

Comments on Draft Data Governance Policy 2026

Submitted by: LIRNEasia

Submitted to: Ministry of Information Technology and Telecommunication,
Government of Pakistan

LIRNEasia welcomes the opportunity to provide comments on the Draft Data Governance Policy 2026 and appreciates the Government of Pakistan's commitment to adopting an open and consultative approach to its development. The Draft Data Governance Policy 2026 represents one of the most significant public policy initiatives undertaken by the Government of Pakistan in support of digital transformation. For the first time, Pakistan has sought to establish a comprehensive framework governing the lifecycle of public-sector data, integrating principles relating to stewardship, interoperability, privacy, citizen empowerment, digital sovereignty, artificial intelligence, open data and the emerging data economy within a single national policy. The draft therefore marks an important transition from fragmented regulation towards a coherent system of national data governance.

The draft policy demonstrates a mature understanding of the role that data plays in modern government. It correctly recognises that data is no longer merely an administrative by-product but a strategic national resource capable of improving public service delivery, evidence-based policymaking, digital public infrastructure, scientific research, artificial intelligence and economic growth. Equally important is its recognition that effective data governance requires not only legal safeguards but also institutional arrangements, common technical standards and coordinated implementation across government.

We therefore welcome the opportunity provided by the government to provide input into this important document. We appreciate the commitment to transparency in policy making shown through this consultation.

In this spirit, we provide the following general and specific observations.

LIRNEasia provides this input as an independent, regional digital policy think tank, based in Colombo, Sri Lanka, and is active across emerging Asia Pacific since 2004. We conduct in-depth, policy-relevant research, using multidisciplinary approaches in order to collect robust data and document regional good practices. We then communicate independent, actionable knowledge, and work with policy makers, regulators, industry, civil society and the media, to improve the services and facilities available for the poorest citizens in Asia. We are currently conducting policy-relevant research in 15 countries in South, South-East and East Asia and have been working in Pakistan since our inception.

1. Defining “sovereignty by default” vs. “openness by default” when it comes to government data.

There are contradictory objectives that are outlined for “sovereignty by default” and “openness by default”.

In clause 3.8, “government data” is defined very broadly as “any data processed by, for, or on behalf of a public body in the discharge of public functions, including records, derived data, metadata, and operational or analytical data.” This covers open data that is released by the government or any public sector entity.

Clause 5.4 states: “Pakistan asserts sovereign control over Government data. Sovereignty is the condition under which the asset is governed. It is preserved by default and yielded only by specific authorisation, with safeguards.”

“Sovereign control” is defined in Clause 3.25 as “the governance, legal, technical, and operational conditions that ensure Government data remains subject to Pakistan’s lawful authority, defined accountability, and effective control.”

Clause 6.2 on “Sovereignty by default” states “Government data shall remain under the lawful authority and effective control of Pakistan. Cross-border transfer is an exception requiring specific governance, justification, and safeguards.” Clause 3.5 on “Cross-border transfer” defines it as “any movement of Government data, or any material remote accessibility to it, from within Pakistan to another jurisdiction, or by persons operating under another jurisdiction.”

Clause 8.1 on data sovereignty repeats: “The Government of Pakistan asserts sovereign control over Government data. Sovereignty is a function of legal authority, residency, technical control, and operational practice acting together.”

Clause 6.5 on “Open by default, closed by exception” states: “Public-sector data shall be open by default, subject to lawful classification. Restrictions shall be applied by exception, on the narrowest grounds permitted by law, with reasons recorded and reviewable.”

Clause 9.1 reiterates this “openness by default” position: “Public-sector data is open by default. Data shall be made available for public use under lawful terms, in machine-readable formats, with metadata, and through the National Open Data Portal, except where lawful classification or specific statutory restriction applies.”

But then Clause 13.1 states: “Cross-border transfer of Government data is permitted only where it is lawful, justified, governed, and safeguarded, and where the sovereign control of Pakistan over the data is not compromised.”

And Clause 15.8 states: “The realisation of value from public-sector data shall not compromise sovereign control over authoritative-source data of national importance, nor shall it permit the transfer of effective control over such data to entities operating under foreign jurisdiction.”

What it means to “transfer effective control” is unclear, and the term “effective control” is never defined. What does it mean for the Government of Pakistan to retain “effective control” over open data that is published by it under a permissive licence but has been copied onto the laptop of a researcher in a Pakistani university, or even say, a British university? This is never spelled out.

Data, unlike tangible property, is non-rivalrous. Data can be copied. Data in digital form is non-excludable because digital technology not only allows copying but requires it — after all, when a person views a

Pakistani government website, the contents of that website are transferred (copied) to the web browser of the viewer. Clauses 6.2, 8.1, read with 3.5 essentially means that no “data processed by, for, or on behalf of a public body in the discharge of public functions” shall be published on a government website; and Pakistani websites should become inaccessible from outside Pakistan if they contain any “government data” unless they have been approved (under Clause 13.2). That is the direct implication of the idea of “sovereignty by default” and the clauses on “cross-border data transfer”. That is not a sustainable position, and cannot have been what the MOITT intended.

Recommendations

Government data ought to be seen as data that belongs to the public unless otherwise required. It should be open by default, with limitations placed on it as required for reasons of integrity, security or defence of Pakistan, or public order, or privacy and personal data protection. The concept of “sovereignty by default” as provided for in the policy is far too wide and cannot coexist with the existence of publicly accessible government websites leave alone the existence of open data. Hence, we recommend dropping that principle.

2. Sections 5 & 6: The National Data Doctrine & Foundational Principles

These two sections mostly provide a sound basis for a data governance policy, and we commend them. There is considerable overlap between the doctrine laid out in Section 5 and the principles laid out in Section 6. They seem to fall at the same level of hierarchy. Hence we recommend that they be merged.

While the doctrine laid out in Section 5 are well thought-out and are useful for any data governance policy, we believe that there is a conflict between the idea that “public bodies are custodians, not proprietors” (5.2) and the invocation of “sovereignty as a condition” (5.4). Further, as explained below, the principles of “sovereignty by default” (6.2) and “openness by default” (6.5) / “data as a national asset” (6.1) are irreconcilable.

Recommendations

We recommend that the concept of “data sovereignty” be defined more narrowly, and the principle of “sovereignty by default” give way to “openness by default”, for the reasons explained above.

3. Section 7: Scope of Application

Clause 7.5 on extraterritorial reach is in practice unenforceable, and further, there is no justification for it to be applicable to all “government data”. This would imply that Pakistani law applies to any reporting of, say, the population of Pakistan as reported in any newspaper outside of Pakistan. The population of Pakistan, as reported by the Pakistan Bureau of Statistics, falls under the definition of “government data” (3.8, “any data processed by, for, or on behalf of a public body in the discharge of public functions, including records, derived data, metadata, and operational or analytical data.”). There is no reason to hold that reporting Pakistan’s population statistics or GDP numbers ought to be governed by this policy.

Recommendations

Remove Clause 7.5, and amend Clause 3.8 to be more limited.

4. Section 8: Data sovereignty and residency

This Section has significant problems. While a multi-dimensional definition of sovereignty is provided in Clause 8.1 and 8.3, the actual requirement in Clause 8.2 completely ignores this. While 8.1 & 8.3 note that sovereignty is legal, technical, and operational (including key custody), not just geographic, the operational rule (Clause 8.2) imposes a strict geographical notion of sovereignty by requiring that government and personal data be "Hosted, stored, and processed within the territory of Pakistan," thus rendering 8.1 and 8.3 nugatory.

Clause 8.2 seeks to enforce data localization under Tiers 1 & 2 (unless excepted) for all "government data" (very broadly defined in clause 3.8), except if classified as "open data"; as well as all "personal data" (again, very broadly defined in clause 3.17); and all "authoritative-source data" (3.2) and "internal data" (undefined).

Tier 1 covers: "RESTRICTED, CONFIDENTIAL, and personal data, including all sensitive personal data and authoritative-source data of national importance". Tier 2 covers: "INTERNAL data not within Tier 1, where a credible operational case exists for offshore processing". (Note: "Internal data" is not defined in the policy.)

When any Pakistani resident uses a non-Pakistani service, or even a Pakistani service running on servers outside Pakistan, their personal data (such as their name, username, gender, and other "data relating to an identified or identifiable natural person" (3.17) are relayed outside of Pakistan. **As drafted, this provision would effectively prohibit Pakistani residents from using any non-Pakistani services. This cannot have been the intent of the MOITT.**

While sensitive governmental data should be subject to data localization requirements, this inverts that and makes data localization the default.

Recommendations

- Restrict Tier 1 strictly to classified national security records, central biometric registers (e.g., NADRA), and other limited databases of national importance.
- Rewrite Tier 2 requirements to permit processing in global cloud environments, provided that the data remains under Pakistani legal jurisdiction and is protected by the technical safeguards (e.g., local encryption key custody) described in Clause 8.3.
- Clarify that open data is the default and restricted data is the exception.

5. Section 9: Open data

This is very well drafted, and the emphasis on openness by default, and the emphasis on machine-readable formats and metadata is appreciated, along with the requirements for open licensing, proactive disclosure, and citizen feedback.

The term "public-sector data" is not defined in the policy, and it is unclear if it is intended to mean something different from "government data". Further, while clause 6.8 ("Interoperability by standards") speaks of non-proprietary formats, this section should also hearken back to that principle. Additionally, the term "machine-readable", might not be understood by those public sector bodies who do not have any familiarity with open data. Given the amount of scanned-image PDFs published by public sector bodies, it might be worthwhile to define the term.

One important provision missing from this section is that of open API access to the open data. In addition to open data, the policy should also require provision of access to the open data via well-specified open API, and maintain a catalogue of APIs on a single site. This is especially important in the world of agentic AI.

Recommendations

Use the term “government data” instead of “public-sector data”. Add a definition of “machine-readable”. Add references to interoperability and usage of open standards for the open data. Add requirements for open APIs for access to the open data, and a catalogue of open APIs alongside the datasets.

6. Section 10: Data sharing and reuse

The National Data Exchange proposed in the draft policy is very interesting as it seeks to promote four very important principles: (a) interoperability (6.8); (b) once-only principle (6.6); (c) authoritative source discipline (6.4); (d) federation of custody (6.3).

However, it fails to provide for (a) citizen empowerment and informational self-determination (6.11); (b) privacy by design (6.10); (c) accountability and transparency (6.15); (d) minimum disclosure (6.7); and the principle of “trust, not ownership” (5.2). Clause 12.6 lays out what seems like a sensible position: “The citizen shall have the right to obtain personal data held about them in a structured, commonly used, and machine-readable format, and to have such data transmitted directly between public bodies where technically feasible and lawful.” In 12.6, the direct transmission of data between public bodies is under the control of the citizen. That is not reflected in Section 10, nor are the other clauses under Section 12.

Recommendations

To remedy this, a mechanism is needed to allow citizens to be in control of any sharing of personal data about them in the data sharing exchange. Citizens should be empowered to have control over (a) what data is and is not shared; (b) when it was accessed by any governmental entity; (c) why it was accessed (not just the lawful basis, but also the precise purpose); (d) whether it will be used as personal data (with the individual being identifiable) or only in aggregated form, for statistical purposes; (e) what rights and safeguards accompany the usage (such as data subject rights, data retention limits, etc.)

Given that citizens rely on the government to provide them services, it is not the *consent* that is most empowering, but the *transparency* we are recommending for inter-sharing of personal data that is important for citizens to know that their data are not being misused. Thus Section 10 should explicitly be subordinated to Section 12.

7. Section 11: Personal data and privacy

This Section is welcome, but it should be moderated keeping in mind state capacity (e.g. to carry out privacy impact assessments) as well as Section 12 (citizen data rights).

Recommendations

“Performance of a public task” (11.2) should only be used as a lawful basis for processing when access to personal data is necessary and the amount of personal data accessed is proportionate.

8. Section 12: Citizen data rights and empowerment

This section is commendable and is necessary to ensure that citizens aren't seen as "passive subjects of data governance". It is important to ensure that Section 12 and Section 10 do not conflict with one another.

Recommendations

Section 12 should explicitly have priority over Section 10.

9. Section 13: Cross-border data transfer

There is significant overlap between Section 8 ("data sovereignty and residency") and Section 13. All the feedback we have provided to Section 8 are also applicable to Section 13. A prohibition on cross-border data transfer should be an exception and not the default. While sensitive governmental data should be subject to data localization requirements, this inverts that and makes data localization the default.

Recommendations

For the reasons we've laid out under section 8, we believe it is not prudent to retain Section 13 in its current form.

10. Section 14: AI and emerging technology

The principles laid out in Section 14 are good, but are difficult to put into practice. For instance, given state capacity in most emerging economies including Pakistan, it is may be ambitious to require risk-tiered controls for automated decision-making. Given the current state of interpretability of machine learning systems, it is not possible to demand "explainability commensurate with use" as a requirement.

Furthermore, this section does not cover AI decisions which aren't about "legal or similarly significant effects", but say, effects on health care provision.

Lastly, while this Section focusses on AI and ADM use by the government, it could also speak to how excessive copyright protection and personal data protection can harm the very emergence of AI and other such technologies. This can be avoided by allowing for a broad text and data mining exception in our Copyright Act, and by reducing the scope of personal data protection to those instances of data processing where personal data is not processed incidentally from publicly-accessible data sources, thus excluding activities such as web scraping for the purpose of creating a search engine or training of a large language model, while still covering web scraping for building a database of people.

Recommendation

Given the potential consequences of automated decision-making, including over-reliance of humans on inscrutable AI even when there is human oversight, we recommend the policy take a position of caution. While the adoption of AI to increase governmental efficiency and productivity is inevitable, any incorporation of automated decision-making should be done in small pilots, rigorously studied, and effects monitored. To the greatest extent possible, citizens should not be subject to automated decisions which officials cannot effectively explain and which citizens cannot effectively contest.

We also recommend providing a broad text and data mining exception in copyright law, and restricting incidental processing of personal data from publicly-accessible data sources.

11. Section 15: Data value and the data economy

The policy focuses almost entirely on government or public sector. However, in practice even the government data is routinely processed by private sector institutions such as banks, telecom operators, cloud providers, hospitals, universities, research institutions, digital payment companies, fintech and health-tech. Private sector access to data (in legal, privacy and national security preserving ways) will help catalyze innovative business and citizen services that are currently unavailable and even unimaginable.

Recommendation

The policy could provide a clearer governance framework for trusted private-sector participation beyond the obvious contractor obligations.

12. Section 16: Institutional framework

Pakistan Digital Authority (PDA) Powers are Extremely Broad: PDA is assigned extensive policymaking, regulatory, operational and supervisory responsibilities. While central coordination is desirable, good governance requires institutional accountability, independent review/oversight, transparent decision-making, and effective appeal mechanisms. The PDA has wide-ranging regulatory and enforcement powers, including issuing standards and binding directions, approving cross-border data transfers and offshore data processing, designating primary registers, conducting audits, requiring compliance, issuing certifications, maintaining registries, and enforcing corrective actions. Though the Policy talks about many functions but nearly every national function resides in PDA. This may overwhelm one institution.

Recommendations

The Policy should therefore strengthen governance safeguards without weakening national coordination. This can be done through independent appeals mechanism, judicial review, independent review panel, annual Parliamentary report, public consultation, external audit and / or Provincial Advisory Committee. Additionally, It is recommended to delegate some operational responsibilities to sector regulators and major ministries.

13. Section 17: Implementation

Many operational matters are deferred to future supporting instruments. The policy frequently states, "detailed requirements shall be prescribed...". Nearly every operational requirement is deferred to future standards, this provides flexibility but also creates uncertainty. It is feared that PDA could substantially change policy without wider consultation.

Implementation of the Policy certainly requires considerable resources but there is no estimate of cost or identified source of funding. Implementation requires significant resources and investment in institutional capability, workforce development, legacy system modernization, metadata management, cybersecurity, interoperability, organizational change management, data stewards, agency Chief Data Officers (CDOs), metadata catalogues, API development, master data management, security upgrades, identity federation, privacy systems, citizen portals, and AI governance.

The policy assumes rapid implementation, which may not be realistic. Pakistan has hundreds of agencies, thousands of legacy databases, national archives, paper records, non-standard data, duplicate citizen records and manual workflows. Implementation with such legacy data could take 5 to 10 years. The success of the Policy will ultimately depend not on the quality of its drafting but on the quality of its implementation. International experience demonstrates that data governance reforms are among the most complex public-sector transformation programmes. They require legal reform, institutional coordination, organisational change, digital infrastructure, workforce development, cybersecurity, sustained political commitment and long-term investment.

Recommendations

It is recommended that Policy should contain mandatory principles into the policy itself to create certainty and provide guidance to PDA. While subordinate instruments should focus on technical implementation.

The Policy would benefit from recognizing the implementation realities more explicitly. We recognize it's not the realm of a Policy to include detailed budgeting consideration (e.g. if the funding sources might include federal PSDP, development partner financing, public-private partnerships, Digital Pakistan Fund; whether innovation grants and provincial programmes will be enabled) and detailed operational timelines e.g. if the implementation will be phased in any way). But referring to how such details will be decided upon might help frame next steps.

We once again thank you for the opportunity to provide our comments on the draft Data Governance Policy 2026. We remain at your disposal to answer questions at any time. We hope that our observations and recommendations will be useful in further strengthening the policy. We remain at your disposal should you wish to discuss any of the comments or recommendations in this submission. Please contact us at info@lirneasia.net.

Thank you.

Yours sincerely,

Aslam Hayat
Senior Policy Fellow

Pranesh Prakash
Senior Research Fellow