CH 9 DATA GOVERNANCE

📘 DATA GOVERNANCE – Summary Notes

🔹 *Introduction*

- DG = Framework to ensure **data security, privacy, accuracy, availability, usability**
- Involves **people + processes + tech** across **data lifecycle**
- Ensures **data integrity, consistency, compliance**, avoids misuse
- Key due to **data privacy laws + data analytics reliance**
- Requires: **Governance Team + Steering Committee + Data Stewards**
- Success = alignment with **business goals**, not just IT
- **Evolution**:
  - Old: IT-centric, data cataloging only
  - DG 2.0: Collab, accountability, data-driven ops
  - 2018 onward: Data breaches → need for robust DG

🔹 *Core Functions*

1. **Organizing** – Identify sources, centralize data
2. **Securing** – Ensure compliance (laws + policies)
3. **Managing & Presenting** – Define how data is shared internally
4. **Tech Use** – Deploy DG platforms & tools

📘 DATA GOVERNANCE PRINCIPLES

1. **Integrity**
   - Ethical data usage (means & goals)
   - Honest decisions, clear impact
2. **Transparency**
   - Disclose how/why/whose data is used
   - Ensure clear comms on control decisions
3. **Accountability & Ownership**
   - Assign **ownership**
   - Define **access rights**
   - Governance covers **decisions, processes, controls**
4. **Auditability**
   - All data decisions/processes must be **auditable**
   - Maintain documentation for **compliance**
5. **Standardization**
   - Uniform format across teams
   - Rules: **definition, access, security, privacy**
6. **Change Management**
   - Control **proactive/reactive** data changes
   - Covers: **reference values, metadata, master data**
7. **Stewardship**
   - Assign **data stewards**
   - Ensure compliance + best practices
   - Responsibility = **data integrity + ethical use**

📘 IMPORTANCE OF DATA GOVERNANCE (DG)

- In **info era** (big data + tech + analytics) → DG = **critical**
- DG = ensures **rules, quality, compliance, interpretation**
- Helps in **business planning, risk control, data-driven decisions**

🔹 *Key Benefits*

1. **Better Analysis**
   - ↑ Data quality & findability
   - Faster, compliant analytics
2. **Clear Business Goals**
   - DG = roadmap for achieving **defined objectives**
3. **Regulatory Compliance**
   - Ensures adherence to laws
   - ↓ Risk of **fines + reputational damage**
4. **Efficient Data Management**
   - Avoids duplication
   - ↑ Operational efficiency
5. **Standardization**
   - Uniform **systems + data policies**
   - Builds **ethics & data awareness**
6. **Improved Data Quality**
   - Accurate data = better **business outcomes**
   - ↑ **Customer trust**

🔹 *Conclusion*

- DG = **essential** for:
  - Healthy data culture
  - Fast, reliable decisions
  - **Trust, transparency, integrity**
- Without DG: **inaccurate, slow, unreliable** data usage

📘 DATA GOVERNANCE FRAMEWORK – Exhibit 1

Core Components:

- **Ownership** – Define data owners/stewards
- **Knowledge** – Metadata, documentation

- **Technology** – DG platforms/tools
- **People** – Roles, responsibilities
- **Accessibility** – Controlled, role-based access
- **Process** – Policies, workflows, compliance
- **Quality** – Clean, consistent, reliable data
- **Security** – Data protection, privacy laws

📘 DATA GOVERNANCE CHALLENGES

1️⃣ Understanding Business Value of DG

- Biz mgrs prefer **faster access** > data **quality/control**
- Legacy orgs = **slow digital adoption**
- Need to show DG = **foundation** for data democratization
- Role of **CDO** evolved → peer to CIO, ensures DG architecture
- **Analytics teams** suffer most from low data quality

2️⃣ Perception: IT Owns the Data

- Past view: DG = IT job → seen as **non-critical**, gets **budget cuts**
- True: **Business owns data**, IT = enabler (e.g., DBA ≠ data owner)
- Need: Shift **ownership to departments** (e.g., Sales owns sales data)
- Appoint **data steward** (links business + DBA)
- Case: EU utility → 30% ↓ external cost, 50% ↓ discovery time, €8M benefit

---

3️⃣ Limited / Misallocated Resources

- DG needs **Data Owners** (decision-makers) & **Stewards** (execution)
- Common: **part-time roles**, lacks focus
- Use **business analysts / BI experts** as temp stewards
- Long-term: Hire **full-time stewards**
- Alt: Hire **consultants** (DG experts, not generic IT firms)

4️⃣ Siloed Data

- Legacy systems + tech gaps = **data silos**
- Biz units = siloed → ops ≠ strategy
- Hybrid DBs (e.g., NoSQL + relational) = ↑ silos
- Solution: **Metadata mgmt** + tools to **map & link systems**
- View data **uniformly**, update architecture **dynamically**

5️⃣ Poor Data Quality & Trust Issues

- Poor quality → ↓ trust → ↓ use → ↓ ROI
- Examples:
  - 75% blank fields
  - Duplicate vendor rows
  - Dummy data not deleted
- Solution:
  - **Data intelligence tools** → metrics + thresholds
  - **Feedback loop** for quality issues
  - Ensure **transparency** in data quality

6️⃣ Poor Data Context

- Not quality issue, but **understanding gap**
- E.g., "Final sales" = Gross or Net? Leads vs. MQL?
- Misinterpretation = ↓ trust
- Fix via:
  - **Clear labeling**
  - **Feedback loop** for clarification

7️⃣ Lack of Proper Data Control

- IT control > DG = blocks access
- True control = DG-driven, **context-aware access**
- Aim: Say "**yes with control**", not "no by default"
- Tools help **map assets**, define use, prevent misuse

📊 Data Governance vs. Data Management

S. No.	Basis	Data Governance (DG)	Data Management (DM)
1️⃣	**Definition**	Process to define & enforce **business rules** for data	Process to **collect, manage, use & dispose** data properly
2️⃣	**Scope**	Ensures **data quality** & proper **DM processes**	Ensures **data storage, access, & availability** to authorized users
3️⃣	**Benefits**	Framework → ↑ **efficiency**, consistency, ↓ errors	↑ **efficiency, accuracy, security**, ↓ human errors
4️⃣	**Challenges**	Hard to **start**, employee **resistance** to new rules	No standard format, difficult **data retrieval** across sources
5️⃣	**Best**	**Data**	Good **naming,**

S. No.	Basis	Data Governance (DG)	Data Management (DM)
	Practices	**classification**, **metadata use**, **access control**	**labelling**, **backup**, **security**, data quality
6	**Technology**	Tech for **automation**, **monitoring**, process enhancement	Tools: **DBMS, Data Mining, Visualization, Big Data Tech**

📌 Implementing Effective Data Governance Framework

✓ *Objectives*

- Stdzn. of data definitions across enterprise
- Align data governance with **business drivers**
  - e.g. Healthcare → patient privacy, compliance, access

🔑 *Key Components*

- **Covers:** Strategic, tactical & operational roles
- Ensures:
  - 📌 Trusted, documented, discoverable data
  - 📌 Secure, compliant, confidential handling

⭐ *Advantages*

1. Consistent data view + business glossary
2. Ensures: data quality, accuracy, completeness
3. Tracks location of critical data assets
4. "Single version of truth"
5. Org-wide std. methodologies & best practices
6. Secure + compliant access per legal/regulatory needs

🏛️ 3 Pillars of Data Governance Framework

Pillar	Description
i) **All Data Assets**	Includes dashboards, code, models → **Analytics Governance**
ii) **Bottom-Up, Practitioner-led**	Decentralized; all creators responsible (data mesh model)
iii) **Daily Workflow Embedding**	Integrated into practitioner tasks, not an afterthought

---

🔄 Traditional Approaches

i) Top-Down Approach

- **Focus:** Control
- Centralized; IT governs → business consumes
- Pros: Strong data modelling & governance
- Cons: Not scalable, slow access, gatekeeper issue

ii) Bottom-Up Approach

- **Focus:** Access
- Starts with raw data → schema-on-read
- Pros: Scalable, agile
- Cons: ↓ Control, ↑ Regulatory Risk, Data Sprawl

🔄 Need for Modern Approach

- Balance **control + access**
- Establish governance **early**
- Empower **users** as data owners & curators

📌 Steps to Create Data Governance Framework

i) Revisit Definition of Data Governance

- Re-evaluate purpose & scope before framework design
- Key Qs:
  a) Purpose?
  b) Covers all data assets?
  c) Promotes sharing & collaboration?

---

ii) Identify & Define Data Domains

- Stdzn. of domains (e.g., Finance, HR, Marketing)
- Key Qs:
  a) Major domains?
  b) What data generated?
  c) Data location?
  d) Who uses it?

---

iii) Identify Data Owners & Consumers

- Assign ownership & responsibility to each domain
- Shared responsibility model
- Key Qs:
  a) Who creates data?
  b) Who consumes? Daily workflows?
  c) Access dependencies?

---

iv) Validate & Document All

- Stdzn. of:
  - Definitions, data flow, access policies, workflows
- Key Qs:
  a) Data source?
  b) Meaning?
  c) Flow path?
  d) Goal alignment?

---

v) Conduct Security & Risk Assessment

- Ongoing reviews of access & risks
- Key Qs:
  a) Existing policies?
  b) Who accesses what, why?
  c) Risk mitigation vs accessibility?

IN Data Governance at Govt. Level – Indian Scenario

Evolution & Initiatives

- 1881: First Census
- 1950: NSSO; 1951: CSO
- Manual data collection → MIS + Dashboards (last 20 yrs)
- Emphasis on Open Data →
  - *National Data Sharing & Accessibility Policy*
  - *data.gov.in* portal for public use

Current Govt. Frameworks

- MIS for schemes at all levels (village → state)
- Dashboards, basic analytics → quick insights
- Platforms for integration:
  - **DISHA**, **Prayas**, **OOMF**

Data Governance Quality Index (DGQI)

- Purpose: Assess & improve data preparedness of Ministries/Departments
- Encourages:
  ✔ Peer learning
  ✔ Healthy competition

3 Pillars of DGQI

1. **Data Strategy** – Guidelines
2. **Data Systems** – Collection, mgmt, usage
3. **Data-Driven Outcomes** – Decision support & sharing

- **DGQI 1.0**: Focus on Data Systems
- **DGQI 2.0**: Covers all 3 pillars

Current Scenario – Data Governance in India

1 Existing Govt. Systems

- Most schemes have MIS (e.g., **HMIS** for NHM) → tracks state data for planning, mgmt., decisions
- MIS → generate **standardized analytical reports**
- **MoSPI** monitors infra projects via **TPP-2006** & **IPMD**

---

2 Digital India Initiative (2015)

- Managed by **NeGD** → supports e-Governance
- States using dashboards (e.g., **Pratibimba** – Karnataka)
- Promotes **accountability + transparency**
- Govt. active in tech adoption → improving **outputs & outcomes**

---

3 Need for Review

- Gap: Programme monitoring & decision support still weak
- Solution: Comprehensive **data preparedness review** via **DGQI Toolkit** by **DMEO (NITI Aayog)**

Objectives of DGQI

- Self-assess data systems on maturity scale
- Tool for **internal diagnostics** for M/Ds
- Enable **comparative assessment** & share best IT practices
- Goal: Create **Overarching Dashboard** (online, API-based) for all CS/CSS schemes

---

DGQI Methodology – 6 Key Themes

1. Data Generation

- Checks **digitization level**, **data frequency**, **granularity**
- Use of **mobile, GIS, location tracking**

2. Data Quality

- Covers:
  ✔ Profiling
  ✔ Cleaning
  ✔ Pipeline design
  ✔ Schema stdzn.
  ✔ Mobile tech use

3. Use of Technology

- Linkage with **PFMS, JAM (Jan Dhan, Aadhar, Mobile)**
- Use of:
  ✔ Remote sensing

✔ Social media
✔ AI/ML, Blockchain, IoT

4. Data Analysis, Use & Dissemination

- From basic to **predictive analysis**
- Use of dashboards
- **Social media dissemination**, multilingual support
- **GIGW compliance** (Govt. website stds)

5. Data Security & HR Capacity

- Anti-virus updates, internal audits
- Dedicated data teams (HR)
- Not exhaustive → just basic capacity checks

6. Case Studies

- Highlight best practices not in structured Qs
- Focus: Scheme-level MIS
- Enable inter-ministerial learning & innovation

National Data Governance Policy – Union Budget 2023-24

- Focus: Enable access to **anonymised (non-personal) data** for start-ups, academia
- **IDMO** to be set up under **Digital India Corporation**
- Draft **Data Protection Bill 2022** doesn't cover **non-personal data**
- Separate from **Personal Data Protection Bill**
- Plan: Launch **India Datasets Programme**
  - Govt. & private orgs to share anonymised datasets
  - Boost to **research + innovation**

Banking Sector – Data Governance

Key Focus:

Compliance, Innovation, Risk Mgmt., Cost-efficiency, ESG integration

---

Importance of Data Governance in Banking

i) Regulatory Compliance

- Meet **state + federal norms**
- Locate + tag data → **cloud migration + digital transformation**
- Enforce **security controls** effectively

ii) Cost Cutting

- Manual mgmt = High IT cost
- **Centralized governance tools** = ↓ 3rd-party tools + ↓ IT burden
- **Self-service access** = Cost-effective secure data use

iii) Market Insights

- Competitive edge via **data analysis**
- Supports org-wide data access + innovation
- Promotes **data democratization**

iv) Data-Driven Culture

- Better biz goals, CX, innovation
- Culture shift → **data-centric ops**

v) Collaboration & Risk Mgmt.

- Create **data catalogs** → quality + discovery
- Helps in:
  - ✔ New customer acquisition
  - ✔ Fraud detection
  - ✔ Risk analysis
  - ✔ Decision-making

vi) Compliance + Customer Experience

- Enhance **secure CX** (e.g., Fifth Third Bank)
- Use platforms (e.g., **NCBA**) to track **customer journey**, improve personalization

---

ESG Data Governance – Need in Banking

Why ESG?

- New ESG regulations = ↑ Transparency, disclosures
- Stakeholders → demand **climate + social impact** data

Challenges

- Lack of **integrated ESG data systems**
- Fragmented, outdated IT infra

Requirements

- Revamp **IT architecture** → ESG-ready
- Tools:
  - ✔ ESG scorecards
  - ✔ Climate models
  - ✔ Financed emissions tracking
  - ✔ Climate-adjusted ratings
- Integrate ESG in:
  - ✔ **Credit approvals**
  - ✔ Risk mgmt.
  - ✔ Strategic decisions

Steps for ESG Integration

- **AI in decision-making** (credit, risk)
- **Org-wide ESG awareness** → change mgmt.
- Update data processes → ↑ frequency, real-time
- Add ESG certificates to investment frameworks
- Design data architecture → ESG-centric

---

2. Automobile Sector – Data Governance

Focus: Big Data, IoT, Connected Vehicles, AI, Sensors, GDPR, IP, Predictive Models

---

Industry Stats & Evolution

- 2021: $4,500M → Projected $15,800M by 2030 (CAGR 17%)
- Driven by: **Big Data**, AI, Sensors, M2M, IoT
- Auto sector adapting to **Industry 4.0**

German Association – 4 Data Categories

a) Traffic Data

- Anonymized → Shared with **police/fire**
- Used in **real-time alert systems**

b) Usage Data

- Anonymized → 3rd parties
- Enables **new biz models**, telematics-driven innovation

c) IP Data

- Exclusive to OEMs + partners
- Used for **product lifecycle improvement**, brand-specific services

d) Personal Data

- Requires **user consent** (GDPR)
- Enables **personalized services**, needs strict privacy governance

---

Connected & Automated Vehicles (OECD View)

- **Connected Vehicles**:
  ↳ Communicate via Internet (V2X, V2I, V2V)
  ↳ Share data w/ 3rd-party systems
  ↳ Use **cellular + non-cellular tech**
- **Automated Vehicles**:
  ↳ Use sensors, AI, partial-to-full autonomy
  ↳ Rely on **connectivity** for real-time updates + navigation
  ↳ Connected = part of **IoT ecosystem**

Key Data Types Generated

i) Locational Data

- GPS coords, speed, direction
- Used w/ **HD maps, weather, traffic**

ii) Sensor Data

- Inputs from LIDAR, radar, cameras
- Detect infra, obstacles, road signals

iii) Diagnostic Data

- Fuel use, emissions, battery/engine status

iv) Driving Behaviour Data

- Speeding, braking, seatbelt usage

v) Identity + Biometric Data

- Names, IDs, fingerprints (w/ consent)

Data Governance Implications

- OEMs act as **Data Gatekeepers** (via user T&C)
- Real-time **2-way data exchange** = key for autonomous function
- Needs **interoperability** across:
  ✔ OEMs
  ✔ Regulators
  ✔ Infra providers
  ✔ 3rd parties

Uses of Vehicle Data

- **Public Policy**: traffic mgmt, infra planning, safety
- **Ops Efficiency**: predictive maintenance, smart routing
- **New Biz Models**: usage-based insurance, dynamic pricing
- **Customer Experience**: personalization, proactive services

---

Emerging Data Governance – Connected & Automated Vehicles

Focus: Data Collection, Processing, Access, Sharing, Reporting, Privacy

---

a) Data Collection & Mgmt at Vehicle Level

- Vehicle data often **personal** (driver/passenger info)
- Existing **privacy risks** → sectoral principles supplement general laws
- **EU**: ACEA's *Principles of Data Protection*
- **US**: Alliance for Automotive Innovation → *Consumer Privacy Protection Principles*
- Example: Otonomo's *Privacy Playbook* (2019)
- Some **Data Authorities** also publish guidelines

---

b) Data Processing Within/Near Vehicle

- Processing choice = mix of **technical needs + legal obligations**
- Stakeholders: OEMs, regulators, cloud/network providers, 3rd parties
- **EDPB Recommendation**:
  ✔ Prefer **local (in-vehicle) processing**
  ✔ Avoid excessive **external cloud use**
  ✔ Enhances **user control** + reduces **cloud-related risks**

---

c) Access to In-Vehicle Data

- OEMs hold **effective control** post user consent (usually mandatory for vehicle use)
- Justification: **privacy + security**
- **Issue**: Owner ≠ sole user → data control disputes
- **3rd parties (repair, insurance, etc.)** demand access for competition
- EU: Ongoing efforts for **GDPR-compliant access**
- Countries w/ set rules: **Austria, France, Germany**

---

d) Models of Data Sharing & Access

- Data = **non-rivalrous, reusable** → complex to commoditize
- Key users: OEMs, 3rd parties (e.g., repair, entertainment, insurance)
- Trials: Data-sharing among innovators = **faster experimentation**
- Current gap: No mature **data markets** due to info asymmetry
- Need for **new frameworks** to incentivize sharing across the value chain

---

e) Data Reporting & Mandatory Collection

- Potential: Data → **public value** (road safety, infra planning, transport policy)
- Some OECD jurisdictions mandate **vehicle data reporting**
- **Europe**:
  ✔ *Data for Road Safety Initiative* → OEMs share hazard data w/ road authorities
  ✔ Tech already detects **dangerous conditions**
- **International Transport Forum** → *Mobility Data Reporting Principles*
- **Mandatory Devices**:
  ✔ *Event Data Recorders* required in some regions (crash data storage)
  ✔ **Similar systems** under discussion for **automated vehicles** (since 2020)

⚡Data Governance in Energy Sector

🔑 Focus: Innovation vs. Compliance | Data Quality | Digital Maturity | Strategy Alignment

---

1) Data Surge in Energy Sector

- Data from **smart meters, IoT, sensors** increasing rapidly
- Use of **BI, AI, ML, Predictive Analytics** → risk of breaching **privacy/regulations**
- Balance needed:
  ✔ **Too much regulation** → hinders innovation
  ✔ **Too little** → breaches trust, non-compliance
- Regulatory frameworks must evolve with tech

---

2) Limitations in Current Governance

- Most power/utilities cos. not fully using data nor enabling data sharing
- **Existing strategies ≠ meet today's needs**, let alone **future needs** (e.g., emissions, sustainability)
- Data vital for:
  ✔ Compliance
  ✔ Grid reliability
  ✔ Financial reporting
  ✔ Disaster response
  ✔ Peak demand planning

---

3) Data as Strategic Asset

- Quality data → **operational efficiency + customer experience**
- Enables **self-service analytics + innovation**
- Need for **agile data marketplace**

---

4) EY Survey Highlights (Electricity Industry)

i) Data Strategy Gaps

- No respondent had **Chief Data Officer (CDO)**
- < ⅓ had a **formal data strategy** (covering IT + Business)

ii) Leadership Link

- Strong link between **senior leadership's commitment** to data governance ↔ success in **digital maturity, data accuracy**

iii) Oversight & Ownership

- CDO **not mandatory**, but:
  ✔ Need **formal oversight**, centralized data ownership
  ✔ Define "**data governor**" roles
  ✔ Goal: **Single Source of Truth (SSOT)** for all stakeholders

iv) Future Compliance

- Governance must adapt for:
  ✔ **Renewable integration**
  ✔ **Climate change impacts**
  ✔ **Disaster mitigation**

v) Sector-wide Info Flow

- **Free flow** of info essential to:
  - ✔ Meet customer expectations
  - ✔ Handle **renewables from customers**
  - ✔ EV adoption

---

5) Key Questions for Strategy Formulation

a) How to measure **data quality**?
b) Are **key stakeholders** (leaders, IT) involved?
c) Are **ESG & sustainability** trends identified?
d) Is there a **right strategy** to give right data at right time?
e) Does it support **self-service analytics**?
f) Are **all stakeholder needs** identified + addressed?
g) Can **data be monetized** as key product?
h) Who is responsible for **data quality reviews**?

🛏 4. Data Governance in Hospitality Sector

🔑 Focus: Guest Experience, Revenue, Data Security, Compliance, Insider Threats

---

a) Analytics Adoption

- Goal: ⬆ Guest experience, revenue, ops efficiency
- Strategy: Targeted marketing + strong guest DBs
- Learnings from: Retail, CPG, Banking

---

b) Data Security = Major Concern

- Data collected: **PII + PCI** → name, address, card data
- Target for cybercrime: identity theft, card fraud
- Multi-database, multi-device → increases risk

---

c) Key Security Challenges

i) Complex Ownership Structures

- Franchisor, owner, mgmt co. → different systems
- Info flows across systems = more vulnerable
- Ex: **Wyndham breach (2008–10)** → 619K customer records stolen
  - Cause: Weak passwords + system spread

---

ii) Card Payment Dependence

- Hotels need card info for booking + final billing
- POS malware → scrapes card info
- 20/21 major hotel breaches (2010–17) = **POS-based**
- Malware spreads across hotel branches, often unnoticed

---

iii) High Staff Turnover

- Seasonal workers = ⬇ training consistency
- UK turnover in hospitality = 90%
- Poor training = easy victim to social engineering
- One untrained staff = major breach point

---

iv) Compliance Requirements

- **GDPR (EU, 2018)** → stricter control of PII
  - Global impact → pushes stronger compliance
- **PCI DSS** → credit card data protection
  - Non-compliance fine = $500K/incident
  - Risk = loss + reputational + survival threat

---

v) Insider Threats

- Employees may sell guest preference/behavioral data
- Data touchpoints: websites, forms, bookings, reviews
- Risk: data sold to competitors → unfair advantage

🔐 Best Practices in Hospitality Sector

a) Encrypt all **card info**
b) **Train staff** regularly in cybersecurity
c) Strictly follow laws: **GDPR, PCI DSS**
d) Tech protection: firewalls, anti-malware, traffic filters
e) Conduct **penetration tests** (simulate hacker behavior)
f) Use **least privilege** principle for data access

💻 5. Data Governance in IT Sector

🔑 Focus: Security, Privacy, Data Access, Strategic Asset

---

a) General Trends

- Data = "New Oil" → ⬆ need for strict **data governance**
- Focus: Restrict access + define who gets what data
- Drivers: ⬆ breaches, hacks, ransomware
- Shift: From **IT issue → Strategic boardroom concern**

---

b) IBM

- Robust data governance model

- Offers enterprise-level **data governance solutions**
- Core disciplines:
  • Data Quality Mgmt
  • Info Security & Privacy
  • Info Life-Cycle Mgmt
- Supporting disciplines:
  • Data Architecture
  • Audit & Logging
  • Classification & Metadata
- Enablers: Org. structure, awareness, policy, stewardship
- Goal: ⬆ Data quality, integrity via inter-org collaboration

---

c) SAS

- Data Governance = **Cultural shift** (Biz + IT align)
- Focus: Define data elements + rules org-wide
- History: Emerged early 2000s → evolved with **IoT, Hadoop**
- Models matured into **BPM tools** with automated rules
- Challenges:
  a) Data seen only as IT issue
  b) Fragmented/siloed culture
  c) Poor resource allocation

---

d) Oracle

- Focus: **Data as enterprise asset**
- Tools: Visual dashboards, metadata discovery, data security
- Policy includes:
  a) Enterprise data strategies
  b) Assigning stakeholder accountability
  c) Reporting status of data initiatives
- Functions: Identify, secure, manage sensitive data

---

e) Microsoft

- Shift: Data Mgmt → **Strategic Advantage**
- Data = Core asset → high risk + value
- Governance framework:
  • Aligns **data strategy with biz goals**
  • Focus on **cloud security protocols**
- Key message: Data = boardroom-level concern
- Leader in driving digital innovation + cloud-based governance

---

📡 6. Telecom Sector – Data Governance

🔑 Focus: Data Usage, Forecasting, Business Strategy, Compliance

---

a) Need for Governance

- Rapid ⬆ in data due to 5G, remote work, digital content
- **Big Data = key** to cost reduction, ⬆ sales, faster reporting
- Missed biz opps if data is unutilized or unmanaged
- **Centralized, clean data** → visibility, real-time forecasting, ⬆ decisions
- ⊖Stats:
  • 80% DG projects fail (Gartner)
  • 65% users unclear of DG impact (HBR)
  • 74% DG leaders can't measure ROI (Forbes)

---

b) 4 Pillars of Successful DG in Telecom

1. Link DG ↔ Business Goals
2. Prioritize high-value data sets
3. Engage strategic + operational teams
4. Leverage culture + people for daily DG reinforcement

---

c) TRAI Recommendations (India)

- Apex Body: **DDMC (Data Digitization & Monetization Council)**
- Role:
  • Oversee data sharing, monetization, storage
  • Set ethical use framework (Govt + Corporates)
- Legal Backing: New law/amendments needed
- Representation: DoT + MeitY
- Objective: Regulate India's data economy (centres, CDNs, exchanges)

---

d) Case Study: Airtel

- Privacy safeguards → built-in dev stage (app sec + compliance review)
- 3rd party non-compliance → disciplinary action/termination
- B2B cybersecurity services = **Airtel Secure**
- Airtel handles huge personal data → prioritizes **privacy + protection**
- Policy: **BIPP (Bharti Airtel Information Privacy Policy)**
  • Based on IT Rules 2011 + GDPR
  • Covers: Collection, processing, retention, dissemination, destruction
  • Applies to: All Airtel employees + 3rd parties (all locations, functions)

🛒 7. E-Commerce – Data Governance

🔑 Focus: Data Quality, Security, Insights, Compliance

a) Why DG is Vital

- E-comm relies on **quality data** → accurate, fast decision-making
- Collects sensitive consumer data: payments, behavior
- Sources: Web analytics, email tools, transactions, surveys, etc.
- DG ensures stakeholder access to reliable, consolidated, governed data
- Benefits:
  - Gap analysis (market trends)
  - Better customer engagement
  - Pricing/inventory/labour optimization
  - Innovation + new market exploration

b) Key Significance

1. **Visibility & Consistency**
   - Data flows across platforms: purchases, shipping, inventory
   - Poor governance = silos, obsolete data
   - Unified DG = seamless updates, cross-platform consistency, improved ops
2. **Limit Data Exposure**
   - Circulation needed but risky
   - Breaches hurt brand-customer trust
   - DG tools: 2FA, encryption, tokenization for secure access
3. **Fix Data Inconsistency**
   - Changes in one repo need syncing across all
   - Impacts: sales, productivity, strategy
   - DG uses pipelines: curation, validation → better analytics + visualization

c) Merits of DG in E-Comm

1. ⬆ Overall performance: ⬇ time, ⬆ efficiency, quick insights
2. 📊 Data quality tracking: usage metrics → team-level data understanding
3. 🎯 Better business insights: ID weak areas, gain edge, find new rev streams
4. ✅ Improved decision-making: Fast + accurate + compliant
5. 🤝 Ownership & Accountability: Clear roles, contact points, responsible usage

d) Product Data Governance Strategy

- Crucial for brand image, revenue → treat product data as strategic resource

Key Elements:

1. **Ease of Use**: For all stakeholders (incl. non-dev users), avoid over-restriction
2. **Transparent Lifecycle**: From item attributes → listings → 3rd-party sharing
3. **Pre-validation + Adaptation**: Clean outdated info, boost buyer confidence
4. **Controlled Delivery & Monitoring**:
   • Format-check per endpoint
   • Recheck listings, images, completeness post-distribution

⚖ DATA GOVERNANCE – Regulatory Dimensions

📜 1. IT Rules, 2011
[SPDI Rules under Sec 43A, IT Act, 2000]

🏢 Applicability: Body corporate/authorized personnel handling personal info

🔹 **Rule 4 – Privacy Policy**
• Must publish privacy policy (website)
• Covers:
(i) Practices & policies (clear, accessible)
(ii) Types of personal/SPDI collected (Rule 3)
(iii) Purpose of collection/use
(iv) Disclosure of info (Rule 6)
(v) Security practices/procedures

🔹 **Rule 2(i) – Personal Info**
• Any data capable of identifying a person (direct/indirect)

🔹 **Rule 3 – SPDI Includes**
• Password
• Financial info (bank/card details)
• Physical/mental health
• Sexual orientation
• Medical records/history
• Biometric data
• Any detail related to above, given for service
• Info collected/processed under contract

📌 *Excludes*: Info in public domain / RTI / under any law

⚖ 2. Digital Personal Data Protection Act, 2023

📅 Assent: 11 Aug 2023

🔹 Applicability

(a) **In India** – Digital data (collected directly or digitized later)
(b) **Outside India** – If linked to offering goods/services to Data Principals in India

✗ Not Applicable to:
(i) Personal/domestic use
(ii) Data made public by:
• Data Principal
• Legal obligation

🔍 Key Focus:
• Obligations of Data Fiduciary
• Processing children's data
• Significant Data Fiduciary: extra duties
• Rights: Info access, grievance redressal
• Duties of Data Principals
• Org areas affected: Legal, IT, HR, Sales, Finance, InfoSec

🔗 For full act: MeitY Website

---

3. AI Complementing Data Governance
AI + DG = Efficient, Secure, Compliant Data Handling

📈 Benefits of AI-Powered DG
1. **Improve Data Quality**
   • Auto error detection/correction
   • Standardization → structured, comparable data
   • ML finds hidden patterns/errors/missing data
2. **Automate Compliance**
   • Real-time monitoring of flows
   • Detects violations/anomalies, sends alerts
   • Auto classification, labeling of sensitive data
   • Generates compliance reports
3. **Strengthen Data Security**
   • Real-time access pattern analysis
   • Suspicious behavior alerts
   • ML-based malware detection
   • Patch mgmt, policy compliance automation
4. **Democratize Data**
   • Easier data access → boosts data culture
   • AI search engines = fast, relevant results
   • Auto dashboards → user-friendly insights

EU General Data Protection Regulation (GDPR) – EU
📅 **Effective**: 25 May 2018
📜 **Replaced**: 1995 Data Protection Directive

🔹 Objective
• Safeguard **personal data**
• Uphold **privacy rights** of persons in **EU territory**
• Unify EU under 1 data protection law

🔹 Applicability
• Any org (even outside EU) processing EU persons' data
• Covers all data activities – collection, storage, transfer, analysis
• Applies to **non-profit** & **for-profit** orgs

🔹 Personal Data = Info identifying a person
e.g. Name, Email, IP, Eye color, Political views, etc.

🔹 Compliance Measures
• **GDPR Assessment** – data inventory & security
• Appoint **Data Protection Officer (DPO)** (if required)
• Update **Privacy Notices**
• Obtain **Valid Consent**
• Ensure **Data Portability**
• Implement **Tech & Operational Safeguards**

🔹 Key Provisions
7 Principles: Lawfulness, Fairness, Transparency, Purpose Limitation, Data Minimization, Accuracy, Storage Limitation, Integrity & Confidentiality, Accountability

8 Rights of Data Subjects:
1. Right to be Informed
2. Right to Access
3. Right to Rectification
4. Right to Erasure (Right to be Forgotten)
5. Right to Restrict Processing
6. Right to Data Portability
7. Right to Object
8. Rights in Automated Decision Making

🔹 Enforcement & Penalties
• By national **Data Protection Authorities (DPAs)**
• Penalties: €20 million or 4% of global revenue (higher)
• Also: processing bans, public reprimands

Regulatory Trends in AI

🌍 Global Approach → 6 Key Trends
1. **Core Principles (OECD, G20)**
   • Human rights, Transparency, Sustainability, Risk mgmt
2. **Risk-Based Regulation**
   • Proportional obligations:
   – Low-risk → less regulation
   – High-risk → strict control
3. **Sector-Agnostic & Sector-Specific**
   • Hybrid models to address unique AI use-cases
4. **Policy Alignment**
   • AI rules aligned with Digital Policy (Cybersecurity, Data Privacy, IP, etc.)
   • EU → most comprehensive AI strategy
5. **Private Sector Collab – Regulatory Sandboxes**

• Safe space to test innovations + co-create ethical rules

6. **International Collaboration**
   • Cross-border coordination → handle risks from General Purpose AI & GenAI

Other Key Points

• Regulators need AI domain expertise
• Clear policy scope: regulate **tech** itself or **use-case**?
• Define responsibility for AI use by **3rd-party vendors**
• Avoid **regulatory arbitrage** – promote global rule interoperability

---

OECD Principles on AI

◆ A) Value-Based Principles (Para 1.1 – 1.5)

1.1 Inclusive Growth, Sustainability & Well-being
• AI → augment human skills, creativity
• Reduce inequality (eco., soc., gender)
• Protect environment → inclusive growth + SDGs

1.2 Human-Centred Values & Fairness
• Respect law, rights, democracy, diversity
• Ensure dignity, freedom, privacy, equality
• Safeguards for fairness, social justice, labour rights
• Human control mechanisms per context

1.3 Transparency & Explainability
• Disclose AI usage → users must know when interacting
• Explainable outcomes → reasons, logic, factors
• Right to challenge decisions
• Contextual, understandable info

1.4 Robustness, Safety & Security
• AI safe in normal, misuse, adverse use
• Ensure **traceability** (data, decisions)
• Apply **continuous risk mgmt** (bias, privacy, security, etc.)

1.5 Accountability
• Devs/Deployers accountable for functioning
• Align with OECD principles based on role & context

◆ B) Policy Recommendations for Governments (Para 2.1 – 2.5)

2.1 Invest in AI R&D
• Public-private funding (tech + legal/ethical issues)
• Promote **open, representative datasets**
• Ensure **privacy** & avoid bias

2.2 Foster Digital AI Ecosystem
• Build infrastructure, access to data/tech
• Enable **safe data sharing** → e.g., data trusts

2.3 Enable AI Policy Environment
• Promote **agile R&D to deployment** transition
• Use **experimentation (sandboxes)**
• Update laws & policy for innovation

2.4 Build Human Capacity & Labour Transition
• Skill development across life
• Fair transition: training, upskilling, reskilling
• Encourage safe AI use in work, promote job quality

2.5 International Cooperation
• Global collab → info sharing, standards
• Support **interoperability**, multi-stakeholder efforts
• Use **common metrics** for AI R&D & policy progress

Data Protection Seal (DPS) – Data Security Council of India (DSCI)

◆ Overview (Para 1–2)
• DPS = Cert. mark like ISI → for **data privacy compliance**
• Assures app/site/product follows **basic privacy stds**
• Aids compliance with **DPDP Act + future regulations**

◆ Features (Para 3–5)
• Pilot in **Delhi & Bengaluru** with partner orgs
• Verifies: Responsible data use, privacy expectations met
• Increases **user trust** + transparency

◆ Tackling Deepfakes (Para 6–7)
• Deepfake = Major digital threat
• Solution: **Certified DPOs** (Data Protection Officers)
• DPOs → trained to identify deepfake issues, preserve data authenticity

◆ Other Cyber Threats (Para 8–9)
• Ransomware, MFA attacks, AI misuse
• DSCI ↔ works with govt., regulators, think tanks → builds capacity, policies

◆ Challenges (Para 10)
• Balancing privacy + authenticity analysis
• Challenge = Detect fake content **without full platform disclosure**

◆ Future Steps (Para 11–13)
• Expand DPS program + DPO training
• Enhance **data protection ecosystem**
• Promotes **responsible data handling** across platforms

Sun Pharma Cybersecurity Breach – Case Study (Mar 2023)

◆ Incident Summary (Para 1–5)
• Reported: **Mar 2, 2023** → InfoSec incident
• **Ransomware group** → breached file systems, stole **company + personal data**

• Immediate actions: Isolation, containment, recovery protocol
• Result: **Operational impact**, revenue drop, unknown long-term risks

---

◆ **Sector Vulnerability – Pharma (Para 6–7)**
i) Valuable IP & R&D
• Clinical data, patents = high-value → target for black market / competitors
ii) Sensitive Data Access
• Patient info, trial results, filings
• Used for **fraud, blackmail, identity theft**
iii) Complex Partner Network
• Multiple vendors = **more access points** → insider threats, breaches
iv) Global Operations
• Single attack = **multi-country impact**
v) Weak Cyber Infra
• Low budget, reactive mindset → poor security measures
vi) Exploitation Types
• Ransomware, phishing, **insider trading**, regulatory data misuse

---

Here's a **concise, exam-focused summary** of the **Way Forward in Data Governance**, following your preferred format (short points, direct keywords, minimal grammar, para-wise structuring):

📊 **WAY FORWARD – Data Governance**

◆ **Para 1: Maturity Levels (3 Stages)**
i) Data Integration → App integration, data loading
ii) Data Integrity → Prep, stewardship, quality
iii) Data Intelligence → Cataloging, lineage, metadata mgmt.
➡ Progress ↑ = ↑ ML use (data profiling, matching) → ↑ value + trust

---

◆ **Para 2: Role of Governance Teams**
• Data governance = **enabler**, not barrier
• Supports: **Access**, **trust**, **better decisions**
• Recognize DG team = **ally in biz ops**

◆ **Para 3: Balancing Governance & Innovation**
• Governance ≠ blocking innovation
• Modern tech = **works in synergy** (e.g., ML + blockchain)
• Example: Chatbots for energy mgmt. using semantic processing
• Developers need **flexibility** in early stages (design/dev)

◆ **Para 4: Future Outlook**
• Tech evolves → **DG must adapt**
• Devs/architects = specialize by **industry/domain knowledge**
• New DG models = must allow **innovation + protection**

📚 GET COMPLETE NOTES – ₹300 PER SUBJECT

✔ Covers all subjects except for cr portion in crvi
✔ Made from 6 coaching notes + module, refined with AI
✔ Crisp, exam-focused & easy to revise

For more free notes join telegram channel - **https://t.me/csprofessionalnote**

👉 To purchase:

1. Pay ₹300 via UPI (**9024770603-2@ibl**)
2. Fill this form with payment proof:
   https://forms.gle/u4k3XxEy8nCJiXVs9

📩 Notes will be delivered to your WhatsApp within 12 hours.