Comments and Critique

Data Governance Framework for Data Platforms (Data Exchange Operators)

for Non-personal Data

Center for Society and Policy
Indian Institute of Science, Bangalore

Draft Data Governance Framework Policy, May 2022 (NDGFP)

Concerns

- Data managed, stored and accessed in differing and inconsistent ways.
- Data spread across different entities, thus preventing potential and innovative use.

Motivations

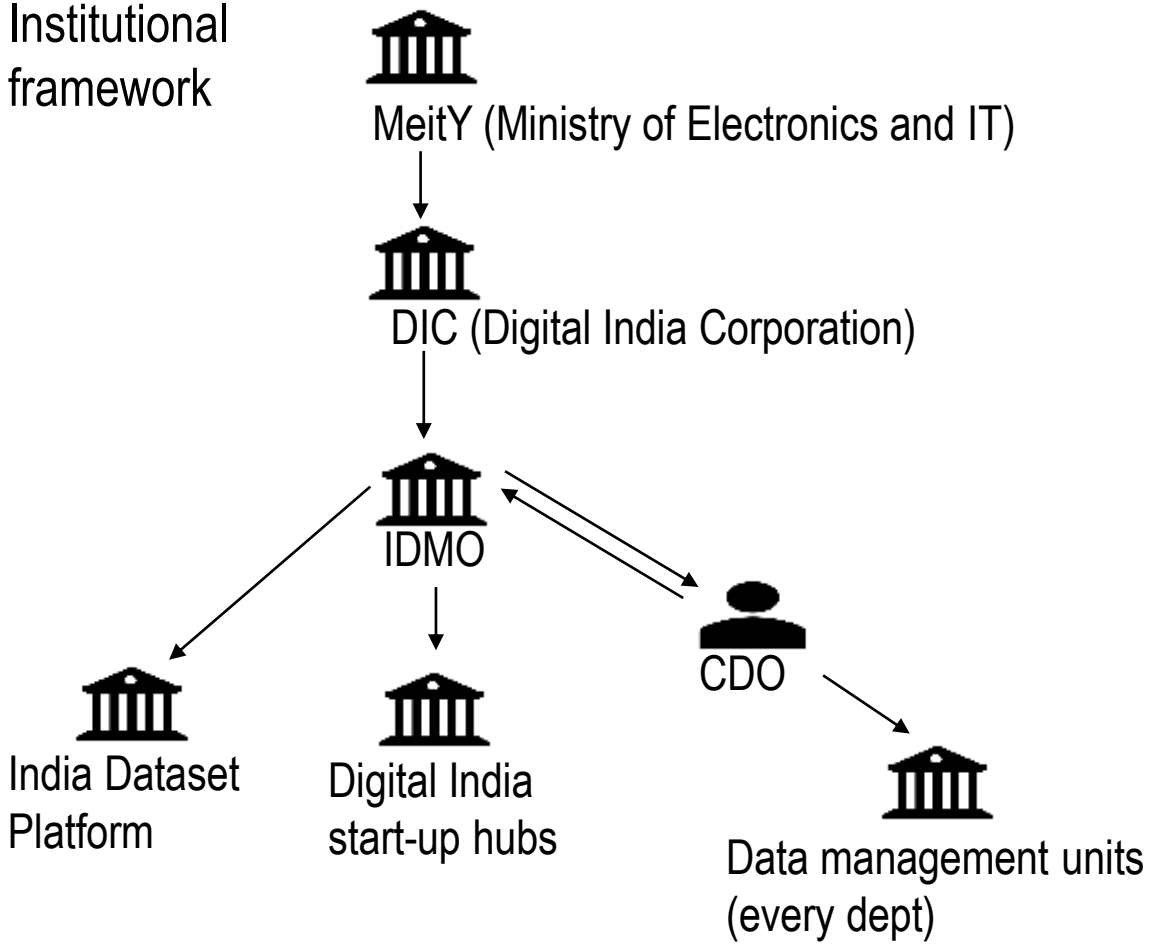
- Power of data must be harnessed for public good and innovation.
  - Transform government services delivery in healthcare, justice, agriculture education.
  - Data based innovations like apps for local soil and crop mapping.
  - Catalyse AI and data led research.

Objectives

- Modernize government data collection and management.
  - Standardize data management and security standards.
  - Accelerate creation of common standards for data platforms.
  - Build a platform for Dataset requests (IUDX)
  - Quality standards for data sets (in progress)
  - Ensure greater citizen participation.

Draft Data Governance Framework Policy, May 2022 (NDGFP)

Institutional framework



What the framework addresses:

- Data Access and Disclosure norms
- Data Quality and Standards
- Usage Rights
- User Charges and Pricing
- Ethical and Fair use of Data
- Policy Monitoring and Enforcement
- Redressal Mechanisms

What are we not addressing in the framework:

- Data Storage & Retention
- Data Anonymisation and Identification of datasets
- Data Security
- Capacity and Skill Building

Examples of NPD: Location of street-lights, public restrooms; real time location of ambulances, traffic, pollution;

The Future of our cities will be shaped by Data

69 cities
Have deployed sensors for data collection across urban sectors

63 cities
Have published 30+ datasets on the Smart Cities Open data Portal

62 cities
Have created GIS layers to view data effectively and support urban planning

32 cities
Are working on use cases using city's data to solve urban challenges

41 cities
Have leveraged their data for development of portal and applications

29 cities
Have published 60+ stories/blogs on the Smart Cities Open data Portal

Questions we deliberated

Guiding Principles:

1. Flexibility to build an ecosystem
2. New age policy directions: public as an equal participant
3. Newer market structures: SEBI vs Gpay vs Amazon vs E-bay

- Access: easy access with minimum documentation versus DXO built controls due to lack of government regulations.
- Quality and Standards: responsible for quality checks – the platform versus citizen driven.
- Roles within DXO: checks for data breach: human intervention versus technological control
- Usage rights: controlled usage or free to use
- User charges: dynamics pricing, structured pricing- Data is a public good!
- Ethics and fair use of data: who controls – platform, government: breach regulations and penalization.

Over to Professor Inder Gopal for

Welcome and Context

Data Access

Protocols for sharing of non-personal datasets while ensuring privacy, security and trust.

Limits of access: Should the platform control the buyer and seller entry? Is legitimacy check by the platform required?

IDMO: Suggests limits to data access

	DXO Controls	DXO Checks	Market Controls
Basic Features	Name/business name	Name/business name	Name/business name
	Aadhar/Certificate of Incorporation	Aadhar/Certificate of Incorporation	Aadhar/Certificate of Incorporation
Advanced Features	PAN/GSTIN	PAN/GSTIN	
	Bank details; For Business: type of business, partners; For Individual: location and address, contact number		
Agreements:	Buy/Sell agreement	Buy/Sell agreement	Buy/Sell agreement
	Lawful processing agreement	Authenticity agreement	Authenticity agreement
	Fair usage policy agreement	Fair usage policy agreement	Fair usage policy agreement
	Continuity agreement		

SEBI **E-bay/ Amazon**

Data Quality and Meta Data Standards

IDMO enabled. Checks and levels as a DXO Service.

Can quality be a service? Can quality standards be defined by the market or we need ISO style standards defined for datasets?

	DXO Controlled	DXO enabled	DXO service
Basic Features	Source or legitimacy Claims of data principal verified	Source or legitimacy Claims of data principal checked internally	Source or legitimacy
	Registered seller/buyer	Registered seller/buyer	Registered seller/buyer
Advanced Features	Quality check: rating or ranking of dataset	Quality: internal quality standard	Quality: consumer ratings

SEBI **E-bay/ Amazon**

DXO certified datasets

- Service provided by DXO: data processing and transacting assistance to Data Owners.
- DXO name associated with dataset provides quality guarantee, and gain consumer trust.

Sub question: What defines quality for difference consumers? Is it standard for all?

If not then can we set standards?

Defining Quality:

- Accessibility: difficulty level for users to obtain data (technical, language etc)
- Accuracy: free of error (values in database correspond to real-world values)
- Completeness: data represents every meaningful state of the represented real world system.
- Consistency: the logical relationship between correlated data is correct and complete.
- Interpret ability: data definitions are clear. Data is not ambiguous if it allows only one interpretation
- Timeliness: time delay from data generation and acquisition to utilization.

Usage Rights

Data usage rights along with permissioned purposes to be with the Data Principal

Data principal

- The natural person to whom the personal data relates
- Individuals, companies, communities
- Exercise key rights, including monetary rights to the data
- Will also provide consent for anonymization and usage of anonymized data

➢ Data Principal has the right obtain any past data from DXO, choose to abstain from transacting with any customer (has to give permission); choose to withdraw where the platform informs the consumer and proceeds with closure; and has the right to object if he/she feels the data is not being utilized for the purpose procured for.

Given that DP has rights where does the Data market and DXO fit in?

Controls DP interactions	Facilitates DP	Collaborates with DP
DXO has right to object to Data Principal requests	DXO has right to object to Data Principal requests	DXO has right to object to Data Principal requests
Right to delete data	Quality check: rating or ranking of dataset	
Choice of abstaining to interact with data principal		
Quality check: rating or ranking of dataset		

User Charges

The DXO charges User charges/Fees for its maintenance and services

Use case reviews

WORK IN PROGRESS

- THEME I- GENERAL PRICING
- THEME II- QUALITY-BASED PRICING
- THEME III- QUERY- BASED PRICING
- THEME IV- PRIVACY-BASED PRICING
- THEME V- SPECIAL CASES

Example: Quality based

WORK IN PROGRESS

❖ *DYNAMIC DATA PRICE* $\Rightarrow P_d = \alpha(Brand\ Value) + \beta\ (Reputation\ of\ Data) + \gamma(Value\ of\ data) + \delta(Quality)$ where P_d stands for price of dynamic data.

= 0.3 (No of years of service) + 0.806 (Feedback of customers) + 0.676 (Reputation of the supplier) + 0.799(Presence of alternative data exchange platforms) +0.361 (Advertising) +0.87 (User Experience) + 0.846 (Turnover Rate) +0.4 (Updation Frequency of dynamic data) + 0.82 (match between actual data and meta data)

❖ *STATIC DATA PRICE* $\Rightarrow P_s = \alpha(Brand\ Value) + \beta\ (Reputation\ of\ Data) + \gamma(Value\ of\ data) + \delta(Quality)$ where P_s stands for price of static data.

= 0.3 (No of years of service) + 0.806 (Feedback of customers) + 0.676 (Reputation of the supplier) + 0.799(Presence of alternative data exchange platforms) +0.361 (Advertising) +0.87 (User Experience) + 0.846 (Turnover Rate) +0.87 (Static nature of data) + 0.82 (Match between actual data and meta data)

User Charges – recommending a public good+ model

Market Involvement	Market plus public good	Public good
Dynamic pricing for data sets.	Platform fixes the base price for a dataset (static pricing).	User fee for maintenance and services (cost based pricing).
1. Base price for user category – researchers, NGOs, administrative groups like panchayats. 2. Dynamic pricing for other user categories. 3. Dynamic pricing for value-added services, and DXO shares an honorarium with the data principal as well.	Platform uses dynamic pricing for its services including collaborating with app developers, hospital services and similar.	

Ethical and Fair use of Data
Monitoring and Enforcement

The IDMO defines the principles for ethical and fair use of data

DXO regulated	DXO monitors	Market regulated
Fair use 1. Monitors fair use 2. DXO takes up internal investigation and enforce legal orders from competent authority.	1. Investigates un-fair use on warning 2. DXO enforces legal orders from competent authority.	1. Reports un-fair use 2. Penalties of prevailing Acts can apply; compensation through legal mechanisms.
Enforcement 1. Enforce appropriate data protection policies and sniff out vulnerabilities and data breaches. 2. Regular audit & assessment of internal controls	1. Enforce appropriate data protection policies. 2. Regular audit and assessment of internal controls.	1. Regular audit of internal controls
Consumer protection Active monitoring 1. Data transactions detrimental to consumers: platform takes up suo-moto action or investigation.	Call upon competent authority 1. Data transactions detrimental to consumers: action based on formal complaints.	Buyer beware! Data transactions detrimental to consumers: cooperates during legal proceedings.


THANK YOU

- Buying Agreement & Selling Agreement: agreements for transaction.
- Authenticity Agreements: Seller claim on authenticity of the dataset.
- Fair Usage policy: Buyer agrees data transacted would be used in a fair manner.
- Additional Service Agreement: agreements on other services provided by the DXO.
- Continuity Agreement: indicates time duration for which data is guaranteed to the consumer by the buyer.
- Lawful processing agreement indicates that the data be processed as per rules and regulations of DXO including authenticity.