aapti institute

COMMENTS TO THE COMMITTEE ON NON PERSONAL DATA
SUBMITTED BY: AAPTI INSTITUTE

Overview

The Report by the Committee of Experts on a Non Personal Data Governance Framework is part of a larger trend across the globe on efforts to understand the nature of the need to regulate data, apart from the basic protections to individuals over their personal data. For instance in the EU, the free flow of non-personal data describes an effort to build a Digital Single Market to unlock the value of data for member states.[1]

In order to rationalize markets and policy-making around data, the report seeks to establish data as an economic resource that should generate value for all stakeholders. To that end, the premise of the NPD report, that data must not be monopolized by a few entities and must be deployed to benefit society is correct. However, the report fails to fully acknowledge that data has unique characteristics, which makes it difficult to determine its principles. The existing model exists around "ownership" of data is insufficient and potentially problematic and yet, our laws and policies tend to reflect this model of ownership.

The fundamental challenge with NPD is that specifying a beneficiary or clearly identifiable set of interests to protect is a complex task. Being non-personal, NPD contains within it the interest of the community, and not only of an individual. A hurdle in framing good regulatory approach lies in identifying the communities that may be impacted and how regulation can protect their interests – especially since these communities can be overlapping, and their interests in conflict with one another. A clear identification of interests as a problem statement is essential to framing regulation to protect it, and the report fails to do so. Community rights and their protection still lack clear articulation, and preventing collective harm requires further definition and precision.[2]

Any policy tools for exercising the rights and will of the community pertaining to data would need to take into account existing infrastructures to create compatible governance structures. The proposed solution in the report for instance is a form of data stewards (called data trustees), to try to address this issue. The formulation leaves open questions of accountability and grievance redressal mechanisms.

The policy measures need to move away from being coercive and disruptive. We need greater ideological grounding in our approach to regulating data – a clear articulation of how we perceive data in specific contexts and our larger objectives that govern our regulatory approach.

---

[1] A Digital Single Market Strategy for Europe, European Commission, 2015, accessible at https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52015DC0192&from=EN.
[2] Siddharth Manohar, "Data Sharing for Public Good: Theoretical Bases and Policy Tools", Aapti Institute, 2020, accessible at: https://thedataeconomylab.com/2020/07/31/the-basis-for-data-sharing/.

The NPD policy seems to be premature – it is vague, and possibly a worrying articulation of the state of play in top-down data governance measures. Therefore our comments below have been framed in order to highlight the gaps in the policy document. The committee must organise broader consultations to ensure that the objective of unlocking data in public interest and through collective consent does not end up creating structures that exacerbate the problems of the data economy and are susceptible to regulatory capture.[3] This is especially essential since the discussion on NPD in India is happening while we still await a personal data protection law.

1. Public interest

The idea of "public interest" is at the central to the NPD report. One of the main purposes of data sharing is public interest; however, this also provides space for the government to acquire data for public interest, and for data trustees and custodians to similarly request data. One of the problems with this approach is that public interest has been left broadly defined, and needs to be more context and case specific to avoid abuse of the power to acquire data. The balance between state power, private investment in data collection and processing, and "public interest" is unclear.

One such approach to to usage of resources for public interest is that of the commons, a public resource that is freely accessible to all members of the community. In Elinor Ostrom's seminal work, '*Governing the Commons*', the term 'common-pool resource' is said to refer to "a natural or man-made resource system that is sufficiently large as to make it costly (but not impossible) to exclude potential beneficiaries from obtaining benefits from its use".[4] The premise of a Data Commons is that data is sourced from and is relevant to certain specific communities from which it arises. Data on people living in a particular region for instance, would be most relevant to those individuals and people and issues connected to the same set of people; data generated by a company through its services, would be relevant to the company as well as its customers from whom the data has been sourced. Data in this framework is seen as part of a "digital intelligence system" that has tangible social and economic impact through its role in organizing the distribution and access to resources.[5] Fundamentally, this approach recognizes the value of data beyond the individual, and focuses of the collective – the data of an individual is only valuable when combined and analysed along with the data points of a millions of others.

2. Consent in Non Personal Data

Another issue that on concern is that of consent in processing of non-personal data. The policy specifies that consent should be collected for processing

[3] "Wanted: Data Stewards: (Re-)Defining the roles and responsibilities of Data Stewards for an age of data collaboration", GovLab, 2020. http://www.thegovlab.org/static/files/publications/wanted-data-stewards.pdf.
[4] Elinor Ostrom, "Governing the Commons: Evolution of Institutions for Collective Action", 1990.
[5] Parminder Jeet Singh, "Data and Digital Intelligence Commons", Data Governance Network, 2020, accessible at http://datagovernance.org/files/research/ITFC_Parminder_Data_Commons_-_Paper_2.pdf.

anonymised data at the point of collection of personal data, therefore providing consent for processing non-personal data that may pertain to a Data Principal. How this operates along with the framework of Data Trustees is an interesting question – it is not evident that Data Trustees may be able to revoke consent on behalf of their representative community. The policy does explain the concept of a "policy switch" that allows Trustees a platform to set access controls, but no mention of revocation of permission/data has been addressed in the document. How an individual or community would exercise consent through a Trustee when data already exists within the control of a data fiduciary/custodian is not clear. Further, the assumption of informed consent for anonymization does not take into account the ways in which individual consent is broken[6].

3. Community Data Rights

Communities must be allowed to play a prominent role in articulating their interests and preferences. They may come up with incentives that match up to their priorities. Nominally, users have been understood to be data principals, similarly as under the Personal Data Protection Bill, and groups of users with common interests have been understood to be 'communities' which have their own joint common interest in data that pertains to the specific group. These communities are expected to articulate their interests and preferences over their data through 'Data Trustees' – and these Data Trustees are a representative body meant to represent this collective interest – however it is not entirely clear how it is ensured that the community interest is in fact expressed in the data trustee's submissions, and how the trustee is protected from regulatory capture. What ensures that a Data Trustee does not go rogue or serve interests other than those of the Data Principal?

The Data Custodian is more straightforward in its scope, tasked with the actual implementation of the interests of the Data Principal, itself being a data fiduciary and therefore undertaking data processing and sharing, once again in the best interests of the Data Principal. The distinguishing factor here however is the presence of a legal duty of care – being a fiduciary, the Data Custodian can be held liable for failure to carry out its duty to protect the Data Principals' interests.

The idea of community-driven efforts for data governance are explored in the works of Sylvie Delacroix and Neil Lawrence, who discuss bottom-up data trusts as a way for people to take the reins of their data, and move away from the current top-down regulatory frameworks that are "one-size fits-all" and do not account for the variation in the abilities and desires of people to participate in their own data governance.[7]

---

[6] Jeni Tennyson, "Community Consent", accessible at http://www.jenitennison.com/2020/01/17/community-consent.html.

[7] Sylvie Delacroix and Neil Lawrence, "Bottom-up data Trusts: disturbing the 'one size fits all' approach to data governance", International Data Privacy Law, 2019, accessible at https://academic.oup.com/idpl/article/9/4/236/5579842.

For bottom-up data trusts, there are challenges to implementation such as uptake by people, usability by governments and private sector, and governance and technological architectures, and these need to be designed and tested multiple times before implementation.

Collective harm from data has not been sufficiently addressed in the policy. There is no option available to communities to escalate violations of collective privacy.[8] The principle is acknowledged in the policy but not sufficiently reflected in the recommendations – questions of how to enforce collective rights to privacy lead back to accountability structures that are in place for the protection of user groups. Communities need to be able to engage with other stakeholders in the data economy to enforce their collective claims over how their data is used by third parties. The lack of decentralised accountability structures that leads to this situation is further explored below.

4. Accountability

There is room for the extension of competition law, IP regime, etc. in existing frameworks, and room in the PDP bill to incorporate further notions of group privacy. Accountability must be embedded at the level of the data processes. Us-based tests specific to sectors and particular levels of risk may be applied. Accountability and agency must work hand in hand. Accountability in the system is what enables agency for the users who interact with and are subject to it. Fiduciary responsibility plays a role here, to set up institutional processes to protect their rights and maximize user interests through participation. The aim of accountability is to overcome the failures of the notice and consent system where users are passive and choose from a limited set of limited advantages, and instead state preferences outright, and are provided a platform to do so.

The framework created by the policy is backed up with directions to prescribe incentives, remuneration, and rules for the operation of the Data Trusts, Trustees, Communities and Custodians. What is less clear however, is how Data Trusts will ensure that the purpose for which they are set up will actually be served. Another possible complication arises when a Data Principal community may disagree with the measures of the Data Trustee, or when there is a lack of consensus within the identified Data Principal community. The parameters for identification of the community itself contain some ambiguity within the policy. It is not clear how a conflict between the Data Principal Community and the Data Trustee would be resolved, or how the Data Custodian remains accountable to the Data Trustee. Without such a system, the decision making process is open to abuse and external influence.[9]

---

[8] Martin Tisne, "The Data Delusion: Protecting Individual Data is Not Enough When the Harm is Collective", Stanford Cyber Policy Center, 2020, accessible at https://fsi-live.s3.us-west-1.amazonaws.com/s3fs-public/the_data_delusion_formatted-v3.pdf.

[9] Mark Bunting and Suzannah Lansdell, "Designing decision making processes for data trusts: lessons from three pilots", Involve UK, 2019, accessible at http://theodi.org/wp-content/uploads/2019/04/General-decision-making-report-Apr-19.pdf.

The framework of Principal-Trustee-Custodian requires the presence of agile actors as part of the system to ensure smooth operation. The mechanism of data sharing itself also leaves further clarity to be desired – data requests may be made to Data Businesses (i.e. companies that work with large proprietary datasets), or alternately made to a Data Custodian. An important complication here is that the Data Custodian is also empowered to offer value-added services based on data at their disposal – this presents a direct conflict of interest if Data Custodians are in effect competing with Data Businesses. The report, while mentioning that there should be incentives and remunerations for Data Custodians, does not explicitly prohibit a for-profit operation, and nor does it indicate a mechanism to prevent third-party influence.

The also report presents the possibility of the Data Trustee setting up a Data Trust to enable easier management of the data to be shared – as per the report, the Data Trust is described merely as an institutional 'data infrastructure' that ensures the control of the Data Trustee of data usage. Although it has been stated that a Data Trustee would deploy Data Trusts for 'public' use of non-personal data, the conditions where Data Trusts may or may not be used have not been completely addressed in the policy. The Report must go into the implications of the use of a data trust and what this entails. A definition of Data Trust must include a comprehensive discussion of its functions, its powers, where it sits in the data value chain, and how it is governed, and who exercises influence over the activities of the Data Trust.[10]

5. Data as an asset

Personal data as an asset has an existing body of work outlining the interests and tensions at play.[11] On a similar note, the Report mentions the movement towards coneptualising non-personal data as an asset, but does not cite an attempt at doing so, beyond an acknowledgement of their value in the acquisition and operation of companies. It also does not engage in depth with how data can be viewed beyond merely financial (or other) value – the choice to rely solely on valuation through specific operations of a certain class of companies remains a premise that demands explanation. There are a number of arguments for data as body[12], or exhaust[13] that the report does not engage with. There is no reasoning provided as to why the asset-conception of data is the only one applicable to a policy for governing the sharing and usage of all non-personal data. Engagement on this issue would provide better direction on what specific rules are required for various stakeholders to better align incentives, maximizing user interest and value from data, and minimizing community harm and conflicts of interest.

---

[10] Siddharth Manohar, "Trust Law, Fiduciaries, and Data Trusts", Aapti Institute, 2020 (upcoming, October 2020).
[11] "Personal Data: The Emergence of a New Asset Class", World Economic Forum 2011, accessible at http://www3.weforum.org/docs/WEF_ITTC_PersonalDataNewAsset_Report_2011.pdf.
[12] Anja Kovacs, Nayantara R., "Data sovereignty, of whom? Limits and suitability of sovereignty frameworks for data in India", Internet Democracy Project, 2020, accessible at https://internetdemocracy.in/reports/data-sovereignty-of-whom/.
[13] Martin Tisne, "Data isn't the new oil, it's the new CO2", Luminate, 2019, accessible at https://luminategroup.com/posts/blog/data-isnt-the-new-oil-its-the-new-co2.

6. Mandatory data sharing and state powers

The NPD Policy, in the footsteps of Section 91 of the Personal Data Protection Bill (PDP Bill), prescribes mandatory requests for data by state authorities. Given the immensely broad non-exhaustive definition of Non Personal Data under the PDP Bill as well the policy, categorizing data as non-personal cannot provide a basis for lesser protections on its use and collection. Like personal data, non-personal data also has harms attached to its use and rights enforceable against its use in a manner that adversely impacts citizens.

Protection against these harms require recognition of the interests of the state as separate from those of the users – the policy seems to conflate the interests of the state as necessarily in line with user interests, which is not necessarily the case. This is also reflected in the nature of the examples given whilst explaining Data Trustees, who represent interests of user groups – they have been described as a role possibly played by government departments. This is an arbitrary inclusion of powers within state functions without application of administrative principles that would require a separate body to possess the power to acquire and manage data. Exercise of such powers by the Non Personal Data Authority requires rules and specific cases where the exercise of such power is justified with proportionate cause and a clear legal basis for doing so. Such rules are required irrespective of whichever methods are used for enabling data sharing, including Data Trusts, cooperatives, etc.[14]

7. Creating trust between state and citizens

This means robust legislation and policies that offer social protection and comply with constitutional principles and judicial pronouncements in incursions upon data rights, while simultaneously ensuring that there are adequate public dispute mechanisms that allow the enforcement of these rights. Making rights central to the conversation on health crisis management in particular, will ensure that all stakeholders – the government, private sector, communities, and individuals – remain accountable to each other.[15]

It is of utmost importance to create mechanisms and processes that build trust between the public and the government. This trust can be inculcated through greater transparency in communication, improved access, ensuring that systems are voluntary, such as the downloading of contact tracing apps. Further, social solidarities are a critical tool for working effectively together to address the pandemic and can reduce the need for harsh policies such as curfews and surveillance. This can be done through communication of common objectives, targeting behaviours and not people, and creating effective systems of financial support. Solidarities create a space for workers

[14] Sean McDonald, "The Fiduciary Supply Chain", Centre for International Governance Innovation, 2019, accessible at https://www.cigionline.org/articles/fiduciary-supply-chain/.
[15] Richard S. Whitt, "Old School Goes Online: Exploring fiduciary obligations of loyalty and care in the digital platforms era", Santa Clara High Technology Law Journal, 2020, accessible at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3427479.

and migrants to mobilise and negotiate change in their relationships with the platform economy.[16]

[16] Elinor Ostrom, Christina Chang, Mark Pennington, and Vlad Tarko, "The Future of the Commons - Beyond Market Failure and Government Regulation", Institute of Economic Affairs, 2012, accessible at https://papers.ssrn.com/sol3/papers.cfm?abstract_id=2267381.

aapti institute

Annexe

Aapti Institute is a public research institution that works on problems at the intersection of technology and society. We hope that our work informs the development of technology, programs, and policy. Aapti works in three areas: 1) Futures of workers, where we examine labour relations in the platform economy; 2) Governance and citizenship, where we look at technology deployed by the state, and citizens' use of tech to reach the state; and 3) the Data Economy, where we take a rights-centric perspective to understand individuals' and communities' relationships with data. Aapti has extensively worked on the idea of data stewardship – unlocking value of data while safeguarding rights. Aapti has followed a problem-led, use-case driven approach to data stewardship, exploring models and their applicability in different sectors. Aapti is also exploring the value of stewardship in rebalancing power in the data economy and the ways in which models can help communities negotiate better on their data rights. Aapti's body of work on stewardship can be found in the data economy lab.