aapti institute

Gates Foundation

Data Sharing and DPI

This is a work of independent research, commissioned by **Gates Foundation** in 2025.

It was built by **Aapti Institute**, a public research institute that works on the intersection of technology and society. Aapti examines the ways in which people interact and negotiate with technology both offline and online.

The team consisted of Ameya Thachappilly, Astha Kapoor, Avani Airan, Rakshitha Ramesh, Vignesh Shanmugam, Vinay Narayan.

ACKNOWLEDGEMENTS

In addition to contributions from the wider Aapti team, this report also draws upon the expertise of numerous academic and industry experts, as well as practitioners. We are grateful for their input during interviews and feedback sessions.

Report design: Meher Rajpal | Cover illustration: Ananya Broker Parekh

Glossary & Abbreviations

ABDM	Ayushman Bharat Digital Mission
AI	Artificial Intelligence
API	Application Programming Interface
APEX	API Exchange
ASEAN	Association Southeast Asian Nations
AU	African Union
CAR	Cadastro Ambiental Rural
CBDF	Cross-border data flows
COWIN	Covid Vaccine Intelligence Network
DECODE	DEcentralised Citizens Owned Data Ecosystem
DFFT	Data Free Flow with Trust
DGA	Digital Government Development Agency
DIAL	Digital Impact Alliance
DPI	Digital Public Infrastructure
e-GIF	e-Government Interoperability Framework
eKYC	e-Know Your Customer
ESG	Enterprise Service Bus
EU	European Union
FAIR	Findable Accessible Interoperable Reliable
FHIR HL7	Fast Healthcare Interoperability Resources Health Level 7
GoT-HoMIS	Government of Tanzania - Health Operations Management Information System
GSB	Government Service Bus
HIE	Health Information Exchange
HIPAA	Health Insurance Portability and Accountability Act
IEC	International Electrotechnical Commission
ISO	International Organisation for Standardisation
IUDX	India Urban Data Exchange
LMIC	Low and Middle-Income Countries
MIS	Management Information System

MoU	Memorandum of Understanding
MSME	Micro, Small, and Medium Enterprises
NEHR	National Electronic Health Record
OECD	Organisation for Economic Cooperation and Development
OGD	Open Government Data
ONDC	Open Network for Digital Commerce
P2P	Peer-to-Peer
PET	Privacy Enhancing Technologies
PKI	Public Key Infrastructure
PPP	Public Private Partnership
REDATA	Regime Especial de Tributação para Serviços de Data Centers
SDE	Shanghai Data Exchange
SICAR	Sistema de Cadastro Ambiental Rural
SOAP	Simple Object Access Protocol
TEFCA	Trusted Exchange Framework and Common Agreement
TGDEx	Telangana Data Exchange
TLS	Transport Layer Security
VC	Verifiable Credential
WEF	World Economic Forum
XML	eXtensible Markup Language

Table of Contents

01

Introduction

TABLE OF CONTENTS

With the boom in data-driven insights, data sharing as the third DPI is the call of the moment globally

Data runs the world through decision-making, innovation, policy and production processes

Some countries have achieved cross-sectoral implementation with high adoption rates while others perform specific, siloed functions

The need of the hour is systems that complement this abundance of data and allow wide access and sharing

The push for data sharing as DPI has various instantiations and is layered, pointing to a need for global alignment

Data sharing is an opportunity for global collaboration to achieve greater public good by traversing institutional and technical boundaries through interoperability

Though definitional conundrums persist, the need of the hour is to establish alignment on components and operations of the ecosystem. We use the term "data sharing" since:

- It is the most expansive term to capture the interoperable frameworks being considered for this study , as opposed to data exchange or ecosystems.
- With the addition of trusted data sharing models, there is opportunity to assess ambiguities through a wide web of frameworks that capture flow and exchange of data across various architecture styles and data types

Data sharing recognises the welfare potential of data and is a catalyst for cross-sectoral innovation

Source: Aapti analysis

Data sharing is a hotbed for investment geared towards public good

$100 billion+ investment expected for data centre expansion by 2027

Public-facing digital systems are increasingly funded and maintained collaboratively by PPPs like IndiaStack, X-Road, and Brazil's gov.br

Over half of global venture funding in 2025 is directed towards AI. Larger investments into data sharing means understanding capital and information flows are important.

85% of OECD states deploy government-backed data sharing systems leading digital infrastructure rollouts in Australia, Belgium, Korea, Latvia, and others

Global spending on AI data centres is set to exceed **$1.4 trillion by 2027**

Private sector investment in global cloud and data infrastructure is expected to grow to **$6.7 trillion by 2030** in cloud architecture, compute power, and storage

Standardisation within the data sharing ecosystem optimise investment and collaboration opportunities

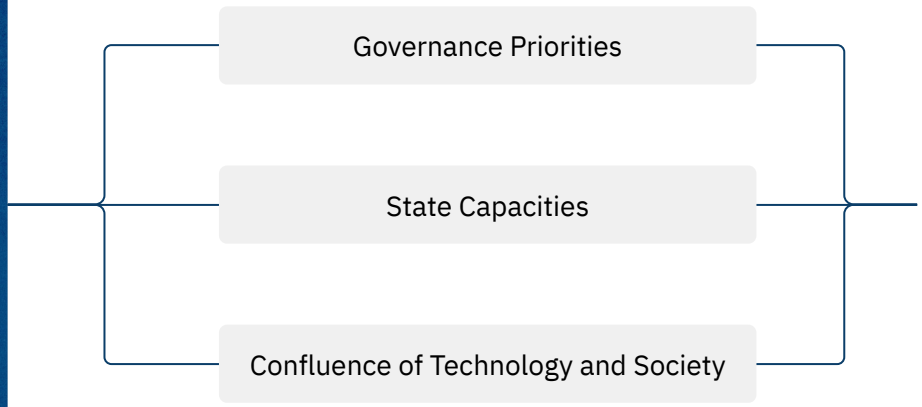
Source: Forbes, The Economic Times, World Bank 2025; Aapti analysis

Mapping country priorities and implementation logic to use cases identifies functional elements that influence the ecosystem

Our work takes the socio-technical approach to ensure value generated by data flows back to society

The Socio-Technical Approach

"Technologies do not exist in a vacuum: they influence and are influenced by the social contexts in which they are deployed." (Ada Lovelace Institute)



- Frames the two-way relationship between technologies and the people who are affected by them
- Ensures LMIC representation
- Helps governments and model developers can place themselves on the map through use-case study
- Aids countries in understanding what approaches can work to achieve their specific motivations
- Enables sound, well-informed decision-making by focusing on the implications and trade-offs of data sharing

The report identifies foundational considerations in the ecosystem to increase government buy-in and provide functional guidance on implementation strategies

Source: Aapti analysis

02

Objectives and Expected Outcomes

TABLE OF CONTENTS

For foundational global alignment and mitigation of ecosystem complexities, Aapti focused on two objectives...

Global alignment on data infrastructures

This work recognises the need for a broader global alignment on definitions and principles and the building of a taxonomy on the data infrastructures, particularly data sharing infrastructures

This will benefit in enhanced resource mobilisation, allocation, and coordination

Facilitating easier design and deployment of data sharing infrastructures

Many LMICs depend on DPI for implementing data strategies (identity, financial services, data sharing, etc.)

Aapti is building out a self-assessment tool and a comprehensive taxonomy

This will enable better understanding of data as a DPI and aid with the implementation of policy measures to build effective DPIs

Aapti aims to develop a holistic governance framework for data sharing systems that can enable ethical, high-impact data sharing

Source: Aapti analysis

...resulting in a self-assessment tool, functional taxonomy as a design guide, and use case repository

Taxonomy for functional design and decisions

Provides a structured way to navigate available models, match them to needs, and make informed adoption choices with layers like:

- **Core architecture styles:** centralised federated, gateway
- **Model functions & characteristics:** CBDF, citizen-first, private sector
- **Tech components:** with tiers on the essential components and add on features
- **Governance components:** with priority ordered as must/should/could have
- **Data sources & flows:** sector registries, legacy data etc.
- **Enabling structures:** policy environment, capacity building, financing, operational models

Use case repository to guide decision making

Aapti is building a repository of ~8 (around 8) use case ecosystems to complement the abstraction in the functional taxonomy

- The repository maps the stakeholders, infrastructures, processes and flows in the use cases to recommend regulatory or oversight measures.
- Aapti has mapped common features/ models with particular use cases instead of exhaustive prescriptions on the possibilities

Self-assessment tool for subjective diagnosis

Aapti is building a platform to assist countries to build DPIs through identifying their context, gaps, priorities, and constraints

- **Identify current state:** gaps in digitisation, legal, governance, and technical capacity
- **Map readiness:** existing assets and missing capacities
- **Diagnose risks & enablers:** socio-political, institutional, legal considerations
- **Clarify priority use cases:** problem statements matched to data sharing entry points
- **Articulate adoption priorities:** functional, architectural, and institutional preferences

These outcomes are intended to guide purpose discovery, influence and support decision making while building out data-sharing systems

Source: Aapti analysis

03

Methodology

TABLE OF CONTENTS

Key focuses of this work: components of data sharing, ecosystem processes, and nature and role of stakeholders

Theoretical Review

Aapti analysed:

- **50+ reports by international organisations** (primarily World Bank, DIAL, OECD and WEF)
- **20+ scholarly articles / opinions**
- **50+ practitioner / technical reports**

The theoretical review was anchored in:

- Understanding global **policy aspirations and ecosystem influences** around data sharing infrastructures
- Identifying **on-ground considerations** while building out data sharing systems as a DPI
- Uncovering the **existing gaps** in the ecosystem, to arrive at a comprehensive governance framework

Expert Interviews

- Aapti interviewed **10+ experts** for insights
- This included core-tech practitioners, academics, data policy experts, private-sector professionals, global development consultants, technology executives, government spokespersons, sectoral experts, development sector leaders, and pioneering voices from the Global South

Model Analysis

- Aapti analysed **50+ models across 9 sectors** (public administration, education, agriculture, energy, citizen and population, business and trade, healthcare, banking and finance, private sector integrations)

The review was anchored in:

- Ensuring intentionality in **diversity across the geographies, sectors, and use-cases** for the models
- Identifying common **core architectural styles** and their implications
- Early mapping on the long lists for the **repeating elements within technology and governance** layers and their primary functions

To address the gaps from theoretical and model reviews, the study relied on expert interviews to enrich understanding and provide insights that were not covered in literature

Source: Aapti analysis

04

Challenges and Mitigation Strategies

TABLE OF CONTENTS

Challenges to the study were at the ecosystem level and marked by lack of alignment on different fronts

Absence of a Taxonomy	A taxonomy serves as a tapestry of the various societal, technical, regulatory and architectural, aspects of developing data sharing models, to assist various stakeholders in the public and private sector. Discord around the functions of key technical components also adds to confusion in starting points and design strategies.
Global North Centricity	The Global North hosts most sophisticated data sharing systems. Therefore, existing literature that studies and documents debate and discussion on data sharing is largely Global North centric. They do not account for the unique considerations and requirements of Global South nations looking to implement DPI through data sharing. This leads to both, alienation from the technology and policy shaped for and by the Global North.
Model Overlaps	The nascent state of most data sharing models prevent sophisticated and complete analyses of the effects of various governance and technical strategies. Overlaps in the functions of models and the presence of multiple systems performing similar tasks within a country diluted categorisation.
Discorded Motivations on Building Data Sharing	Rooted in interoperability and scalability, data sharing systems serve various functions and evolve with time, making the process of categorising motivations tricky. Further, countries with the same needs could opt for diverging methods of implementation based on preferences in other components (architecture, regulation, etc.).

The lack of a unifying language for comparison in the ecosystem was reflected in our study and accounted for by engaging with diverse stakeholders

Source: Aapti analysis

Despite efforts to mitigate challenges, there are some limitations to the study

Expert Outreach	Access to experts that were part of the process of building data sharing systems across countries could have better informed the mapping of architectural, regulatory and technical design to government motivation.
Access to Building Documents	Understanding the blueprints of various technical aspects employed in data sharing models could provide insights into the future capacities of data sharing models along with a full view of the associated implications.
Understanding Motivations Around Building Data Sharing	Identifying explicitly stated push factors of data sharing projects as responses to sectoral or societal needs can help governments identify areas for implementation. Technical experts can also build more fluid and adaptable systems that are not intended to address singular, specific issues.

The challenges found in the data sharing ecosystem reflected in the research as well

Source: Aapti analysis

05

Literature Review

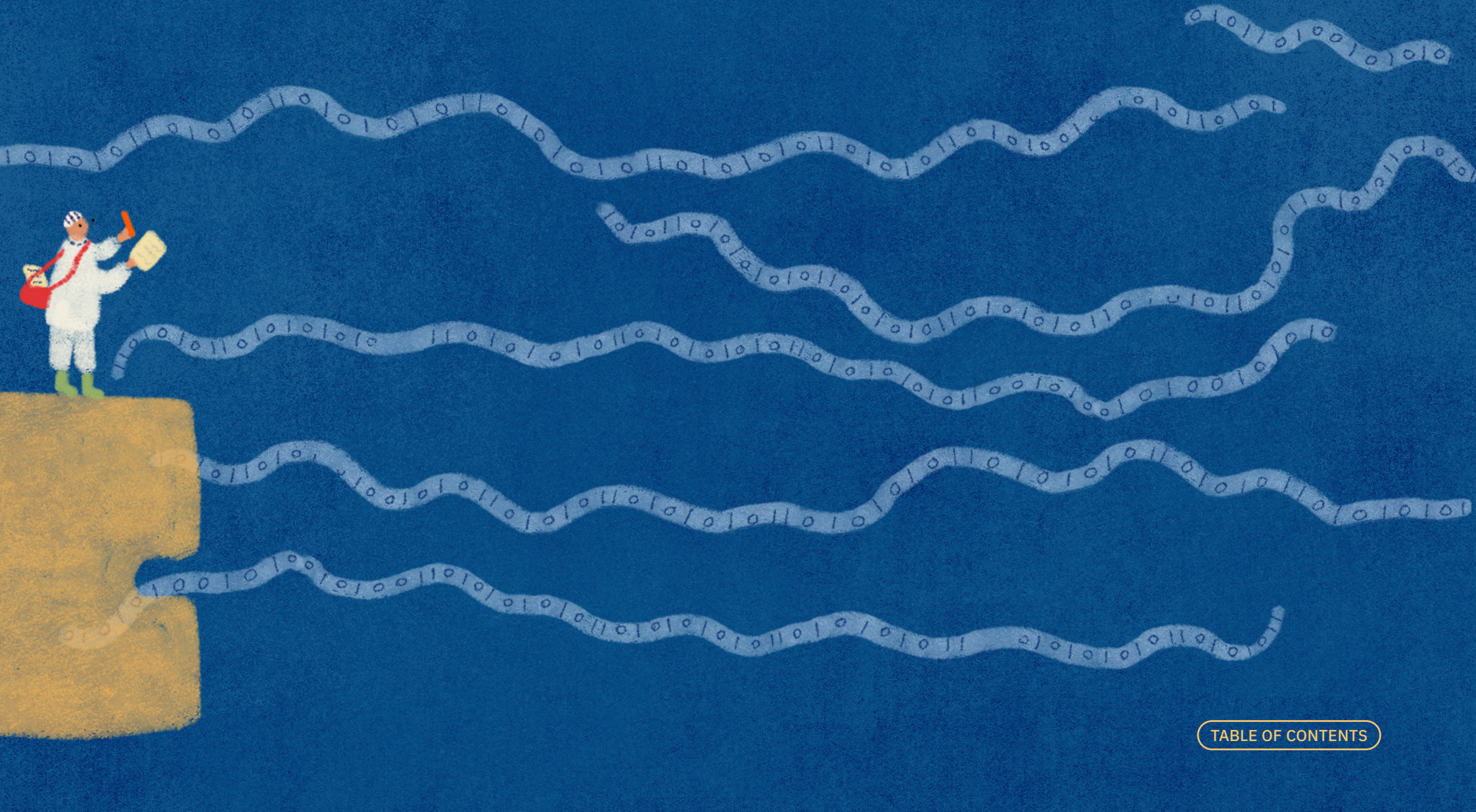
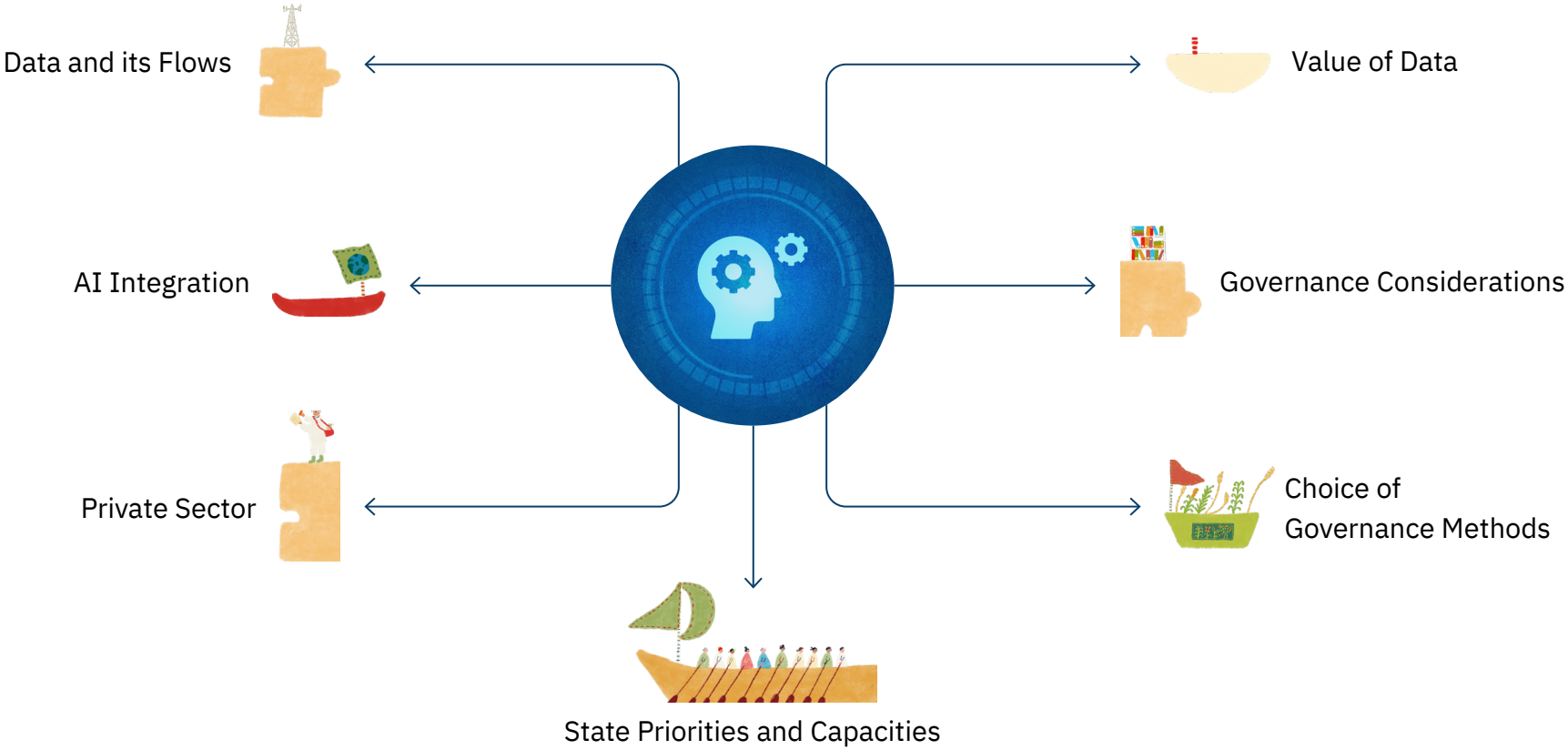


TABLE OF CONTENTS

Aapti's literature review brought out key considerations that guided its framing towards the taxonomy, repository and self-assessment tool



The review brought out different foundational considerations to help build out data sharing infrastructures

Source: Aapti analysis

Data sharing systems handle various kinds of data, making its governance complex

Within data sharing ecosystems, a fundamental challenge lies in defining what constitutes "data" itself. Unlike the identity and payments layers of DPI, where the units of exchange and its implications are relatively standardised, data in exchange models resists such categorisation.

Data in sharing models is

Unbounded

Data can be personal, non-personal, open, transactional, aggregated, etc.

The units within which they are made useful depends on the operational needs of the user

Reusable

Data is reusable and does not provide singular, end-to-end solutions

Various users may receive different transformed versions of the same data

Indeterminate

Data is used in a specific context to affect a specific decision

However, it is not fungible and cannot be replaced with other data within the ecosystem

Responsive

Data responds to the interests of the user

It can be collected and stored beforehand but is typically transformed into information that is useful to the user by an intermediary

Enabling wide stakeholder access whilst ensuring ethical use demands a purpose-specific governance approach given the variety of data flows

Source: World Bank, The New Hanse Institute; Aapti analysis

Data flows are determined by the nature of the actor and function of the data

- Data is complex as its significance and flows transforms depending on who uses it and why
- The channels through which data flows are not uniform since several datapoints collected at once can be used differently and separately
- For data to flow, it must be verified and authenticated, establishing the trust relationships that make exchange possible
- Without functioning data flows, DPI cannot develop the robust networks, interoperate, or generate insights and analyses that make it valuable
- Basis the use and function of the data, data lifecycles involve commitments to privacy, security, interoperability, and regulatory compliance at various levels
- Effective management at each lifecycle stage increases efficiency and reduces the costs of data sharing across users and may involve external data intermediaries

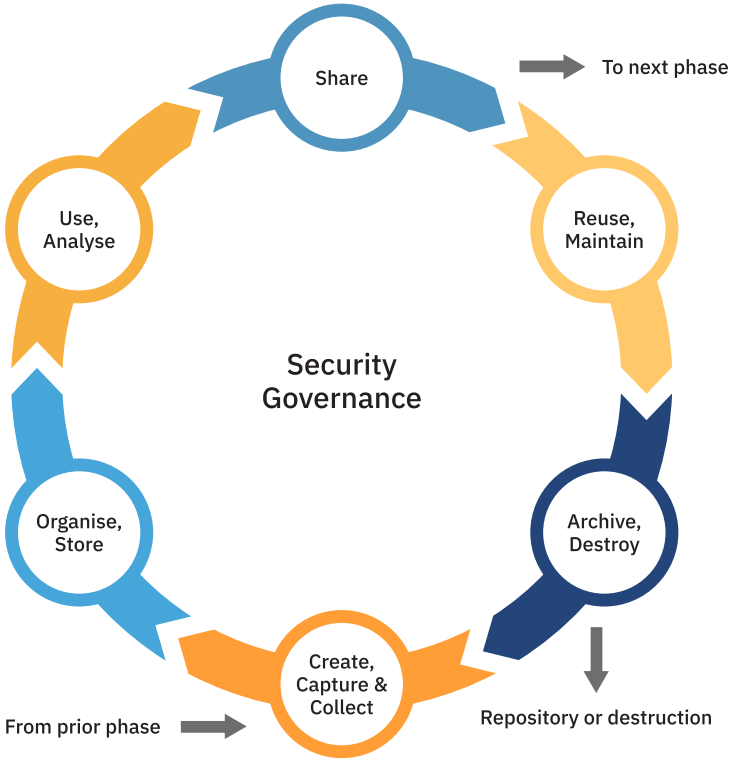


Image Source: World Bank

Interoperability transforms fragmented networks into a cohesive, evolving system where information flows, generates value, and serves multiple contexts and stakeholders

Source: World Bank, OECD; Aapti analysis

Various architectural approaches represent equilibrium points between different functional priorities

POOLING
Combining multiple data sources for reuse by third parties

Examples:
- Data warehouses, lakes, lakehouses
- Open data portals
- Data brokers
- Data marketplaces

DIRECT
One entity shares data directly with another

Examples:
- Systems integration
- Interoperability platforms or services
- Periodic syncing/updating

INTERMEDIARY
A third party facilitates or manages data sharing on a user's behalf

Examples:
- Trusted data intermediaries/fiduciaries

DECENTRALIZED
A user manages their verified data, sharing directly with third parties

Examples:
- Verifiable credentials
- Personal data vaults or wallets

DIFFUSION
Data is sent to multiple users or subscribers simultaneously

Examples:
- Message queues
- Broadcasting

Image Source: World Bank

- While some data is undoubtedly sensitive, other forms might be personal and sensitive only in some contexts
- The appropriate data sharing system depends on the purpose and type of data being shared
- All data sharing architecture is faced with striking a balance between accessibility, security, broad utility and privacy protection of data
- Systems that enable various categories of actors - individuals, businesses, and governments - to seamlessly and securely share or exchange data across (and within) sectors are key for many types of exchanges
- Data at risk of oversharing particularly by private sector actors requires privacy and security checks with strict access controls and, in the case of personal data, gives control and transparency to the subjects over how their data is used
- However, models designed to widen the reach of data through open data repositories should be designed with accessibility and usability in focus

Choosing an architectural approach is a balancing act that does not have a one-size-fits-all solution

Source: The New Hanse Institute, World Bank; Aapti analysis

The value of data changes depending on the stakeholder in question

- Data's economic value is a key push factor in policy dialogues around data unlock
- Being non-rival in nature, it acquires value once transformed into intelligence and actionable guidance for decision-making
- Tendencies to use data exclusively due to consumer privacy dimensions and competition interests differentiate it from other intangible assets
- As a result, conversations around unlocking the value of data are often locked in a binary of privacy and innovation

Value of data for governments

- Efficiency
- Regulatory oversight
- Evidence-based policy making
- Building foundational governance layers
- Improved service delivery

Value of data for international organisations / multi-laterals

- Digital trade
- Benchmarked global standards
- Open knowledge flows

Value of data for the private sector

- Open access to anonymised/consent-based data for innovation
- Reduced entry barriers
- Valuable insights
- Better risk assessments
- Open APIs and better interoperability

Value of data for individuals

- Access to public services
- Privacy, security, accountability, transparency
- Empowerment and agency over data
- Access to open data for public good (research, community initiatives)

The value of data needs to be anchored beyond the binary of 'privacy versus innovation' to help enable better data sharing for the public good

Source: World Bank, OECD; Aapti analysis

Building data exchanges as DPI requires a framework that equitably encompasses value for all

Out of all the frameworks Aapti examined, the World Bank Social Contract brings out how to anchor value in a larger stakeholder ecosystem

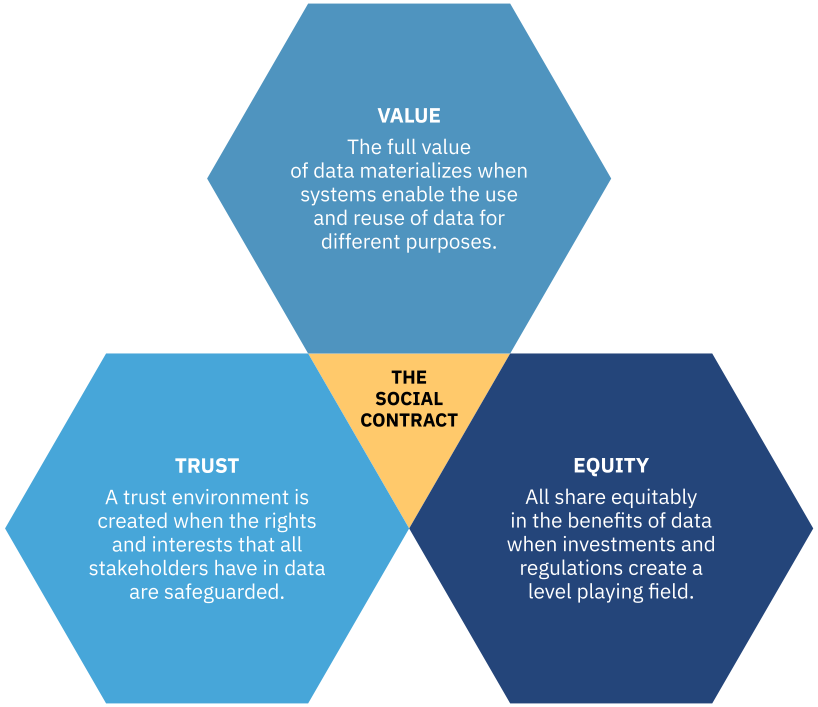


Image Source: WDR 2021

The World Bank Social Contract:
An agreement among all participants in the process of creating, reusing, and sharing data that fosters trust that they will not be harmed from exchanging data and that part of the value created by data will accrue equitably.

Open flows of data must be underpinned by ownership and equitable control that acknowledge and respect the different values data holds for diverse actors

Source: World Bank, OECD; Aapti analysis

While data sharing as DPI is gaining momentum, the current policy ecosystem is largely localised with staggered implementation

POLICY DEBATE

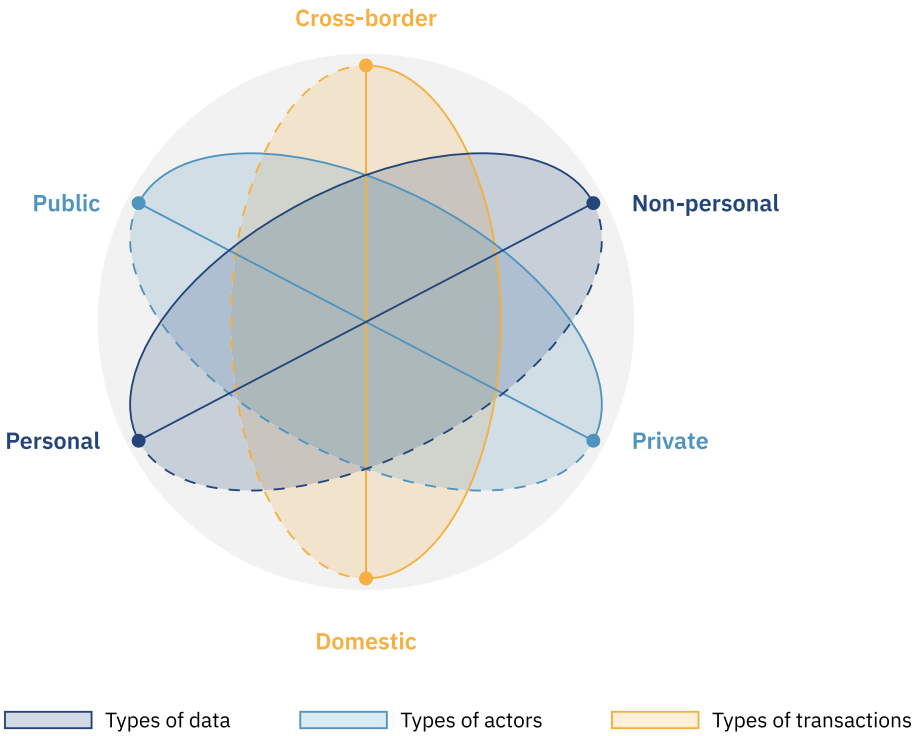


Image Source: WDR 2021

Lack of suitable laws: Existing data governance laws have failed to protect technology users and "subjects" from the harms of data extraction and sharing

Human-centricity: Need to place privacy and autonomy of data subjects at the center of the policy-making process

Socio-economic centrality: Data governance laws so far do not account for the socioeconomic and normative centrality of data relations

Context-specificity: DPI projects must be context-specific and undertaken only if there is a political will to support and sustain them

Exclusion through protection: Data encryption and protection mechanisms are necessary but have the potential to be exclusionary

Data governance is a three-dimensional space which attempts to balance access and protection

Source: World Bank, OECD; Aapti analysis

Data governance is crucial but complex, due to the maze of cross border data flow policies

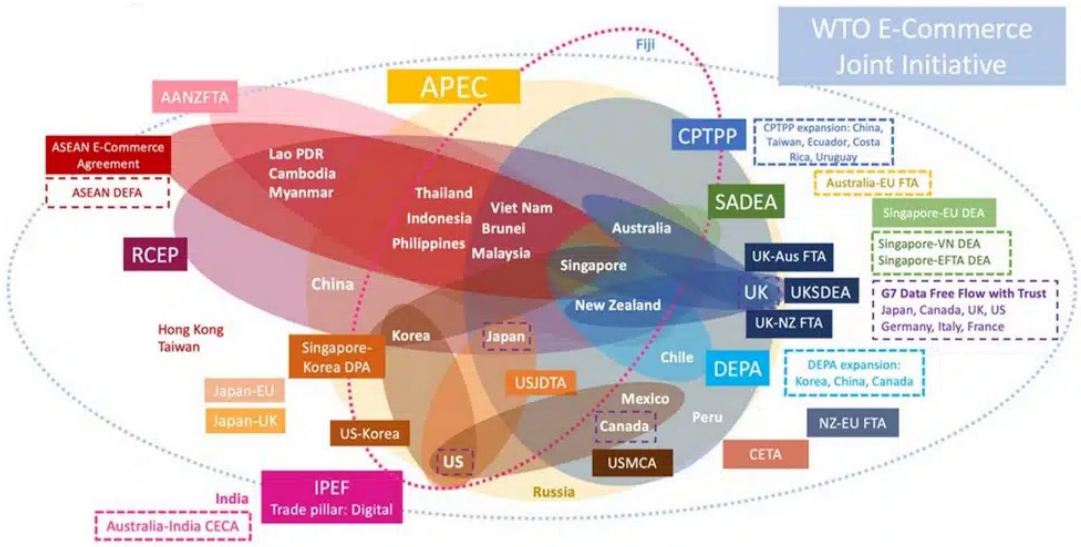


Image Source: 'Noodle bowl' of digital trade provisions / CITP

- The localisation versus free flow of data debate is complex and omnipresent
- There is under-participation despite technical readiness due to conflicting interests in the ecosystem
- Global free flow of data will remain complex due to geopolitical tensions and differing priorities
- Emergence of a maze of digital trade provisions that has been called a "digital noodle bowl"
- Sustainability and fostering consumer trust require mechanisms for cross-border enjoyment of data rights, enforcement, and redress options
- Trusted data sharing systems will have to balance out this reality of data sovereignty and global interoperability
- Data sharing frameworks must also balance against reinforcing inherent asymmetries in the international ecosystem

Global governance solutions need to be layered, adaptive and must address diverse ecosystem needs cohesively

Source: Teasdale, Centre for Inclusive Trade Policy; Aapti analysis

The complexities of data sharing governance bring out varying tussles within sectors – health, finance and mobility demonstrate this friction

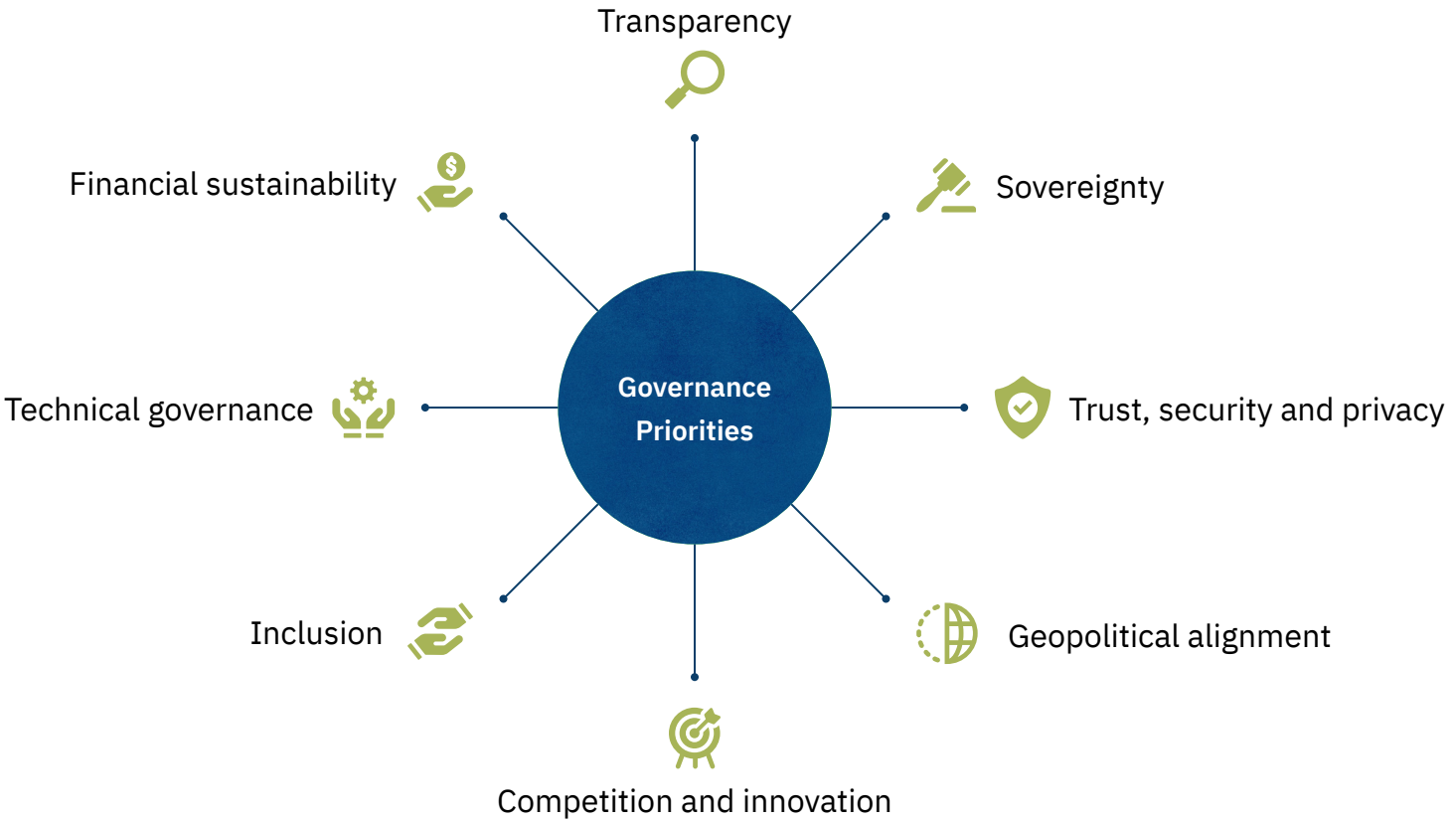
	TUSSLE	GOVERNANCE PRIORITY	EXAMPLES
Health	Protecting highly sensitive data versus sharing for public good	• Quality consent mechanisms • Revocation of consent • Transparency over data use • Individual and community rights • Quality control and accuracy of data (patient research)	Global Digital Health Partnership, EU Health Data Space, Act on Secondary Use of Health and Social Data (Finland)
Finance	Financial inclusion versus monetisation, protecting highly sensitive data versus sharing for public good	• Open data ecosystems • Bridging informational asymmetry, financial inclusion, community representation • Highly interoperable architecture • Strict regulatory oversight over intermediaries/aggreggators • Industrial support • Bank participation	Open Banking, Sahamati, Agristack, ONDC
Mobility	Market growth versus privacy considerations and surveillance	• Interoperable architecture for Smart city development • Ownership questions • Transparency over data use • Individual and community rights • Quality control and accuracy of data (patient research)	DECODE, IUDX

While there can be no one-size-fits-all approach, trust underpins different tussles, emphasising the need for a governance-first approach to data sharing

Source: World Bank, Pretzsch, Dreez & Rittershaus, Center for Financial Inclusion; Aapti analysis

Data sharing systems are driven by a plethora of governance priorities

Through this research, Aapti arrived at 8 priorities that drive governance choices. *For eg., enacting data protection laws falls under trust, security and privacy priorities.*



Choice of governance instruments and methods can speak to multiple priorities

Source: Aapti analysis

These examples illustrate how a governance method can speak to different priorities

GOVERNANCE METHOD	GOVERNANCE PRIORITIES							
	Trust, security and privacy	Geopolitical alignment	Inclusion	Competition & innovation	Financial sustainability	Transparency	Sovereignty	Technical governance
AU Digital Transformation Strategy		●			●			
ABDM data localisation rules	●						●	
Brazil Open Finance Model		●		●				
Data Protection Authority	●					●		

Choice of method and priority also reflects the capacity of the country (economy, political ecosystem, welfare, etc.)

Source: Aapti analysis

These examples illustrate how a governance method can speak to different priorities

DATA EXCHANGE	GOVERNANCE PRIORITIES							
	Trust, security and privacy	Geopolitical Alignment	Inclusion	Competition & innovation	Sustainability	Transparency	Sovereignty	Technical governance
Ghana.gov	1 2	A C	B	3	E	4 5		6 D G F

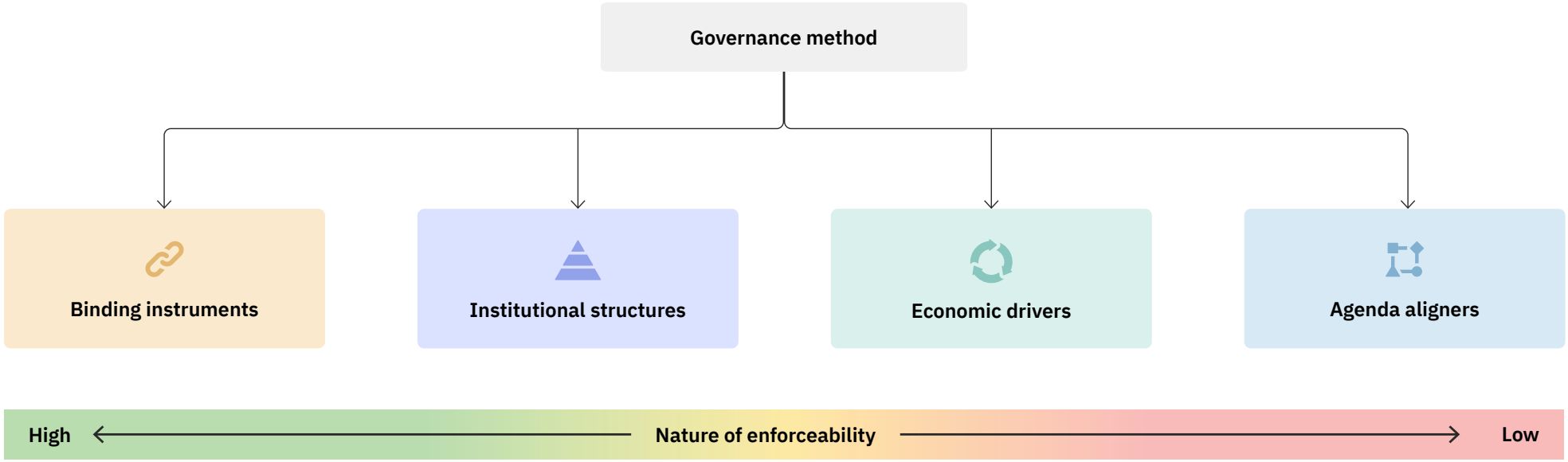
1. National Identification Authority Act
2. Data Protection Act
3. Payment and Services Act
4. Right to Information Act
5. Data Sharing Policy
6. eGIF Interoperable Framework Guidelines

A. Ghana Open Data Initiative
B. Open Government Partnership
C. National Open Data Action Plan
D. National Data Strategy 2024
E. National Financial Inclusion and Development Strategy
F. eGIF Implementation Guide
G. Ghana governance enterprise architecture framework

Inefficient revenue collection led to Ghana.gov, with agenda aligners and binding instruments being prioritised first and data protection frameworks emerging later

Source: DIAL; Aapti analysis

This research looks at governance methods through four lenses based on their nature of enforceability and value add

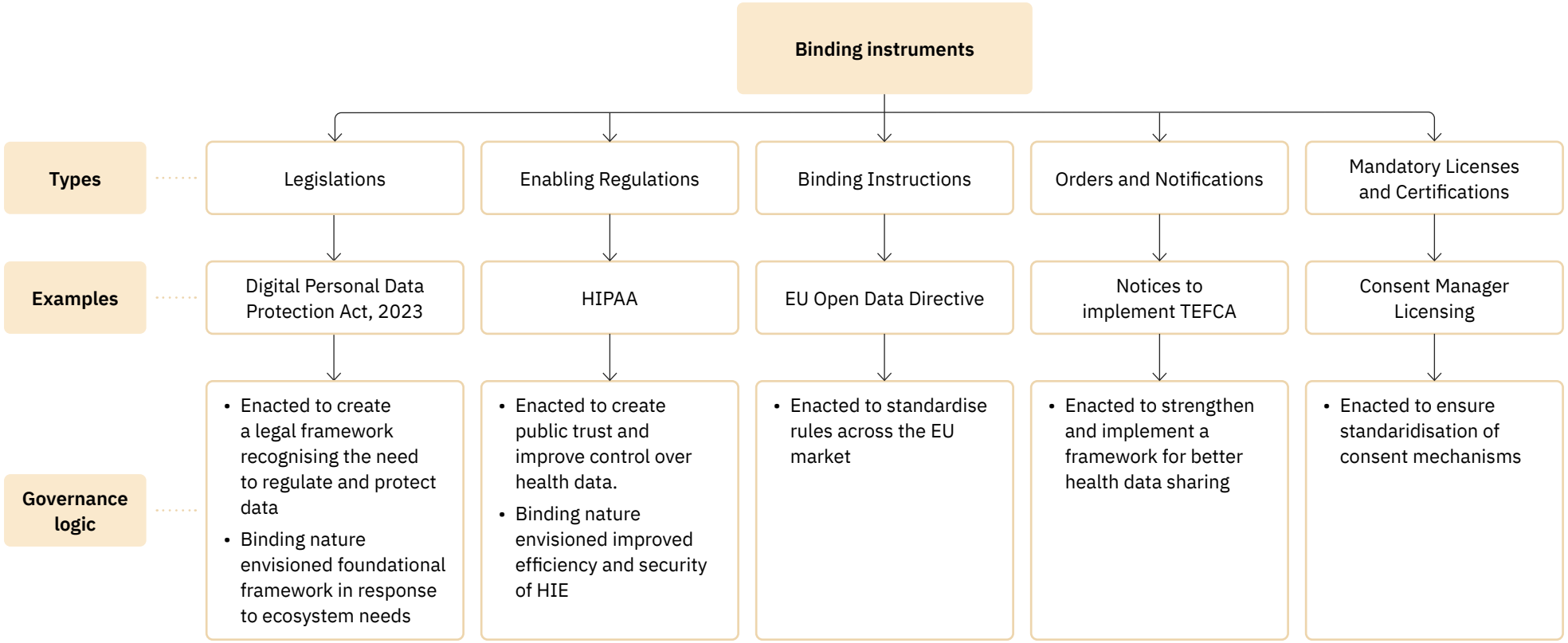


While these methods are not unique to DX systems, they hold true for DX as well

Beyond managing governance priorities, countries optimise choices of methods based on their need for sovereignty, to aid regulation, improve market friction or align with global standards

Source: Aapti analysis

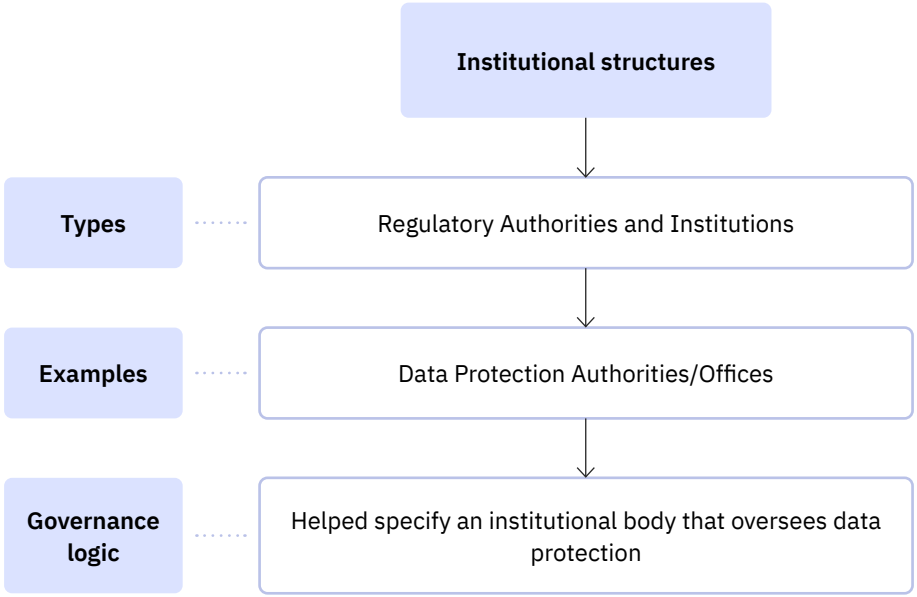
Binding instruments create a clear assertion of sovereignty, enabling a strong governance foundation



These instruments help create an ecosystem of trust through defined penalties, demarcation of rights and market enablers

Source: Aapti analysis

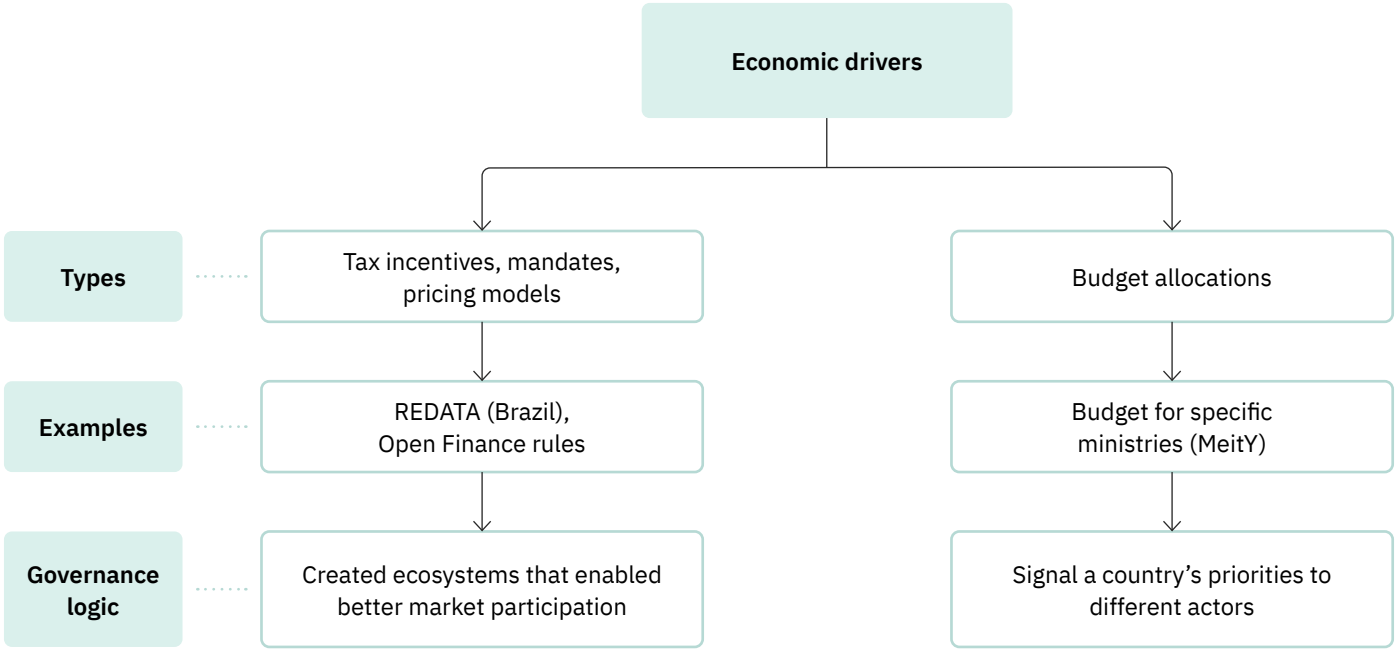
Institutional structures form building blocks for regulations



By complementing binding instruments, this method enables coordination by creating ecosystems that aid implementation

Source: Aapti analysis

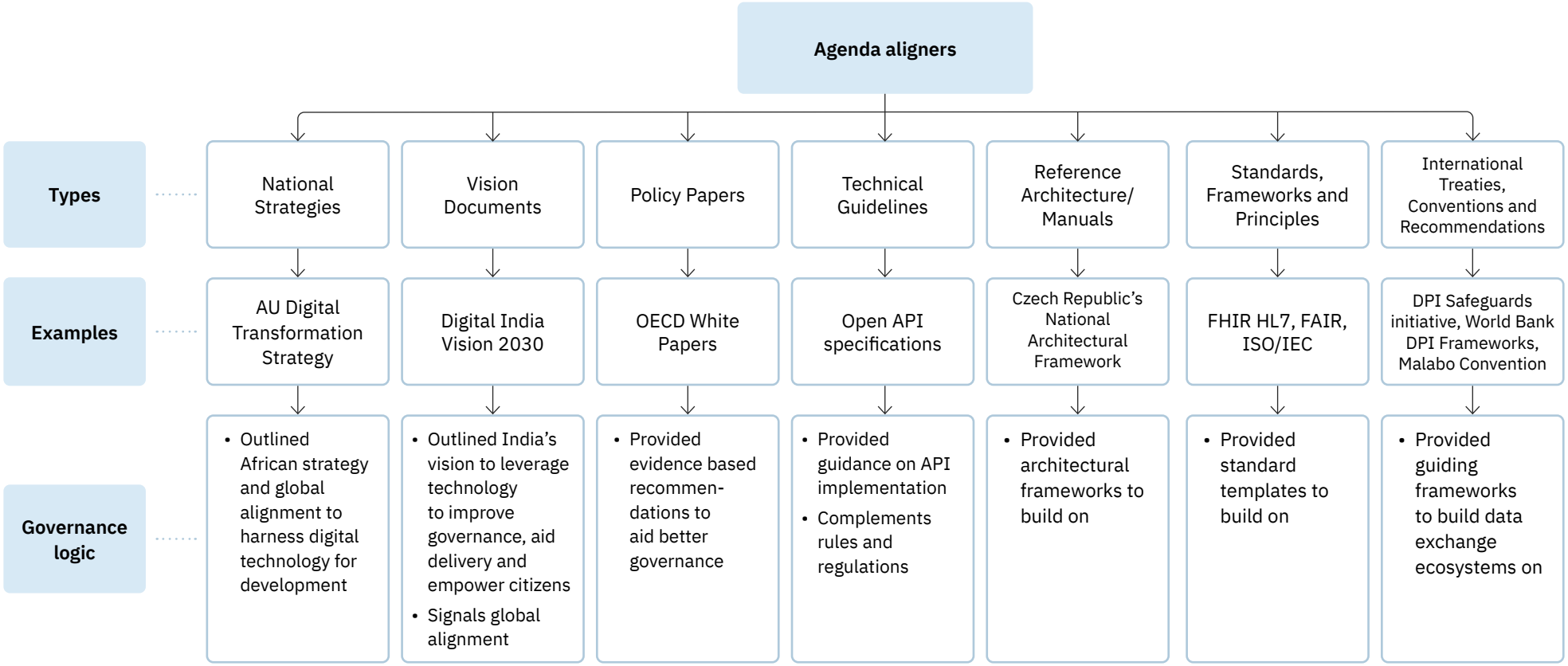
Economic drivers act as carrots for better rollout and less ecosystem resistance



The semi-binding nature of this method encourages competition and innovation and ensures financial sustainability

Source: Aapti analysis

Agenda aligners are usually used when governance has lower institutional capacity



Choosing to use agenda aligners creates an ecosystem of adaptability while also aligning on global vision, standards and principles

Source: Aapti analysis

Along with governance methods, coherence and transparency function as critical de-risking variables to sustainability and longevity of models

STATE CAPACITY

- Governance maturity and institutional stability are prerequisites for implementing interoperable data ecosystems
- Fragmented jurisdictions and weak coordination mechanisms slow adoption and create economic risk that discourages long-term investor engagement
- World Bank's GovTech Maturity Index, and the University of Oxford's Readiness Network measure whether countries have the capacity to participate in the global economy by harnessing technology
- UGhub represents how institutional collaboration enables functional data exchange: 47 public entities and 66 private entities have exchanged data securely more than 100 million times in the two years of its operation

POLITICS OF IMPLEMENTATION

- The role played by actors within the model fundamentally shape implementation outcomes – who controls data flows, how data circulate through exchange systems, and whose interests are prioritised
- Sharing models reflect distributions of power - between domestic and international, central and local government, and public and private sector competition
- Different formats of governance – authoritarian, democratic, centralised, and participatory – solve for different issues but exacerbate others. For instance, centralised frameworks promote quick rollout but are often not scalable
- India's National Digital Health Ecosystem how digitisation of information into data concentrates power and service delivery

Bureaucratic incentives and regulatory frameworks determine adoption trajectories and long-term viability in data exchange models

Source: World Bank 2021, World Bank 2025, Chandrasekhar 2024, OECD 2019; Aapti analysis

Stakeholder and financing networks determine governance standards, compliance incentives, wide adoption and financial viability

FINANCING NETWORKS

- Data sharing decisions include development finance and digital trade regimes
- Development finance Institutions link lending to governance reforms and standards
- Technical assistance packages shape which standards and platforms become defaults. Regional digital economy frameworks create compliance incentives that determine data sharing architecture
- Strong creditworthiness or strategic geopolitical positioning attract favourable financing terms, creating a stratified landscape where institutional capacity itself depends on access to international capital

STAKEHOLDER AND GOVERNMENT DYNAMICS

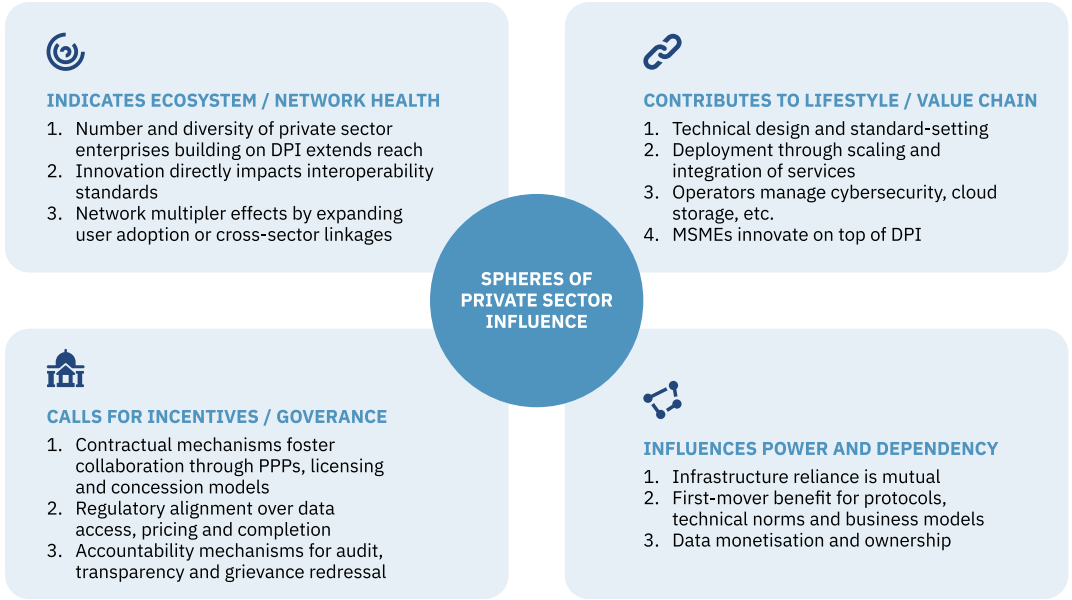
- Multi-actor coalitions influence both model design and confidence in investment
- Meaningful participation leads to greater public trust and adoption rates, reduced political risk for investors, improved data quality through stakeholder buy-in, and enhanced legitimacy that survives electoral transitions
- Technocratic or elite-driven processes of implementation face resistance, workarounds, and political contestation that undermine functionality and investment returns
- Smart Africa's initiatives and similar regional programs demonstrate how participatory governance frameworks align domestic legitimacy with external investor comfort, directly linking governance inclusion to financial viability

Involvement of external actors brings perspectives that can improve functionality, adoption and customisation in the process of data sharing implementation

Source: Ada Lovelace Institute, DIAL, IISD; Aapti analysis

The nature of private sector participation impacts data sharing DPI's functioning as infrastructure, marketplace or a platform

Private sector participation in data exchange DPI includes public-private partnerships and other collaboration over technical design, operational management, governance, and service delivery. Drawing on implementations across Estonia, Singapore, India, Thailand, Brazil, and Jordan, the analysis identifies how private sector engagement influences interoperability, inclusiveness, sustainability, and governance structures.



- The Sahamati case demonstrates an industry-led collective that enforces technical and operational standards
- Government regulators provide oversight and legal framework
- Transparency in governance is fundamental to success in data exchange
- Sahamati and the UK's Open Banking Implementation Entity shows how making critical performance data accessible to all participants builds trust and enables effective collaboration

Private sector participation influences its inclusiveness, sustainability, and governance and the speed of deployment of data sharing systems

Source: Aapti, UCL, DCO, ISEAL Alliance; Aapti analysis

Private sector involvement necessitates determining the value and licensing of data

Private sector participation brings technical capabilities that may not exist within government agencies, market knowledge that informs user-centred design, innovation capacity that extends infrastructure utility to new use cases, and scaling expertise that accelerates adoption.

> Singapore's APEX platform charges fees for certain advanced features and integration services, balancing sustainability with access.

Data and the processes applied to it become profitable and thus subject to commercialisation.

If data becomes a commodity, valuation processes, guidelines on ownership and licensing of data and databases must develop alongside to protect data subjects and retain the public facing imperatives of DPI.

> DaaS models look to roll-out the DPI approach at scale and speed, and involve significant private sector participation in both technical implementation and governance.

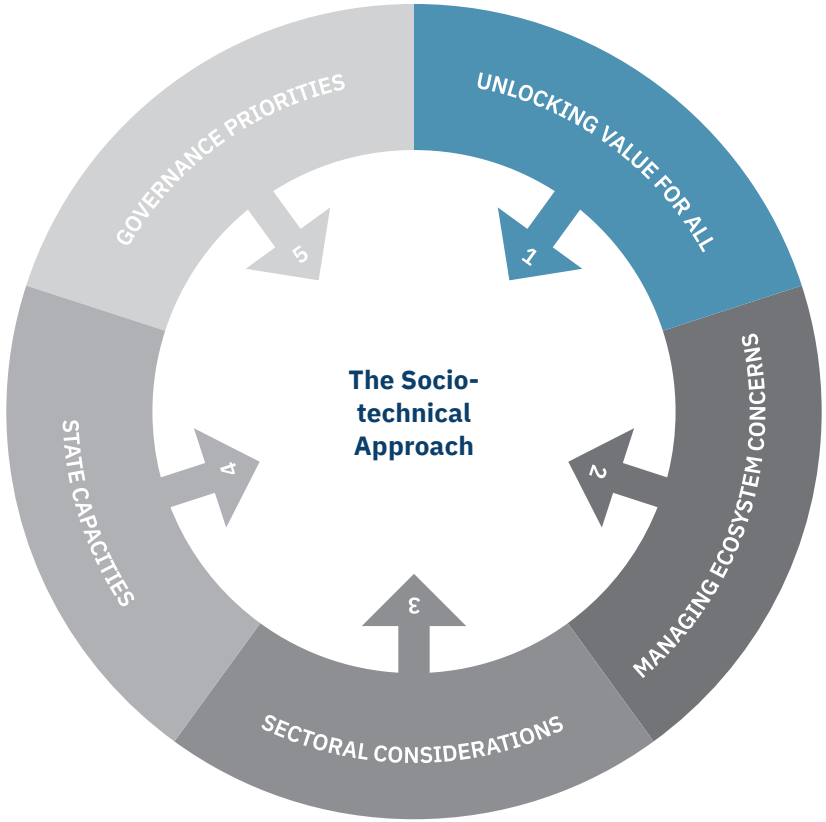
Auditability and provenance through technical traceability and governance measures could provide data subjects with information of how their data is being transformed and used.

> Estonia's licensing model for X-Road allows Cybernetica to generate revenue from international implementations, separating sustainability from commercialisation of underlying technology.

New market competition requires governance and technical mechanisms that assign value to data

Source: Aapti, UCL, DCO, ISEAL Alliance; Aapti analysis

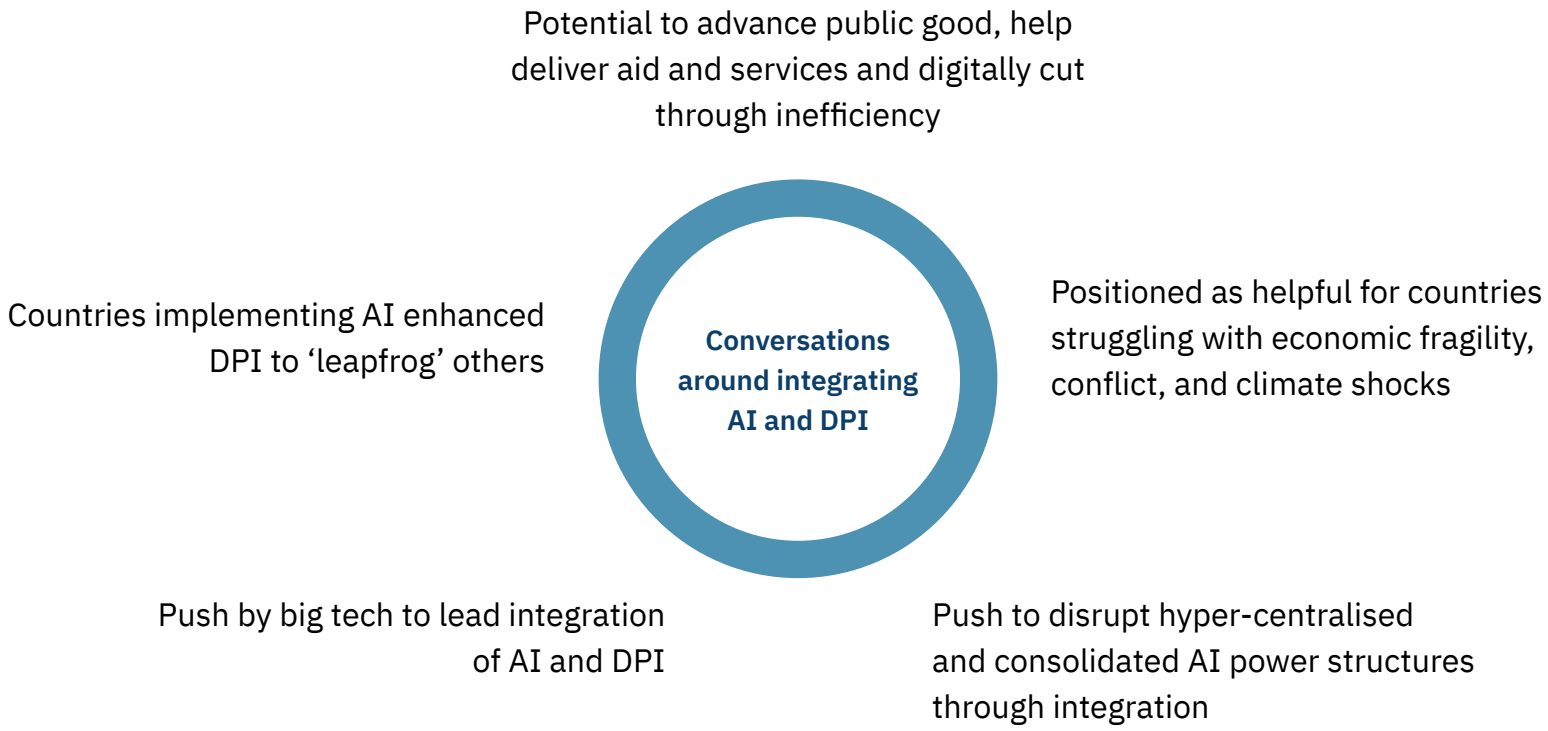
Governance, therefore, needs to be all-encompassing to address a multi-faceted ecosystem



A sociotechnical approach to governance can help address societal contexts, economy considerations, varying capacities and influences

Source: Ada Lovelace Institute; Aapti analysis

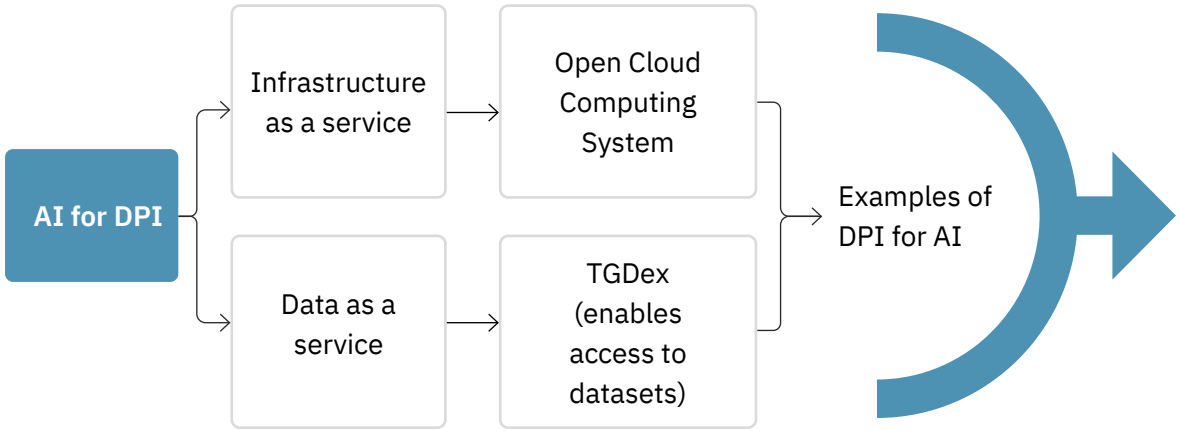
An additional emerging consideration for data sharing is the strong international push towards integrating AI and DPI



The growing vision is for AI to enhance DPI, and DPI to form a foundation to build better frontier AI, emerging as AI for DPI and DPI for AI

Source: UN, DPI Africa, Pandey; Aapti analysis

Foundational DPI can help a country build out its own AI systems and disrupt dependencies on established powers



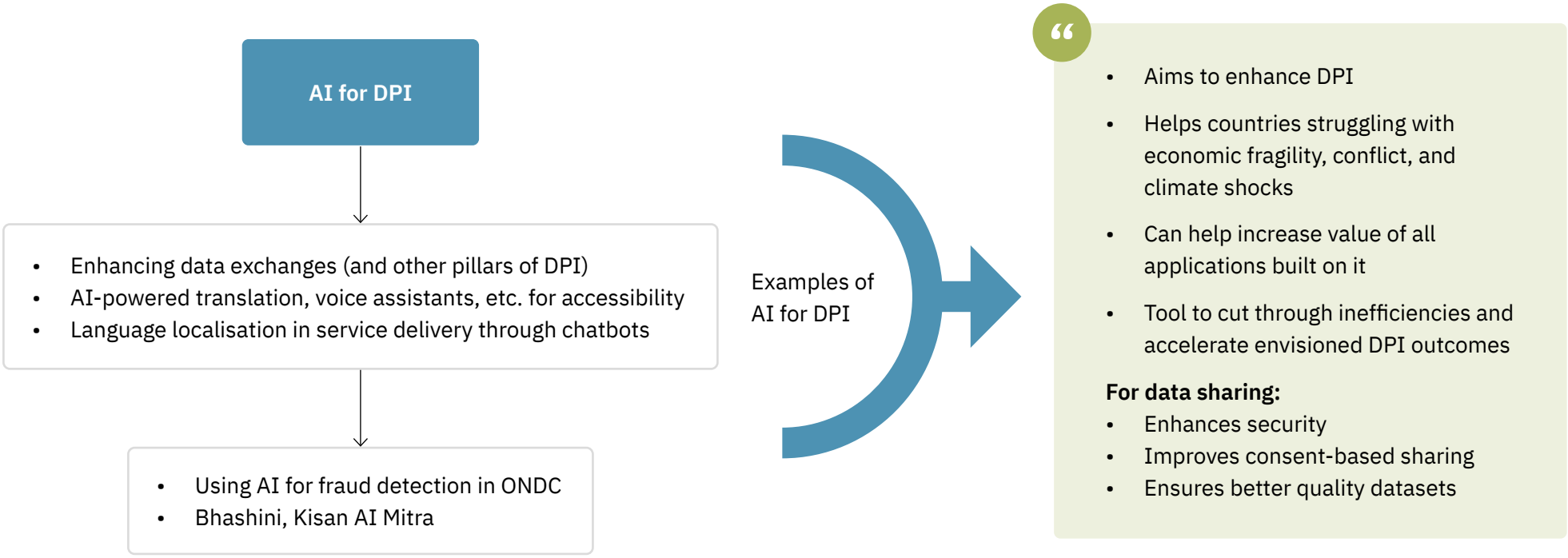
Incorporation of data exchanges with AI through:

- ***Infrastructures as a service***
  stores data, enables internet access through servers and performs cloud computing
- ***data as a service***
  access to large high-quality open datasets for AI training and testing, enabling data sharing on demand, access, management and analysis of datasets without owning infrastructure
- Allows for AI models built on local contexts, to help reduce bias and better citizen welfare

DPI can be a foundational tool to help prevent the exploitation of the Global South and push for sovereignty

Source: People + ai, Sengupta et al; Aapti analysis

There is a global push to implement AI into existing data exchange systems to ‘leapfrog’ others in the global market

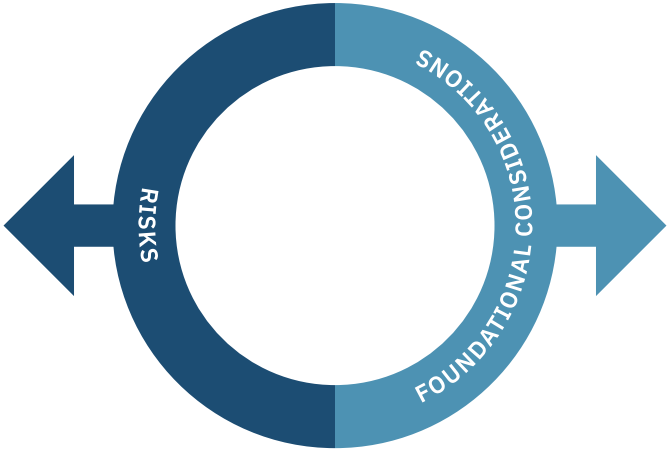


The pull towards integration of AI emerges from the need to enhance security, improve consent-based sharing and ensure the creation of higher quality datasets

Source: ACET, Nagar et al., Nigam; Aapti analysis

The integration of AI with DPI is an additional factor driving a country's choice, providing an opportunity to push for better governance

- **Emergence of immediate concerns:**
  - Compatibility with legacy systems
  - Regulatory alignments
  - Lack of knowledge around financial stability
  - Data security, privacy and trust
- Data's fundamental role in informing global digital inequalities runs the risk of being ignored
- Nascent stage of integration can multiply existing exclusion and bias



- Integrate renewable energy sources
- Finance based on lifetime costs for sustainability
- Transparency and accountability
- Embed resilience
- Ensure compatibility with legacy systems
- Prioritise universal access and design
- Service delivery pilots before wide-spread implementation
- Make trust a non-negotiable
- Collaborate across sectors

AI and DPI need to be integrated thoughtfully, to ensure effectiveness, scalability and relevance for whom it's being built out for

DPI for AI and AI for DPI can be leveraged as an opportunity to ensure better risk management and a governance-first approach

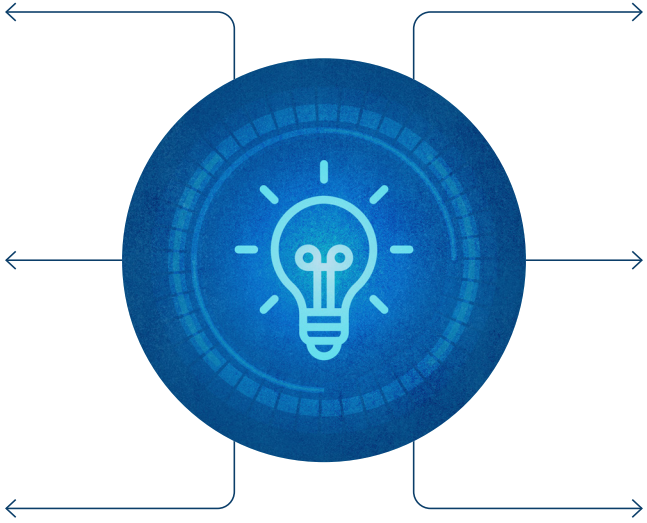
Source: UNDP, Sengupta et al.; Aapti analysis

Our literature review reinforces the need for a governance-first approach to data sharing

There continues to be ambiguity around the definition of data sharing itself. We find more value in understanding the functions of the elements of the ecosystem and the consequences of their interactions

A comprehensive taxonomy can help countries engage in more long-term planning around building data sharing DPI

The various elements of the ecosystem have compound effects and must be scoped out more fully



Data in sharing models are flexible in form. A useful exercise could be to consider provenance, ownership or value to categorise data type

There is a substantial lack of literature on the various considerations (state capacity, limited capital, lower digital literacy) that are unique to building data sharing DPI in the Global South. These considerations need to be understood on a more practical level

Even following the deployment of data sharing models, governance and oversight play a key role in the evolution of the model, and its interaction with other existing infrastructures

Source: Aapti analysis

06

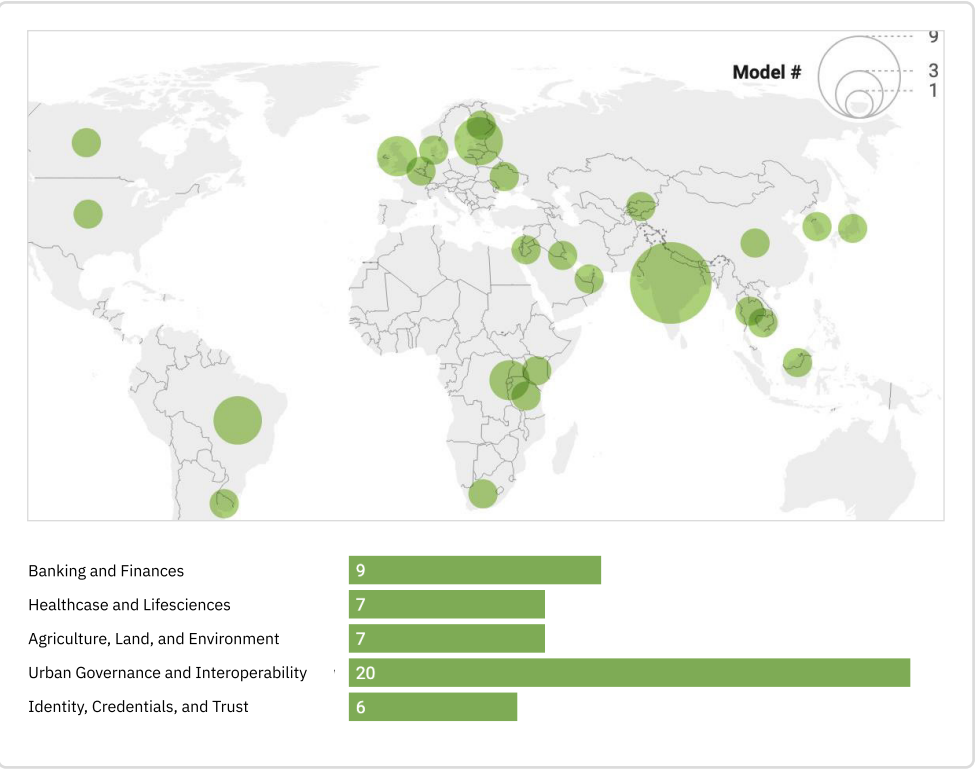
Lessons from Models Landscaping

TABLE OF CONTENTS

Review of 50+ models across ~26 countries to build an actionable taxonomy

Our review reveals observable patterns around **regional geopolitics, sectoral imperatives, and dominant functional use-cases** that influence how data-sharing ecosystems take shape

- **Despite diversity, a set of recurring core architectural styles is evident**, each with characteristic governance structures, technology layers, and institutional enablers
- Variations in state capacity, regulatory maturity, and **underlying motivations drive distinct adaptations of otherwise similar core models** (e.g., Open Banking, X-Road, OpenHIE), with different governance choices and implementation pathways
- Sector-level mapping reveals **functional overlaps and nuanced sub-sector categorisations,** particularly across trade, public administration, urban and citizen services, pointing to converging needs but divergent instantiation patterns



This led to the adoption of **a socio-technical approach with opportunities to further examine the factors that influence** the choices in adopting the different architecture types and governance approach for the data exchanges.

Source: Aapti analysis, Research data base (detailed excel sheet)

Aapti undertook a socio-technical approach to landscaping and analysis, opposed to a more singular technology first study of the infrastructures

Accordingly, we study the driving forces that set the motivations and purpose definition of a model by a state, and map the primary architecture style adopted along corresponding axis around the data processing, the coordinator function, the trust anchors etc.

Motivations and use-cases identified to prioritise a socio-technical framing

- What constitutes state motivations (service delivery, compliance, market-making, or sovereignty etc) and why they are the anchor to the direction for architecture, governance, and mandate design
- How motivations offer the strategic “why” that guides system design as opposed to the use-cases which are the tactical “how” reflected in domain applications and workflows
- How motivations can guide the evolution of the architecture choices along institutional capacity, ecosystem maturity, and political institutional capacity (e.g., movement from centralised control to more federated or user-centric models).
- How motivations and use-cases carry diverse contextual variation across states while pointing to recurring tendencies for structured analysis and anticipation of likely model pathways

Common core architectures mapped to functional design and governance implications with key questions

- Where is the data hosted and processed? Centrally, distributed across nodes, or retained by the source organisation or user?
- How are participating entities authenticated? A federated trust mechanism, centralised IDs, or cryptographic verification?
- How are standards defined and evolved? Is it via government mandate, stakeholder consortia, or open community process?
- What are the functions of the central coordinator vs. the nodes? Does the central unit only manage trust and registry, or does it also host data, enforce policy, or run analytics?
- How are data access permissions and consents managed? Through user-mediated consent layers, institutional MoUs, policy-based access controls, or none (open data)?
- What are the prominent mechanisms of anchoring trust? technical, legal, or social?

Source: Aapti analysis

State motivations to adopt such systems can be mapped to observable use cases and point to examples which necessarily prioritise to adopt the same

The following mapping traces the **Motivation (+description) → Use case / application → Examples**

Administrative efficiency & state system modernisation	**Integrated service delivery & social outcomes**	**Regulatory compliance, and sector coordination**	**Economic competitiveness, innovation & market creation**	**Sovereignty & national resilience**
To make government systems interoperable, reduce duplication, and modernise legacy platforms.	To enable whole-of-government services centered on a person, community or business.	To enable real-time compliance, reporting, supervision, and secure sectoral exchanges.	To stimulate digital economy growth, create fair markets, and enable innovation.	To ensure jurisdictional control with interoperability, cross-border verification, and continuity under crisis.
Governance interoperability	Social protection	Financial account aggregration	Digital trade	Early warning systems
Coordinated decision making/ policy fromation	Integrated business registration	Health HIEs	Open finance/ open insurance	Cross-border mobility + ID credentials
Core e-governance transactions	Welfare delivery/ inclusion	Tax/business registries	AI/ language datasets	Public integrity systems against corruption
Estonia X-Tee	India ABDM	Account Aggregator India	Shanghai Data Exchange	Trembita Ukraine
Belgium Federal Service Bus	CAR SICAR Brazil	Rwanda HIE	Open Banking UK	EU Digital ID Wallet

Source: Aapti analysis

Countries adopt models based on core architectural styles to optimise for unique implementation contexts and priorities

API GATEWAYS

A central portal publishes & manages APIs for consumers to authenticate or connect

- Scaled by governments to unify digital services for GovTech
- Ideal for ecosystems with many agencies needing a consistent developer experience

CENTRALISED DATA HUBS

Requests flow via central hub, which enforces consent, access & interoperability

- Emerged where sector-specific exchanges were needed: energy, agriculture, logistics.
- Efficient when one dataset is reused

P2P EXCHANGES

Participants communicate directly often via federation and encrypted "security servers"

- Any central hub merely guide on trust anchors, member lists etc
- Better fail safes and sovereign with easy federation structures and cross border models

OPEN DATA PORTALS

One-way publication of datasets for reuse, not consent-based or transactional exchange

- Typically accessible by way of bulk downloads or APIs
- Focus on transparency and innovation with non-sensitive datasets

ACCOUNT AGGREGATORS

Account/ consent manager mediates access to user's data across institutions only via explicit consent

- Relevant in sensitive data contexts like health, finance etc
- Often alongside API gateways as in ABDM and open finance

GOVERNMENT SERVICE BUS

Message-oriented central middleware (bus/queue) that routes messages, often in XML/SOAP formats

- Based on ESB architecture, used to integrate fragment gov systems
- Popular in 2000s before API-first design matured or in legacy systems

Source: Aapti analysis

Models studied frequently integrate features from multiple architectural styles beyond their base technology design choice

API Gateways	P2P Exchanges	Government Service Bus	Open Data Portals	Central Data Hubs	Verifiable Credentials	Consent mediated flows
A central portal publishes & manages APIs for consumers to authenticate or connect Scaled by governments to unify digital services for GovTech Ideal for ecosystems with many agencies needing a consistent developer experience	Participants communicate directly often via federation and encrypted "Security Servers" Any central hub merely guide on trust anchors, member lists etc. Better fail safes and sovereign with easy federation structures and cross border models	Message-oriented central middleware (bus/queue) that routes messages, often in XML/SOAP formats Based on ESB architecture, used to integrate fragment gov systems Popular in 2000s before API-first design matured or in legacy systems	One-way publication of datasets for reuse, not consent-based or transactional exchange Typically accessible by way of bulk downloads or APIs Focus on transparency and innovation with non-sensitive datasets	Requests flow via central hub, which enforces consent, access & interoperability Emerged where sector-specific exchanges were needed: energy, agriculture, logistics. Efficient when one dataset is reused	Digital credentials cryptographically signed by an issuer, shared securely to prove authenticity Decentralised exchanges that do not require any central database of personal data Has cross-sectoral utility in public and private sectors (ex. wallets).	Account/ consent manager mediates access to user's data across institutions only via explicit consent Relevant in sensitive data contexts like health, finance etc Often alongside API gateways as in ABDM and open finance
Open Banking UK	-	-	-	-	-	Open Banking UK
-	Sahamati	-	-	-	-	Sahamati
-	-	-	-	GoT-HoMIS (Tanzania HIE)	-	-
-	-	-	-	-	OpenCerts	-
-	MyGDX	-	-	MyGDX	-	-
-	MUNI Chatbot	-	-	MUNI Chatbot	-	-
-	-	-	-	-	Singpass	Singpass
-	-	-	OGD Platform India	-	-	-
-	-	UAE GSB	-	-	-	-
API Setu	-	-	-	-	API Setu / Digilocker	-
-	CamDx (with higher central control)	-	-	-	-	-

Source: Aapti analysis

Finally, motivations were mapped against architectures to guide early state choices on system design and governance while keeping evolution in mind

Administrative efficiency & state system modernisation	Integrated service delivery & social outcomes	Regulatory compliance, and sector coordination	Economic competitiveness, innovation & market creation	Sovereignty & national resilience
To make government systems interoperable, reduce duplication, and modernise legacy platforms	To enable whole-of-government services centered on a person, community or business	To enable real-time compliance, reporting, supervision, and secure sectoral exchanges	To stimulate digital economy growth, create fair markets, and enable innovation	To ensure jurisdictional control with interoperability, cross-border verification, and continuity under crisis
API gateways	Consent flows	Central data hubs	API gateways	P2P exchanges
ESB/GSB	API Gateways	Consent flows	Open data portals	VCs and decentralised identity
P2P exchanges	Verifiable credentials	Open data portals	Verifiable credentials	Open data portals

While the taxonomy provides the comprehensive overview of all the choices available and long list of various tech gov data and enabling components for each, detailed guidance on this and useability of the taxonomy shall come from the complementing outputs of the use case repository and tool

Source: Aapti analysis

07

Taxonomy

TABLE OF CONTENTS

A comprehensive taxonomy is urgently needed to navigate the complex and largely unchartered data exchange ecosystem

Why is a taxonomy necessary?

- A majority of the scholarship and the developers of data sharing models do not have a common set of definitions on key terms in the data exchange ecosystem
- There is a significant overlap of different architectural styles in the building of data exchange models, making it difficult to categorise and understand how these models work, and what role each architectural or technical component plays in its functioning
- There are no comprehensive guidance documents to assist governments and other key stakeholders in familiarising themselves with the data exchange ecosystem, and building specific data exchanges to fulfill their needs

How do we approach building a taxonomy?

- The taxonomy is built using a socio-technical approach to examine data exchanges as DPIs, as opposed to the commonly used (and limiting in impact) techno-legal approach. This socio-technical lens primarily focuses on the impacts of these models, and the value created by them over their technical efficiency
- The key architectural, technical, and governance elements to be included in the taxonomy are based on the literature review of existing scholarship. Further, the various elements of the taxonomy layers are populated based on the review of over 50 different data exchange models across the world, at various stages of their implementation

Source: Aapti analysis

Our work mapped various components and elements of the data sharing ecosystem into a comprehensive taxonomy from a socio-technical lens

The taxonomy has been prepared as a comprehensive analysis of global developments in data exchanges and trusted data sharing by cataloguing and synthesising the diverse and complex data exchange ecosystem into clearly defined, distinct components. The purpose of the taxonomy is three-fold.

Firstly, it aims to provide clarity to readers on the various societal, regulatory, architectural, and technical components of developing data exchanges

Secondly, it aims to assist governments, innovators, investors, policy professionals, and other interested stakeholders to gain a broad understanding of the various factors that influence the development of a data exchange

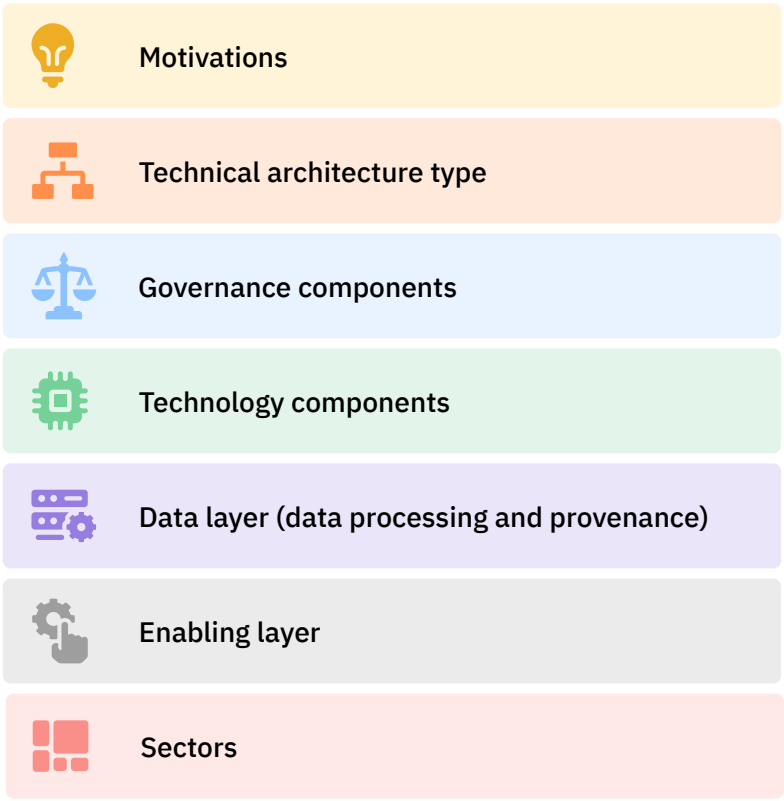
Thirdly, it acts as a foundation to firmly establish data exchange as the third pillar of digital public infrastructure (DPI)

Source: Aapti analysis

The taxonomy is a deconstruction of the trusted data sharing ecosystem, not a decision-making matrix

The taxonomy is a comprehensive deconstruction and classification of the various elements that constitute data exchanges and the key factors that shape these models. The taxonomy is structured based on the insights from secondary research and expert interviews. These elements are divided into seven distinctive categories.

The seven categories of the taxonomy are arranged based on the specific importance of each category for the states and operators intending to build a data exchange model and for those interested in understanding how trusted data sharing function. As such, it is recommended to peruse the taxonomy in a top-down manner.



The elements in the different categories of the taxonomy do not interact with or flow into each other

Source: Aapti analysis

The core architectural model types can be better understood by examining their motivations, main functions, and other key aspects

The seven core models provide a long list of the base architectural design that can be adopted for a data sharing system depending on the motivations, priority functionalities, and technical/social contexts of the implementing state body. These carry a varied set of implications on the larger data, governance, and enabling layers that support its operations.

By mapping the seven core architectural types against six major questions to consider as a model operator, the spectrum of design choices of each architecture type becomes evident.

- Where is the data hosted and processed? Is data stored centrally, distributed across nodes, or retained by the source organisation or user?
- How are participating entities authenticated and authorised? Is there a federated trust mechanism, centralised identity management, or cryptographic verification?
- How are standards defined and evolved? Are technical and semantic standards set by a government mandate, multi-stakeholder consortium, or open community process?
- What are the respective functions of the central coordinator vs. the nodes? Does the central unit only manage trust and registry, or does it also host data, enforce policy, or run analytics?
- How are data access permissions and consents managed? Through user-mediated consent layers, institutional MoUs, policy-based access controls, or none?
- What is the mechanism of anchoring trust, such as institutional trust (government registry), cryptographic trust (public key infrastructure), or social trust (community reputation)?

Source: Aapti analysis

The taxonomy maps the seven core architectural model types against key questions to differentiate the functioning and priorities of each type

Differentiator	API Gateways	P2P Exchanges	Government Service Bus	Open Data Portals	Central Data Hubs	Verifiable Credentials	Consent mediated flows
Data hosting							
Centrally hosted	-	-	-	Centralized public hosting or storage of published datasets (South Korea data.gov.kr)	Data centrally stored/ curated in a domain-specific repository (e.g. Brazil SICAR)	-	-
Node level	-	Each participant retains its data on the node level security servers (e.g. X-Tee, CamDX)	Central middleware (message bus) routes and sometimes temporarily buffers messages (e.g. UAE GSB)	-	-	-	-
Retained at the source org/user	APIs are published in a central catalog; the data itself usually remains at the provider (e.g. API Setu)	-		-	-	Issued credentials are typically stored in the holder's wallet (device/ cloud) (DigiLocker)	-
-				-	-		Access mediated by a consent manager, hosting remains provider-side (Account Aggregator)
Authentication / authorization processes							
Cryptographic verification	-	Federated trust via certificates; mutual TLS and signing handled at node level	-	-	-	-	Two-step auth: institutions authenticate (certificates/keys) + user passes consent
Cryptographic verification	Centralized registration and API keys/tokens management	-	Gateway issues credentials and enforces role-based access centrally	Usually public access (no auth) for open datasets; maybe managed centrally for API usage or rate limits	Central gatekeeping on participant onboarding, API credentials, contractual access etc.	-	
Cryptographic verification		-				Cryptographic verification (signature validation) and revocation checks	-

Source: Aapti analysis

Taking a closer look at the different layers of the taxonomy

Motivations

The motivation clusters and are not mutually exclusive as priorities often overlap. This layer lists the primary motivations to which the design of a data exchange may be optimised. Motivations may be foundational which enable other purposes, such as interoperability, social protection, or registries, while others may be advanced or sectoral priorities that models evolve into.

Administrative efficiency & state system modernisation	Integrated service delivery & social outcomes	Regulatory compliance, and sector coordination	Economic competitiveness, innovation & market creation	Sovereignty & national resilience
To make government systems interoperable, reduce duplication, and modernise legacy platforms	To enable whole-of-government services centered on a person, community, or business	To enable real-time compliance, reporting, supervision, and secure sectoral exchanges	To stimulate digital economy growth, create fair markets, and enable innovation	To ensure jurisdictional control with interoperability, cross-border verification, and continuity under crisis
Governance interoperability	Social protection	Financial account aggregation	Digital trade	Early warning systems
Coordinated decision making/ policy formation	Integrated business registration	Health HIEs	Open finance / open insurance	Cross-border mobility + ID credentials
Core e-governance transactions	Welfare delivery / inclusion	Tax/business registries	AI / language datasets	Public integrity systems against corruption

Source: Aapti analysis

Taking a closer look at the different layers of the taxonomy

Data architecture types

The core technical architectures adopted is essential in ensuring that the data exchanges can achieve their intended purpose. However, although the taxonomy lists the different architecture types separately, many data exchanges are built using two or more architecture types to achieve different aspects of the exchanges purpose.

P2P Exchanges	Government Service Bus	API Gateways	Central Data Hubs	Open Data Portals	Consent mediated flows	Verifiable Credentials
Participants communicate directly often via federation and encrypted "Security Servers" Any central hub merely guide on trust anchors, member lists etc. Better fail safes and sovereign with easy federation structures and cross border models	Message-oriented central middleware (bus/ queue) that routes messages, often in XML/SOAP formats Based on ESB architecture, used to integrate fragment gov systems Popular in 2000s before API-first design matured or in legacy systems	A central portal publishes & manages APIs for consumers to authenticate or connect Scaled by governments to unify digital services for GovTech Ideal for ecosystems with many agencies needing a consistent developer experience	Requests flow via central hub, which enforces consent, access & interoperability Emerged where sector-specific exchanges were needed: energy, agriculture, logistics. Efficient when one dataset is reused	One-way publication of datasets for reuse, not consent-based or transactional exchange Typically accessible by way of bulk downloads or APIs Focus on transparency and innovation with non-sensitive datasets	Account/ consent manager mediates access to user's data across institutions only via explicit consent Relevant in sensitive data contexts like health, finance etc Often alongside API gateways as in ABDM and open finance	Digital credentials cryptographically signed by an issuer, shared securely to prove authenticity Decentralised exchanges that do not require any central database of personal data Has cross-sectoral utility in public and private sectors (ex. wallets)

Source: Aapti analysis

Taking a closer look at the different layers of the taxonomy

Governance components

Governance plays a crucial role in shaping data exchanges. Governance mechanisms often impact or even determine the additional technology components or features that exchanges will need to incorporate. Good governance measures are also key to the sustainability of data exchanges- ensuring that exchanges comply with required standards.

Legal or policy foundations	Trust and compliance mechanisms	Institutional or organizational governance	Operational and security governance	Contractual or economic governance
Data protection laws	Certification authorities & PKI	Central coordinating authority	Access control rules	Data sharing agreements
Sectoral regulations	Time-stamping & signature auth	Community/consortium governance	Encryption & data security mandates	Standardized participation contracts
Cybersecurity laws	Audit & logging requirements	Membership & onboarding rules	Incident response & liability rules	Consent frameworks
Digital ID & eKYC regulations	Standards & interoperability	Oversight & accountability bodies	Resilience & continuity planning	Cost recovery & pricing models
National strategy/policy docs	Testing, sandboxes & certification			Bilateral/multilateral agreements
Open data policies	Open data policies			Trade and digital economy agreements
Adequacy standards				

Source: Aapti analysis

Taking a closer look at the different layers of the taxonomy

Technology components

The technology components determine how users, operators, and service providers interact with the data exchanges. These components also help enhance the intended function of the core architecture, and granting flexibility to data exchanges.

API & integration components	ID, authentication, access mgmt	Data & metadata components	Onboarding & administrative tools	Monitoring, logging, and operations	Security & trust components
API registry / catalogues	Authentication & authorization	Metadata / service descriptions	Registration / onboarding portals	Logging & audit trails	Security servers
API gateway / orchestration	Role / permission management	Dataset & AI model repositories	Configuration proxy / services	Monitoring APIs & operational consoles	Digital signatures & time-stamping
API management / generation tools	Consent management and token	Standardized data formats / schemas	API discovery & publishing tools	Load balancers & traffic management	Encrypted & signed traffic / TLS
Standardized APIs & Adapters	User profile management	Data processing & transformation	Cloud / hosting infrastructure	MIS dashboards	PKI integration
Data Flow APIs & Notification APIs	Secure multi-party computation	Data mapping tools			PETs
					Encryption tools
					Confidential clean rooms

Source: Aapti analysis

Taking a closer look at the different layers of the taxonomy

Data processing and data provenance layers

The data layer provides useful context on the manner of data processing undertaken by the data exchange models. Whereas, the data provenance layer provides significant insights in the functioning of data exchanges when viewed through use-case illustrations (such as choice of architecture for personal vs. non-personal data).

Data format	Databases	Metadata	Data registries	Middleware	
Data Processes	Data transformation	Data normalisation	Data security	Data mapping	Data abstraction

Government origin data (Collected via administrative systems, registries, or compliance mandates)	Institution origin data (Generated by regulated intermediaries but harmonized through common standards)	Individual origin data (Collected through user action or consent, often portable)	Sensor / environmental data (Captured through connected sensors, satellites, or networks)	Derived / analytical data (Created through AI/ML, or derived indexes)
Civil registries (X-Road)	Banking transaction data (Sahamati)	Self-uploaded documents	Farm or climate data (Agristack)	ESG Scores / Compliance ratings (CAR)
Taxation systems	Hospital EHRs standardised under national Schema (Singapore NEHR)	Trusted credentials (EU ID, Digi-locker)	Transport telematics	Metadata with traceable transformation (Shangai SDE)
Administrative records / land cadastres (CAR, Agristack)	Telco usage data	Consent-based health data (linked to ID, eg. COWIN)	Health indicators	Open banking / credit scoring etc.
			Utility grids (Estfeed)	

Source: Aapti analysis

Taking a closer look at the different layers of the taxonomy

Enabling layer

The enabling layer plays a key role in ensuring robustness or sustainability of the data exchanges by providing supporting infrastructure for the exchange. This includes community contribution, community engagement, funding initiatives, etc. Although these factors are not a component of the data exchange, they should be considered for the taxonomy for the role they play in supplementing data exchanges.

Connectivity & foundational readiness	Institutional and capacity building	Ideation, pilots, and model definition	Legal and governance set-up	Sustainability and scaling
Digital connectivity infrastructure	Cross-ministry integration	Base digital ID & trust infrastructure	Cybersecurity posture	Reference architectures
Cloud hosting services	Digital transformation strategies	Developer sandbox environments	Institutional alignment mechanisms	
Digital acceleration plans	Capacity development programs	Use case co-design workshops/hackathons		
	Change management & institutional buy-in	Open Source databases		

Source: Aapti analysis

Taking a closer look at the different layers of the taxonomy

Sectors

Similar to the data layer, the specific sectoral applications of data exchanges provide interesting insights, but a comprehensive taxonomy of sectoral usage would not be appropriately captured in a taxonomy of data exchanges.

Source: Aapti analysis

The governance layer of the taxonomy includes a comprehensive list of governance approaches for data exchanges

Legal or policy foundations	• **Data protection laws:** National or regional legislation that governs personal data processing (GDPR EU, PDPA Singapore, DPDPA India) • **Sectoral regulations:** Domain-specific mandates (financial regulations in Open Banking, energy market rules in Estfeed etc.) • **Cybersecurity laws:** Security compliance obligations (e.g., Thailand's Cybersecurity Act 2019) • **Digital ID & e-KYC regulations:** Laws enabling identity verification and authentication for participation (KYC regulations in India, Thailand etc.) • **National strategy/policy documents:** Country or model level vision documents (Digital Transformation Strategies, ADeX Booklet etc.) • **Open data policies:** Legal/policy rules on public sector data availability and re-use
Trust and compliance mechanisms	• **Certification authorities & PKI:** Management of cryptographic credentials (mandatory in MyGDX, CamDX, X-Road) • **Time-stamping & signature authorities:** Integrity verification tools for transaction traceability (X-Road global config, CamDX security server) • **Audit & logging requirements:** Rules or tools for traceability, accountability, and redress (Estonia, Finland, Ukraine) • **Standards & interoperability:** API/metadata standards, message formats, and security standards (Singapore APEX's API standards manual etc.) • **Testing, sandboxes, & certification:** Pre-deployment testing for APIs/systems (e.g., Thailand's DGA pre-deployment security/ performance tests)

Source: Aapti analysis, X-road

Institutional or organisational governance	• **Central coordinating authority:** A state entity managing trust, oversight, and membership (MAMPU for MyGDX, GovTech for APEX) • **Community/consortium governance:** Multi-stakeholder arrangements where industry agencies co-govern (Sahamati in India, NIIS for X-Road) • **Membership & onboarding rules:** Contractual/administrative rules for who can join, including vetting, obligations, exit protocols • **Oversight & accountability bodies:** Independent or semi-independent actors ensuring compliance and trust (e.g., EU regulators under Estfeed)
Operational and security governance	• **Access control rules:** Organisational-level or user-level access policies (X-Road, CamDX) • **Encryption & data security mandates:** Encryption in transit/storage, often PKI-based (CamDX, X-Road, MyGDX) • **Incident response & liability rules:** Rules for breach handling and accountability (less explicit in some LMIC models, stronger in EU) • **Resilience & continuity planning:** Provisions for redundancy, caching, fallback mechanisms (X-Road caching; MyGDX redundancy)
Contractual or economic governance	• **Data sharing agreements:** Bilateral/multilateral agreements between data providers and consumers • **Standardised participation contracts:** Common templates to reduce friction (Estfeed onboarding contracts) • **Consent frameworks:** Mechanisms ensuring user-controlled data access (DEPA in India, MyInfo/Singpass in Singapore) • **Cost-recovery & pricing models:** Economic rules for platform use (GovTech service charges for APEX APIs)

Source: Aapti analysis, X-road

The technology layer lists the key technologies that sustain data exchanges

API & integration components	• **API registry / catalogues:** Centralised or federated directory of APIs and services; supports discovery, subscription, and governance • **API gateway / orchestration:** Handles request routing, traffic management, and security enforcement • **API management / generation tools:** Tools to create, publish, and manage APIs for participating agencies • **Standardised APIs & Adapters:** Ensures interoperability across heterogeneous systems • **Data Flow APIs & Notification APIs:** Supports push/pull of encrypted data and real-time updates between providers and consumers

Central to API-first architectures; with variations on cloud hosted, private sector focused, legacy integrated, or citizen facing models

Identity, authentication, and access management	• **Authentication & authorisation:** Manages consent artifacts and lifecycle for data principals; includes token issuance for secure API access • **Role / permission management:** Defines who can do what with which data • **Consent management and tokens:** Manages consent artifacts and lifecycle for data principals; includes token issuance for secure API access • **User profile management:** Tracks system and user credentials, linked datasets, and access history

Essential in all architectures; with variations for federated models or consent-based models

Source: Aapti analysis

Data & metadata components	• **Metadata / service descriptions:** Enables discovery of datasets and services via standardised documentation • **Dataset & AI model repositories:** Stores and documents curated, de-identified, and tagged datasets and AI models for reuse • **Standardised data formats / schemas:** Ensures consistent representation of financial, healthcare, or registry data • **Data processing & transformation:** Normalisation, enrichment, and transformation between systems
Universally relevant; critical for interoperability with variations in AI/data intensive architectures or for legacy integration	
Onboarding & administrative tools	• **Registration / onboarding portals:** Allows agencies or private actors to join and configure access, including self-service and testing tools • **Configuration proxy / services:** Centralised configuration for nodes and participants • **API discovery & publishing tools:** Enables API providers to register, document, and expose services for subscription • **Cloud / hosting infrastructure:** Provides scalable, secure hosting; often includes regulatory compliance certifications
Essential for API-first and private-sector-integrated models with variations for distributed or cloud hosted models	
Monitoring, logging, and operational tools	• **Logging & audit trails:** Immutable tracking of data exchange, access, and modifications • **Monitoring APIs & operational consoles:** Real-time health monitoring, traffic analytics, and dashboards • **Load balancers & traffic management:** Ensures scalable, fault-tolerant operations • **MIS dashboards:** Management information systems for administrators to track performance, usage, and issues
Essential in distributed/federated systems for traceability with API first multi-stakeholder exchanges	

Source: Aapti analysis

Security & trust components	• **Security servers :** Node-level security for encryption, signing, and secure message transport • **Digital signatures & time-stamping:** Ensures data integrity and non-repudiation of exchanged data • **Encrypted & signed traffic / TLS:** Ensures secure data exchange channels across all nodes • **PKI integration:** Used to authenticate participants and secure transactions • **Privacy-enhancing technologies (PETs):** Optional mechanisms for anonymisation, differential privacy, or secure multi-party computation

Critical in federated structures with node level exchange, non-repudiation, and sensitive data contexts

Source: Aapti analysis

There are unique governance mechanisms for cross-border data sharing models that should be included in the taxonomy

The taxonomy highlights certain specific elements across all its layers which play a crucial role in building a data exchange for cross-border data sharing. This specific emphasis is to spotlight the unique governance, technology, and enabling layer requirements for cross- border data sharing models.

National legislations and adequacy standards

The primary regulatory requirements for cross-border data sharing is local data protection laws within country jurisdictions. These laws must contain provisions for ensuring safe cross-border data flows, including provisions on adequacy standards for jurisdictions with which the data is shared.

- **The adequacy decisions** undertaken by the EU is a prime example of using jurisdictional data protection laws and adequacy standards in order to permit cross-border data sharing with other jurisdictions.

International frameworks and agreements

International bodies play a much larger role in the framing of governance approaches for cross-border data sharing

- **The Asia-Pacific Economic Corporation** published the APEC Cross-Border Privacy Rules to provide a comprehensive data protection framework for institutions to adopt, to be certified for cross-border data sharing in certain countries.
- **Data Free Flow with Trust (DFFT)** is another similar framework that aims to promote the free flow of data while ensuring trust in privacy, security.

Contractual governance measures

In addition to national legislations and international frameworks, contractual compliance measures play a key role in governing cross-border data sharing. Bilateral and multilateral agreements, trade and digital economy agreements, and other contractual models ensure that necessary compliance is undertaken by data exporters and importers in international data transfers.

- **ASEAN Model Contractual Clauses** and **EU Standard Contractual Clauses** are key examples of model clauses that can be incorporated by data exporters and importers in contractual agreements to ensure safe data sharing.

Highlighting these unique elements and observing the overlapping elements with data exchanges that do not permit cross-border data sharing further bolsters the comprehensiveness of the taxonomy.

Source: Aapti analysis

The need for a separate taxonomy on AI integrated data exchanges

Previous sections of this report have examined in detail the importance of understanding the interplay between AI and DPI, and the notable increase in interest in governments and data exchange builders to integrate DPI with AI systems. The specific uniqueness of the governance, technical, and enabling factors for AI integrated data exchanges has, therefore, necessitated the building of a separate taxonomy. Key insights from the taxonomy are provided below.

Governance approaches

- National AI strategies are unique strategic plans to set-up AI initiatives, regulatory approaches, and strategies for growth in the AI sector.
- AI governance also includes special regulatory approaches like risk assessments, impact assessments, AI sandboxes and AI pilot projects.

Core technologies

- AI integrated data exchange models also require specific technology components, including AI frameworks and libraries, model training algorithms, language processing algorithms, model and dataset repositories.

Enabling factors

- The development of AI integrated data exchanges also requires supporting infrastructures like public databases, AI research initiatives, knowledge and data consolidation initiatives, and developer toolkits.

The separate AI taxonomy provides additional clarity on existing elements in AI integrated data exchanges in addition to providing a foundation for the future AI for DPI and DPI for AI initiatives.

Source: Aapti analysis

aapti institute

Aapti is a public research institute that works at the intersection of technology and society.
Aapti examines the ways in which people interact and negotiate with technology both offline and online.

contact@aapti.in | www.aapti.in