory Council (CMRETAC) was a significant step to ensure our preparedness for short-term and long-term development objectives. In the year 2021-22, the Council prepared nine (09) policy studies on areas as diverse as Fiscal Management; Managing Urban Informal Sector; Integrated Agro-Business Infrastructure; Sustainable Agriculture; Doing Business; Quantifying Intangible Cultural Assets; Education and the New Paradigm (bridging digital divide); Medical Services; and Public Private Partnership in Infrastructure.

In the year 2022-23, the Council undertook six (06) new policy studies. These pertain to Building Energy Transition Roadmap; Financing Green Infrastructure; Urbanization of Rural Areas; Recalibrating Institutions to meet Climate Challenges; Using Data for Better Policy Formulation and Evidence-based Decision Making; and Redesigning Trade in the Era of E-commerce.

These policy areas may appear to be separate and discreet but one commonality that binds them all is that they are truly geared towards a bottom-up approach to the development of the state while embracing and addressing new challenges. I urge my colleagues in the state government to also focus on inter-linkages in these policy areas for the best possible outcomes.

I am confident that these fifteen (15) path breaking studies would be a valuable input for the state and I am happy to state that the present policy study is very much part of this endeavour.

I am grateful to the Members of the Council, my Ministerial colleagues, officers of the Government, all collaborators and organizations who have worked tirelessly to make this possible. My special acknowledgement of Dr. Arvind Mayaram, Vice Chairman, CMRETAC, whose leadership and contribution towards this endeavor have been extremely valuable. My appreciation is also to the entire team at CMRETAC which has diligently worked to put these reform-oriented studies together.



(Ashok Gehlot)



सत्यमेव जयते

**Economic Advisor to CM
& Vice Chairman**

**CM's Rajasthan Economic
Transformation Advisory Council**



VICE CHAIRMAN'S MESSAGE

The challenges confronting the economy are several. On the one hand, there are issues that underline structural sluggishness in growth while on the other hand, tech and climate change have fundamentally redefined several aspects of managing and running the economy.

Therefore, governance too has to rise to the challenge and has to be much more responsive and agile. This is possible if decision making is faster, accurate and quicker. Therefore, what is really needed is a framework for evidence based policy making premised upon usage of data.

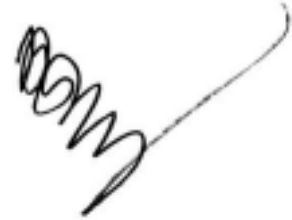
Even advanced countries like United States, with sophisticated governance systems, felt the need to enact a law on Evidence-Based Policy making. In India, both union and state governments have been collecting data on various parameters across diverse sectors for several decades. Despite this, there is no overarching legal framework driven by legislative foundation mandating use of data for evidence based decision making. While the data is collected through various channels, their use case is extremely disoriented, lacking structural and strategic direction for their application in evidence-based decision-making.

With this problematique in the background, the Chief Minister's Rajasthan Economic Transformation Advisory Council (CMRETAC) in collaboration with Centre for Policy Research (CPR) carried out this study for providing actionable recommendations for harnessing data to enhance policy formulation and promote evidence-based decision-making.

The study has taken a case study approach and looked at two sectors – Health and Education, and delves into aspect such as data collection methods, usability, and contribution to policy enhancement. To enhance evidence-based decision-making, the study recommends both short term and long term strategies. Similar studies need to be

carried for other sectors and integrated with each other to arrive at a holistic architecture of data integration.

I express my gratitude to Hon'ble Chief Minister for providing continuous support and guidance. I also express my gratitude to Hon'ble concerned Ministers, esteemed members of CMRETAC for their valuable guidance, concerned secretaries, other officers in the government, Technical Support Organization to CMRETAC and all other collaborators.

A handwritten signature in black ink, consisting of several loops and a long, sweeping tail that curves upwards and to the right.

(Dr. Arvind Mayaram)

ACKNOWLEDGEMENT

At the outset, we would like to express our deepest gratitude to the Chairman of Chief Minister Rajasthan Economic Transformation Advisory Council (CMRETAC), Hon'ble Chief Minister, Shri Ashok Gehlot and Hon'ble concerned Ministers.

We are indebted to Dr. Arvind Mayaram, Vice-Chairman for his valuable guidance and generosity with timely feedback on our work. We are grateful to the members of CMRETAC who have guided this study from time to time.

We are immensely grateful to Smt. Veenu Gupta, Additional Chief Secretary, Department of Industries; Shri Pawan Kumar Goyal, Additional Chief Secretary, School Education Department; Smt Shubhra Singh, Additional Chief Secretary, Medical & Health Department; Shri Alok Gupta, Principal Secretary, Administrative Reforms and Tribal Area Development; Smt. Manju Rajpal, Secretary, Department of Rural Development; and Shri. Ravi Jain, Secretary, Panchayati Raj Department.

We deeply appreciate the continuous and close engagement of Shri Inderjeet Singh, Commissioner, Department of Information Technology and Communications; Smt Anupama Jorwal, Managing Director, Rajasthan Medical Services Corporation Limited; Smt Suchi Tyagi, CEO, Rajasthan State Health Assurance Agency (RSHAA); Smt Rachna Bhatia, Additional Director, Department of Secondary Education; Smt Nalini Kathotia, Managing Director, RUDA; Smt Manisha Arora, MD, Rajasthan Small Industries Corporation; Shri Sudhir Sharma, MD, National Health Mission, Rajasthan; Dr. Anil Paliwal, State Project Director, Samgra Shiksha Abhiyan (SMSA); Shri Ashok Meena, Additional Director, Education Department, Rajasthan; Dr. Mohan Lal Yadav, Former State Project Director, Samagra Shiksha and Additional Commissioner, Rajasthan Council of School Education; Shri Rakesh Kumar Gupta, Additional Project Director, Samagra Shiksha; Shri Ashok Sangwa, Additional Director, Department of Elementary Education;

This study would not have been possible without the support of Dr. Kalpana Vyas, Executive Director (Logistics) Rajasthan Medical Services, Corporation Limited (RMSCL); Dr. Sushil Kumar Parmar, Additional Director (Hospital Administration), Department of Medical, Health & Family Welfare; Dr. Lokesh Chaturvedi, State Programme Manager, National Health Mission (NHM), Dr. Govind Singh, OSD to Secretary, Medical Health & Family Welfare Department; Shri Kushal Yadav, CEO, State Health Assurance Agency; Shri Vinod K Jain, Joint Director, National Health Mission, Rajasthan; Dr. Dilip Singh, Special Officer, Department of Economics and Statistics; Shri Sitaram Swaroop, Joint Director, Rajasthan Jan Aadhaar Authority;

Shri Vinod Jain, Scientist, Rajasthan State Centre - National Informatics Centre; Smt Toolika Dani, Deputy Director, Shaala Darpan Cell; Shri RS Meena, Additional Commissioner, MDM; Shri Maninderjeet Singh, Executive Engineer, RCSE, and Dr. Lokesh Chaturvedi and colleagues(State Programme Manager, NHM).

This study would not have been possible without the proactive support of Shri Bhawani Singh Detha, Principal Secretary, Plan; Shri Bhanwar Lal Bairwa, Joint Secretary, CMRETAC & Director, Directorate of Economics & Statistics; Shri Munshi Singh Rajput, OSD to Economic Advisor to Chief Minister; Shri Vijay Kumar Sharma, Assistant Director, CMRETAC; Shri Ashutosh Sharma, Assistant Director, CMRETAC; Dr. Devanand, Assistant Statistical Officer, CMRETAC; Shri Yograj Pacherwal, Assistant Programmer, CMRETAC; and Smt Saroj Saini, Informatics Assistant, CMRETAC.

We are immensely thankful to the team at Centre for Policy Research (CPR) and take this opportunity to thank Shri Ritwik Shukla, Senior Research Associate; Shri Sharad Pandey, Senior Research Associate; Shri Sidharth Santhosh, Senior Research Associate; Ms Poonam Chaudhary, Senior PAISA Associate; Shri Ram Ratan Jat, Senior PAISA Associate; Ms Mallika Arora, Learning and Development Associate, Ms Anwesha Mallick, Research Associate; and Ms Mugdha Mohapatra, Research Associate.

Indicc Associates provided technical support throughout the study period. We thank Shri Akshat Mishra, Shri Kartikeya Gupta, Shri Dhvaj Khattar, Ms Sowmya Kolla and Shri Krishn Mathur at Indicc Associates who constantly worked hard, sometimes on short deadlines, to ensure timely delivery of all policy studies.

Finally, any error or omission that may have remained is solely ours and should not be ascribed to any of the above acknowledged person or institutions.

Avani Kapur
Senior Fellow & Lead
Accountability Initiative
CPR, New Delhi

Abhishek Kumar
Technical Advisor, CMRETAC
& Partner, Indicc Associates

LIST OF ABBREVIATIONS

ANC	-	Antenatal Care
ANM	-	Auxiliary Nurse Midwives
ASHA	-	Accredited Social Health Activist
BCM	-	Block Community Manager (BCM)
BNO	-	Block Nodal Officer
CBEO	-	Chief Block Education Officer
CDEO	-	Chief District Education Officer
CHCs	-	Community Health Centres
CHOs	-	Community Health Officers
CMHO	-	Chief Medical Health Officer
DBT	-	Direct Benefit Transfer
DCF	-	Data Capture Format
DHs	-	District Hospitals
DNO	-	District Nodal Officer
DPM	-	District Programme Manager
DPO	-	Data protection officers
EBPM	-	Evidence Based Policy Making
EDAR	-	Electronic Detailed Accident Report
FGD	-	Focussed Group Discussions
FLWs	-	Front Line Workers
GDPR	-	General Data Protection Regulation
GoTN	-	Government of Tamil Nadu
IDSP	-	Integrated Disease Surveillance Programme
KPI	-	Key Persons Interviews
LS	-	Lok Sabha
MDM	-	Mid-day Meal
MLAs	-	Members of Legislative Assembly
MOIC	-	Medical Officer-in-Charge
MPs	-	Members of Parliament
NHM	-	National Health Mission
PCTS	-	Pregnant Women and Child Health Tracking System
PCTS	-	Pregnant women and Child Tracking System
PEEO	-	Panchayat Elementary Education Officer
PFMS	-	Public Finance Management System
PHC	-	Public Health Centres
PHCs	-	Public Health Centres
PMJAY	-	Pradhan Mantri Jan Arogya Yojana

PRIs	-	Panchayati Raj Institutions
PSP	-	Private School Portal
RBSK	-	Rashtriya Bal Swasthya Karyakram
RCHO	-	Reproductive and Child Health Officer
RoP	-	Record of Proceeding
RS	-	Rajya Sabha
TB	-	Tuberculosis
TMS	-	Transaction Management System
TPA	-	Third Party Administrator
UDISE+	-	Unified District Information System for Education Plus
VS	-	Vidhan Sabha
WHO	-	World Health Organisation

INDEX

EXECUTIVE SUMMARY	i-ii
SECTION I	
CHAPTER 1: INTRODUCTION	2-6
CHAPTER 2: METHODS	7-9
2.1 Data Mapping and Information Triangulation	7
2.2 Key Person Interviews (KPIs) and Focus Group Discussions (FGDs)	7
2.3 Case Study Type Interviews	9
CHAPTER 3: ANALYTICAL FRAMEWORK	10-12
3.1 Feature of “Data Generation”	10
3.2 Feature of “Data Usability”	11
3.3 Feature of “Data Use”	11
3.4 Feature of “Data for Policy Improvement”	12
CHAPTER 4: RESEARCH FINDINGS AND CHALLENGES OF DATA IN EDUCATION	13-24
4.1 Trends and Challenges in Generation of Data	16
4.1.1 Records maintained both Online and Offline	16
4.1.2 Technology and Infrastructure Challenges in Data Generation	18
4.2 Data Usability	18
4.2.1 Need to Strengthen Accessibility and Usability: graphs, visualisations etc	18
4.2.2 Data Quality and Verification: Felt Need for Randomised Checks	19
4.3 Use Cases for Data	20
4.3.1 Innovative Approaches at the State Level	20
4.3.2 Limited Usage at the Frontline	21
4.3.3 Interoperability of Data: 3 case studies	21

CHAPTER 5: RESEARCH FINDINGS AND CHALLENGES OF DATA IN HEALTH 25-46

5.1	Introduction to Data Use in the Health Department	25
5.1.1	Examples of data systems in Rajasthan	27
5.1.2	Examples of data use in the health department	29
5.2	Data Generation	33
5.2.1	Data is collected both online and offline	33
5.2.2	Duplication of entry	34
5.2.3	Missing data restricts efficient targeting and tracking	35
5.2.4	Lack of outcome data	36
5.2.5	Lack of real-time data	36
5.3	Data Usability	37
5.3.1	Errors and verification	37
5.3.2	Data discrepancies	38
5.3.3	Capacity gaps	39
5.3.4	A lack of human resources	40
5.3.5	Issues with software affects both citizens and functionaries	41
5.3.6	Infrastructure	42
5.4	Data Usage	42
5.4.1	Multiple applications and siloed functioning	42
5.4.2	Applications are not created in a participatory manner	43
5.4.3	Examples of successful interoperability, as well as potential interoperability	44

CHAPTER 6: RESEARCH FINDINGS AND CHALLENGES OF DATA IN THE PANCHAYATI RAJ DEPARTMENT 47-51

6.1	Introduction	47
6.2	Data Generation	49
6.2.1	Capacity gaps in data entry	49
6.2.2	Data stored both online and offline due to a lack of trust in online systems	49
6.3	Data Usability	49

6.4	Data Use	50
6.4.1	Data primarily used for monitoring	50
6.4.2	Mixed views on integration across portals	50
6.4.3	Redundancy	51
CHAPTER 7: RECOMMENDATIONS FOR THE EDUCATION, HEALTH, AND PANCHAYATI RAJ DEPARTMENTS		52-71
7.1	Data Generation	52
7.2	Data Usability	55
7.3	Data Usage	56
7.4	Towards a Stronger Data Cycle Framework	61
7.4.1	Institutional Support: The creation of a data stewardship body	62
7.4.2	Ensuring an Effective Legal Framework	64
7.4.3	The architecture needed to facilitate optimum data use for decision-making	65
CHAPTER 8: CONCLUSION		72-74
REFERENCES		75-77

SECTION II

1	INTRODUCTION	79
2	INTERNATIONAL PRACTICES	79
3	WHERE DOES INDIA STAND?	82
4	LEGAL CHANGES REQUIRED TO OPERATIONALISE EPBM	84
4.1	Changes required in laws that allow the collection and storage of data by the government.	84
4.2	Changes in laws that require the dissemination of data to citizens	85
4.3	Changes to ensure the veracity and accuracy of information	85
5	COMPARING THE PERSONAL DATA PROTECTION BILL 2019 AND THE DIGITAL DATA PROTECTION BILL 2022	86
6	STATE LEVEL DATA POLICIES	87
REFERENCES		90

LIST OF TABLES

1	Different lenses to see data	3
2	Respondents	8
3	Recommendations for data generation	52
4	Recommendations for data usability	55
5	Recommendations for data usage	56
6	Data Sharing Matrix	68

LIST OF FIGURES

1	Purpose of data governance frameworks	2
2	Types of decision making	4
3	Framework for Evidence-based Policymaking	10
4	Objective of School Infrastructure	14
5	Examples of Data Systems in Education	15
6	Examples of Data Use in Education	20
7	Health Data Governance Principles	25
8	Data check process for PCTS	37
9	EDAR as a cross-departmental platform	44
10	Potential for integration across the spectrum of health and education departments	46
11	Portals maintained by the state government (different departments)	48
12	Portals maintained by the union government	48
13	The Policy Cycle	61
14	Potential Data Stewardship in Rajasthan	63
15	Types of Data Stewardship	64
16	Data management cycle	66

EXECUTIVE SUMMARY

Welfare governance in sectors such as education, health and Panchayati Raj involve a comprehensive exercise of collecting and using data in innovative ways. The use of data can enable making decisions based on evidence (what is referred to as ‘evidence-based decision making’) rather than biases and assumptions. For this, an effective data governance framework is required that can enable setting standards for collecting data, sharing data across departments and encouraging a data-based comprehension of policy challenges.

Rajasthan has taken important steps towards strengthening use of data in everyday governance. However, there is further strengthening required to achieve smooth use of data at all tiers of the government and across different departments.

In this report, we assess three sectors of education, health and Panchayati Raj to unpack how data is collected, its usability, use cases of data and how data can feed into policy improvement. On conducting in-depth interviews and focussed group discussions with functionaries across all tiers of governance in these three departments, we identified that each sector is at a different stage of adoption of evidence-based decision making. For example, the education sector uses a unified Shala Darpan portal for school management and generation of report cards while health uses over – applications for similar management practices. Both sectors along with Panchayati Raj required stronger capacities at the frontline to collect and use the data. Publication of anonymised data in the public domain (as open data) remains limited. Large governing principles such as a clear identification of what data is required at which stage in the policy cycle was missing.

We propose the following recommendations to improve evidence-based decision-making in Rajasthan.

Short-term solutions to existing systems that can be made include:

- Generation of data: adoption of a policy or guideline on e-office practices, strengthen infrastructure and recruit Data Entry Operators at the frontline, adopt citizen identifier data from Jan Aadhaar
- Data usability: re-evaluate data verification systems, improve capacities on use of software.

- Data use: work on combining datasets where there is fragmentation, introduce participatory data collection portals, define clear data utilisation practices at the frontline and improve use of predictive analyses.

Long-term changes that are required include:

- Linking data requirements to each part of the policy cycle for clearer targets
- Creation of a data stewardship body that can facilitate dialogue between departments on sharing and use of data.
- Ensuring an effective legal framework for governance
- Strengthening core areas of data use that include a clear recognition of the data management cycle, metadata, standards for data use and secure storage of data.
- Creation of Data Analysis and Response Teams (DART) to analyse data consistently, and regularly feed insights to decision-makers, to deliver services more effectively.

SECTION I

CHAPTER 1: INTRODUCTION

An effective and efficient governance system ensures that its decision-makers have access to and use high-quality information to make informed policy decisions (*Lugo-Gil, Jean-Baptiste, & Jaramillo, 2019*). With the generation of data growing in an unprecedented manner, governments around the world are seeking ways to leverage this data to accelerate economic growth, improve the efficiency and transparency of government, and tackle socioeconomic challenges (*Feigenbaum et al., 2022*).

Data Governance has been referred to as “diverse arrangements, including technical, policy, regulatory or institutional provisions, that affect data and their creation, collection, storage, use, protection, access, sharing and deletion across policy domains and organisational and national borders.” (*OECD, 2022*). In India, the Information Technology Act 2000 defines data as a “representation of information, knowledge, facts, concepts or instructions which are being prepared or have been prepared in a formalised manner”.

Data governance frameworks have proven to be extremely valuable for enhancing the reliability and quality of data, bolstering data security and adherence to regulations, boosting the utility and worth of data, facilitating data-centred decision-making, and nurturing a collaborative culture around data. These frameworks serve multiple purposes described in the diagram below.

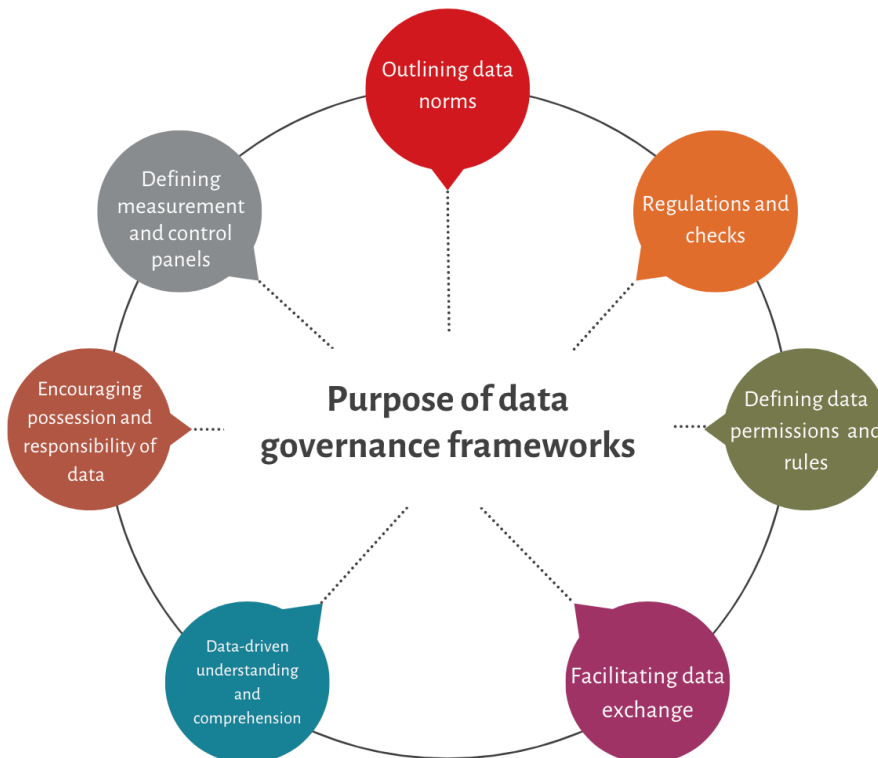


Figure 1: Purpose of data governance frameworks

With governance becoming increasingly complex, there is a need to move rapidly towards practices such as evidence-based policymaking, data-oriented decision-making, and open public data. However, most countries across the world do not have a robust data governance framework (*World Bank, 2022*). In fact, only 40% of the countries across the globe are on the path towards implementing good data governance practices (*World Bank, 2021*).

Broadly, there are 3 different lenses through which data can be seen. First is sectorally. Typically, each department, and in fact, sub-units within departments have their own mandates. Second, is with respect to the specific function which cut across departments. These can take the form of service provision (such as education, health, pensions, etc.), finance related matters (revenue, tax, etc), administration (general administration, home department, law and justice, etc.), and utilities and infrastructure (public works, transport, urban development, etc.). Lastly is with respect to the types of decisions made in the government using data can be understood as strategic, tactical, and operational.

Table 1: Different lenses to see data

Lenses	Types of Data	Description
Sectoral	Education, Health, Sanitation, Tribal	Usually linked to specific line departments
Function	Service Provision Finance Related Administrative Utilities and Infrastructure	Crosscuts across departments
Decision-Making Type	Strategic Tactical Operational	Crosscuts across departments

Strategic decisions are long-term plans such as the creation or revision of new policies that align with the state or union's big-picture goals. Tactical decisions are those made to materialise these long-term plans into action through mid-term or annual strategies. Finally, operational decisions are those made on a daily basis to operationalise these annual strategies.

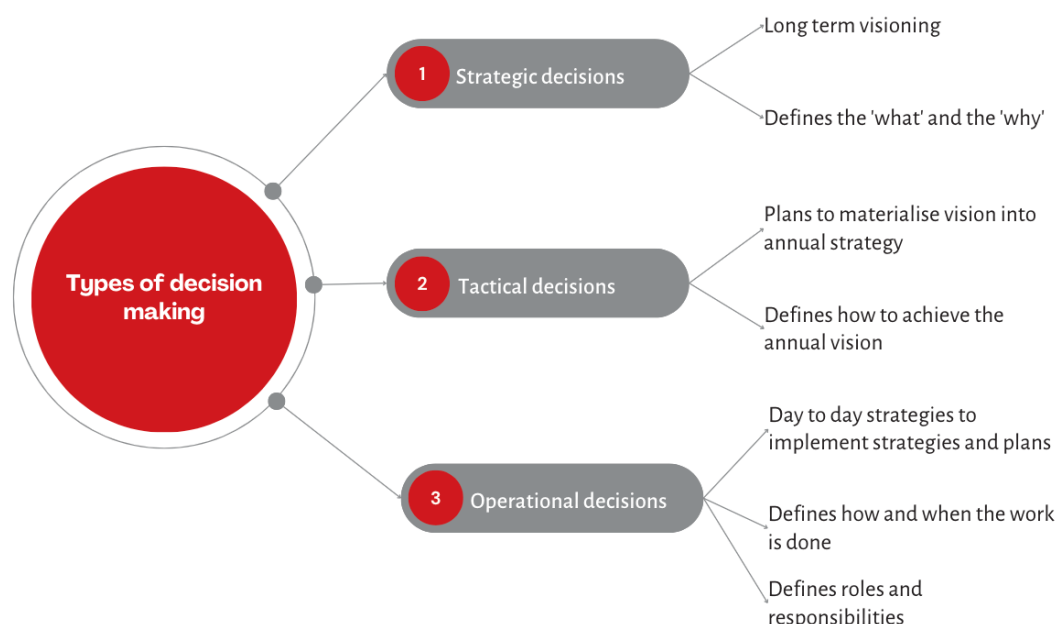


Figure 2: Types of decision making

Rajasthan's Data Policy Evolution

Recognising the need for building a data governance framework that moves away from the siloed department approach to a more holistic, multidimensional collection and usage, in recent years, the Government of India has focused on several policies related to data accessibility and storage via their National Data Sharing and Accessibility Policy 2012. A few states in India have also made a move in a similar direction. States such as Punjab, Odisha, Karnataka, Tamil Nadu, Sikkim, Telangana, and Chandigarh formulated state-level data policies similar to the National Data Sharing and Accessibility Policy of 2012.

Rajasthan too has made several attempts towards a data-driven policy-making space. The State government has been actively undertaking initiatives to collect data across a wide range of categories in different sectors including education and health. In the last five years, the state has made significant commitments towards making target-oriented schemes easily accessible for eligible citizens through e-governance solutions. This commitment includes ensuring citizens have access to different portals where they can view their entitlements and file grievances if any. To make these systems possible, the state has invested significantly in digital infrastructure to ensure data made available to citizens is collected and maintained in convenient systems.

The policies that regulate and protect data collection in the state include: - The Rajasthan E-Governance IT and ITeS Policy 2015, the National Data Governance

Framework Policy and the draft India Data Accessibility and Use Policy of the Ministry of Electronics and Information Technology, Government of India. These policies were developed considering the recent digitisation of governance in India. They highlight the need for regulation in data collection, storage and use.

The state has also invested in various aspects of data governance. While the state began with scoping exercises to understand how data storage needs to be structured, it soon invested in constructing robust systems that collect and verify data as well. This can be seen in systems such as e-aushadhi for healthcare and Shala Darpan for school education. The data systems involve designing processes that functionaries in the line departments need to follow to collect data on a routine basis.

The processes of data collection and verification are supported by the National Informatics Centre of the Government of India by creating systems to store and use the data. Line departments work together to follow processes that are made together. The state has also realised the need to constantly improve data systems for which either high-level meetings are organised, or studies (such as this) are commissioned. To understand data systems in more detail, it is necessary to unpack the system goals and how decisions are made using data.

In 2022, as part of the Chief Minister’s Economic Transformation Advisory Council (CMRETAC), the Government of Rajasthan sought to understand the current status of data availability and its use under the project “Using data for better policy formulation and evidence-based decision making”, undertaken by the Accountability Initiative, Centre for Policy Research through case studies of two departments: Education and Health along with a preliminary understanding of data collected by the Panchayati Raj Department.

The focus of the study is on evaluation of inter and intra sector data flow for optimum results. Thus, two case studies namely health and education are used to unpack the following questions: -

1. What data is collected, at what level, in what detail, and the quality?
2. How can government offices and officers use data for decision-making at all levels and all stations?
3. What are the various barriers and facilitators in using data?
4. What are some national and international practices on data use and evidence for policy making?

The findings from the study are posed against systemic goals to identify avenues for improvement. The study proposes comprehensive actionable recommendations that the state can implement to improve data collection, usability, and use in governance.

The recommendations identify outcomes that can be achieved if implemented based on the state's needs. These include (and are not limited to) practices within each line department and interoperability of data between departments.

This report is structured as follows. Chapter 2 describes the methods followed, and Chapter 3 lays down a framework for the analysis. Chapter 4 covers findings from the education sector. This covers data generation, data use and usability, details challenges with existing systems, and provides recommendations. Chapter 5 does the same for the health sector and Chapter 6 describes findings from the Panchayati Raj department. Chapter 7 contains recommendations for the departments and Chapter 8 concludes the report.

CHAPTER 2: METHODS

The report uses qualitative data to unpack the journey of evidence-based decision-making, from identifying and collecting data points to analysing the generated data and translating the information into policy formulation and implementation.

2.1 Data Mapping and Information Triangulation

To develop a base-level understanding of the situation of data-oriented decision-making, a preliminary data mapping exercise was conducted. This exercise provided information about the following: -

1. The different data portals that collect information about education in the state of Rajasthan, which are owned and operated by different tiers of the government/departments and other external agencies.
2. The quantum of data that is available as open-source data.
3. The frequency with which the data are collected for the different portals.
4. The different categories/ themes of data that are collected.
5. The granularity of the collected data.

For better triangulation of data, a form was circulated among the health and education departments. The form was intended to give further insights into the information collected through the mapping exercise, along with additional information on the objectives tied to the collection of data and the usability of the data generated. As of February 2023, the Department of Education and the Department of Health had not returned a completed form, leaving the researchers with a secondary review of the aforementioned points. Therefore, the study pivoted to having in-depth conversations with relevant officials across levels of government.

2.2 Key Person Interviews (KPIs) and Focus Group Discussions (FGDs)

Semi-structured qualitative questionnaires were prepared, with questions from the identified themes. The questionnaire was used for conducting in-depth interviews and FGDs with state-level and mid-level bureaucrats. A mix of purposive and snowball sampling techniques was used for identifying respondents. It was ensured that the sample was selected in a manner that would represent officials of different sub-departments within Rajasthan at the state level.

For health, interviews were conducted with officials in major working bodies of health. These include the Rajasthan State Medical Corporation Limited, the Rajasthan

State Health Assurance Agency, the National Health Mission, as well as officials working on other aspects in the health department.

For education, apart from the interviews conducted with state and mid-level educational policy decision-makers, interviews were also conducted with other stakeholders which included officers managing the state-level government data portal, Unified District Information System for Education Plus (UDISE+) database, and the Private School Portal (PSP). These can be seen in Table 2.

Table 2: Respondents

Mode of investigation	Respondents	Dates
Department of Education		
Focussed Group Discussion	10+ Senior bureaucrats from Department of Education	5th December 2022
Key Person Interviews	9 Senior bureaucrats from Department of Education and National Informatics Centre	6th - 7th December 2022 and 17th May 2023
Case Study Interviews	4 Headmasters, 2 School Teachers, 4 Panchayat Elementary Education Officers, 2 Chief Block Education Officer, 1 Block Resource Person, 1 Chief District Education Officer, 2 District MIS Managers	15th - 20th December 2022 and 27th to 29th April 2023
Department of Health and Family Welfare		
Key Person Interviews	10+ Senior bureaucrats from Department of Health, 1 Senior bureaucrat from the Rajasthan Jan Aadhaar Authority	4th - 5th January 2023 and 17th - 18th May 2023
Focussed Group Discussion	20+ senior bureaucrats working on NHM, Department of Health	4th January 2023
Case Study Interviews	3 ASHA workers, 3 Auxiliary Nurse Midwives, 1 LT, 2 Computer Operators, 2 Information Assistants, 1 Medical Officer-in-charge, 1 Chief Medical Health Officer, 3 Block Programme Managers, and 2 District Nodal Officers	7th- 20th February 2023; 18th May 2023
Department of Rural Development and Panchayati Raj		
Key Person Interviews	1 Senior bureaucrat, 2 Senior MIS Manager	17th to 18th May 2023
Case Study Interviews	2 Village Development Officers, 1 Block MIS Manager, 1 District Programme Manager, 1 Gram Panchayat Sarpanches	17th to 18th May 2023

2.3 Case Study Type Interviews

Qualitative questionnaires were used to capture the perspective of frontline workers and data entry operators. For education, this included the Headmasters (HMs), PEEO (Panchayat Elementary Education Officers), School Teachers, Chief Block Education Officers (CBEOs), Chief District Education Officers (CDEOs) and Block Resource Persons (BRPs) among others. It was ensured that data generators interviewed during the study belonged to different levels of schools. For health, it included the Accredited Social Health Activist (ASHAs) and Auxiliary Nurse Midwives (ANMs), the computer operators and Information Assistants, and block and district officials like the Medical Officer-in-Charge, Block Programme Managers, the Chief Medical Health Officer, and a District Nodal Officer. For Rural Development and Panchayati Raj, this included officers such Village Development Officers (VDOs), Sarpanches and MIS Managers (at block, district and state level).

CHAPTER 3: ANALYTICAL FRAMEWORK

This report proposes a framework that supports a holistic understanding of how data is generated and used in policymaking. The framework has been designed by conducting an extensive review of existing frameworks proposed by international organisations, academia and other research organisations. The framework is unique in its characteristic features as it brings to fore certain aspects of data governance that can guide Evidence-based Policymaking (EBPM). It intends to nudge the state’s policy makers towards more effective reflection of priorities and actions taken. The framework proposes four features of ‘data generation, data usability, data use and the data for policy formulation’. These elements have been designed for convenient use by government functionaries at different tiers of the government. The framework allows for easy replication for further investigative exercises across sectors such as health, education or social justice (see Figure 3 below).

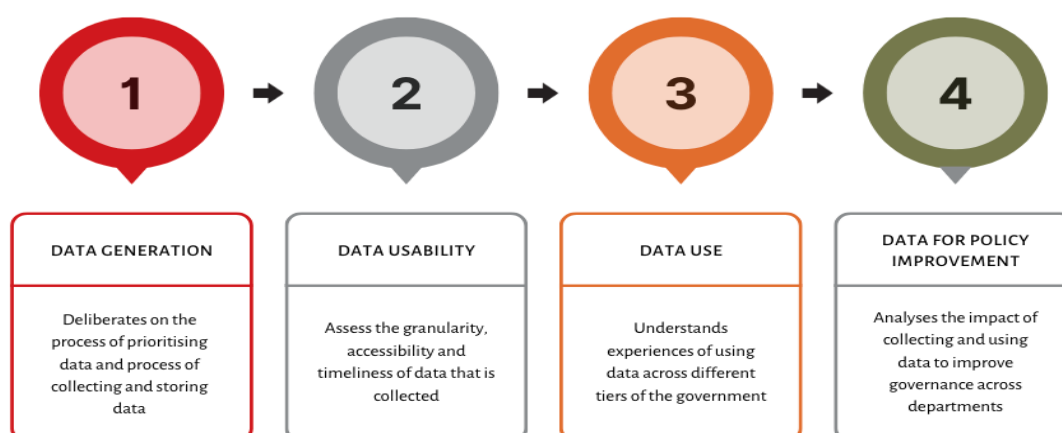


Figure 3: Framework for Evidence-based Policymaking

3.1 Feature of “Data Generation”

Data generation can be looked at through two main processes: -

- The process of prioritising data that needs to be collected.
- The process of how this data is then collected and stored across different tiers of the state.

In the first process of understanding which data is collected, different stakeholders from the state were interviewed. By triangulating evidence from different tiers of the state, the framework enables a thorough understanding of how it was decided to collect the data. It unpacks the methods through which the state decides what type of data is necessary for evidence-based policymaking and which functionaries get access to this data for further use.

In the second process of data generation, we analysed the different processes through which the data is collected. The aim of this feature of the framework is to map different routes through which data is collected – including all platforms maintained by different departments of the state. The feature seeks to identify and suggest methods to avoid duplication of effort in collecting similar data across different platforms.

3.2 Feature of “Data Usability”

The second feature of the framework tries to understand the usability of the data that is collected. Data usability looks at three core aspects of granularity, accessibility and timeliness.

- Granularity which is a measure of the level of detail with which the data collected is presented to functionaries. A high level of granularity is desirable as it enables understanding which circumstances or characteristics of the data collected affect a change in others.
- Accessibility of data that unpacks the different formats in which a functionary can visualise the data when using it. For example, it understands whether the dataset can be downloaded as EXCEL files, PDF files or scans of hard copies of the data collected. It also unpacks whether the platform where the data is collected provides the user with visualisations such as graphs for selected data fields.
- Is the data collected verified and reliable?
- Lastly, timelines that reflects how often the data is collected and whether the functionaries are provided with this updated data on a regular basis. A core aspect of timelines is whether the platforms are updated in real-time or at other time intervals.

3.3 Feature of “Data Use”

This feature of the framework analyses how data is used in the everyday governance of education and health in the state. It thus looks at how functionaries use the data including for making decisions such as planning interventions or informing other functionaries of developments. It also looks at if there is necessary convergence between departments with regard to collecting and sharing data. Analysing this aspect of data in the state allows us to recognise best practices in using data and challenges that have emerged. Considering the state pivoted towards using data a decade ago, the feature provides an opportunity to reflect on how these challenges were resolved.

3.4 Feature of “Data for Policy Improvement”

This last feature of the framework ties together the generation of data and the use of data to provide policymakers with a reflection of how impactful the exercise has been. It reflects on the core question of whether functionaries receive the right kind of data for making high-impact decisions. It also tracks whether citizens and civil society have access to data and the rules that protect civil rights in data generation and use.

These questions aim to deliberate on the intent of data use in evidence-based policymaking. It provides policymakers with an opportunity to revisit the foundations of why data is used, how functionaries can be trained to use this data and how the exercise can reach other departments.

CHAPTER 4: RESEARCH FINDINGS AND CHALLENGES OF DATA IN EDUCATION

The use of evidence-based policymaking in education has increased over time. A data-oriented governance system has the potential to hold the system accountable by tracking gains in students' achievements through increased public investment in education (*UNESCO, 2023*). Data can also help flag issues with resource provision and use, enabling decision-makers to target policy actions. Research evidence shows that data-based decision-making in education by different actors has the potential to improve the quality of education (*Schildkamp, 2019*); (*Williamson, 2016*); (*Jornitz & Engel, 2021*). However, the use of data by different actors varies across the system. For example, teachers look at more granular data than the aggregated data at the national/sub-national level (*UNESCO, 2023*), (*Breiter & Light, 2006*).

For education, Rajasthan's commitment towards data governance is reflected in portals like Shala Darpan where citizens can access anonymised information on schools and government functionaries can access more granular data on indicators such as teacher attendance. With Rajasthan having one of the highest number of government schools, it becomes crucial for the government to continue monitoring and tracking the quality of education provided through such digital initiatives.

To understand data generation, it is important to first map the different moving parts of the state tasked with ensuring every child receives education. The Department of Education within the state is tasked with implementing education policies. This includes, and is not limited to, developing and maintaining different categories of schools, designing and implementing interventions to promote school education and ensuring adults have the opportunity to learn as well. Within the Department of Education, the Rajasthan Council of School Education (RCSE) functions as the Single State Implementation Society. This society is tasked with implementing Centrally Sponsored Schemes (CSSs) like Samagra Shiksha that improve school effectiveness for pre-primary to class 12.

In addition to this Council, elementary education is implemented by the Directorate of Elementary Education and secondary education, by the Directorate of Secondary Education. These directorates maintain autonomy in deciding how schools function and the interventions that are planned. These directorates receive support from the Council and other departments of the state such as the Department of Rural Development and Panchayati Raj, the Women and Child Development Department and functionaries of the general administration such as the District Magistrate. Elected

representatives such as Sarpanch and Members of the State Legislative Assembly also have the authority to support the state in implementing education policies.

Effective use of data in evidence-based policy making requires data to be shared across different tiers within each department and across departments as well. An example of ideal inter and intra sectoral decision making between education, health and local governments (or Panchayats) is demonstrated in the diagram below where the common objective is of improving infrastructure in schools. To achieve the objective, effective implementation of a data sharing structure is critical. The understanding through a step-by-step process can be useful to understand how data is generated and what its potential uses could be.

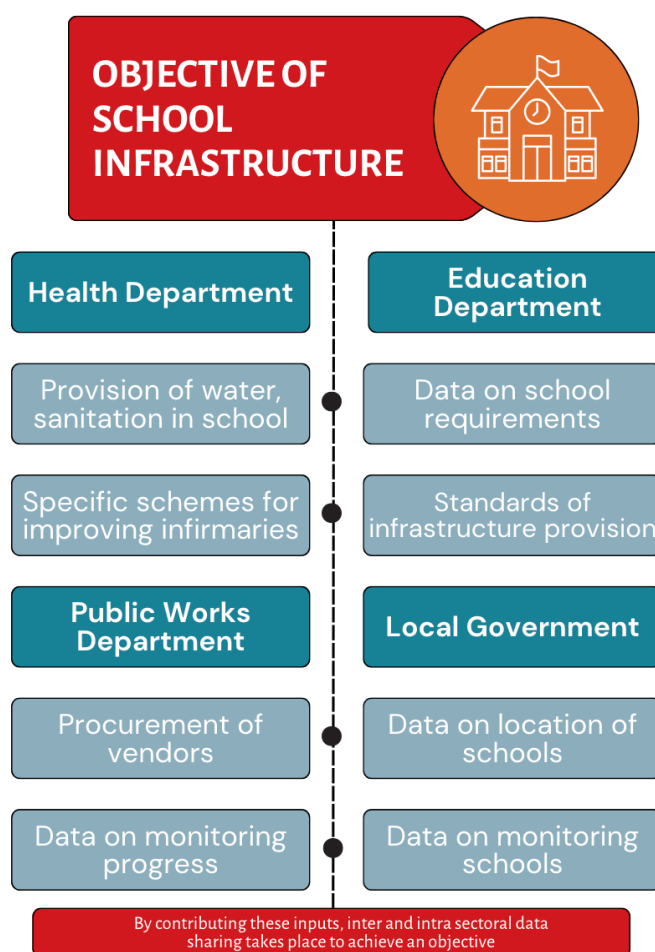


Figure 4: Objective of School Infrastructure

In the next section we will look more closely at the different sources of education data available and their usability and use.

SOURCES OF EDUCATION DATA

Broadly the study found 3 main sources of data in the form of platforms used by the state. These were the Shala Darpan Portal, the Private School Portal and the UDISE portal. See Figure 5 for more details.

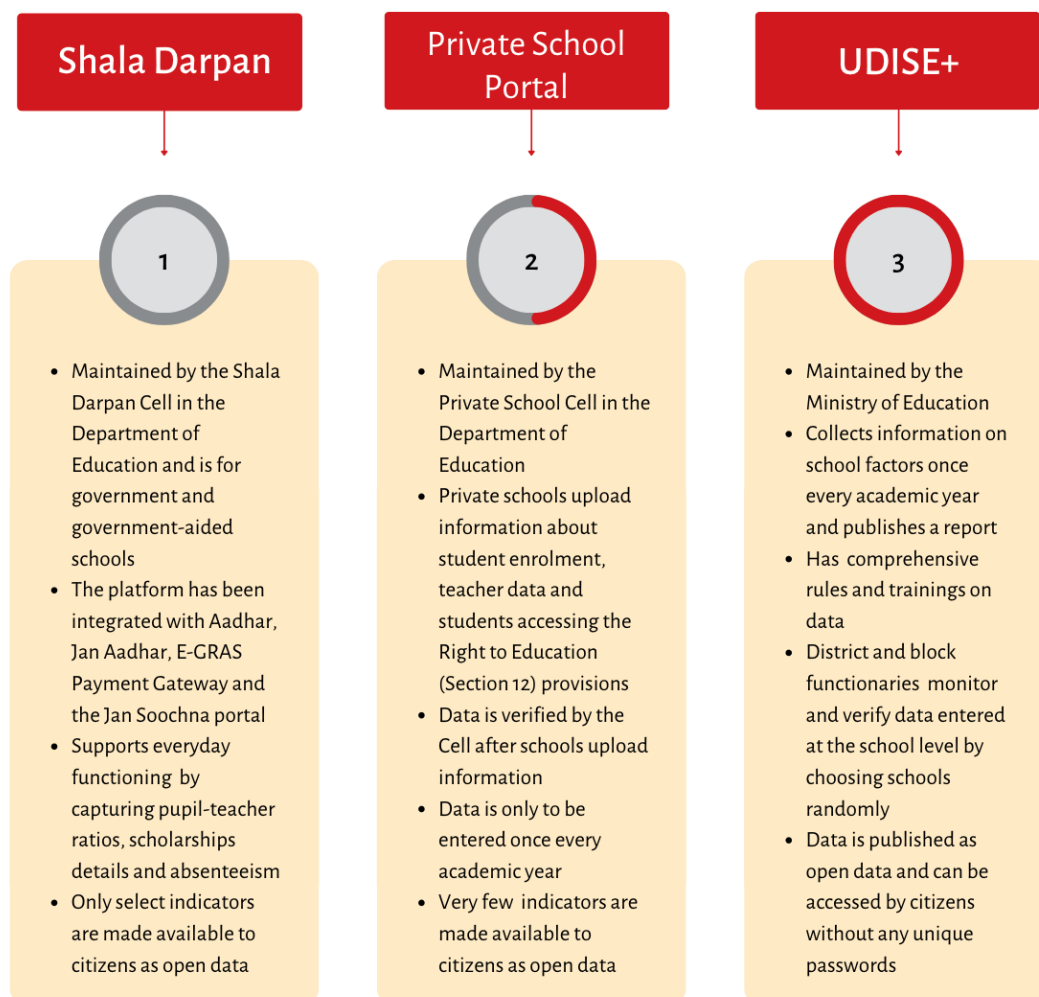


Figure 5: Examples of Data Systems in Education

Overall, functionaries in the state shared that the comprehensiveness of the Shala Darpan data generation system has made identifying bottlenecks in any implementation challenges simple. As a functionary at the state-level shared:

“Shala Darpan portal, since its launch, has improved a lot. It now has details of different types from beneficiary schemes and student assessment performance to school management. The platform allows us to access lots of information in one place...”

The bilingual platform provides a dashboard that every school can access. It provides information about infrastructure shortages and has recently begun issuing student report cards as well. The portal has received many accolades for its design from the Government of India, Skoch Foundation, etc.

As Shala Darpan collects data across each school, every student and teacher are included in the data. These data sets also collect a wide range of data on indicators such as attendance or receipt of welfare provisions. Similarly, UDISE+ also collects data across each school for several indicators on school factors as well as indicators such as attendance or receipt of welfare provisions. Functionaries were in support of this high level of granularity of data collected. It was felt that the visualisations available as well as the time series data allows for a comparison in developments across a period of years. The platform allows for such comparisons to be made and downloaded accessible EXCEL or PDF files.

Unlike Shala Darpan, however, data is not collected in real time and is entered once every academic year and thus Shala Darpan was the preferred data for most decision-making.

The PSP portal was found to be the least used as it is yet to achieve this degree of granularity as non-state actors do not upload data on a regular basis.

The remainder of the report focuses on the data framework for the Shala Darpan portal as it is the most comprehensive education portal in the State. This report does not explore challenges and improvements required for UDISE+ portal as it is out of the state's scope for change.

The following sub-section presents findings through the lens of the framework where data, that is intended to be used for evidence-making, is analysed from when it is generated, to how it is used and the impact it has on policy making.

4.1 Trends and Challenges in Generation of Data

4.1.1 Records maintained both Online and Offline

In specific, the Shala Darpan portal collects data for more than **100 data points** with the process of data generation beginning at the school level. Interviews with frontline officials however pointed to the data generation exercise posing an additional burden on schoolteachers. It was mentioned by most of the HMs during the interviews that entering data takes away a significant amount of their teaching time.

As the school collects data not only related to education but also household data of the students, teachers are often required to trade off their hours in the classroom to collect household data. The HMs mentioned in the interviews that data entry was either done at the cost of missing classes or working post-school hours. Nevertheless, the process of collecting household data has eased over the years as the required data is now collected by the schools during the admission process of the students. However, in remote regions with poor digital penetration, teachers must find locations with adequate internet facilities to upload the data on the portal.

“Collecting household data takes away a lot of time. We end up missing our classes to collect the data. Also, when we go to collect the data, we find it difficult to get hold of the people”

- a school HM

Further, the shift of data from the offline to the online portal has resulted in duplication of efforts at school-level. It was indicated during the interviews with HMs that despite all the data being available online on the Shala Darpan, the schools maintain a separate offline database of the same data fields.

We found that the main reason behind maintaining records online as well as offline was for different uses of the data. Offline data was used for audits while online data was used for monitoring by functionaries at the block, district and state levels. Teachers and HMs shared that this offline data was trusted more and used for all audit purposes when officials would visit their schools. A PEEO shared that while data entered online could have errors, data entered in the offline books are thorough and cannot be changed.

“We keep offline records of almost all the data. Mostly we first generate an offline record of the data like the MDM register, and attendance register. Then with the help of these records, we upload the data online. We also keep printouts of the data”

- a school HM.

As a senior functionary at the state level shared, there is a need for the system to reflect on which methods it prefers to use – online or offline. Choosing one over the other would have not been possible in the early years of introducing Shala Darpan. But now that the data systems have been in use for the last five years with consecutive improvements, there is a need to make a choice. Functionaries across the system believed that this choice rests with decision-makers at the state level who can ensure all use of data is routed through online databases rather than offline records.

4.1.2 Technology and Infrastructure Challenges in Data Generation

Functionaries using Shala Darpan to enter data at the school level shared that they found the portal challenging at times. The causes for concern were as follows:

1. Technology at the school level: HMs and teachers shared that they did not have the required technology, such as laptops or computers, to enter data onto the platform. While large schools and model schools enjoyed an abundance of technology, smaller schools (especially those with 2 or lesser teachers) suffered more. To solve this, teachers have been visiting e-Mitra centres or working from their homes to enter data in the stipulated window for data entry.
2. Technology at the state level: Functionaries across all levels of the state shared that they suffered immensely from server crashes. While these crashes did not last long, the short stints in crashes can affect functioning and push functionaries to re-enter large amounts of data that was not saved. Functionaries at the state level shared that these server crashes have reduced with the support of NIC that manages the portal's web development. Functionaries agreed that there is a need to improve more which is an on-going process.
3. Learning about changes to the online system: As there is currently no structured method of training that reaches the frontline functionaries, they are forced to learn about changes to the system on their own. Frontline functionaries use WhatsApp to receive updates (which are either text messages or scanned letters) about new data that is to be entered. But this is sufficient to navigate the system which forces them to learn through other avenues such as using YouTube videos that other teachers have made or visiting the closest e-Mitra centre. Functionaries at the block and district level that receive training shared that there is scope to use the Shala Darpan platform itself to conduct capacity building initiatives rather than using WhatsApp or YouTube. This perception was not shared by functionaries at the state level that instead believed that functionaries should learn on their own.

4.2 Data Usability

4.2.1 Need to Strengthen Accessibility and Usability: graphs, visualisations etc

We find that the usability of the data collected can be strengthened. Usability includes features such as ease of accessing the platform for generating data, downloading data and visualising it in innovative ways.

Functionaries at different levels we interviewed shared that they can download the data in EXCEL and PDF files, but this was challenging to peruse. They believed the data in its raw form did not allow for analysis that could support making decisions. For this, functionaries shared they **would benefit from visualisations in the form of graphs. A ‘click’ feature on these graphs would enable easier access of the data.** The platform with its high degree of granularity could also consist of features which **allow users to choose different indicators** they can compare across years.

4.2.2 Data Quality and Verification: Felt Need for Randomised Checks

As of April 2023, the only method in use for data verification was that of block-level functionaries checking if the data entered is correct during their monthly visits. When asked about the need for better robust methods, such as randomised checks or expert visits to schools, functionaries at the state level shared that they believed in their frontline functionaries and felt that they did not require such a system.

“We trust our teachers and HMs entirely. We do not think they will make any mistakes in entering data. The data that is entered online is final and ready to use by anyone”

- a state-level officer.

However, at the block and district level, functionaries felt that there was a need for randomised checks that should be conducted at different intervals during the academic year without informing the HMs or the teachers. They believed this would allow for better quality data on the systems.

“I can’t always tell if the data entered at the school is correct. Sometimes they get very short time to enter data, which is not enough, and they may enter wrong details. We need to check this”

- a block-level functionary.

As was mentioned earlier, functionaries at the frontline emphasised that they did not entirely trust online data yet as it could have been entered in a hurry or before verification in the school itself. They trusted offline data maintained in records more.

This finding brings to light the urgent need to ensure that a common trust is built across all levels of governance. This trust can enable efficient use of the data collected and can be supported by random data verification checks to catch errors.

4.3 Use Cases for Data

4.3.1 Innovative Approaches at the State Level

The study found several great examples of the data collected on Shala Darpan being used in different ways at the state level. Figure 6 shows use cases that highlight the growing use of the portal.

One of the primary purposes served by the real-time data generation on the Shala Darpan portal was indicated as monitoring the present status of education in Rajasthan and tracking its progress over time. At the state level, the officials use the data to get a holistic picture of the different indicators of education. It was also indicated by mid-level educational officers in Rajasthan that currently the data collected on the portal is primarily used for monitoring purposes at the block level. The data is disaggregated at the school level and enables the policymakers to draw comparisons among the different districts, blocks, and schools. This helps the state-level authorities in identifying the different challenges faced by the school related to the availability of human resources, infrastructural status, examination performance of students, educational qualification of the teaching staff, attendance of teachers and students, etc.

PLANNING FOR TARGETED INTERVENTIONS

Data has been used to plan interventions such as the distribution of scholarships and incentives to students. Data helps with identifying students through their family details.

MONITORING EXPENSES OF SCHOOLS

Data is used to track how much a school is spending. Schools are given notices on unspent balances and releases of new transfers which are useful for the school administrators

GENERATING REPORT CARDS

Report cards for students in Grades 5,8 and 10 are generated on the portal. The portal also collects data on the Rajasthan Ke Shiksha Mein Badte Kadam programme.

RESPONDING TO PARLIAMENTARY QUESTIONS

Data is used to respond to parliamentary questions at Lok Sabha, Rajya Sabha and Vidhan Sabha on education.

LOBBYING FOR BETTER STATE BUDGETS

Data has been used to lobby for better state budgets by showing areas that require more funds such as infrastructure.



Figure 6: Examples of Data Use in Education

Identifying the schools which require additional assistance and support from the state, assists the state authorities in defining their priority areas which would include schools with a poor enrolment of girls, poor infrastructural facilities, and secondary and higher secondary schools with insufficient laboratory equipment or computer facilities. In an interview with a state-level official, it was mentioned that special focus is being given to the aspirational and special priority districts in education. These observations reveal that data is enabling the state to make targeted planning for areas that lag in certain key performance indicators. Similarly, in an interview with a Chief Block Education Officer (CBEO), it was indicated that the educational data collated on the Shala Darpan portal enables them to create block-level plans aimed at creating targeted improvements in education.

What was missing however in the current way data is being used in data is comprehensive data sharing (also referred to as interoperability) across departments. Functionaries in the education department currently did not have access to data from other departments such as health or social justice. They also did not share their data in a systematic manner with other departments. There is a need to improve data sharing as it can feed into making decisions that are transversal and cross departmental boundaries. This is elaborated further in section 4.3.3.

In addition to this, while the Shala Darpan portal has a significant amount of data, the same is not available for public use in an anonymised manner. This practice of open data can allow citizens and other interested actors to make a difference to the education sector by contributing to identifying challenges and building solutions. This is described in recommendation 7.4.2.

4.3.2 Limited Usage at the Frontline

Teachers however, shared that they believed functionaries above them in hierarchy did not use this data well. Functionaries at the block and district level resonated with this perception and shared that they used only limited parts of the data uploaded for monitoring progress daily and used other data to create annual plans for their blocks and districts.

4.3.3 Interoperability of Data: 3 case studies

With education being a multi-faced public service, the department realises that it is important to link educational data with data from other departments such as health. Convergence in planning and implementing provisioning of welfare schemes has many benefits that include saving time, reducing financing through pooled budgets and increasing the reach of the welfare scheme. It was communicated during the project investigation that recently few health indicators have converged with the educational data on the Shala Darpan portal for better provisioning of health services.

Case Study 1: Provisioning of Sanitary Napkins

The sanitary napkins are acquired by the health department and distributed by the schools. This proved challenging for the government to keep a last-mile track of the sanitary napkins and to ensure that napkins are being received by the targeted beneficiaries. To solve this challenge, the data on sanitary napkins from e-aushadhi and Shala Darpan portals have been integrated to keep a stepwise check of the distribution. Further, with the Shala Darpan portal providing data beneficiary-wise, an efficient tracking system can be ensured by keeping a per-unit account.

Senior functionaries from the education department shared that designing a system of convergence comes with many challenges. The first challenge is ensuring that officers responsible for different parts of the converged systems are present in common meetings. The second challenge is ensuring that a workflow process is designed such that each element of the service delivered is reflected on the portals. The third challenge is ensuring this convergence takes place in short periods of time that are decided by the senior officers at the state level. These challenges require officers to move past the silos of departmental functioning which requires a clear redefinition of roles and responsibilities. Officers were of the impression that this process could be refined more going forward – from planning at the state levels to implementation at the schools.

In the case of distributing sanitary napkins, functionaries from the education department shared that they find the process of defining workflow complicated. The lack of a clear workflow has resulted in lost packages that never reach the schools. After letters are issued by the health department about the packages that are meant to be delivered, schools end up receiving a different number of packages. The lost packages are not recorded on the portal as they are not tracked based on their batch code. Officers felt helpless as they could not deliver what was intended for the schools. **This case shows us that a clear identification of the workflow can enable identifying where the circuit of responsibilities break – which in turn helps in identifying and improving accountability in delivering outcomes.**

Case Study 2: Distribution of folic acid tablets

The folic acid tablets are procured by the health department, collected by the school authorities from the Public Health Centres (PHCs), and then distributed to the students. The procurement data is available with the health department while distribution data is being collated by the educational department – this often creates information silos with packages getting lost as a result.

It was stated during an interview with a state-level functionary that while tracking the distribution of folic acid tablets, it was discovered that large groups of students in a few schools did not receive tablets. This was because data entered on Shala Darpan did not have details of how many tablets were received and distributed. It was found that with the PHCs running out of folic acid tablets, few schools would not receive enough tablets. However, post-convergence, student-wise data has been collected, ensuring accountability per unit of the tablet and appropriate targeting of supplies. Now, a letter is issued by the health department about how many tablets are assigned to each school. Once the tablets are delivered to the school, teachers enter details on how many were received and distributed on their e-aushadhi login which is linked to Shala Darpan through API. Teachers also maintain this record offline in registers.

However, here too packages are still not tracked by their batch codes which can result in lost packages. The letter issued by the health department is also still offline and not integrated into the portal till the teachers enter the details. An automated system of entering the number of tablets assigned to each school would enable both the health and education department to track the tablets.

“With the procurement data being collected by the health department and the distribution data being collected by the school. It was very difficult to track if all the tablets reached the girls. Now by integrating, it is easier to monitor”.

- a state-level officer

Case 3: Best practice of improving interoperability of Shala Darpan with Mid-day Meal

The data collected on the Shala Darpan portal has been beneficial in provisioning COVID-19 vaccines to school children during the recent pandemic. Apart from the educational data, the portal contains household data related to school students such as Aadhar numbers, addresses, parents' occupations, and mobile numbers. With students often forgetting the required information at the vaccine centres, as pointed out by a state official, the information was used from the portal to provide vaccines to students in the vaccine centres.

Apart from vaccines, the Shala Darpan data was also used to provide dry ration post the COVID-19 pandemic. With schools reopening after the pandemic, a sudden spike was experienced in the enrolment numbers in government schools. This would often create a mismatch in the Mid-day Meal (MDM) provisions provided to the schools. Thus, a decision was made that the MDM portal would be converged with the Shala Darpan portal, which enabled the state in keeping a track of the increase in enrolments and subsequently provide an adequate amount of MDM provisions to the schools.

“Earlier MDM data was collected through MDM message counter. When schools reopened after covid, there was a sudden increase in the enrolment number. We were finding it difficult to keep a track of the MDM provisions. Then enrolment data was reverified with Shala Darpan portal. Now, we have integrated both portals for better tracking”

- a state level officer.

Functionaries at the state level were cognisant of API enabled convergence that could happen with other departments such as Department of Labour or Social Justice depending on joint work that is currently underway. According to them, it is important to first define the flow of goods and services which can then be tracked using the different platforms integrated through API. They also recognised the need to monitor and evaluate the functioning through random checks at regular intervals which does not currently take place.

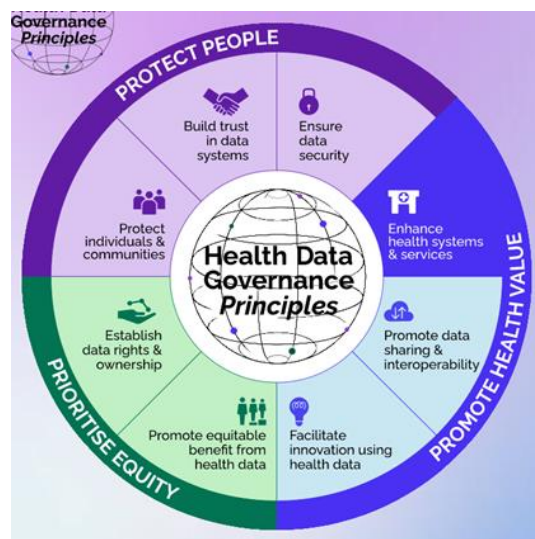
PSP Portal

There were broadly 2 main challenges that came out from the interviews with respect to the PSP portal. First, since non-state actors do not upload information regularly the platform has incomplete data sets. Second, challenges such as server crashes were common. The cell shared that they felt short staffed to manage these glitches while also completing their other daily work. As a result, functionaries from PSP Cell shared that they are currently unable to use the platform for efficient decision making. Users of the platform believed it had potential for improvement along the lines of making the user interface more accessible and ensuring there are better on-boarding programmes for non-state actors.

CHAPTER 5: RESEARCH FINDINGS AND CHALLENGES OF DATA IN HEALTH

5.1 Introduction to Data Use in the Health Department

Digital initiatives across a wide range of publicly provided services are rapidly gaining importance for tracking progress, monitoring the performance of workers, assigning resources where needed, etc. However, the potential benefits of strengthening health systems and improving health outcomes through digital initiatives come with a concomitant need to implement strong data governance structures and ensure the ethical use of data ([Tiffin, George, and LeFevre, 2019](#)).



Source: Health Data Governance Principles

Figure 7: Health Data Governance Principles

The [Health Data Governance Principles](#) present significant benefits of health data, while also providing guidance on protecting citizens from data misuse, as well as prioritising equity, something that Rajasthan has focused on through its various state-level initiatives. International evidence has shown that health information systems are fundamental to delivering quality health care ([Chaudhry et al., 2006](#)).

Reliable health data are the foundation of credible evidence of the health status of patients, and can assist policymakers in making decisions and plans to improve healthcare services delivery for better health ([Haux, 2006](#)).

Unfortunately, developing countries face challenges in implementing and utilising health information systems, including concerns about the quality of data and inadequate communications technology, human expertise, and financial resources ([Williams and Boren, 2008](#)).

In Rajasthan, the main goal for the health department is to provide quality health services at the lowest possible cost with the widest coverage.

To this end, the health department is wide and operates on a range of services from the provision of free medicines, diagnostics, health insurance, and regular screenings and check-ups. The state has made significant strides towards making data collection simpler and more comprehensive to improve scheme coverage as well as quality. This has necessitated investments in digital architecture.

The Department of Health and Family Welfare in Rajasthan is implementing various health policies, schemes, and programmes. This includes, among others, expanding and maintaining health facilities, the provision of medical supplies to them, and implementing various schemes to improve access such as the provision of health insurance and free medicines. The Department has significant breadth and has several different bodies working on various aspects of health service provision. These include the Core Health Department (responsible for facilities and provision), the Rajasthan State Medical Corporation (responsible for procurement), the Rajasthan State Health Assurance Authority (responsible for health insurance), and the officers working on the National Health Mission (NHM), a Centrally Sponsored Scheme. NHM alone has 22-23 divisions. There are 14 vertical divisions, 6 Reproductive and Child Health (RCH) divisions, HSS divisions, NUHM, NQAS, the civil wing which oversees 14,000 sub-centres. Then there is the Finance division part of NHM as well.

At present, Rajasthan operates at least **48 different applications for the health department**. Each generates a large amount of data to track progress on various schemes and indicators. Two things have necessitated this widespread data collection. First, the health department itself has tremendous breadth as mentioned before. This breadth itself has led to the proliferation of applications, given the range of issues covered. Second, Rajasthan is a large state, and has taken the lead in launching new programmes such as the Mukhyamantri Nishulk Nirogi Rajasthan Yojana for free medicines and the Chiranjeevi Yojana, a precursor to the Pradhan Mantri Jan Arogya Yojana, the Government of India's health insurance scheme.

Data is used for everyday governance across levels, especially at the department level. Data is used for tracking scheme implementation progress, the status of new and old initiatives, coverage, tracking treatment received by patients, monitoring the performance of functionaries, payments received by functionaries and eligible people under various programmes, outcomes, etc.

There is a clear need for systems within the health department to be integrated in a way that allows policymaking and implementation to achieve the goals of the health system i.e., universal health coverage. Given the multiple divisions and silos working within the health department, data systems should be designed in a way that allows officials to track an individual and their health-related history, or the extent of all

services provided at a facility or a given geographical area, or the performance of functionaries across the spectrum of their responsibilities.

As the remainder of the report will show, that while a significant amount of data is generated and used by the health department, what is missing is a data architecture that is able to integrate and generate insights using multiple datasets across departments.

Data generation can be understood by following various departments working within the wider ambit of the health department. These include Health and Family Welfare, the National Health Mission, Medical Education, the Rajasthan State Medical Corporation which handles procurement of medicines, and the Rajasthan State Health Assurance Authority, in charge of Chiranjeevi.

5.1.1 Examples of data systems in Rajasthan

A few examples of data systems in the state include:

Pregnant women and Child Tracking System

The Pregnant women and Child Tracking System (PCTS) is the main software which covers Reproductive and Child Health, including pregnancy Antenatal Care (ANC), birth, 45 days follow-up, immunisation, etc. Data is primarily collected through Frontline Workers (FLWs) and functionaries at villages and health facilities. There are 3 forms for reporting – form 6 for Sub Centres, 7 for PHCs, and 8 for Community Health Centres (CHCs) and District Hospitals (DHs). Data for each month is updated by the 10th of the next month. By the 5th the data is entered. Officials and functionaries in the Demography Cell analyse the data and check for errors, with an aim of fixing these within 3 days. The data is locked after that.



Chiranjeevi



For the health insurance scheme, in fact, all data is online. Rajasthan started its programme in 2015, before the launch of the Pradhan Mantri Jan Arogya Yojana (PMJAY) by the Government of India (GoI). Rajasthan has been continuing with its own systems, which as per officials, have matured. All structures, tenders, packages, etc. are different from GoI.

This ranges from the pre-authorisation request, hospital admission, documentation required while in the hospital, discharge-related information, and payment-related information.

The architecture has many data layers – the presentation layer, service layer, data access layer, etc. The presentation layer includes the Beneficiary Identification System (used for registration and identification), the HEM or hospital empanelment module, the Transaction Management System (TMS) which is used to identify beneficiaries and manage claims, and the SR or service request module for hospitals, doctors, or for packages. These help in making a life cycle for the department. There are several systems which are online. What gets reported to higher authorities is also from here. The fraud team also uses this data to check for abnormal or outlier patterns.

In Rajasthan, when a claim is filed, a Third-Party Administrator (TPA) analyses and processes claims, and then the payments are processed. The TMS is used to identify beneficiaries online. This process is linked to a citizen's Jan Aadhaar¹. The MO can also authenticate beneficiaries if the above process doesn't work. At hospitals, eligibility is checked using the BIS. After diagnosis, the packages are booked in the TID. Once the citizen is discharged, then the claims are booked.

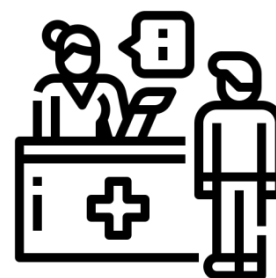
For treatment, many records are kept. The process involves the use of an MDP or Minimum Document Protocol. This captures the details of the treatment. The details of the doctor providing treatment are also captured. These documents are verified by the TPA. Typically, most claims are approved. The rejection ratio is only 2-3%.

Rajasthan uses the Jan Aadhaar (based on Aadhaar) database to bring people into the scheme. Every citizen has a Single Sign On (SSO). People can use this to log in and even sign up for the scheme if they are not already identified through PMJAY.

This linkage allows all details to auto-populate the database. Data is also pushed back to the Jan Aadhaar database. This is data on the packages claimed, etc. If people are not able to do it themselves, they can go to an e-Mitra service centre for assistance. Citizens are sent alerts using SMSs, which are also based on the data system.

e-Aushadhi

Rajasthan's main procurement agency, called the Rajasthan Medical Services Corporation Limited, oversees procuring medicines under the Mukhyamantri Nishulk Nirogi Rajasthan Yojana. The corporation collects data on the expiry of medicines, and regularly checks stocks.



The entire process pertaining to procurement under e-aushadhi is online. Right from getting the demand (right from PHC and everyone fills

¹ Jan Aadhaar is a unique identification number for each resident in Rajasthan. It is based on a person's Aadhaar number. Having a Jan Aadhaar number is mandatory to avail any publicly provided service in Rajasthan.

in the annual demand, compilation is done by the department). The provision for supplementary demand is online as well now.

Technical evaluations are online as well. Suppliers upload documents which are then verified by the procurement section. This application provides detailed information from procurement of drugs to consumption by the end user. Complete tracking is done through the application.

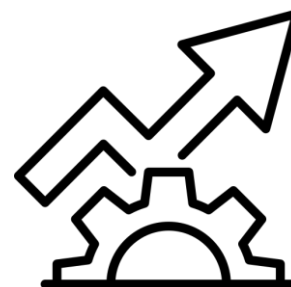
One critical learning from these databases is that it is possible to create fully online systems, which can accommodate data from many stakeholders including citizens, functionaries, and facilities. These systems have also created monitoring mechanisms that allow decision-makers to have an overview of progress made, as well as rapidly respond where required. The sharing of data as well as analysing multiple interoperable datasets is also facilitated using fully online systems.

5.1.2 Examples of data use in the health department

There are several ways in which data is used. Some examples are given below:

Improving programmes

Data collected under Chiranjeevi is used to update benefit packages². Data on the utilisation of various packages are analysed and this data is available to other stakeholders as well such as the insurance company and the TPA. Some examples of using this data include a) checking if a package needs to be added or removed based on usage; b) reserving certain packages for only government facilities or making certain packages open as well. For example, c-sections are reserved while Total Knee Replacements have been made open for hospitals accredited by the National Accreditation Board for Hospitals; and c) if a hospital feels the rate for a particular package doesn't cover their costs, then this data is used to change package rates.



Utilisation data has been used to increase coverage of the scheme. PMJAY covers a set number of people³, and coverage has been expanded by Rajasthan under Chiranjeevi using scheme utilisation data.

² Health benefit packages are a list of what is covered under the scheme including which surgical procedures and treatments.

³ This was based on the Socio-Economic Caste Census or the Rashtriya Swasthya Bima Yojana (the predecessor to PMJAY). As per the [PMJAY dashboard](#), 58,95,363 families in Rajasthan were covered. An additional 75,02,474 families have also been covered by the state.

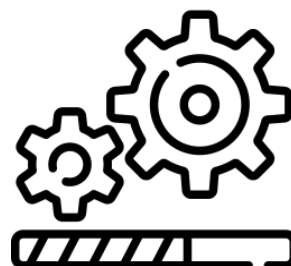
Detecting fraud



Various data triggers are used to detect fraud under Chiranjeevi. For example, if one package is used more in one hospital compared to other hospitals; if eligible citizens are accessing a hospital from only one adjacent area instead of all nearby areas; if the same family is visiting repeatedly, etc. Officials check for patterns to detect anomalies in data to pinpoint potential fraud cases.

Tracking progress

Due lists are generated under PCTS for each ANM, which have lists of registered women and children to check up on. This functions like a to-do list. These can be viewed by people operating above the ANM in the departmental hierarchy, which allows officials to track pregnant women. This is validated using Aadhaar cards currently, and there are tentative discussions around linking this with Jan Aadhaar for the next financial year. This would allow the collection of each person's history, which can be integrated with the database.



An example of data being used to plan activities, especially at the district and state levels, is as follows. An official stated that the data may show that children born and mothers giving birth during a particular time period may die more than children in other time periods. Officials use such analysis to then take decisions to reverse this trend.

There are discussions in sector meetings, and inter-block meetings as well. These include checking up on data collection and entry targets. The role of monitoring is played by the MOIC, DNO, and CMHO, among others.

Assigning human resources

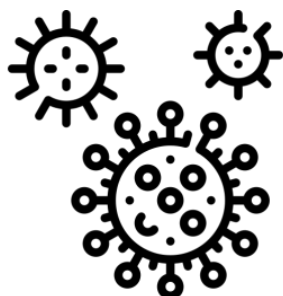
Data are also used to check manpower needs such as requirements for specialists like gynaecologists. Officials said they used data to shift people to facilities based on needs. From this financial year, the department is trying to collect data from private facilities as well, to map human resources with greater completeness.



Tracking disease outbreaks

The Integrated Disease Surveillance Programme (IDSP) has a portal called the Integrated Health Information Platform. This is a GoI portal that collects data in 'near

real-time'. Data is collected from various facilities to track diseases. The goal is to track outbreaks and stop their spread.



Similarly, officials tracking vector-borne diseases take some data from ISDP. This data is used to monitor and act. The data is also used to check if the activity has happened. If there is a mismatch between 'action taken' in the data and officials find out the team hasn't gone, then they send the team again.

Data is also collected and monitored to check for potential sources of vector-borne diseases. There are checks for larvae, number of containers, etc. as well as symptoms.

Monitoring patients

For Tuberculosis (TB), an application called Nikshay is used. A patient who is suspected to have TB (not just those diagnosed), is enrolled in the database. Personal information is gathered, and relevant investigations required are mentioned. If any TB is detected, the patient is moved from the enrolled stage to the notification stage, as TB is a notifiable disease.

Other information like comorbidities like HIV and diabetes are also listed. Family members' details are also added to prevent TB. GoI scheme – for Nikshay POSHAN yojana – where money is given to patients (₹500 per month). There is a system to monitor patients daily.



“We should have diabetic status of 100% [all] patients – we talk to them [district officers] if the percentage is below 90%.” – Official, Health Department

Similarly, for Rashtriya Bal Swasthya Karyakram (RBSK), monthly reports are generated to track screenings, diseases, on-spot treatment, operations, follow-up treatment, etc. Children are identified and 40 different diseases (congenital, childhood, and adolescent diseases) are tracked.

Tracking stocks

The Rajasthan Medical Services Corporation Limited collects data on the expiry of medicines, and regularly checks stocks as a part of the e-aushadhi system. Data is available at both warehouses which stock medicines and facilities where the medicines are distributed to citizens. Official's direct warehouses to not only check expiry dates of their own stock but also for other warehouses, in case there is a



shortage that can be filled from an adjacent warehouse. When one warehouse raises that they have an excess of stock, and one raises a shortage, then it is approved by the supply section and then an inter-warehouse transfer happens for the stock of particular medicine.

Tracking performance

Data are also used to track the performance of various FLWs and functionaries. One software used for this is ASHASoft. In Rajasthan, ASHA workers are paid approximately ₹3,625 per month (with ₹10-20 variances based on experience). Several activities conducted by ASHAs are incentivised, and ASHASoft is a DBT (Direct Benefit Transfer) platform which releases incentive funds. Officials use ASHASoft to analyse the number of activities conducted in a given time period to understand ASHA performance.



Another example of facility performance is when the medical education department takes data from the SHA. They take data on medical colleges and medical education hospitals. This is data on claims, the number of cases covered, the payments made to the hospital from the claims processed, and so on. This data is used to assess the performance of the hospital on two counts. Revenue generation for the hospital, and coverage at hospitals.

Feedback for programme managers

Under NHM, data for gap assessments from time to time is collected real-time. Assessments are done by the demography cell, and they do RCH analyses monthly. There is a report generated every month, analysed, and feedback is given to district-level programme managers. Each of the vertical programmes under NHM also analyses data actively and carries out innovations.

One example of this is the use of Nikshay, the application for TB. Data is collected on patient bank details, payments made, and so on. Officials at senior levels track the completion of data entry. So, if not all bank account details have been entered, or not all payments are made, officials can reach out to officials at lower levels and check in on issues. The data is also used to



resolve issues such as a perceived funding crunch. Some district officials mention that there is no money to send for patients, but officials at the state level ask them to

upload the details on the PFMS system regardless and let other officials realise there is a funding crunch.

“This is used as a progress tracking, monitoring, and feedback tool. If district data is poor, we open block data, and talk to them.” – Official, Health Department

Health-specific messaging and communication with citizens

As per an official, there is a lot of good data available – area-wise use, trends, etc. During the COVID-19 pandemic, the department checked village by village the number of vulnerable people who had comorbidities. Officials sent SMSs to people as well based on this.



The next section looks at challenges for the health department, across generation, usability, as use, as well as policy improvements.

5.2 Data Generation

5.2.1 Data is collected both online and offline



A vast majority of data collection is online. There are several applications and dashboards which are used to generate, verify, collate, and use data. The main push for online data entry via applications, especially mobile applications, is to reduce effort. As per officials, Front Line Workers (FLWs) using applications such as ASHAs will be able to collect and upload data and save time previously spent on collecting information on paper and making a trip to deliver the data to the sector level or PHC level

to the data entry operators. There is also an ongoing attempt to integrate identity fields (personal information of citizens availing services) which is expected to reduce effort further as well.

“Paper requires two people to do this, but an app saves manpower.”

– Official, Health Department

Data is mainly collected by FLWs such as the ASHA and ANM. Some examples are given below:

FLW	Applications used
ASHA	ASHA soft, Nirogi Rajasthan, PMJAY registration, and Aadhaar face ID
ANM	Health Digital App, Aayushman Bharat, PCTS-ANM, IHIP- daily reporting, Antara Portal, the Non-Communicable Diseases (NCD), Nirogi Rajasthan, and FPL MIS (family planning)

However, offline records are still maintained across levels, despite the move to online data collection. As per FLWs, offline records are kept in case older data is requested, for when a senior functionary visits for monitoring, and entering data in applications at a later point of time if the FLW cannot enter it immediately. An FLW did mention that online collection has started recently, but the shift has not been completed yet.

An official also stated that still some kinds of data are entered offline, and such a hybrid system does not help. Procurement is done at the district level or state level, and there is both online and offline registration at the moment, which officials questioned the need for.

When inquiring further, some officials mentioned that offline data entry is necessary as there could be internet issues, or servers could be down, and it functions as a backup. Officials also stated that offline data is used to match with online data and check its veracity. There is also a cultural reliance on offline data in the form of papers and registers.

“Hum kagaaz ko maante hain”

- Official, District level

5.2.2 Duplication of entry



Another challenge is that of duplication of data collection and entry. One key example of this is the repeated entry of identifying fields. These are data of citizens accessing schemes, which must be filled across multiple portals. For instance, if a woman is pregnant, her details will be entered in PCTS. But if she develops Tuberculosis later, her details will have to be filled again in the application for tracking TB patients and payments for them, called Nikshay (run by GoI).

“...issue not of many apps for ASHAs. But every time she goes to the field, she has to enter identifying fields again and again across apps. For example, delivery and birth and both get TB – so has to be entered both times in Nikshay.”

– Official, Health Department

“ASHAs and ANMs collect similar data – like ANC and immunisation” - ANM.

“The target is that there should be no data redundancy and that there should be accuracy. Having to fill the same information like personal information again and again is time consuming and can also lead to mismatches and errors. For example, spelling mistakes can happen, or other data entry errors can happen. So, using Jan Aadhaar to pull information is being done.”

- Official, Health Department

Box 1: RBSK identifier data has to be entered repeatedly

Under the programme, children are screened for illnesses, referred to health facilities, and treated. Children are examined in schools and Anganwadi Centres. One challenge is that identity data must be entered again, even though this data is easily available to the education department.

5.2.3 Missing data restricts efficient targeting and tracking

An official described the need to have information on the Couple Protection Rate (CPR), which is defined as the total number of couples using any method of contraception divided by the total number of eligible couples in a given area. However, information on the number of eligible couples is not collected through the PCTS portal, which makes calculating CPR impossible. This implies that measuring progress against targets towards population control is difficult.

**Box 2: A lack of data on Incidence**

Another example is data on prevalence and incidence. The former refers to both new and old cases, and the latter refers to new cases only. While total cases are available monthly, new cases are not. Even subtracting total cases across months does not indicate how many cases are new and how many are old. This holds true for data on the coverage and provision of condoms and oral pills.

ANMs get a target at the start of each financial year for covering a set number of eligible couples. PCTS has data on total prevalence, but not incidence. For example, it shows in April that 40 couples are using condoms. It could be the same next month, but it is possible that 5 have stopped using condoms, and 5 new couples are using condoms. So, one cannot determine how many, and which newly eligible couples need to be reached for discussions on contraceptives.

A related challenge is that of **definitions used by the systems**. Data on hysterectomies conducted is unavailable, largely because GoI does not consider hysterectomy as a form of sterilisation. Therefore, it is not included while calculating CPR. This restricts tracking progress on population control.

“GoI wants the Net Reproductive Rate is 1, or TFR of 2, or CPR of 60%. If we don’t know CPR, then how can we track progress.”

– Official, Health Department.

Another example of missing data is financial data. Officials conduct financial reviews of districts, but month-wise spending data is not always available. However, most

spending is booked in February and March as per officials. So, if there are requests for additional funds, it is hard for officials to determine how much has been spent and on what.



Currently, some kinds of information are not available on e-Upkaran, as it only manages bio-medical inventory. An official stated that even things like furniture should be included, similar to private facilities. The officials further stated that bio-medical equipment should be one set of inventory and non-medical another, and total equipment and others should reflect online with the current status (which should include all inventory available). This restricts the scope of inventory management, and therefore, slows down decision-making on treatment options for patients, if a patient has to be transferred from one facility to another. This can be risky, especially when some patients require urgent care.

5.2.4 Lack of outcome data

Similarly, as per officials, tracking outcomes is not straightforward. Data is generated on service provision, or coverage (how many people are given a certain pill or treatment), but not outcomes, which makes it harder to understand the impact and effectiveness of programmes. Understanding the impact of various provisions would help in prioritising resources and allocating funds in a more efficient manner.

“We can currently track how many patients came in and how many were screened, but we cannot see what the outcome was.”

– Official, Health Department

5.2.5 Lack of real-time data

Data is collected in real-time, daily, weekly, monthly, quarterly, yearly, and on an ad hoc basis. Some examples are given below:

Frequency	Examples
Real-Time	e-Aushadhi medicine stocks, Chiranjeevi coverage and uptake
Daily	PCTS data for RCH, immunisation, line listing, OJAS, ASHA soft, Tuberculosis (TB), etc.
Weekly	IHMS
Monthly	PCTS MPR, family planning
Quarterly	Antara (for TB)
Yearly	Public service guarantee, Citizen IDs

As shown above, some data is not real-time, and officials have stated that having access to real-time data is absolutely necessary. Some data comes in within 1-3 days,

or in some cases, monthly. For instance, officials said that during the COVID-19 pandemic, it was found that many hospitals only had 30 beds available but 60 patients. If the data had been available in real-time, officials could have sent patients to hospitals with available beds. The absence of real-time data, therefore, affects citizens directly.

Big real-time data can also help by shortening supply chain steps, reducing risk, and the sustainable planning of smart healthcare investments. Flexibility, driven by timely big data feedback, contributes to reducing the intrinsic rigidity of long-term healthcare investments (Visconti and Morea, 2019). *“All data should also be real time. Every single bit of data. The policy is there for this, implementation is the issue.” - Official, Health Department*

5.3 Data Usability

5.3.1 Errors and verification

Errors and unclear verification reduce trust in the system, and can lead to poor decision making. The lack of trust stems from the fact that there are errors in data despite an extensive verification process.

While officials clarified that processes are in place to verify data, some felt not all data is checked, and there are concerns about quality. The process for PCTS data checks is shown in Figure 8 below.

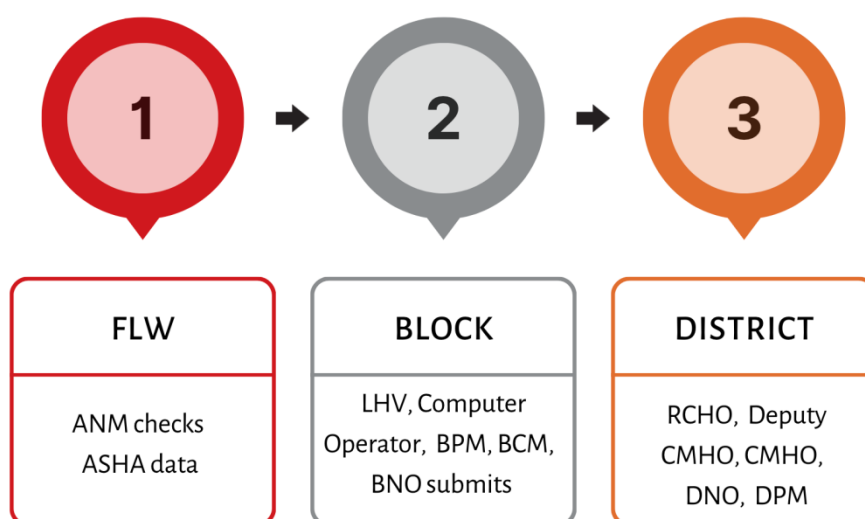


Figure 8: Data check process for PCTS

“Monitoring is good. Blocks and districts see their own area’s data and can track progress.”

– Official, Health Department

Officials keep a track of how much data is submitted, and if it meets targets set for collection. For instance, if an ANM had to submit data on 25 people and has not, officials reach out to her if less than 90% has been entered. An ANM said a WhatsApp group was created to facilitate checking data entry targets, and a notice is only issued if there is a bigger issue. An ASHA reported that there is pressure to submit data on time and a computer operator said that if ASHAs don't submit data on time, their payments are halted.

Functionaries reported that officials at higher levels do check the data and get back if there are any errors found. Data trends are also checked to see if there is a sudden jump, or an unexpected change, which is investigated. Officials say that 600 indicators are checked, including major indicators which are checked every 10-20 days. As per PCTS officials, there is a spot check of 10% of data for discrepancies.

One official stated that sometimes the officials filling PCTS data and even other officials don't always trust the data. An example was mentioned for this, which is that when others ask for data at the district level, requests are always redirected to the CMHO, which could imply a lack of trust.

“If any question is asked in the Vidhan Sabha regarding that data, then officials say that this data is not ours. The CMHO hesitates to answer the question, because we do not trust the data. Which means data is not correct.”

– Official, Health Department

Similar verification issues hold true for e-Upkaran⁴, which keeps a record of inventory at hospitals. As per officials, some entries are not correct. For example, the X-ray machine specification might be 100 mA, but written as 300 mA. This may not have been verified by the technical person, as per officials. Even though these are minor issues, they must be addressed.

5.3.2 Data discrepancies

Despite robust processes, there are implementation gaps, and therefore there are discrepancies in data. One official stated that data checks are not thorough enough, and often what is entered is uploaded directly. An official pointed out that Antenatal Care (ANC) data did not match the number of ANCs and said that it is possible that ANC data is faked to achieve targets.



⁴ A web-based application which deals with the management and maintenance of equipment and instruments to be used at health facilities from one platform.

Another official pointed out that a person may have registered their pregnancy in the district in which they currently reside but went to another district for delivery. This could be because they may have visited their home district or visited a hospital with better facilities in another district. Since the birth is not registered in the same district as the pregnancy, it means that data for some districts show a higher number of deliveries than pregnancies (for example, 100 pregnancies and 103 deliveries). This is an issue across districts.

Box 3: Duplication in reporting

An example of a different kind of error is duplication in reporting. For instance, sterilisation is done through static centres and fixed-day services (camps for sterilisation at PHCs). Fixed-day services cannot be provided below the PHC level. However, there are instances of ANMs reporting sterilisation in form 6, which is for Sub Centre level.

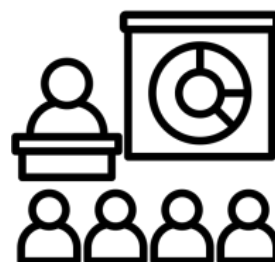
In form 6 for PCTS, there is an option for the number of cases done through fixed day services, other than fixed day services (or static centres). If the ANM has supported two couples with sterilisation, she lists the same in form 6 for Sub-Centres. The facility where this was done, also reports 2 couples as sterilised. Therefore, there is a duplication. The payments have been made for 2 couples, and data is uploaded for 4.

“There are checks at every level. There might be sly behaviour somewhere.”

– Official, Health Department

5.3.3 Capacity gaps

Even if data is verified, and ANMs are told about this, some ANMs will rectify it, but many will not. Efforts via training to avoid this situation have limited success. An official pointed out that it is hard to stay very focused in a training session, and a person’s mind can be easily diverted and therefore miss out on important pointers.



*“The problem is that people don’t understand this.
The ANM is probably 8th pass, she’s doing her best.”*

– Official, Health Department

A lack of resources, both human and physical, affect the generation and use of data. Furthermore, glitches in software affect service provision directly. These are discussed below.

5.3.4 A lack of human resources

Vacancies were a challenge listed as a lack of human resources can increase workload at times. Given the scale of data collection, subsequent analysis, and report generation for 48 different applications, there is also a dearth of Data Entry Operators. An official also pointed out that the lack of functional machines is also a challenge, which hampers work.

Another example of this is data for the Electronic Detailed Accident Report (EDAR). This application has only 1,100 hospitals registered in Rajasthan, and the process is still ongoing, as per officials. Officials mentioned that the delay in hospital registration is due to a lack of awareness and the need for dedicated human resources for this task.

The lack of human resources and increasing workload with new programmes and applications has increased the strain on FLWs. An official stated that functionaries simply don't have enough time. For instance, when FLWs have meetings, or when a supervisor or mid-level manager visits, their other work (such as counselling, door to door visits, referrals, providing services at Mother and Child Health and Nutrition Days, etc.) takes a backseat. Often, this happens as different officials visit on different days, meaning the FLW in question has even less time for regular activities, including data collection.



“FLWs don't have time. They are caught up in VCs, meetings, and filling application data. My position for monitoring is that all officials should go together, not on separate days. They go on separate days, so when will the ANM work.”

– Official, Health Department



Another instance is that ASHA workers, who are entering a lot of data on applications, may even go beyond their immediate mandate. For instance, the PMJAY application used for e-KYC (Know Your Customer) for people, is operated by ASHAs for which they receive some financial incentives. However, financial incentives are limited to some applications and not all. This is substantiated by incentives for regular activities, such as institutional deliveries, sterilisation, etc. but the workload has only increased with the addition of more applications.

“In some areas, ASHAs have to climb a mountain to go to 10 houses, and then we are giving them another application. Their lives are being made difficult by so many applications.”

– Official, Health Department

There are time limits to enter the data, after which data entry is not possible. An ANM described the burden as follows: *“I can either do field visits or enter data, but I cannot do both”*. FLWs said that while the collection is easy, data entry is hard and time-consuming. Some ASHAs request their husbands to help out with data entry as well. One ASHA said that she maintains offline records, while her husband enters data on the phone.

A Computer Operator also said that sometimes they have to work extra to get all the data entered. Officials like the Medical Officer-in-Charge (MOIC) also must take out time from their schedules for monitoring.

5.3.5 Issues with software affects both citizens and functionaries

Glitches and issues with software can have a real impact on citizen entitlements. For instance, part of the TB control programme is payments for citizens, which are routed through the Public Finance Management System (PFMS). Each patient gets ₹1000 when they enrol, and ₹500 every month. As per an official, one challenge is that sometimes PFMS shows that payments have happened when they have not. This leaves some payments stuck for months, and officials at the state level must contact GoI. However, as per the same official, some progress has been made towards resolving this issue in the second version of the application.



Another example of a software issue is that several people’s payments happened to a single bank account. Now while each account can only receive 1 payment. But this has raised another issue, which is that if a husband and wife share a joint account and both have TB, only one person’s payment can be made to that account. Similarly, if someone has recovered from TB, and gets TB again, then making payments again to them is an issue.

Another software issue is that the RBSK portal (as on 4 January 2023) did not allow for data on more than 120 children at a time. There is a plan to expand this to 170, but at a time some children get missed out.

5.3.6 Infrastructure

Network issues were reported by both FLWs while submitting data, and by computer operators and other functionaries at the block at the district level. The server of the website also goes down sometimes, causing delays.



Another challenge is that the government has not issued phones or provided network plans for the same. An ASHA worker said that given the salary (around ₹3,500), it is difficult to both buy a phone and a network plan. These additional costs on the FLWs further the feeling of the work being burdensome. One ASHA said that with the low salary and having to buy a phone, making ends meet is difficult.

Another mentioned that the government had discussed providing phones to FLWs. This process must be hastened, and network charges should also be covered by the government. Another issue is the lack of network or software glitches that arise in day-to-day functioning. These can be resolved via regular checkups and improving existing software.

5.4 Data Usage

Interoperability across datasets has great value, especially in a vast data system such as that of the health department, which is spread out. There is great scope not just to link data across datasets within the department, such as patient health records to form EHR as mentioned above, but also with other departments. There are several examples of data being made interoperable, but there is also scope for a lot more, which can streamline systems, and substantially improve decision making and delivery. One challenge to building an interoperable system, however, is the sheer number of applications in use, and the fact that they do not speak to each other.

5.4.1 Multiple applications and siloed functioning

The health department, as mentioned earlier, has over 48 applications. Some officials feel there are too many applications. There are 4 reasons for the breadth of applications. First, the health department itself has a large number of programmes. This requires the use of different portals, as per an official. Second, both GoI and the Department of Health run parallel applications, and that has increased the number of applications. Often, these run parallel to each other. For example, e-WIN by GoR, and CoWIN by GoI, or PCTS by GoR and HMIS by GoI. Third, needs constantly emerge and evolve, with new schemes and programmes getting launched. As per an official, there is a rush to implement, especially for flagship programmes. So as a matter of

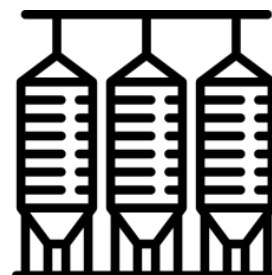
convenience, new applications are made, without necessarily looking at the bigger picture and needs across the department. Lastly, multiple officials stated that there are many applications as development partners make policies and accompanying portals, and convince senior officials that they are useful, and new applications are launched.

“We can't do everything one system, and we need some separate portals.”

- Official, Health Department

There are two challenges that emerge with this spread:

First, officials and functionaries operate in silos. Each sub-department or team uses a particular set of applications, which don't speak to each other for more comprehensive analysis. In this situation, ways to combine applications for efficiency are missed. As per an official, DoIT and NIC are there to provide support but still everyone directly reaches out to them directly and we work in silos.



Second, given the extensive nature of data collection, there is an increasing burden on FLWs. Multiple officials felt that the frontline is too burdened with data collection.

5.4.2 Applications are not created in a participatory manner

A related issue is that of building applications in a participatory way. Applications are not always made with the complete involvement of users – both at the frontline and the department level. Some officials pointed out an incongruence between the needs of the department and the current array of applications and data collected.



An example of this is that applications and formats are in English, even though FLWs are unfamiliar with the language and struggle to adapt. This was mentioned by FLWs and acknowledged by mid-level managers such as officials and functionaries at the block level.

“The formats are in English, and I find it very difficult to understand. Data collection is not an issue but entering it is”

– ASHA.

An official suggested that there is a need to solicit suggestions from field staff the next time an application is made, to understand the challenges they face.

Some software's have also not been user-friendly, as per officials. An example that was cited was that of the GoI NCD software (2018-19), which was for ASHAs,

ANMs, etc. This application required data to be filled in a particular way which proved to be time-consuming. As per one official, the computer expert team took 30 mins to enter data on one patient. It is likely that other states faced a similar issue with this application. Rajasthan ranked 4th in feeding information on this software despite only entering 10% of data on patients. Later, the World Health Organisation (WHO) made an easier-to-use application which was used in Churu and Bikaner, which has been adopted by GoI.

“Software should be user-friendly.”

– Official, Health Department

5.4.3 Examples of successful interoperability, as well as potential interoperability

Officials acknowledge the challenge with the lack of interoperability and have moved towards building systems which link data both within and across departments. Some examples are given below:

- Data from the PCTS, for example, are directly linked with the Government of India’s Health Management Information System (HMIS).
- Across departments, some data are linked, such as the e-Aushadhi portal with the Department of Women and Child Development for tracking the delivery of sanitary napkins.
- Another example is the Electronic Detailed Accident Report (EDAR). If any accident case comes up, the data is entered and helps officials find out which location is accident-prone and try and determine why accidents are happening (due to a blind turn, etc.). The e-Detailed Accident Report portal is used by the police, the Public Works Department, and the Medical Health Department. The Public Works department uses this information to put up better signs or improve road safety, and the health department can assess requirements for medical facilities like trauma centres, ambulances, etc.



Figure 9: EDAR as a cross-departmental platform

- The Handicapped portal is linked with social welfare and with CMHO data. For COVID ex gratia, data is linked between social welfare and medical health. Access is based on a user ID and password which has been provided.
- Birth certificate provision has been linked with PCTS data. The citizen is discharged with a birth certificate without a name, and they can use their mobile to update the name anytime in the first year of the child.
- Rajasthan has an electronic health record system, which is linked with [IHMS](#)⁵. Data from the SHA feeds into the electronic health record database as well.

The goal of data interoperability is to make work easier for officials. However, this has not always been the case for the Rajasthan health department. Some examples of a lack of interoperability and their implications are given below:

Box 4: Scope for interoperability

- An official said that for certain tasks, they need to open many different datasets. One of the issues with having several applications and dashboards while reviewing and monitoring data, is that officials have too many different IDs and passwords which they need to use, which is a tiring exercise. The official said that ideally, access should be in one place.
- For HWCs, officials stated that both progress tracking and performance tracking of functionaries is important. However, there are multiple portals which remain unintegrated, which implies that officials must collect data from multiple sources. Some examples of this are Nikshay and PCTS. Furthermore, there are 15 indicators on which CHOs (Community Health Officers) receive incentives. Officials said they need to look at 5 different portals to assess performance, increasing workload.
- Some datasets are unlinked, such as PCTS (which has data on children below 6 years of age) and Shala Darpan (the school education portal). This implies that data on children must be re-entered across datasets, increasing workload through duplication of entry.
- For DBT payments to citizens (for sterilisation, etc.), there is a separate portal. This portal is not linked with the physical progress portal under PCTS. This increases workload, and officials described this as a “major problem”.

⁵ Integrated Health Management System (IHMS) is a comprehensive software solution of Government of Rajasthan consisting of three components – Hospital Management System, Telemedicine and Wellness and Screening – that captures all health-related data and events of a person in the form of an Electronic Health Record (eHealth Record) which any individual / hospital would be able to access / generate using his / her Aadhaar / Bhamashah Card.

One example of interoperability is shown below (figure 10). Data from the education and health departments can be linked to create a unified database linked with personal identifiers with the use of Jan Aadhaar. This can serve both citizens in accessing all their information in one location, and integrating datasets within the health department can be useful for services providers like doctors as well as decision makers to obtain a wider view of health status of various groups and sub-groups, across the varying geographies of Rajasthan to better inform policy making and planning.

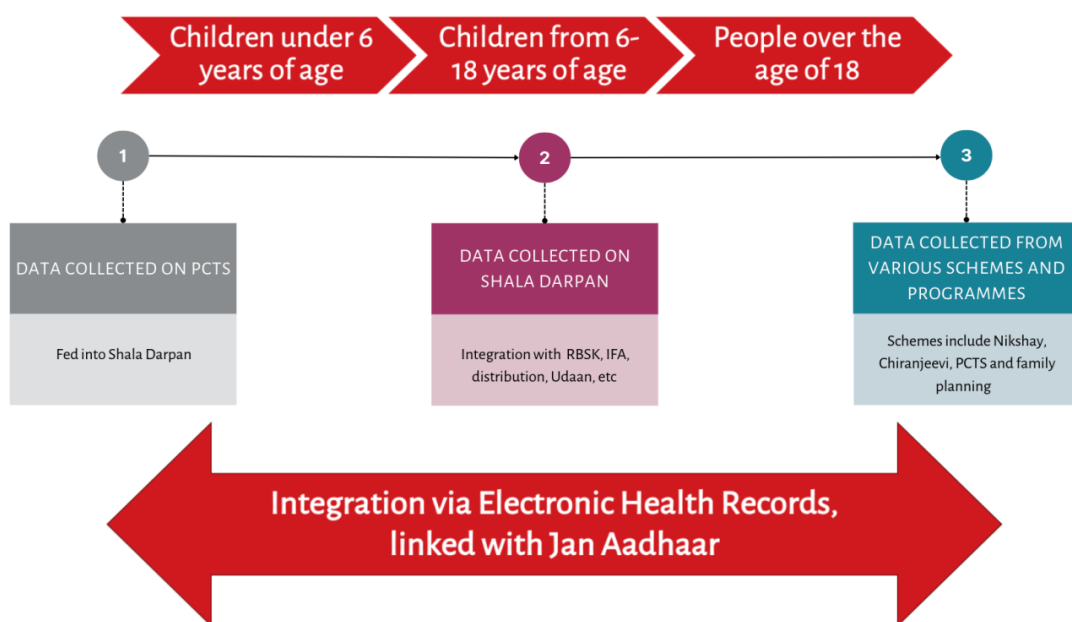


Figure 10: Potential for integration across the spectrum of health and education departments

CHAPTER 6: RESEARCH FINDINGS AND CHALLENGES OF DATA IN THE PANCHAYATI RAJ DEPARTMENT

6.1 Introduction

In Rajasthan, the Department of Rural Development and Panchayati Raj (RD&PR) oversees all rural development and local governance related undertakings of the state government. The department implements large centrally sponsored schemes such as the Mahatma Gandhi National Rural Employment Guarantee Scheme, Pradhan Mantri Awas Yojana (Grameen), National Rural Livelihood Mission and District Rural Development Plan among others. State schemes implemented by the department include the Mahatma Gandhi Jan Bhagidari Vikas Yojana, Regional Development Plans and Members of Legislative Assembly Local Area Development fund.

The RD&PR department is in-charge of overseeing the smooth functioning of the three tiers of decentralised governance, namely the Zila Parishad, Block Panchayat and Gram Panchayat. Like education and health departments, the RD&PR department consists of a secretariat at the state level and implementing functionaries such as Chief Executive Officers (at the district), Block Development Officers (BDOs) and Village Development Officers (VDO) at the village level. These functionaries are supported by clerks, Data Entry Operators, MIS Managers.

To ensure digital transitions, RD&PR has taken multiple initiatives in the last decade. Two parallel tracks of work have been undertaken. The first is work designed and funded partially by the union government and the second is that of the state government. The different schemes and programmes undertaken have led to the creation of a wide variety of applications and portals that are expected to support the monitoring of action taken.

The image below details the portals maintained by the state government:

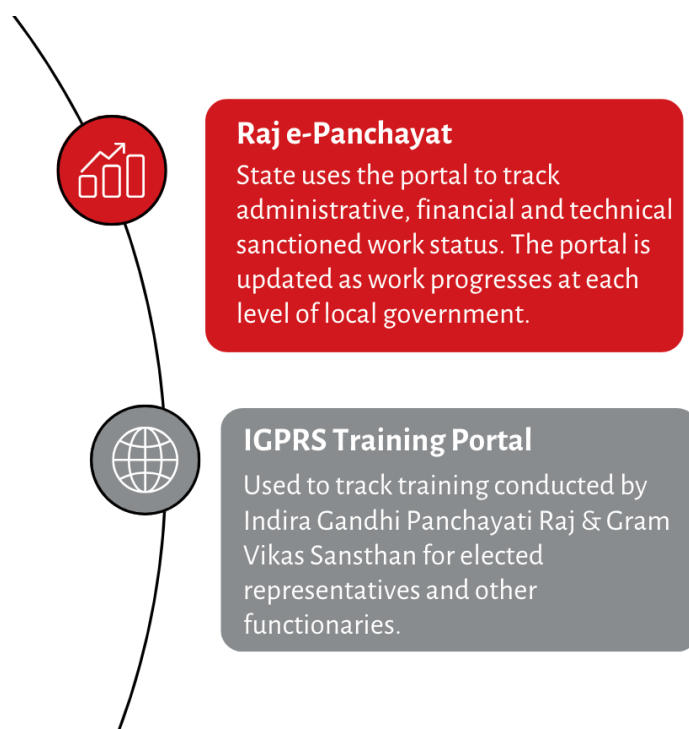


Figure 11: Portals maintained by the state government (different departments)

And this image below shows the different portals maintained by the union government:

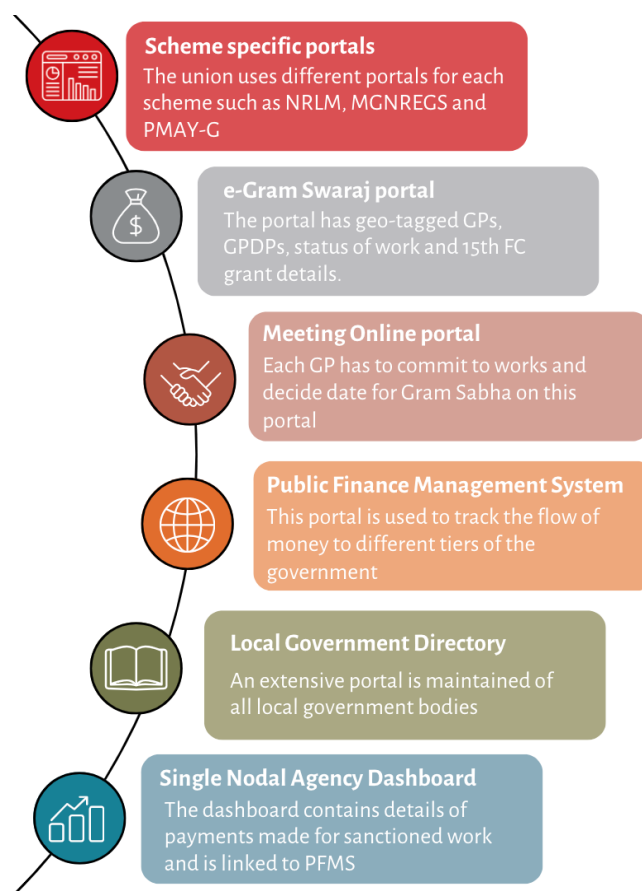


Figure 12: Portals maintained by the union government

6.2 Data Generation

6.2.1 Capacity gaps in data entry

Data is entered on planned works by the Village Development Officer (VDO) on Raj e-Panchayat and e-Gram Swaraj. Currently, data can only be entered on laptops **as the platform is not user friendly enough for data to be entered via phone.**

LDCs and Panchayat Sahayaks help in this entry at times, indicating capacity gaps. VDOs also utilise the plethora of e-mitra centres in the state to enter and upload data. As per an official, basic training on using computers is currently not delivered by the department.

Capacity gaps can be seen elsewhere as well. Some applications, like the National Mobile Monitoring System for MGNREGS which is run by the Union government, are updated frequently. The state has limited control over such applications. One official reported that functionaries struggle to fill the application as they have studied till class 8th or 10th only.

There is also the hindrance of poor internet connections, similar to data collected by other departments as well.

6.2.2 Data stored both online and offline due to a lack of trust in online systems

Hardcopies are maintained for a lot of data, which are the standard for functioning. As per an official, “data is not safe, and the portal is not successful. We still use hardcopies for everything.” For example, audits require hard copies. However, a functionary at the state level indicated that all data is only collected online, and audits are also conducted using online data.

6.3 Data Usability

Data usability issues are like those seen in the education and health departments: These are as follows:

- There are server issues with using state portals, slowing down work.
- Some information cannot be viewed in real-time, such as information on pending payments, which again affects the accuracy and efficiency of data use.
- An official mentioned manpower related challenge. Specifically, there is a lack of manpower who can help fix existing systems.
- A state level official specified that there is no dedicated verification process for data collected. The official specified that there should be a system put in

place for this. District-level should do the check. Block and GP-level should be doing data entry.

- **There are no graphs.** State level functionaries shared that they take PDF files that they get from the portal and convert it to EXCEL files. As a functionary shared “this doesn’t take too much time. The data already has percentages, and we don’t need to do any calculations”. However, one official said that the applications have a provision to view data graphically and download raw data in EXCEL. The official felt that graphical analysis is not useful, and that functionaries download EXCELS and use only that for monitoring.

6.4 Data Use

6.4.1 Data primarily used for monitoring

Currently, data is used primarily for monitoring. The function of each application mentioned in figures 11 and 12 is to track or monitor progress.

An example is of sanctions issued for planned works. For works at Panchayat level, three sanctions are made - administrative, technical, and financial. Data on these sanctions is used to track payments and monitor progress. Gram Panchayats are the primary implementing agency for most works, and officials at block, district, and state level stated that the “only use” of the data is for monitoring.

Similarly, physical progress is tracked using a Google spreadsheet which includes details on the number of sanctioned works, works yet to start, ongoing, and completed works. At the time of interviewing officials, this information was not on a portal, and the spreadsheet was uploaded once a month.

*“Through the portal, I can check how much work has begun after it was sanctioned.
If 60 Rs is sanctioned, I can track where this 60 Rs is being spent.
I can track the status”.*

- Official, Panchayati Raj Department

Officials also mentioned reaching out to specific GPs in case there were delays in work visible on various tracking platforms.

6.4.2 Mixed views on integration across portals

E-Gram Swaraj covers 15th Finance Commission funds and is not integrated with Raj e-panchayat which covers State Finance Commission (SFC) funds. Both these portals focus on funds that are devolved to GPs. An official stated that these should be integrated.

As per one official at district level, integrating both would make work simpler as officials would have to visit only one portal instead of two. In fact, as per the official, functionaries working on e-Gram Swaraj have stated that they can display SFC works as well on the same portal, but this has not been done as the Government of Rajasthan wants to maintain a separate portal for state activities. Currently, there are no API or other processes for data to be shared across applications.

“Right now, we have to track and monitor across two portals. If they were mapped on the same portal, or if SFC works were mapped onto e-gram swaraj (the 15th FC funds portal), things would be better.”

- Official, Panchayati Raj Department

Integration could also lead to avoiding the rare scenario of work duplication. An official explained this using an example. If a road must be built from point A to point B, a sanction may get raised by the Gram Panchayat, and the Public Works Department (PWD) may also raise a sanction. It is possible that the PWD may find the road has already been made, despite them receiving funds for it. Therefore, integration could help with reducing duplication and improving fund utilisation.

Another example of possible integration is the NREGA application with work approval applications like Raj e-Panchayat and e-Gram Swaraj which have financial sanction information. An official said that the integration of both can be useful.

One official stated that within the department there was no need for the data. When there is, a demand is placed, and data is made available. Across departments, the official stated that data is not requested and therefore there was no need for interoperability across datasets and departments.

6.4.3 Redundancy

Redundancies with data collected were mentioned with only one example. The example is of commitments each GP must make on the Meeting Online portal. Each GP must show 50% of their planned activities on the portal which, as a functionary shared, are rarely completed or followed once the GPDPs are made, and sanctions are released. The functionary shared that there is no need for this right now as it leads to uploading dummy plans on the portal before the sanctions are approved.

CHAPTER 7: RECOMMENDATIONS FOR THE EDUCATION, HEALTH, AND PANCHAYATI RAJ DEPARTMENTS



Rajasthan has made great strides in collecting granular data across several applications for both the health and education department. Steps have also been taken into strengthening interoperability of data by linking different datasets as well as data use for decision-making. The building blocks for evidence-based decision-making are thus in place for these three departments.



This section presents a series of recommendations on strengthening data usage and usability for policy improvement from the frontline till the state level based on the existing challenges identified with the 3 departments studied. These are relatively short-term immediate solutions to existing challenges and includes suggestions from officials themselves.

This is then followed by a discussion on the way forward for the Government of Rajasthan towards a stronger governance framework for data use in decision-making.

7.1 Data Generation

Table 3: Recommendations for data generation



	<p>A Policy/Guideline on E-Office i.e., moving from offline to online data</p>	<p>Officials felt that the government must commit to not using registers and pass an order for the same. Given a paper driven culture and lack of trust with online systems at some levels of governance, this can be done by higher level decision-makers, with requisite sensitisation at the lower levels. This would in turn lessen the burden on the frontline on data collection.</p>
	<p>Strengthen Infrastructure at the Frontline</p>	<p>Slow internet connections, or their absence, has prolonged the use of offline data. This must be reviewed across the state and improved, to facilitate a shift to online entry only. Another option is to enable offline entry on applications and ensure local back-ups on</p>

		<p>phones which can facilitate entering data in the absence of an internet connection.</p> <p>Similarly, FLWs in the health department have not been provided with mobile phones yet by the government. This process must be hastened, and network charges should also be covered by the government. The lack of network or software glitches that arise in day-to-day functioning can be resolved via regular checkups and improving existing software.</p>
	<p>Recruiting Data Entry Operators and filling vacancies</p>	<p>The state has currently recruited data entry operators in large schools. However, the challenge of data entry was also reported by schools of smaller sizes.</p> <p>For health, while DEO posts do exist, filling those vacancies will be critical to ensure that data is collected regularly and reliably.</p> <p>Similarly, for the Panchayati Raj Department, functionaries at village level often need additional support to generate and upload data, and there is a need for additional human resources or capacity building of existing personnel.</p>
	<p>Single Source of Truth via Linkage with Jan Aadhaar Database</p>	<p>Across all platforms, citizen identifier data should be used from the Jan Aadhaar database.</p> <p>As per officials working on Jan Aadhaar, Jan Aadhaar is mandatory to avail government services, and approximately 97 per cent citizens in Rajasthan have Jan Aadhaar numbers. The Jan Aadhaar database can serve as a ‘single source of truth’, i.e., ensure that data once collected by the government at the primary unit of activity serves as the single and only source of data that all parts of the government work with. This single source of</p>

		<p>data can then be used across interoperable information systems in government, including workflow systems, public financial management systems, management information systems, and performance appraisal systems. Therefore, integrating Jan Aadhaar with applications in the health department can serve two purposes: -</p> <ul style="list-style-type: none"> ● One, the same identifier data need not be entered again and again, thereby reducing workload. ● Two, it can help build Electronic Health Records (EHR) of citizens. For instance, if a woman is pregnant, her information can be recorded on PCTS. If she develops an illness and accesses Chiranjeevi for inpatient treatment later, that can be recorded on the Chiranjeevi systems. Both can be accessed later in the same file, with Jan Aadhaar serving as a common link. This can assist with having accurate and complete records of each citizen, which can be used by doctors later, easing the burden on both doctors and patients. This is in line with the Ayushman Bharat Digital Mission of GoI as well. This would also solve the challenge of a pregnant woman who is registered in one district but delivers in another, as the birth can then be linked to her personal details and be recorded in the right district. <p><i>At present the use of Jan Aadhaar has started for PCTS and Chiranjeevi, and Aadhaar is used if the person is from outside the state.</i></p>
--	--	---

7.2 Data Usability

Table 4: Recommendations for data usability

	<p>Errors and verification: re-evaluate data verification processes</p>	<p>There is a need to re-evaluate how data is verified to improve both quality and trust in the data, given concerns from officials about the extent of checks. This can be via independent checks, data audits, and the use of image data.</p> <p>For the Panchayati Raj Department as well, there is a need to create robust verification processes (such as district level verification following entry by block and gram panchayat level), which are absent at present.</p> <p>A trust deficit in data platforms often leads to weak engagement with data when it comes to decision making. By building trust in the platform through robust practices, the state can move towards creating stronger evidence-based decision-making systems.</p> <p>For education, one way of doing this is by involving Panchayats a lot more in the collection and verification process. Since education is a devolved subject, it is the responsibility of the PRIs such as the Gram Panchayat to be involved in delivery and improvement of education in the state. With regard to data collection, PRIs can be trained and involved in collecting data in schools at their village-level. They can also be trained to verify data collected by school functionaries, thereby supporting the block and district level functionaries. The data collected can inform better prioritisation of untied funds received by the Gram Panchayat that can be spent on improving the school.</p>
	<p>Improving Capacity</p>	<p>Improving capacity is another way to improve quality. FLWs expressed the need for better training on the use of software. There is an attempt to bridge capacity gaps by using performance data of various FLWs and functionaries.</p>


Box 5: A potential capacity building solution




One software used for performance assessment is ASHAsoft. Officials use ASHAsoft to analyse the number of activities conducted in a given time period to understand ASHA performance. Based on this, officials have identified the best and poor-performing ASHAs and intend to organise a workshop of all ASHAs with 2 best performing ASHAs, and 8 poor performing in each group, in batches of 70-80 with their supervisors.

The training is meant to focus on programmes and activities, how ASHAs can claim incentives, and get poor and best performing in the same group discussion and get the best performing ASHAs to tell the others how to claim more incentives. Similar efforts are needed to improve capacity across levels.

7.3 Data Usage

Table 5: Recommendations for data usage

	<p>Integration across datasets: work on combining and streamlining applications</p>	<p>One of the main demands from government officials themselves is to integrate datasets and make the data landscape for the Rajasthan Health Department far simpler.</p> <p><i>“What will really help us is to combine all 48 apps into 1 application – which can collect all data and then we can give it to the field. We should have an application which has all the data.”</i></p> <p><i>– Official, Health Department</i></p> <p>Another official stated that there are three types of services: hospital based (IPD and OPD), outreach (health camps, MCHNDs, etc.), and home-based services and datasets can be structured in this form. if an ASHA is going for work, then all the different applications or portals should show in one place, making entry easier. Some of these have begun. See Box 6 on IHMS 2.0 for more details.</p> <p>Similarly, for the Panchayati Raj department, to reduce the multiple applications for tracking works at the same Gram Panchayat one</p>
--	--	---

		integrated platform could reduce workload and make monitoring easier.
	Participatory Data Collection and Portals	<p>Applications, therefore, should account for context and needs for FLWs. There are several examples listed above, of existing software being incongruent with user needs. For instance, applications are in English, when FLWs are most comfortable with Hindi. However, now software such as PCTS are in both Hindi and English, but this is not the case with all applications</p> <p>Another example is that software building doesn't account for the needs of mid- and upper-level management. One set of solutions entails including both data collectors and users during the design and implementation of data applications and software, and collecting feedback as applications are rolled out. This should be framed around the usability of the applications, and its impact on workload, as well as challenges faced in implementation.</p>
	Data Utility Should be Clear for the Frontline	To build a culture of strengthening both data collection and building ownership is to enable data collectors to be users as well. FLWs should be able to benchmark their area's performance with others and have opportunities to access data to understand ways in which they can improve.
	Use of Predictive Analysis	In addition to using data for a retrospective understanding of how a scheme performed, there is greater merit in using data for policy improvement and formulation. This implies that the use of data throughout an academic or financial year must strengthen policies made in the moment for immediate benefits. There are many benefits of completing the policy loop by using data for making future decisions such as planning decentralised learning outcome improvement programmes or water improvement in targeted schools. These benefits include informed decisions where outcomes are defined

		<p>and measured, action-oriented functionaries willing to make agile changes depending on data patterns among others. This practice of orienting the department towards benefits of predictive analytics can be conducted through occasional workshops by experts within the department data collections cells such as Shala Darpan or UDISE+.</p> <p>See Box 7.a and 7.b for more details</p>
--	--	--

Box 6: IHMS 2.0

The Integrated Health Management System 2.0 (IHMS 2.0) is in the works. It aims to act as an umbrella application, to merge some applications, and use APIs to push data and integrate datasets. IHMS 2.0 will also aim to maintain all records in one place and link it to Ayushman Bharat Health Accounts (ABHA). IHMS 2.0 also aims to build provisions for new features to be added. For example, for outreach applications like camps, surveys, etc, and new things that the department needs, IHMS 2.0 will have a user specific form which can be tailored to needs.

Current status: The department and IT teams are analysing what can be integrated. Information has been sought from various teams and verticals running applications and portals on their Standard Operating Procedures for each application. The team working on IHMS 2.0 are trying to understand the workflows and process flows. An RFP will be made for IHMS 2.0.

Box 7.a: Examples of predictive analysis: School Performance Dashboard

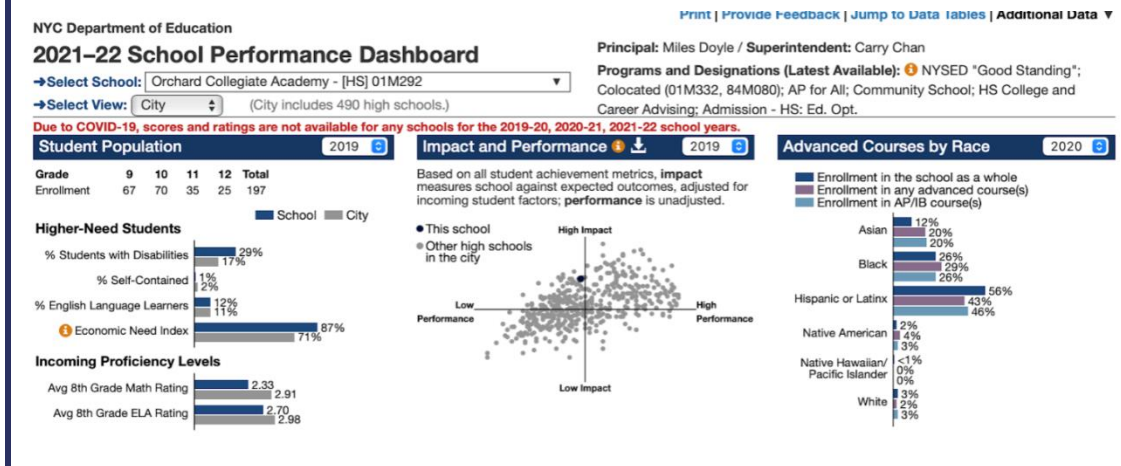
Countries across the world have invested in strengthening their education data that is generated and utilised on Management Information Systems. The example of Sierra Leone proves insightful. The Centre for Global Development’s report describes how Sierra Leone’s Directorate of Science, Technology and Innovation have trained Data Scientists who use population-level data to analyse administrative trends in education governance (Rossiter, 2020). Prominent enabling features of the system include the integration of enrolment and exam data, data use partnerships across different departments and open data made available to the general public. The data analysis allows for officials to understand how they should allocate resources that directly affect student performance.

Another example is that of the New York City Department of Education that is using student performance data to create innovative visualisations (in image below). The graphs allow each user to choose an academic year and then visualise how the school is performing in comparison to other schools in the city. The dashboard also provides details of how students have fared in particular subjects or for variables such as graduation rate, persistence rates for students from specific races. These visualisations allow for Head Teachers, administrators and parents to plan better for their students.

Source: New York City Department of Education Dashboard

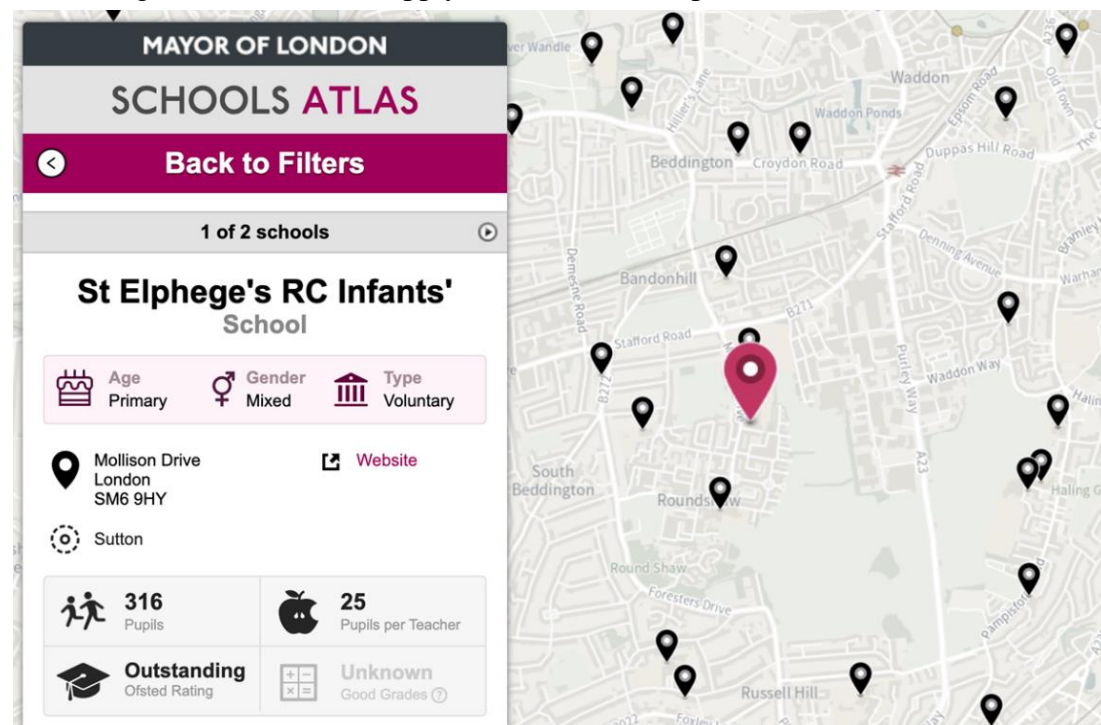
Accessible here:

https://tools.nycenet.edu/dashboard/#dbn=01M292&report_type=HS&view=City



Box 7.b: Examples of predictive analysis: School Atlas

Another example is using location data and pairing it with indicators, as has been done in London, United Kingdom. A platform like this – where district, block functionaries can zoom in by hovering/typing pin codes can help plan targeted school interventions for learning outcomes or assessments, as well as plan monitoring visits, as well as supply routes for school provisions.



Source: London Department of Education Dashboard

The exercise of using predictive analysis can also support making data more accessible. For instance, for a functionary working with school infrastructure graphs where they can understand which district lags in particular aspects such as boundary walls is useful. Whereas, for a functionary developing curriculum visualising learning outcomes on a spatial map with schools across a block or district would be useful. It is also useful to have time series data that enables a comparison between multiple academic years – which the system currently does not allow for.

Currently, the Panchayati Raj department does not use predictive analytics and that can be considered.

While these steps can help in the immediate, given Rajasthan's commitment towards data governance, a comprehensive data policy in the future that looks at the generation, usability and use of data through the policy cycle will be critical.

The next section suggests some longer-term approaches towards achieving this vision.

7.4 Towards a Stronger Data Cycle Framework

Data is a crucial part of every stage of the policy cycle.

1. The first stage is **policy formulation**, which entails defining the problem that must be solved and exploring how that can be done. This requires outcome, coverage, and administrative data on resources and capacities.
2. The next stage is **policy design**, which requires integrating measurable indicators for progress. This includes creating a plan on what indicators are needed, to how they will be generated, validated or verified, and finally, used.
3. The third stage is **policy implementation**, which requires monitoring progress. This requires carefully analysing incentives of data collectors, as well as robust quality checking mechanisms.
4. The last stage is **policy evaluation**, where data from implementation, design, and formulation can be used as part of a feedback loop.

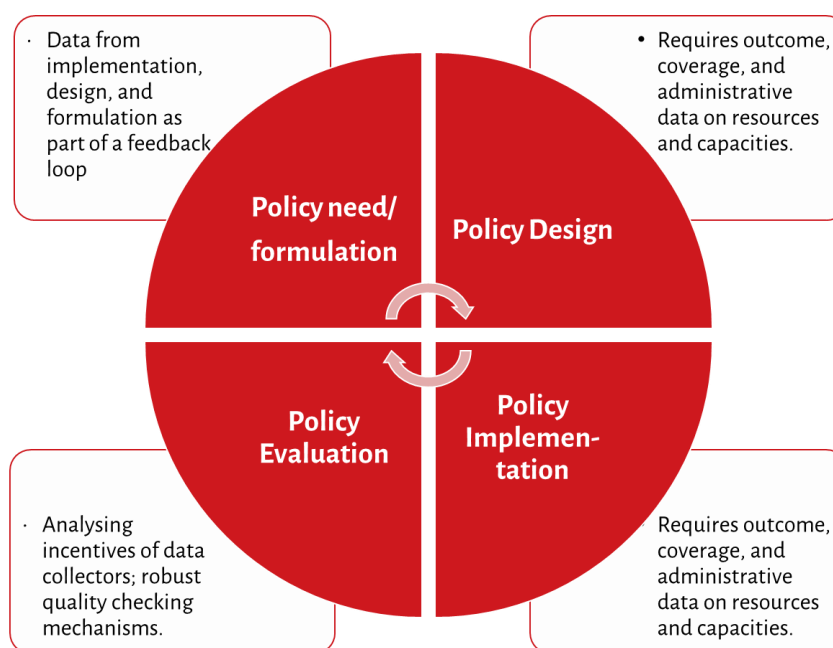


Figure 13: The Policy Cycle

A robust data cycle that is linked to the policy cycle is critical to be able to get the right data to the right stakeholder at the right time for evidence-based decision-making. This requires rethinking and reevaluating the data cycle from a bird's eye view. The treatment of data in high quality research can be most useful in priority setting and visioning, planning for data generation along with preparing annual plans, data analysis, data use and interoperability, and finally evaluating important data and discarding redundant data.

This can help the Government of Rajasthan integrate data into regular processes for the year and can contribute to a more careful and measured use of data.

Using the examples of education and health, the first task should be to identify goals that the health or the education system wants to achieve, and the best source and type of data that can help assess progress towards those goals. Say, for instance, if population control is a goal, then tracking the Couple Protection Rate is important. If the goal is to make faster decisions, then data should be in real time. After such a mapping exercise has been completed, the department must work backwards and analyse what data is required to track and manage the flow of information, and therefore improve efficiency in decision making.

7.4.1 Institutional Support: The creation of a data stewardship body

Currently as mentioned in the introduction, data usually rests with individual government departments. However, one of the first steps in developing a robust data governance framework is the creation of an institutional body that has the capacity and mandate to do this.

Data stewardship is a concept with deep roots in the science and practice of data collection, sharing, and analysis. The concept of a data steward is intended to convey a fiduciary (or trust) level of responsibility toward the data. Data governance is the process by which responsibilities of stewardship are conceptualised and carried out ([Rosenbaum, 2010](#)). A data steward role could include defining big data governance policies, defining data standards, defining flexibility and automation strategies for real-time decision-making, encouraging data sharing between private and public, and working towards identifying big data trends ([Tse et al., 2018](#)).

Furthermore, a data steward could analyse the use of data and improve ways in which it can be done can be useful. In more detail, a data steward should enshrine stronger data collection and verification processes for improving quality, should analyse the needs of the department, and create frameworks that support innovation to strengthen decision-making and implementation. For instance, certain kinds of data are not collected, and some data are not collected in real-time. A data use overseer could analyse gaps such as these and suggest improvements at each level, especially as data use is part of daily governance.

This data steward should be able to respond to specific requests from officials. For instance, senior officials working on Chiranjeevi said that it would be useful to have data and information to learn and improve particularly information on how other states and even other nations are performing and operationalizing health insurance

schemes. Maharashtra and Tamil Nadu were used as examples of states from which officials in Rajasthan want to speak to and learn from.

A data steward could reach out to the private sector to access missing data or form policies to collect the same. For instance, data on health seeking at private facilities is missing, and officials say that having it would be very useful. An example of this is how much people spend out-of-pocket on outpatient and inpatient treatment at each facility. While data on outpatient treatment is available on the IHMS, data on inpatient treatment is unavailable. This, therefore, precludes certain analyses that are useful in improving scheme coverage and outreach. For instance, analysing the proportion of people availing inpatient treatment through Chiranjeevi (Chiranjeevi inpatient treatment sought/total inpatient treatment sought).

The data steward could work on improving processes as well. For instance, given that some types of analysis are still done manually, there is a desire for greater automation with existing data, especially for fraud detection.

The data steward can work with government officials and functionaries, private sector stakeholders, patients, and civil society organisations to define the public interest in certain data uses, improving participation in public programmes, and enabling the creation of knowledge through research, particularly that which is useful for state policy.



Figure 14: Potential Data Stewardship in Rajasthan

Specifically for the health department, the data steward can work across various verticals and teams in the health department to assess needs, have a ‘zoomed-out’ viewpoint, and help streamline and consolidate data collection, use, and analysis.

Similarly, for education, while a lot of the integration has already been done, creating an **Education Data Steering Committee** with members from different departments and cells concerned with education or providing other services at the school could help strengthen workflows and usage. The committee must agree upon a mandate of

ensuring seamless data collection at the school and agree upon a routine for the same. The committee must work towards reducing the effort it currently takes to access data collected by different departments. This committee can push for greater convergence with other departments, thus ensuring infrastructure is built in schools (through MGNREGS) or water supply is improved in schools (through Jal Jeevan Mission).

A data stewardship body can function in three ways. These are shown below:



Source: [Conference of European Statisticians \(2022\)](#).

Figure 15: Types of Data Stewardship

Ideally, given the scale of operations and the size of Rajasthan, a federated/hybrid model is likely to work best. The state government, in conjunction with departments, can set the framework, tools, and practices, aligned with overall vision and goals. Each department can then be responsible for managing its own data. One important factor in the success of such a model is strong collaboration across departments, and with a central authority, which in this case can be a data stewardship body.

7.4.2 Ensuring an Effective Legal Framework

After creating an institutional body, having a comprehensive legal framework/policies are required. The report titled “Exploring Best Practices and Lessons Learned in Data-Driven Governance from National, Subnational and International Cases” has some of these details but broadly there are two aspects that are critical.

1) **An open data policy**

There is a need to increase the scope of data made publicly available while keeping privacy concerns in check through anonymisation of individual details. By making more data accessible, the state commits itself to higher standards or transparency and provides caregivers and civil society organisations with an opportunity to identify where they believe the system needs improvement. It provides these non-state actors with an opportunity to support the health and education system whenever need be.

2) A comprehensive legislation to protect data and citizen rights

As the quantum of data collected increases with passing time, there is a need for legislation to protect the rights of citizens including students and teachers whose data is collected. This legislation must ensure data remains encrypted and secure on the platforms it is currently stored in. It must decide on penalties for leaking or selling data that is sensitive. The legislation must also provide the state with a detailed understanding of how it must use data for evidence-based decision making. It must develop and solidify a fair and ethical way data can be collected and used in everyday governance.

Such a legislation to protect privacy can only be formulated at the Union level because of constitutional entries of 31 (Posts and telegraphs; telephones, wireless, broadcasting and other like forms of communication) and 94 (Inquires, surveys and statistics for the purpose of any of the matters in this List) that fall under the Union list. However, under entry 12 of the State list, a state government can make laws on maintenance of public records. While making such a law, the state government can ensure the right to privacy of citizens.

7.4.3 The architecture needed to facilitate optimum data use for decision-making

Lastly a data management architecture is required. This section draws on the Tamil Nadu Data Policy (2022) to describe the architecture that Rajasthan can adopt while creating a more robust data system.

Data Management Cycle

Both within and across departments, data is collected from several sources including legacy systems (paper-based data collection), online data, third parties such as private entities, and the Jan Aadhaar data base, which continues to expand. A useful data management framework, focusing on the breadth of data from various sources, and its subsequent use is recontextualised and presented below ([Trianz, 2023](#)). This can help with integration.

The next step is aligning the data with the vision for the state and goals for each department. It is necessary to have determine what the state and departments need to do and the decisions they are likely to take.

This is followed by cleaning data, modelling data, i.e., creating variables and linkages that can be used based on needs, matching data i.e., enabling datasets across departments to be linked using panchayat, block, and district codes, and individual data can be matched using the Jan Aadhaar database. Access must be provided to those who need the data, but processes to make the data available to a wider set of government officials must be created, with privacy kept in mind. These datasets then

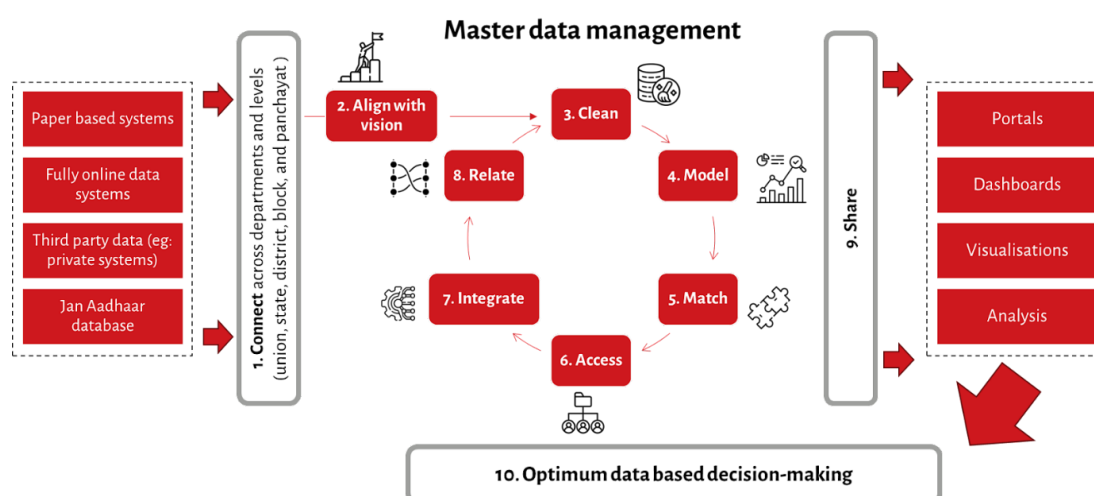
need to be integrated in a usable form. For instance, linking data on motorable roads with the map of health facilities in the state to understand accessibility.

Subsequently, the data can be shared in the form of portals, dashboards, and visualisations. Moreover, it is imperative that the state create a Data Analysis and Response Team (DART) to analyse data consistently and regularly feed insights to decision-makers, to deliver services more effectively. For example, to improve access to a particular scheme, it can be possible to identify areas of weak penetration, specific gaps in service provision, and characteristics of households that are unable to access scheme benefits.

The DART can consist of a small number of data scientists, who can co-develop data analysis requirements, frameworks, and methods, with the government. The work can range from identifying trends and patterns in the data, to more sophisticated techniques, models, spatial analysis, machine learning, and artificial intelligence in large data sets. The DART can also build the capacity of data analysts in each government department.

There is a precedent for this. In 2016, the Government of Tamil Nadu set up a Data Analytics Unit in collaboration between the Department of Economics and Statistics and the Tamil Nadu e-Governance Agency to analyse a large volume of data in real-time to enable improved policy making for better service delivery.

One important point is that datasets need not always translate to dashboards. Data presentation methods should be strongly linked to both needs of various data users across levels, and data should be amenable to presentation in various forms.



Source: Adapted from [Trianz, 2023](#).

Figure 16: Data management cycle

The above framework requires certain principles and practices to be followed. These are discussed below.

Metadata

Metadata – is “data about data.” Essentially, Metadata is the information that describes the data source and the time, place, and conditions under which the data were created. Metadata informs the user of who generated the data, when it was generated, what indicators were collected including their definitions and related explanations, where they were collected and which area, they pertain to i.e., coverage, and the process by which data were generated i.e., methods. Metadata includes the degree to which data can be shared, clearly labelling personal data. Essentially, metadata allows the data to be traced to a known origin and known quality (Tamil Nadu Data Policy, 2022).

Primarily, Metadata enables the use of data across multiple platforms and in multiple ways, by ensuring that relevant data are:

- **Findable:** Metadata makes it much easier to find relevant data including data that is stored in formats like audio, images, and video. The key point is that Metadata provides detailed information on each dataset or relevant documentation, making their use much easier, including by newer hires, reducing overall workload.
- **Accessible:** Metadata clearly indicates how the data can be accessed, possibly including relevant authentication and authorisation. This is crucial in the context of sharing data that may include personal identifiers (see below).
- **Interoperable:** Metadata enables interoperability across datasets and can facilitate a further expansion in the Jan Aadhaar database. It makes it easier to manipulate and slice data to suit various needs, department specific, level specific, region specific, individual specific, and so on.
- **Usability and reusability:** Metadata helps understand the structure and definitions of the dataset, which are crucial to understanding what kind of insights the dataset can and cannot provide. This is particularly useful to an analysis team, a data stewardship body, decision-makers in departments, as well as functionaries at various levels. Metadata is also useful for reusability as government functionaries or researchers can understand the dataset before using them in other settings to generate insights ([GOFAIR, 2016](#)).

Metadata are useful in the context of the expanding welfare state as well. The Government of Rajasthan launches new welfare programmes frequently, and at scale. One challenge as per officials has been data requirements for new programmes as well as subsequent applications are not integrated or connected to existing programmes and datasets. This is where metadata comes in. Metadata is helping in using existing data to improve policy design and implementation, as well as add to the existing datasets in a meaningful way. This helps reduce workload, but can substantially improve the quality of the programme along various aspects such as -

reducing inclusion and exclusion errors, data transparency, assessing progress with indicators in another dataset, etc.

The Government of Tamil Nadu set up the Tamil Nadu e-Governance Agency as the nodal empowered agency for creating the Metadata catalogue ([Government of Tamil Nadu, 2021](#)). Rajasthan, too, can set up a nodal authority to coordinate across departments to collect and integrate data.

Data Sharing Possibilities

As mentioned above, data can be made more open and accessible for citizens. There are, however, concerns around what data can be shared and cannot be shared, which must be clearly delineated in the Metadata to ensure the right data are used by the right people, as well as creating a framework for maximising the use of data. This entails making protocols for sharing anonymised data for research with safeguards.

Table 6: Data Sharing Matrix

Purpose of data sharing	Data requested by	Type of dataset				Access
		Personal identifiers	Sensitive personal data	Anonymised	Aggregate	
Public good	Anyone	N	N	Y	Y	Open/free
Policy making/ monitoring	Government departments/ bodies	Y	Y	Y/N	Y/N	Permission/ Free - basis data sharing policy
Policy research/ Advocacy	Academia/ Civil Society/ Government bodies	N	Y/N	Y	Y/N	Permission/ Free or paid - basis data sharing policy
Value Added Services	Government Agencies/ NGOs/ Private sector	N	Y/N	Y	Y/N	Permission/ Free or paid - basis data sharing policy
Any		Y	Y	Y/N	Y/N	Case by case basis

Source: Tamil Nadu Data Policy (2022), slightly modified by authors.

What is notable in the figure above is the crucial role that civil society can play in strengthening policymaking. Furthermore, there is scope for new requests made as well. The designated authorities for sharing or not sharing data, and framing data sharing policies can be a data stewardship body, or individual departments, depending on the route the state government chooses to take.

Data Standards

Data Standards refer to a framework that defines data handling functions and processes such as data generation, management, transfer, integration, and publication. Data standards also specify the way data can be used. For instance, that data transformations and integration comply with requisite data format and data syntax specifications (Tamil Nadu Data Policy, 2022).

Another example is that data standards can also specify use timelines for faster decision making, which can lead to better service delivery if done correctly. Data standards help in maintaining quality as well, enabling more robust insights,

Furthermore, data standards are critical to breaking existing silos that persist across departments and foster interoperability. Standards are also necessary to create data integrity, ensuring data maintains its value and meaning and is usable, consistent, and accurate ([Balasubramanian, n.d.](#)).

Data standards not only specify the quality of data, but also what data should be collected necessarily. For instance, all procurement of medicines should be indexed against batch number and product codes to ensure veracity and prevent leakages or misuse.

Data Storage

The State Government must choose if the data are to be stored centrally or in a decentralised manner. For instance, all data can be owned, stored, and maintained by the state government at large, or by respective departments. A mix of both is the Government of Tamil Nadu's strategy, where critical master data such as the family database are to be stored centrally with appropriate safeguards, and application specific data are to remain with respective departments (Tamil Nadu Data Policy, 2022).

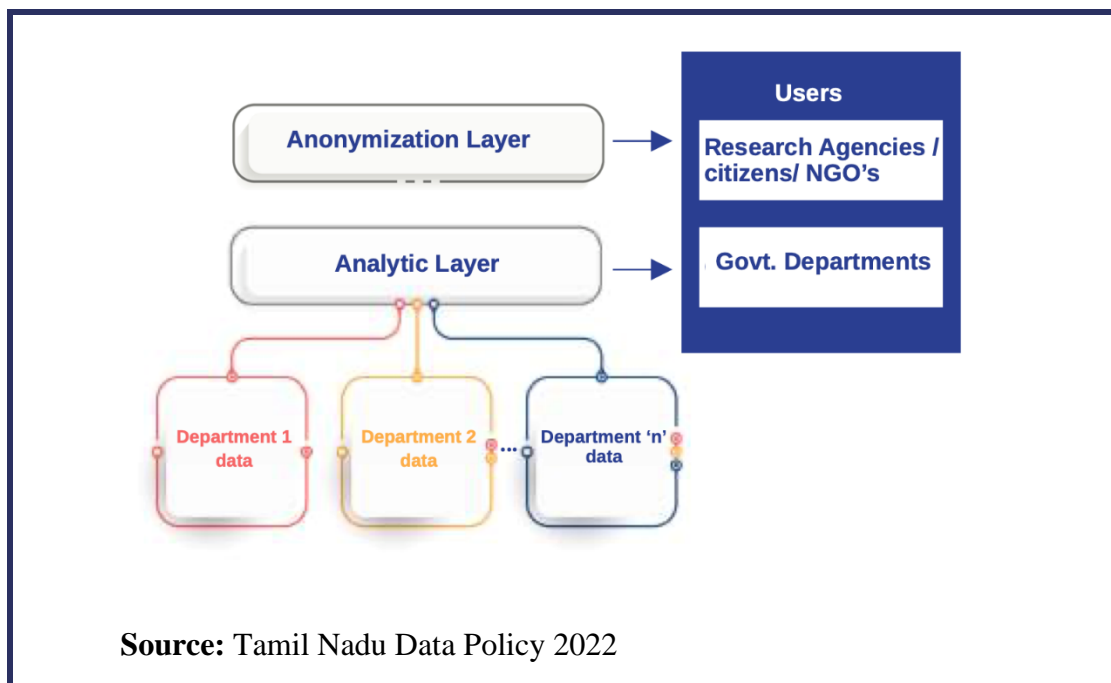
The policy also specifies that latency sensitive data will be stored near where it is needed (Tamil Nadu Data Policy, 2022). Latency refers to the delay that happens between when a user takes an action on a network or web application and when they get a response. This is critical for the management and use of real-time data specifically. This can also be considered by the Government of Rajasthan, given the

gaps in digital infrastructure such as the lack of network in various areas. Datasets can also have retention and destruction guidelines that will be developed (Tamil Nadu Data Policy, 2022). The idea is that data should be stored in a format that can improve the processing time and mitigate the risk of data loss and data leaks.

Box 8: Case of the Tamil Nadu Data Policy

The Tamil Nadu Data Policy was released in 2022 by the state government of Tamil Nadu. The policy highlights important practices that need to be adopted when using data in everyday governance. Key features of the policy include:

1. Adopting a collective goal of “data for public good” to solve challenges of collecting data in a non-standardised manner, lack of sharing data between departments and different extent to which e-governance has evolved.
2. Recognising the steps to be taken to increase collection and improve usage of data across departments. This includes steps such as “targeting beneficiary group, developing rollout strategies and impact assessments” among others (for more, refer to page 4 of TNDP 2022)
3. Establishing a Data Analytics Unit by the Department of Economics and Statistics and Tamil Nadu e-Governance Agency to analyse real-time data for improving public service delivery.
4. Defining key concepts such as open data, privacy, metadata and de-identification among others. By defining these concepts, the state has paved the way for adopting practices that are common among all involved actors.
5. Identifying clear data sharing rules with government departments, academia, NGOs and private sector (for more, refer to page 26 of TNDP 2022).
6. Creating conditions which enable smooth data sharing between departments such as creating an analytic and anonymisation “layer” (refer to diagram below and page 30 of TNDP 2022).
7. Laying the foundation for implementation by establishing positions such as Chief Data Officer and committees for inter department data sharing. This will ensure factors such as data quality stream, compliance and partnerships are managed effectively.



Source: Tamil Nadu Data Policy 2022

CHAPTER 8: CONCLUSION

In the last decade, Rajasthan has committed itself towards rapid digitisation of governance. In the education sector, Rajasthan has worked towards this commitment by introducing platforms such as the Shala Darpan portal. On this portal that is complementary to the UDISE+ portal, functionaries can upload data at the school level, verify it at the block and district level and use it for making decisions at the state level. For health, Rajasthan collects a wide range of data points across 48 different portals. This enables the state to cover all the dimensions of health and oversee the status of different union and state-level schemes along with crucial health indicators.

This report aimed to unpack and suggest recommendations on the everyday governance of education and health through online portals, specifically in Rajasthan. The report proposed the use of a framework to conduct the study and analyse the findings. The framework deliberated on four main features of data generation, usability of the data, use of the data, and data for policy improvements. While data generation focuses on the qualitative aspect, analysing which are the categories for which data is collected and the process undertaken to collect the data, data usability focuses on the quality of the generated data, analyses the granularity, the accessibility, cross convergence, and the timeliness at which the data is produced. Data use analysed how functionaries use data to make targeted policy interventions. Lastly, the feature of policy improvements allowed the policy makers to reflect on the impact of the generated data by examining if the required data was being generated and used for decision-making across all levels. The report applied the case study methodology with a mix of purposive and snowballing interview methods to reach functionaries at different tiers within the departments of education and health in Rajasthan.

The findings reveal that the state is ahead of its time compared to other states in the country as it had introduced real-time data generating platforms when other states were still using platform updates once a year, if at all. As a result of this the Rajasthan government has received its due recognition by the union government. The state has developed cutting-edge methods to collect a wide variety of data that has the potential to enable functionaries to make informed policy formulation and improvements. Despite these efforts championed by the state, the report identified many challenges that are yet to be tackled.

Regarding the generation of data, there is an urgent need to issue a uniform guideline or policy that can regulate the use of online and offline data to reduce duplication of efforts. Frontline functionaries reported spending a huge proportion of

their time uploading data and there is a need to reduce their burden by recruiting Data Entry Operators or filling vacancies for sanctioned positions. There is also a need to strengthen infrastructure at the frontline. In addition to this, there is a need to ensure all students' data is linked to their Jan Aadhar details which can serve as a single truth – a process that is well underway in Rajasthan. Like education, with the government collecting many points to capture all the dimensions of health, it is crucial to reduce duplication, for e.g., data records being maintained online as well as offline and the same identity related data being collected by multiple functionaries on multiple applications. However, FLWs (ASHA and ANM workers) felt that entering data on online portals takes away time for other work. Further, with the lack of training and adequate infrastructure the process of data generation creates an over burden on the FLWs. There is a need to reduce duplication by using a common identifier which exists in the form of Jan Aadhaar.

The useability of data varies largely across different stakeholders. Functionaries across departments have better access to data and have appreciated the availability of real time data and the granularity of data sets on the Shala Darpan portal. However, it was indicated that access to time series data sets and data presented in the form of innovative graphs or visualisations with click features would enable better understanding of the generated data. Thus, the report prescribes exploration and replication of innovative data visualisation practices from across the globe. There is also a need to improve how data is verified once reported on the platforms, an effort that can improve trust in the platform.

In contradiction to the education department, only some data sets are captured real time for health in Rajasthan. Further, there is a need to review data verification processes to ensure better quality. The useability of health data becomes more remote with data points on impact and outcomes not being collected. **Data is currently used extensively** for tracking and monitoring of schemes and crucial indicators across health and education. Thus, there is a need to ensure functionaries use predictive analytics of data rather than only data that reflects the progress of scheme implementation to celebrate its success. Apart from monitoring in the education department, data was used for aiding during the pandemic by simplifying processes in the vaccine centres, provisioning health facilities to the school students such as folic acid tablets and sanitary napkins to girls, identifying students for scholarships and other incentives and answering parliamentary questions. Similarly, in the health department, data is majorly used for improving and updating benefit packages, identifying potential fraud cases, and tracking outbreak diseases. However, the use of data in both departments is highly skewed to the higher tiers of the government. Thus, there is also an urgent need for data use to be taken to frontline and middle-tier functionaries as well, as data use is currently saturated at the elite state level both for

the education and the health department. Another major issue is that of integration across datasets, which is urgently needed to improve the efficiency of data use. Participatory methods of collecting data by engaging elected local governments must be considered as it can save considerable amounts of time and improve the verification process.

Lastly, with regard to policy improvements, the report proposed setting up a Steering Committee across departments that can decide priorities and ensure seamless data sharing across departments, as there is a need for greater convergence in how data is collected and applied in everyday governance.

Despite a large quantum of real-time data being captured by the Shala Darpan portal, only a limited range of data points is available as open data. The state must also work towards ensuring transparency in the data collected by making anonymised data open access for citizens to use as well. By making data accessible and open, it provides caregivers and civil society organisations an opportunity to support the system where need be.

The final recommendation of the report is the urgent need for comprehensive legislation that protects the rights of citizens whose data is collected and stored on the platform. The legislation must ensure data use in governance is built on fair and ethical practices.

While the state faces many challenges in the use of data to make informed policy decisions, there is immense scope for the state to tackle these challenges as it has the necessary political will, human resources and finances. As the state continues to solve these challenges, it will inspire other states in the country to commit to rigorous evidence-driven policymaking.

REFERENCES

- Abraham, R., Brocke, J. v., & Schneider, J. (2019). Data Governance: A conceptual framework, structured review, and research agenda. *International Journal of Information Management*.
- Breiter, A., & Light, D. (2006). Data for School Improvement: Factors for designing effective information systems to support decision-making in schools. *International Forum of Educational Technology & Society*, 206-217.
- Chaudhry, B., Wang, J., Wu, S., Maglione, M., Mojica, W., Roth, E., Morton, S. C., & Shekelle, P. G. (2006). Systematic review: Impact of health information technology on quality, efficiency, and costs of medical care. *Annals of Internal Medicine*, 144(10), 742–752. Available at: <https://doi.org/10.7326/0003-4819-144-10-200605160-00125>
- Court, J (2005) ‘What Political and Institutional Context Issues Matter for Bridging Research and Policy? A Literature Review and Discussion of Data Collection Approaches’, Unpublished Background Paper.
- Department of School Education & Literacy. (n.d.). Retrieved from UDISE+. Available at: <https://udiseplus.gov.in/#/home>
- Donge, W., Bharosa, N., Janssen, M. (2020) Future government data strategies: data-driven enterprise or data steward? Exploring definitions and challenges for the government as data enterprise. *The 21st Annual International Conference on Digital Government Research*.
- Global Partnership for Education. (2019). Meeting the Data Challenge in Education. *Global Partnership for Education*.
- Haux, R. (2006). Health information systems—Past, present, future. *International Journal of Medical Informatics*, 75(3–4), 268–281. Available at: <https://doi.org/10.1016/j.ijmedinf.2005.08.002>
- Jornitz, S., & Engel, L. C. (2021). The Management and Use of Data in Education and Education Policy: Introductory Remarks. In *International Perspectives on School Settings, Education Policy and Digital Strategies: A Transatlantic Discourse in Education Research* (pp. 223-241). Verlag Barbara Budrich.
- Lugo-Gil, J., Jean-Baptiste, D., & Jaramillo, L. F. (2019). *Use of Evidence to Drive Decision-Making in Government*. Mathematica Policy Research.

-
- Luthfi, A., & Janssen, M. (2019). Open data for evidence-based decision-making: Data-driven government resulting in uncertainty and polarization. *International Journal on Advanced Science, Engineering and Information Technology*, 9(3).
 - NIC. (2021, April). *Integrated ShalaDarpan: Decision Support System for Schools Monitoring & Management*. Retrieved from NIC. Available at: https://informatics.nic.in/uploads/pdfs/f5d57a1d_20_22_eps_shaldrpn_compressed.pdf
 - OECD, (2017). *Governing better through evidence-informed policy making*
 - OECD. (2022). *Going Digital Guide to Data Governance Policy Making*.
 - Rosenbaum, S. (2010). Data Governance and Stewardship: Designing Data Stewardship Entities and Advancing Data Access. *Health Services Research*, 45(5p2), 1442–1455. Available at: <https://doi.org/10.1111/j.1475-6773.2010.01140.x>
 - Rossiter, J. (2020). Link It, Open It, Use It: Changing How Education Data Are Used to Generate Ideas. CGD Notes. Centre for Global Development.
 - Schildkamp, K. (2019). Data-based decision-making for school improvement: Research insights and gaps. *Educational Research*, 253-273.
 - *The Principles: Health Data Governance Principles*. (n.d.). Available at: <https://healthdatapinciples.org/principles>.
 - Tiffin, N., George, A., & LeFevre, A. E. (2019). How to use relevant data for maximal benefit with minimal risk: Digital health data governance to protect vulnerable populations in low-income and middle-income countries. *BMJ Global Health*, 4(2), e001395. Available at: <https://doi.org/10.1136/bmjgh-2019-001395>.
 - Tse, D., Chow, C., Ly, T., Tong, C., & Tam, K. (2018). The Challenges of Big Data Governance in Healthcare. *2018 17th IEEE International Conference On Trust, Security And Privacy In Computing And Communications/ 12th IEEE International Conference On Big Data Science And Engineering (TrustCom/BigDataSE)*, 1632–1636. Available at: <https://doi.org/10.1109/TrustCom/BigDataSE.2018.00240>.
 - UDISE+. Available at: <https://www.udiseplus.gov.in/#/home> (Accessed: February 10, 2023).
 - UNESCO. (2023, January 13). *Using data to improve the quality of education*. Retrieved from IIEP Learning Portal: Using data to improve the quality of
-

education. Available at: <https://learningportal.iiep.unesco.org/en/issue-briefs/monitor-learning/using-data-to-improve-the-quality-of-education>.

- United Nations General Assembly. (1948) The Universal Declaration of Human Rights (UDHR).
- United Nations (General Assembly). (1966). “International Covenant on Civil and Political Rights.”
- Williams, F., & Boren, S. A. (2008). The role of the electronic medical record (EMR) in care delivery development in developing countries: A systematic review. *Informatics in Primary Care*, 16(2), 139–145. Available at: <https://doi.org/10.14236/jhi.v16i2.685>.
- Williamson, B. (2016). Digital education governance: data visualization, predictive analytics, and ‘real-time’ policy instruments. *Journal of Education Policy*, 123-141.
- World Bank. (2022, March). Aligning Data Governance with Social Contract for data to promote Value, Trust, and Equity.

Image credits:

Icons made by Freepik, Triangle Squad, Freepik, Smashicons, manshagraphics, Designvector10, Muhammad Atif, Made Premium., Juicy_fish, HAJICON, Uniconlabs, Vitaly Gorbachev, Parzival’ 1997, Monkik, Linector, Eucalyp, Rukanicon, bsd, Lutfix, Reion, Mayor Icons, Canticons, noomtah, Vector Stall, Lagot Design, and justicon from www.flaticon.com

SECTION II

1 INTRODUCTION

The Union, as well as all State Governments in India, have been collecting data on various parameters, across diverse sectors for several decades now. Despite this, there is no overarching legal framework driven by legislative foundation at the national level governing the use of this data in EBPM. Further, the availability of data for public use is challenging - due to the lack of a collated single-platform service, both at the national and the State level for providing easy access to reliable, consistent, and safe datasets for research and advocacy. While this data is collected through various channels, their use case is extremely disoriented, lacking structural and strategic direction for their application in evidence-based decision-making. This report presents reflections on international practices of EBPM and then unpacks India's current state of affairs. The arguments presented are meant to be situated as a legal backbone to the generation and use of data.

2 INTERNATIONAL PRACTICES

When one looks at international practices on data use and evidence for policy making and data sharing, there are cases where EBPM has been successful. EBPM has had a long history of using data for governance, foundation of privacy laws, interdepartmental coordination and necessary infrastructure for collection and management of data. Countries have amended their existing laws, created new authorities or empowered existing ones in order to further the goal of EBPM. They have clear laws on privacy protection which form the main governmental infrastructure behind this.

EBPM had been used in the United States of America by departments such as the U.S. Department of Housing, Department of labour and for the Drug Abuse Resistance Education Program. The U.S. has a long history of use of data in government decision making to revamp collection of data, identify protocols for the use of data while focusing on privacy and confidentiality. This has been done through acts under 3 categories:

- a. Organised Collection of Data by the Government: Federal Reports Act, 1942, Federal Records Act, 1950, Budget and Accounting Procedures Act 1950, Information Quality Act, 2000.
- b. Use of Data keeping in mind Privacy, Confidentiality and Fair use: U.S. The Department of Health, Education, and Welfare developed the Fair Information Practice Principles, 1973, Privacy Act 1974
- c. Efficient Use of Information: Confidential Information Protection and Statistical Efficiency Act, GPRA Modernisation Act, 2010.

In 2018, the US became one of the first nations to introduce the EBPM Act with a key objective to ‘promote the use of evidence to inform decision-making at federal agencies’. The components of EBPM Act include:

- a. Amends Title 5 and Title 44 of United States Code which are existing laws such as the Privacy Act and the code on printing and public documents to include that the government must provide a comprehensive data inventory, define the term data, the manner in which data must be made available publicly and places an obligation to make this information be public by default.
- b. Defines the terms “Agency”, “Director”, “Evaluation”, “Statistical Activities and Purpose”. The agencies are federal government departments such as the department of labour, housing and urban development etc. They are required to prepare strategic plans on the list of data, methods of data collected, an evaluation plan, strategic plans and appoint an advisory committee and evaluation officers.
- c. Efficiency and coordination between departments are sought to be ensured by the Chief Data Officers within departments and by the Director of Office of Management and Budget.
- d. The Act highlights specific rules and regulations for how data can be shared between departments, especially non-anonymised citizen level data. Sharing data with functionaries who do not have the necessary clearances attracts expensive fines and imprisonment in some cases.

Other nations such as Australia, European Union and United Kingdom also have schemes on EPBM. In Australia , components of the legal framework on EBPM include

- a. Bureau of Statistics Act 1975: This act sets up the Australian Bureau of Statistics and places an obligation on the Bureau to ensure coordination of the operations of official bodies in the collection, compilation and dissemination of statistics. Furthermore, the Bureau also has a duty to avoid duplication of data and maximum utilisation of data. This legislation provides an enabling framework for collection and use of statistical data by official bodies.
- b. Freedom of Information Act, 1982: This declares that information held by the Government is to be managed for public purposes, and is a national resource’.
- c. Privacy Act 1988: Addresses concerns of privacy and confidentiality in the use of data by the government.
- d. Data Availability and Transparency Act 2022: This act authorises the sharing of public sector data and seeks to enhance the integrity and transparency of public sector data.

The Australian Parliament has established a new Office of the Australian Information Commissioner (OAIC) to provide advice to the government on policy and practice on how the government collects, stores, manages, uses and discloses information. Australia has created an Open Government National Action Plan (2016-2018). One of the components is to identify National Interest Datasets or designated data sets. These data sets would primarily include public-sector data but may also include private-sector data controlled by the public sector under certain conditions.

The European Union has a strong foundation of EBPM by member states. The manner in which EBPM is operationalised can be understood by looking at policies that enable the use of evidence and policies that act as safeguards against the misuse of data.

The key enabling features can be found in the “Better Regulation Guidelines, 2021”. These guidelines are not legally binding but are instructive for each stage of the policy making process. The guidelines explain that the principles of better regulation require that a comprehensive and coherent approach is adopted while ensuring that the process of policy making is transparent, participative and based on past experience. According to the “The Interinstitutional Agreement on Better Law-Making, 2016” regulation must also be clear in identifying the goal of the institution, and should be clear and effective. This understanding of EBPM emphasises the importance of consultation, expert consensus and coordination between governmental bodies.

The most important safeguard can be found in the General Data Protection Regulation, 2018 (GDPR, 2018). This is the most stringent data privacy and security law in the world. The 7 protections guaranteed under the GDPR, 2018 are:

- Article 5(a)- Data should be processed in a **Lawful, transparent and fair** manner.
- Article 5 (b)- **Purpose limitation**: Data should be collected for a specific goal.
- Article 5 (c)- **Data minimisation**: Only necessary data should be collected.
- Article 5 (d)- **Accuracy**: Data should be regularly updated and corrected.
- Article 5 (e)- **Storage limitation**: Data should not be stored after the achievement of the specific goal.
- Article 5 (f)- **Integrity and confidentiality**: Security measures should be used to anonymise data.
- Article 5 (g)- The entity responsible for collection of data is **accountable** for ensuring that the aforementioned principles are complied with

The United Kingdom scheme on EBPM does not have a law that requires the use of evidence in policy making. This policy document has nine characteristics that are largely aspirational such as “policy making should be forward looking, innovative,

inclusive”. In the UK there is also a Policy Profession Board with Policy Profession Standards which is a competency framework for all UK Civil Servants involved in policy work and operating in a policymaking environment. *The UK focuses to operationalise the use of EBPM by creating incentives for government departments and policy makers to use evidence.* The Cabinet Office, Chief Social Researcher’s Office, HM Treasury release guidelines on how evidence can be incorporated in policy making. The use of EBPM can be seen in the case of the Educational Maintenance Allowance. The UK has implemented the GDPR, 2018 through the Data Protection Act, 2018.

3 WHERE DOES INDIA STAND?

The Indian context of data and evidence-based policy framework is interesting. In the past decade alone India has undergone a massive shift to a more digital economy. There has been a large increase in the collection of data in various aspects of a person's life ranging from Aadhaar, e-SHRAM, Agristack, Arogya Setu, NDEAR etc. Additionally, there has been an increase in digitisation through various policies such as the Digital India Initiative, Bharat Net, Universal Access to Mobile, Smart Cities Mission and the broader framework of India Stack which have all incentivised the push towards going digital, which leads to an increase in the data in circulation and hence there is an urgent need to have a broad framework to govern data. More states are now opening up to the potential of harnessing data in policymaking. Data Governance is gaining momentum and in the past few years, there has been an understanding that a comprehensive policy is the need of the hour.

In India there have been policy documents, such as Acts/ Bills which primarily focus on governance structures related to data privacy, public records maintenance and protection, various classifications and accessibility of government data, etc. -Public Records Act, 1993, Information Technology Act, 2000⁶, National Data Sharing and Accessibility Policy, 2012⁷.

There remain concerns about the asymmetrical knowledge of the citizens, as India does not have a uniform, comprehensive, data protection law in place (*Feigenbaum, et al., 2022*). Data is a double-edged sword; where on one side it has the potential to improve programs and policies, efficiently run economies, and also empower citizens;

6 The Information Technology Act, 2000 also known as an IT Act is the most important law in India dealing with Cybercrime and E-Commerce. The main objective of this act is to carry lawful and trustworthy electronic, digital and online transactions and alleviate or reduce cybercrimes.

7 National Data Sharing and Accessibility Policy (NDSAP), 2012 is to facilitate access to government-owned shareable data or non-sensitive data and information in both 'human readable' and 'machine readable' formats in a timely manner.

and on the other, data accumulation leads to the concentration of economic and political power, raising the possibility of data misuse to harm citizens (*World Bank, 2021*). Having a data protection law at the national level would hold paramount importance. India has also made international commitments such as the Universal Declaration of Human Rights and the International Covenant on Civil and Political Rights which require the protection of the right to privacy. There has been an effort to incorporate individual privacy needs in state-level policies by Punjab, Karnataka, and Tamil Nadu, whereas no mention has been made to individual privacy in the states of Odisha and Sikkim (*Panjir, 2022*).

What is important to look at is that in India the right to privacy is not guaranteed through a statute. Various regulations and rules by entities such as the Reserve Bank of India, Bureau of Indian Standards, Ministry of information and technology etc govern the manner in which privacy can be ensured by placing checks on corporations. The most important directive and guidance on the protection of individual privacy has been provided by the Supreme Court. The case of *K.S. Puttaswamy v. Union of India (2017 10 SCC 1)*, identified the key factors that must be satisfied to justify the collection and use of data while protecting the interest of privacy. These factors are: the existence of a law identifying the goal of the collection of data, the goal should be legitimate such as ensuring proper deployment of resources and that the manner of collection and use of data should have a ‘rational nexus’ with the goal sought to be achieved. These principles form the foundation of the legitimate collection, storage and use of data.

Based on this judgement the following guidance for EBPM emerged:

- a) Privacy is a facet of the right to life.
- b) Privacy may be denied through a procedure established by law which is just, fair and reasonable.
- c) Where the interest of privacy is violated or compromised, the highest standard of scrutiny must be made and such a violation can only be justified for a “social, moral and compelling state interest”
- d) While scrutinising the violation or compromise of privacy, the 3 factors that must be kept in mind are legality: the existence of a law, legitimate aim: like ensuring national security, proper deployment of national resources, and protection of revenue and finally proportionality: there should be a rational nexus between the aim and the means used.

4 LEGAL CHANGES REQUIRED TO OPERATIONALISE EPBM

Based on a review of the legislations in other countries it is apparent that it is not sufficient to enact a new legislation but also ensure consistency with existing laws and update them to ensure that there is a foundation for EPBM. These changes need to be made in accordance with the lessons deduced from international best practices:

4.1 Changes required in laws that allow the collection and storage of data by the government.

In India, the laws that allow the collection and storage of data are Public Records Act, 1993, the Collection of Statistics Act, 2008 and maintenance of records acts enacted by various state governments. The goal of the Public Records Act, 1993 is to regulate the management, administration and preservation of public records at the Central level. The definition of public records includes images and data stored in computers. This act enables the Central government to appoint records officers and provide direction to the national archives on maintenance of records. The Collection of Statistics Act, 2008 facilitates the collection of statistical data on economic, social and environmental aspects. Currently, the Ministry for Statistics and Programme Implementation allows the collection and maintenance of statistical data. At the state level, some states such as Maharashtra have public records acts that allow state governments to manage, administer and preserve state level public records.

These acts provide a strong but insufficient foundation for EBPM. Key elements that are required are

- (a) requiring all government departments and not just the Ministry of Statistics and National Archives to collect and maintain records and data
- (b) identifying officers to ensure efficiency and coordination among various departments to ensure that collected data can be used and transmitted effectively, This can also ensure that data is not duplicated and efficiently collected.
- (c) requiring that government departments prepare plans to strategically collect data and methods for evaluation strategies for existing schemes. One way in which this is currently being done is through the Development Monitoring and Evaluation Office (DMEO) under the NITI Aayog. The DMEO has been at the forefront of preparing the Data Governance Quality Index that is intended to assess the effectiveness of information collected by government departments and facilitate cross learning between departments. Such programmes need to be expanded further with the intent of furthering data driven decision making.

4.2 Changes in laws that require the dissemination of data to citizens

The Right to Information Act, 2005 seeks to promote accountability and transparency in the working of public authorities by allowing citizens to access public information. From the perspective of ensuring the use of EBPM, open data and access to information ensures that citizens are aware of the data that is being collected, they can help correct information that has been collected and rely upon the available data to engage with policy proposals. The RTI Act, places an obligation on all public authorities to maintain records, computerise these records where possible and actively publish these records. Where requested, authorities are also required to answer citizens' queries based on the criteria and timeline specified in the act.

Possible changes based on international best practices that can be included are: firstly, the government should maintain a comprehensive data inventory comprising of information that is of national interest and importance and is not split across various departments and ministries; secondly, information held by the government should be managed for public purposes and is a national resource.

4.3 Changes to ensure the veracity and accuracy of information

In order to ensure EBPM, it is essential that information that is collected and used is accurate and updated. In India, there are rules and guidelines under the Public Records Act, the Collection of Statistics Act and the RTI Act that do make a mention of the requirement of accurate records. However, more comprehensive legislation is required to ensure the accuracy of information and ensure regular updating of data. An example of this is the Information Quality Act in the USA where the act empowers the Office of Management and Budget to issue guidelines to federal agencies on how to maintain records and the manner in which citizens can use grievance redressal to correct records. In India, a key role can be played by the Comptroller and Auditor General's office where auditing and annual performance reports can ensure that government departments are collecting, using and disseminating accurate information.

5 COMPARING THE PERSONAL DATA PROTECTION BILL 2019 AND THE DIGITAL DATA PROTECTION BILL 2022

2019 Bill	2022 Bill
<p>Establishes categories of data (S. 3)</p> <ul style="list-style-type: none"> - Personal Data - Sensitive Personal Data - Critical Personal Data - Financial Data - Anonymised Data 	<p>Establishes fewer categories of data</p> <ul style="list-style-type: none"> - Personal Data (S.2) - Non- personal Data
<p>Applies to personal data processing within India, by the government, companies, persons (S.2) Does not apply to processing of anonymised data for the purpose of enabling better delivery of services (S.91)</p>	<p>Does not require the anonymisation of personal data and allows the Central Government to process personal data for research/ statistical purposes</p>
<p>Specifies the Obligation of Data Fiduciary:</p> <ul style="list-style-type: none"> - Process data for clear and lawful purpose 	<p>Specifies the Obligations of the Data Fiduciary but only requires the Data Fiduciary to try to ensure that</p>
<ul style="list-style-type: none"> - Ensure consent and privacy - Collect only necessary data - Not retain data beyond the period required <p>Personal data can be processed without consent only in case of emergency (Chapter IV)</p>	<p>data is accurate and complete (Chapter II)</p>
<p>Processing Personal Data and Sensitive Personal Data of Children requires consent of parents, verification of age, cannot be used for tracking, profiling or causing harm (Chapter IV)</p>	<p>Processing personal data of a child requires verifiable parental consent and cannot be used for tracking, profiling and causing harm (S. 10)</p>
<p>The data principal has a right:</p> <ul style="list-style-type: none"> - To access information in an accessible format - To ensure correction/ erasure - To be forgotten (Chapter VI) 	<p>The data principal does not have the right to access information in an accessible format</p> <p>The bill also specifies the duty of the data principal to not provide false information (S. 16)</p>

The 2019 bill proposed to establish a Data Protection Authority. The Data Protection Authority specified in the Bill would be a government-established, singular data protection body that would look into breaches of personal data, ensure compliance of data fiduciaries, and ensure compliance of such fiduciaries with the Bill (Chapter IX, Personal Data Protection Bill, 2019).

A data fiduciary as defined by the 2019 bill is an entity or individual who decides the means and purposes of processing personal data (Section 3(13) Personal Data Protection Bill, 2019). It also contained provisions for appointing data protection officers (DPO), who would be appointed by data fiduciaries and would be responsible for adhering to provisions of the Bill (Section 30, Personal Data Protection Bill, 2019).

However, in August 2022, the government withdrew the bill based on Joint Parliamentary Committee review which proposed the need for 81 amendments and 12 recommendations in the direction of a comprehensive legal framework for the digital ecosystem. It had largely been criticised for its severe data localisation norms for storage and processing, with tech stakeholders highlighting the extreme compliance burdens it placed on them. It also drew criticism for falling short on protecting the right to privacy for data principals (users), recognised under Article 21 of the Constitution (*Bhargava, 2022*).

In November 2022, the Ministry of Electronics and Information Technology released the revamped Draft Digital Personal Data Protection. This bill has been criticised for removing the category of sensitive personal data. The data fiduciary does not have an obligation to anonymise data and is only required to make an attempt to safely store data. The current bill also does not contain a right to access the data collected (*Ray, Ajaykumar and Patil, 2022*). In fact, compared to the 90 clauses in the previous version of the bill, this version has only 30 clauses with several components left to the discretion of the Union government⁸.

6 STATE LEVEL DATA POLICIES

On a state level as well, many states such as Punjab ([Punjab State Data Policy, 2020](#)), Odisha ([Odisha State Data Policy, 2015](#)), Sikkim ([Sikkim Open Data Acquisition and Accessibility Policy, 2014](#)), Telangana ([Telangana Data Analytics Policy, 2016](#), [Telangana Open Data Policy, 2020](#)), Karnataka ([Karnataka Data Centre Policy, 2022](#)) and Tamil Nadu have their data policies.

⁸ The phrase “as may be prescribed” (by the Central government) was mentioned in 18 places in the draft bill.

While data protection has found some traction, evidence-based policy development is yet to be designed across Indian States. The Government of Tamil Nadu (GoTN) presented one of the first such policies in early 2022 called the “[Tamil Nadu Data Policy](#)” (TNDP), mentioning that the regulations will ‘support policy-making, scheme implementation, encourage value-added services, improve access to and quality of data’ (TNDP, 2022). The developmental ideas of this policy can be traced back to 2014, with the collaboration between the GoTN and J- PAL, that led to the establishment of a ‘culture of data and evidence used to improve decision making’. Followed by this, and several other initiatives, collaboration with experts, the TNDP 2022 was realised in line with the recommendations of the Economic Advisory Council of TN.

The major objectives of the TNDP include (a) Policy Formulation - through interventions on target groups, cross- departmental data analysis, (b) Scheme Rollout and Implementation - through efficacy and leakage assessment, developing rollout strategies, and (c) Scheme Analysis, Evaluation, and Policy Design – through impact assessments, root cause assessments, redesigning and restructuring policy, etc. (TNDP, 2022, p. 3).

A major key aspect of the TNDP is also its Policy Framework - which mentions specifics relating to data standards, collection, storage, processing, sharing, publishing, maintenance of data quality, evidence-informed decision making, and data security and privacy - specified into a structured format mentioning nodal officers/ departments, along with expected outcomes (TNDP, 2022, p. 43- p. 49). The Implementation Framework specified is yet another critical component of the Policy - indicating implementational departments for key objectives, during each stage of intervention/ action as according to the Act. The Policy is a roadmap document to institutionalise the process of evidence use in policy-making.

While several other State governments in India have hinted at, or ideated upon the use of data for policy making, TN is a standalone State in the development of a formal policy document. Telangana may be considered the next best contender - their “Telangana Data Analytics Policy of 2016”, indicates using data to ‘attract businesses, strengthen smart governance initiatives, encourage data-driven decision making, and nurture requisite talent pool to meet the future demand’ (TDAP, 2016, p. 11). The policy document, however, does not make any mention of modalities of data sharing for policy decisions, implementational challenges, comprehensive stakeholder mappings, etc. - and pronounces itself more as a ‘vision’ document.

The Government of Andhra Pradesh has also made attempts to develop such a policy - in 2020 the Andhra Pradesh Economic Development Board collaborated with the

Indian School of Business, Hyderabad to establish a ‘think tank’ for ‘strategic planning, policy analysis, data analytics and action-research essential for sustaining high rates of growth’ (*Nichenametla, 2020*). However, any policy known in this regard has not been formulated yet. Something similar is true for Karnataka while there is no policy formulation that has been undertaken, the Government of Karnataka has collaborated with institutions like the International Institute of Information Technology, Bangalore, and the Public Affairs Centre (PAC) to use data in ‘policy formulation, program design, and resource allocation decisions’.

The policies in India in its current state focus largely on data procurement and storage structures except for TNDP which also emphasises the focus on data-based policymaking. The aim of data governance laws across the country is not uniform. The American statute of Foundations for Evidence-Based Policymaking Act of 2018, draws a link between these two focus points. The Foundations for Evidence-Based Policymaking Act of 2018 (henceforth Evidence Act) was passed with the aim of better governance, efficiency and transparency. The Evidence Act works on structural governance through the provisions of the statute in itself and addresses policy-making through its yearly annual plans in lieu of a larger decade-long plan aimed at evidenced-based policymaking and a comprehensive database of data available for the same.

There are state level policies on the use of data for governance however none of these are in the nature of a legally enforceable framework or in the nature of laws that create new institutions or add to the power of existing institutions. This is because none of the state governments can enact a law pertaining to the use of evidence as these would be in contradiction to central laws that the state does not have the power to amend.

Apart from the concern of privacy which has been discussed in detail, the state level policies also do not adequately clarify the manner in which interdepartmental sharing of data should be done, the role of agencies that collect data in organising and ensuring the accuracy of data. There is no consistency in the proposed use of the data.

REFERENCES

- Bhargava, K. (2022). Explained: Concerns Over The New Data Protection Bill And How It Is Different From 2019 Draft *Outlook India*. Available at: <https://www.outlookindia.com/national/explained-what-concerns-have-been-raised-over-the-new-data-protection-bill-how-is-it-different-from-the-2019-draft--news-238987>
- European Commission, Better Regulation Guidelines. (2021). Available at: https://commission.europa.eu/law/law-making-process/planning-and-proposing-law/better-regulation/better-regulation-guidelines-and-toolbox_en
- Feigenbaum, E. A., Nelson, M. R., Matthan, R., Nam, T., Park, K. S., Parsheera, S., & Ramann, S. (2022). Data Governance, Asian Alternatives: How India and Korea Are Creating New Models and Policies. *New York: Carnegie Endowment for International Peace*.
- General Data Protection Regulation. (2018). Available at: <https://gdpr-info.eu/>.
- Interinstitutional Agreement on Better Law-Making. (2016). Available at: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016Q0512%2801%29>
- *K.S. Puttaswamy v. Union of India (2017 10 SCC 1)*.
- Karnataka Data Centre Policy. (2022). Available at: <https://itbtst.karnataka.gov.in/storage/pdf-files/Data%20center%20policy%20GO.pdf>.
- Nichanemetla. (Aug, 2020), Andhra Pradesh Govt and ISB policy lab to drive evidence-based decisions for growth. *Deccan Herald*. Available at: <https://www.deccanherald.com/national/south/andhra-pradesh-govt-and-isb-policy-lab-to-drive-evidence-based-decisions-for-growth-869908.html>.
- Odisha State Data Policy. (2015). Available at: <https://data.gov.in/sites/default/files/State-data-policy-Odisha.pdf>.
- Panjir T. (2022), *A Comparison of State Level Policies , Internet Freedom Foundation*, Available at: <https://internetfreedom.in/a-comparison-of-state-level-data-policies/>.

-
- Punjab State Data Policy. (2020).
Available at: https://punjab.gov.in/wp-content/uploads/2021/02/Punjab- State-Data-Policy_Final-1.pdf.
 - Ray T., Ajaykumar S. & Patil S. (2022) The Draft Digital Personal Data Protection Bill 2022: Recommendations to the Ministry of Electronics and Information Technology, *Observer Research Foundation*.
Available at : <https://www.orfonline.org/research/the-draft-digital-personal-data-protection-bill-2022/>.
 - Sikkim Open Data Acquisition and Accessibility Policy. (2014).
Available at: <https://data.gov.in/sites/default/files/SODAAP-Policy-Document-Sikkim.pdf>
 - Tamil Nadu Data Policy. (2022).
Available at: https://tnega.tn.gov.in/assets/pdf/TN_DataPolicy_2022.pdf.
 - Telangana Data Analytics Policy. (2016).
Available at: <https://www.telangana.gov.in/PDFDocuments/Telangana-Data-Analytics-Policy.PDF>.
 - The Constitution of India (1950).
 - United Kingdom Cabinet Office. (2019) White Paper on Modernising Government.
 - World Bank. (2021). *Data for Better Lives*. Washington DC: World Bank Group.