



भारतीय रिज़र्व बैंक

RESERVE BANK OF INDIA

www.rbi.org.in

DoR.ORG.REC.XXXX/XX-XX-XXXX/2026-27

July XX, 2026

Guidance on Regulatory Expectations for Data Governance

Contents

Chapter - I Preliminary	2
A. Introduction	2
B. Applicability.....	2
C. Definitions	4
Chapter - II Governance.....	6
A. Data Governance Framework	6
B. Role of the Board	6
C. Role of Board Committees	6
D. Data Governance Executive Committee	7
E. Data Risk Management.....	8
Chapter - III Organisational structure - Data Roles and Responsibilities.....	11
A. Data Function	11
B. Data Owner	11
C. Data Steward	12
D. Data Custodian.....	12
E. Mapping of Roles and Responsibilities.....	13
Chapter - IV Data lifecycle	14
A. Data Origination and Capture	14
B. Data Processing, Sharing, and Transformation	14
C. Data Retention, Archival, and Disposal	15
Chapter – V Data Architecture	16
A. Single Source of Truth.....	16
B. Metadata and Data Lineage.....	16
C. Data Classification	17
D. Data Quality Management	18
Chapter – VI Third-Party Arrangements	19



Chapter - I Preliminary

A. Introduction

1. Data has increasingly emerged as a critical organisational asset for Regulated Entities (REs), underpinning business operations, customer service, financial reporting, regulatory compliance, risk management, and strategic decision-making. The rapid growth in digital financial services, interconnected technology ecosystems, third-party arrangements, advanced analytics, and automated decision-making processes has significantly expanded the volume, velocity and complexity of data being generated, processed, shared and stored by REs. Consequently, the ability of REs to ensure the accuracy, availability, confidentiality, consistency, integrity, and traceability of data across systems and business functions has assumed heightened importance. Supervisory assessments and stakeholder engagements with REs have indicated that while REs have progressively strengthened their information and data management capabilities, there are still certain weakness which need to be addressed else it may lead to operational, compliance, and financial risks.
2. Against this backdrop, the Reserve Bank has today issued this '**Guidance on Regulatory Expectations for Data Governance**'. It draws upon supervisory observations, stakeholder engagements, and evolving international practices, including the Basel Committee on Banking Supervision's 'Principles for effective risk data aggregation and risk reporting (BCBS 239)'. It sets out broad regulatory expectations for data governance covering various aspects including governance, data roles and responsibilities, life cycle, architecture, and third-party arrangements. It is intended to support REs in strengthening their data governance framework and ensuring reliability, consistency, availability, and security of data across systems and processes.

B. Applicability

3. This Guidance is applicable to following REs of the Reserve Bank of India:
 - (i) Commercial Banks (including Foreign Banks)



Wherein 'Commercial Banks' means all banking companies, corresponding new banks, and State Bank of India as defined under subsections (c), (da), and (nc) of Section 5 of the Banking Regulation Act, 1949.

- (ii) Small Finance Banks;
- (iii) Payments Banks;
- (iv) Local Area Banks;
- (v) Regional Rural Banks (as defined under Clause (ja) of Section 5 of the Banking Regulation Act, 1949);
- (vi) Urban Co-operative Banks;

Wherein, Urban Co-operative Banks mean Primary Co-operative Banks as defined under section 5(ccv) read with Section 56 of Banking Regulation Act, 1949.

- (vii) Rural Co-operative Banks;

Wherein, Rural Co-operative Banks mean State Co-operative Banks and Central Co-operative Banks, as defined in the National Bank for Agriculture and Rural Development Act, 1981.

- (viii) Non-Banking Financial Companies in Base Layer (NBFC - BL), Middle Layer (NBFC - ML), Upper Layer (NBFC – UL), and Top Layer (NBFC - TL);
- (ix) All-India Financial Institutions, viz., Export Import Bank of India ('EXIM Bank'), National Bank for Agriculture and Rural Development ('NABARD'), National Bank for Financing Infrastructure and Development ('NaBFID'), National Housing Bank ('NHB') and Small Industries Development Bank of India ('SIDBI');
- (x) Asset Reconstruction Companies registered with the Reserve Bank under Section 3 of the Securitisation and Reconstruction of Financial Assets and Enforcement of Security Interest Act, 2002; and,
- (xi) Credit Information Companies as defined under clause (e) of Section 2 of the Credit Information Companies (Regulation) Act, 2005.

4. The Guidance shall be read in conjunction with relevant Directions issued by RBI, as amended from time to time, or issued in substitution or succession thereto. In case of any inconsistency, the applicable Directions shall prevail.



C. Definitions

5. In this Guidance, unless the context otherwise requires,

- (1) **‘Aggregation’** means the defining, gathering, and processing of data in accordance with the RE’s risk management and reporting requirements to enable monitoring of data risk, and measurement of performance against defined objectives.
- (2) **‘Data’** means any recorded concepts, facts, information, opinion, or instructions, regardless of form or the media on which it may be recorded, suitable for communication, interpretation or processing by humans or automated means.
- (3) **‘Data Architecture’** means the structural design of data assets, systems, repositories, interfaces, and flows that support creation, collection, processing, integration, aggregation, transformation, storage, and reporting of data. This includes, but is not limited to, technology infrastructure.
- (4) **‘Data Classification’** means the systematic categorisation of data according to the criteria laid out in Data Governance Framework, reflecting its criticality, sensitivity and confidentiality, potential business and financial impact, and governance requirements.
- (5) **‘Data Custodian’** means the individual or function responsible for the technical management of data environment.
- (6) **‘Data Governance’** means the policies, roles, standards, structures, processes, and controls put in place to ensure accountability, integrity, quality, accessibility, traceability, protection, and appropriate usage of data across its lifecycle.
- (7) **‘Data Lifecycle’** means the stages through which data passes, including creation / collection, usage, sharing, transformation, storage, archival, and disposal during its entire span of existence.
- (8) **‘Data Lineage’** means the documented traceability of data from its origin through aggregation, transformation, and usage to its final destination.
- (9) **‘Data Owner’** means the individual or function accountable for data governance outcomes for all the data under her domain.



- (10) '**Data Risk**' means risk of adverse outcomes arising from deficiencies in data, or in its management across its lifecycle, including data architecture.
- (11) '**Data Steward**' means the individual or team under Data Owner supporting in operationalisation of data governance processes.
- (12) '**Metadata**' means 'data about data'. It is structured information that describes, explains, locates, or otherwise enables understanding, retrieval, or management of data.
- (13) '**Personal Data**' means personal data as defined under the Digital Personal Data Protection Act, 2023, as amended from time to time.
- (14) '**Single Source of Truth**' means a single designated authoritative source for specific data element within the RE.
- (15) '**Transformation**' means the process of converting data from one format, structure, or representation, to another.



Chapter - II Governance

A. Data Governance Framework

6. An RE should put in place a Data Governance Framework (DGF) applicable to all data and align it with its risk management framework. It should ensure that the DGF:
 - (i) is proportionate to the RE's size, complexity, business model, and Information Technology (IT) and Information Security (IS) set-up;
 - (ii) is comprehensive and covers all aspects of data governance including organisational structure, policies, and processes, risk management including data privacy and security, technological infrastructure, and audit mechanisms across the data lifecycle;
 - (iii) covers all material aspects of data governance, including its lifecycle, quality, classification, single source of truth (SSOT) arrangements, metadata, lineage, and third-party arrangements; and
 - (iv) complies with the Digital Personal Data Protection (DPDP) Act, 2023, the DPDP Rules, 2025, and all other applicable laws and rules, as amended from time to time.
7. The DGF should be reviewed annually or more frequently as required.

B. Role of the Board

8. The Board should oversee the DGF of the RE, and review the reports and metrics placed before it.

C. Role of Board Committees

9. An RE should establish a Board level Data Governance Committee (DGC) or assign the responsibility to an existing Committee of the Board.
10. The Committee responsible for data governance should:
 - (i) oversee the implementation of the DGF;
 - (ii) formulate policy / policies for data governance which should, *inter alia*, include (a) scope of data governance; (b) data architecture; (c) management of data risk; (d)



ownership, accountability, and responsibility for data across the data lifecycle; (e) data quality management; (f) data classification framework; and (g) arrangements with third-parties;

(iii) review them periodically or on material changes such as in business activities, system architecture, data usage, and regulatory expectations; and,

(iv) place before the Board the reports / metrics and material issues, (e.g., breaches and conflicts), if any.

D. Data Governance Executive Committee

11. An RE should establish an executive level Data Governance Committee or delegate the responsibility to an existing Executive Committee with representation from Data Function, IT, IS, relevant Business Verticals, Risk Management, Compliance, and others, as required.

12. It should be responsible for implementing, operationalising, and effective functioning of the DGF and policies framed thereunder.

13. It should ensure that:

- (i) adequate organisational structures, with defined roles, and capabilities are in place for effective data governance, including coordination across business, risk, finance, technology, and other relevant functions;
- (ii) adequate resources, including personnel, systems, and budget, are allocated for effective implementation of DGF; and,
- (iii) personnel assigned data governance responsibilities have the requisite authority, independence, and competence for effective discharge of their responsibilities.

14. The Committee should, *inter alia*:

- (i) review instances of non-compliance with DGF and related policies, regulatory and legal requirements, and initiate remediation measures;
- (ii) address structural gaps, and conflicts across functions;
- (iii) periodically review data-related metrics, breaches and exceptions, major audit observations, and ensure timely remediation;



- (iv) oversee data classification criteria and methodology, including identification of critical data elements;
- (v) approve material exceptions / deviations, and ensure that they are documented and reviewed; and
- (vi) escalate material Committee of the Board to which powers have been delegated.

E. Data Risk Management

15.(1) An RE should establish processes to manage data risk as a part of its overall risk management framework. It should include identification of data attributes (e.g., quantitative and qualitative), data structure (e.g., structured, unstructured, and semi-structured), data sources (e.g external and internal), data quality (e.g., accuracy, completeness, consistency, timeliness, relevance, validity, and reliability), data classification, and big data perspective (e.g., volume, veracity, velocity, and value), for identification, assessment, monitoring, and managing risks pertaining to data. It should cover *inter alia* the following:

- (i) Accountability (for data domains, data architecture, monitoring and managing associated data risks, and remediation measures).
- (ii) Integrity (protected against unauthorised modification, corruption, or loss throughout its lifecycle).
- (iii) Auditability (data processes, transformation, and controls should be subject to independent review and audit, with appropriate documentation).
- (iv) Transparency (data definitions, sources, assumptions, transformation logic, and limitations should be documented and made available to relevant stakeholders).
- (v) Traceability (data should be capable of being traced to its origin and through material transformations and reporting layers).
- (vi) Proportionality (controls should be commensurate with the criticality, risk, and reliance placed upon the data).
- (vii) Standardisation (data standards, definitions, and business rules should be applied uniformly where the data has same interpretation across systems and functions).



- (2) It should assess inherent and potential risks, across data lifecycle, organisational structure and technological infrastructure (e.g., data architecture, data ownership, data quality, data repositories, data flows, data interdependencies, data privacy, and data security).
 - (3) It should assess and put in place controls to mitigate data-related risks arising from third-party arrangements.
 - (4) It should assess and manage data-related risks in respect of cross-border operations including processing, storage, transfer and usage, and ensure that such cross-border operations do not impair the RE's ability to access, retrieve, and manage its data.
- 16.** It should assess data risks proportionate to the criticality, sensitivity, and legal and regulatory relevance of data.
- 17.** Information on key data risks, incidents, and remediation actions should be placed before the Committee of the Board.
- 18.** An RE should ensure that customer-related data is handled in compliance with DPDP Act, 2023, and DPDP Rules, 2025.
- 19.** An RE should assess the effectiveness of its DGF, and policies, processes, and controls thereunder, on an ongoing basis, which should, *inter alia*, cover:
- (i) timely identification and remediation of weaknesses and emerging risks;
 - (ii) compliance to the internal policies, and legal and regulatory requirements;
 - (iii) adequacy of assurance functions; and
 - (iv) escalation of material issues.
- It should document such ongoing monitoring, remediation measures, and lessons learnt.
- 20.** An RE should ensure that DGF and policies and processes thereunder are subject to periodic internal and external audit, including through CERT-IN empanelled auditors, as applicable. It should adopt a risk-based approach for audit processes, including its frequency, scope, and depth. The audit reports should be placed before the Audit Committee of Board for review.



21. An RE should periodically review and update its processes for data risk management to reflect changes in business activities, system landscape, data usage, and regulatory expectations.



Chapter - III Organisational structure - Data Roles and Responsibilities

A. Data Function

- 22.** An RE should establish a Data Function headed by a sufficiently Senior officer not below the rank of Chief General Manager or equivalent having adequate authority and with necessary competence/ skills for implementation of DGF.
- 23.** The Data Function should act as the central point of coordination to ensure consistent interpretation and implementation of the DGF and related policies thereunder across business, risk, technology, and other relevant functions. It should:
- (i) develop metrics to monitor effectiveness of DGF;
 - (ii) oversee effectiveness of organisational structure and technological infrastructure, established to support data functions;
 - (iii) facilitate resolution of conflicts *inter alia* in respect of data definitions, classification, SSOT designation, metadata management, lineage, standards, consent management, and remediation across data domains; and,
 - (iv) ensure documentation in respect of standards, processes, definitions, taxonomies, and other data related aspects.

B. Data Owner

- 24.** An RE should designate a Data Owner for each data domain accountable and responsible for ensuring that data within the domain is defined, classified, and used in a manner consistent with the DGF. It should also be applicable where a data domain spans multiple business functions, units, or systems.
- 25.** The Data Owner should, *inter alia*, be responsible for:
- (i) business logic, definitions, and interpretation of data;
 - (ii) intended use, reuse, and sharing;
 - (iii) classification of data;
 - (iv) key data rules, including validation, derivation, aggregation, and exception;
 - (v) quality of the data;
 - (vi) oversight of completeness and correctness of metadata and lineage;



- (vii) designation of SSOT and other authoritative sources;
- (viii) maintenance and review of documentation in respect of data including definition, metadata, classification, data dictionary, permissible, sharing and usage, quality metrics, and other relevant data attributes; and,
- (ix) approving changes, exceptions, and remediation plans affecting data within its domain.

C. Data Steward

26. A Data Steward, mapped to specific data domain under a Data Owner, and positioned within the business function or unit of the Data Owner should be designated by the RE.

27. The Data Steward should be responsible for implementation of data governance requirements across systems, processes, and lifecycle stages, in accordance with decisions and standards set by the Data Owner which, *inter alia*, includes:

- (i) day-to-day operations of data governance by connecting business logic and rules with system designs and data structures, and supporting preparation of data definitions, data element specifications, metadata, and related documentation;
- (ii) coordination for data governance issues within the data domain and facilitating resolution across cross-functional data or that spans multiple systems or processes;
- (iii) supporting, documenting, and monitoring data flows as approved by the Data Owner, and undertaking periodic analysis thereof to identify gaps and breaches; and
- (iv) preparing, maintaining, and reviewing domain specific data metrics and report the same, including deviations or material data governance issues, to the Data Owner.

D. Data Custodian

28. An RE should designate Data Custodian, responsible, *inter alia*, for:



- (i) enforcing access controls, and user entitlements in line with data classification and approved usage;
- (ii) maintaining system capabilities to support effective implementation of DGF;
- (iii) maintaining technical capabilities for data metadata management, lineage, SSOT, reconciliation, troubleshooting, and root cause analysis;
- (iv) implementing technical controls relating to data transmission, storage, backup, and availability;
- (v) implementing data retention, archival, and disposal processes and controls; and
- (vi) implementing and maintaining business continuity and disaster recovery mechanisms for the systems and data platforms.

29. The Data Custodian should promptly report the Data Owner of material incidents, control failures, or issues, and escalate the same to Data Governance Executive Committee, as appropriate.

E. Mapping of Roles and Responsibilities

30. An RE should have a structured mapping of data governance roles to processes and data lifecycle stages, which should *inter-alia* include:

- (i) accountability for data governance outcomes;
- (ii) responsibility for execution and coordination;
- (iii) points of consultation where multiple functions are involved; and
- (iv) escalation and decision-making.

31. It should ensure that the mapping is documented, reviewed, and updated periodically and in case of material changes (e.g., internal changes such as data usage, systems, or organisational structure, and external changes such as change in regulation).



Chapter - IV Data lifecycle

32. An RE should establish and implement a lifecycle-based approach to data governance, ensuring that the data is governed across all stages, i.e., origination to deletion, in a consistent manner to identify the associated risks and mitigate them.

33. Management of data should include:

- (i) metadata (including on ownership, permissible usage, and reporting context);
- (ii) lineage (including origin, transformation, aggregation, and downstream usage);
- (iii) rules and logic applied to data, such as validation, derivation, aggregation, and exception handling;
- (iv) dependencies between data elements, datasets, systems, and reports; and,
- (v) consent management for customer data.

A. Data Origination and Capture

34. The RE should ensure that data is created or acquired only for defined and legitimate purposes aligned with approved business, risk, and legal and regulatory objectives.

35. It should ensure that, sufficient foundational attributes of data (e.g., ownership, classification, usage intent, appropriate consent in case of customer data) are established at the point of origination or capture, to enable downstream governance.

36. It should further ensure that systems and processes used for data capture are capable of enforcing validation, completeness, and consistency checks, commensurate with the criticality and classification of the data.

37. The data captured from external sources, including third parties, should be subject to governance standards equivalent to those applied to internally generated data.

B. Data Processing, Sharing, and Transformation

38. An RE should ensure that processing, sharing (internally or externally), transformation, enrichment, or derivation of data is done using approved rules, logic, and standards.



39. It should deploy appropriate security controls (e.g., mechanisms for data loss prevention, data encryption, tokenization, and anonymisation) to protect data both at rest and in transit.

40. It should ensure that:

- (i) accountability is defined for transformed and derived data.
- (ii) transformation logic is subject to impact assessment prior to implementation, and review.
- (iii) transformation preserves the intended meaning of data and does not lead to ambiguity or distortion.
- (iv) transformed or derived data is traceable to its source, and appropriately classified.

C. Data Retention, Archival, and Disposal

41. An RE should have Data Retention and Archival Policy under its DGF. It should ensure that data retention period is justified by business, legal, regulatory, or audit requirements. The policy should be reviewed periodically.

42.A. Under the policy, the RE should ensure that:

- (i) data should be available and usable throughout the retention period;
- (ii) storage and archival systems should preserve linkage to data definition, classification, and lineage, and should not result in loss of context over time;
- (iii) archival of data should not impede timely retrieval of data required for supervisory, audit, or investigative purposes; and,

B. An RE should ensure that the disposal of data is governed by approved policies and carried out in a secure, controlled, and verifiable manner proportionate to the criticality, sensitivity, and classification of data.



Chapter – V Data Architecture

A. Single Source of Truth

- 43.** An RE should establish and maintain a SSOT for all data elements and ensure that no parallel or competing SSOT sources exist for the same data element. All downstream systems, models, and business processes should be derived from the SSOT.
- 44.** For this purpose, the RE may adopt any model for SSOT (e.g., centralised, federated, or hybrid), provided the SSOT architecture, at a minimum, ensures:
- (1) clear identification of the authoritative source for data;
 - (2) consistency of data used across business, risk, compliance, and all other functions of the RE; and
 - (3) traceability of aggregated data and reports.
- 45.** The SSOT designation and any changes to it should be approved by the Data Governance Executive Committee and documented. The same should be put up to DGC for information.
- 46.** The RE should establish reconciliation mechanism to identify and resolve inconsistencies, if any, between the SSOT and downstream data.

B. Metadata and Data Lineage

- 47.** An RE should establish and maintain metadata for all data across the lifecycle and systems. Metadata, at minimum, should have attributes to support consistent understanding, discovery and governance of data (e.g., business meaning, ownership, designated SSOT, approved usage and data flows).
- 48.** The RE should ensure that minimum foundational metadata is created at the time of data capture or generation, which is sufficient to identify:
- (i) the source application or system generating the data;
 - (ii) the purpose of collection of data;
 - (iii) the data owner;
 - (iv) applicable classification; and
 - (v) retention and permitted usage requirements.



- 49.** For system-generated data (e.g., logs, audit trails, cookies, and other high-volume technical data), the RE may maintain metadata proportionate to the classification and criticality of data, through individual records, or aggregated datasets.
- 50.** The metadata should flow with data such that it moves to downstream systems (e.g., master data management systems, data warehouses), without loss of basic attributes created at source.
- 51.** The metadata should be updated when data is transformed or derived, and at minimum should include:
- (i) relationship between source and transformed or derived data;
 - (ii) purpose and nature of the transformation or derivation; and
 - (iii) changes in classification and accessibility of data.

C. Data Classification

- 52.** An RE should establish a well-defined framework for classification of data in a manner that reflects the risk, criticality, and sensitivity of such data. The framework should:
- (i) provide clarity on permissible usage, sharing, and access;
 - (ii) enable consistent application of controls across data lifecycle;
 - (iii) should be consistent with legal, regulatory, and supervisory expectations; and
 - (iv) support prioritisation of monitoring and escalation, remediation, and incident management.
- 53.** The data classification framework, at minimum, should consider either singly or in combination the following aspects:
- (i) criticality to business operations, including impact on continuity, decision-making, financial outcome, and reputation;
 - (ii) criticality to customers, including potential harm arising from misuse, inaccuracy, or unavailability of data;
 - (iii) sensitivity of data, including whether data comprises personal data, or any other data having implications regarding privacy;
 - (iv) confidentiality including potential impact on stakeholders, if data is disclosed or accessed without authorisation; and,
 - (v) legal and regulatory information.



- 54. The framework should include impact owing to data aggregation, enrichment, and derivation such as changes in data materiality, risk profile or sensitivity.
- 55. The RE should ensure that data classification is captured as part of metadata, remains traceable across the data lifecycle, and is reviewed periodically and upon material changes.

D. Data Quality Management

- 56. An RE should establish data quality management processes proportionate to its criticality, sensitivity, and classification, and ensure that data is fit for its intended use and the material issues are identified and remediated in a timely manner.
- 57. It should ensure that gaps in the data quality do not compromise decision-making, risk management, or legal and regulatory reporting.
- 58. It should develop and maintain data quality metrics across multiple dimensions, for enabling identification of trends, emerging risks, and gaps in controls.
- 59. The data quality metrics report and persistent data quality issues should be placed before the DGC at quarterly or more frequent intervals.



Chapter – VI Third-Party Arrangements

- 60.** An RE should be responsible for governance of data shared with third-parties, including group entities. Provided that for the purpose of this Guidance, ‘Group’ shall have the same meaning as defined in the [Reserve Bank of India \(Commercial Banks – Concentration Risk Management\) Directions, 2025](#), as amended from time to time.
- 61.** It should ensure that data is shared with third parties only for defined and approved purposes and by designated personnel.
- 62.** It should put in place systems and controls for access, usage, and deletion of data shared with third-parties by taking into account, *inter alia*, the following:
- (i) the classification, sensitivity, and criticality of the data;
 - (ii) whether the data includes critical data elements, including personal, customer-related, or protected information;
 - (iii) the nature and extent of usage of data by the third party;
 - (iv) potential impact on the RE, its customers, and its operations in the event of misuse, unauthorised access, loss, or unavailability of data; and
 - (v) consent in case of customer data.
- 63.** The RE should ensure that:
- (i) data shared with third-parties remains traceable to the designated SSOT, and metadata and lineage capture the extent of such sharing;
 - (ii) access to data by third-parties is on a ‘need to know’ basis;
 - (iii) non-disclosure clauses are included in its agreements with third-parties;
 - (iv) sharing of data does not result in unauthorised reuse, sharing, or duplication;
 - (v) secure channels and standards are established for sharing of data (e.g., encryption, authentication controls, access controls, auto-deletion, and de-duplication controls);
 - (vi) periodic audit of systems and applications of third-parties is carried out either by itself or through external auditors, including CERT-IN empaneled auditors, as required;



(vii) sharing of data with third parties is subject to ongoing monitoring and periodic reassessment, particularly where data is critical or subject to enhanced monitoring.