The Dialogue
INFORM ENGAGE IDEATE

Virtual Stakeholder Consultation Non-Personal Data Governance Framework

31st July, 2020

1. Introduction

The Dialogue concluded the first in a series of Virtual Consultations on Non-Personal Data Governance on 31 July 2020. It had close to 100 participants including representatives from leading think tanks and civil society organizations, law firms, businesses, industry bodies and the research community.

2. Key Takeaways

- Definitions of the key roles must be articulated with clarity to demarcate how a data trustee would be different from custodian. A future regulation must flesh out how the conflicts of interest would be resolved and the oversight mechanisms must be established
- Inclusion of "Private NPD" within the scope of mandatory data sharing could be disruptive to existing business models and weaken IP protections accorded. This would create second order issues including reduced investment and competition.
- Creating data marketplaces where data is valued and sold as a commodity might further create market imbalances stemming from purchasing power of the organisations with more resources at their disposal
- Ownership, the corresponding rights, and duties must be established to ensure accountability of authorities within the framework.
- Government is one of the largest custodians of data. To address issues of information asymmetry and alleviate fears of surveillance, a surveillance law reform is essential

- Considering the widespread overlaps with the PDP framework and other frameworks of law such as the competition and intellectual property, it would be effective to mandate MoUs between regulators and pre-empt the clashes.
- A relaxed regulatory regime that allows data sharing on a voluntary basis could be envisaged as jurisprudence on data regulation is still nascent.
- Accountability of regulators must be to the Parliament and not to particular ministries, to reduce conflict of interest and to ensure independence.

3. Overview

There were five panels that were thematically arranged for deeper discussions on the challenges within the report. The first session was on **"Definition of NPD and key roles in the NPD ecosystem"**. This session went into the definitions envisaged in the report. The panelists also dove into issues pertaining to the key roles defined in the NPD ecosystem, particularly on the role of trustees and custodians as key decision-makers in the ecosystem. The second session was on **"Rights over data and data ownership"**. The panelists discussed nuances of community rights over data, the role of a data trustee in the framework of rights, the conceptualization of community rights, basis of legal rights over data amongst other things. Panelists discussed the possible conflicts with existing IP laws and how future regulation must deal with such issues.

The third session was on **"Data sharing and data business"**. The panelists discussed the proposed data sharing models, the purposes of data sharing as envisaged, and mechanisms of data sharing. The panel also discussed the potential impact on investments and the possible disruption it will bring to the market if the report is applied in the form it is today. The fourth session was centered around **"Interaction with the PDP framework and surveillance concerns"**.

The panelists opined that with the advent of the report, there is a greater need to discuss the modes of government access to data and the checks and balances that need to be established. In the absence of surveillance law, unfettered access to government bodies for access to data can lead to legitimate surveillance concerns. The session also had discussions based on the privacy concerns that arise in the regulation of NPD derived from personal data, market effect of unfettered government access to data.

The fifth session was on **"Regulatory Challenges and Harmonization of laws"**. This session discussed harmonization of the powers of the regulators such as DPA, NPDA, and CCI. It also had some thought-provoking conversations surrounding the regulatory challenges that could crop up

with the establishment of the NPDA. The next section of this report will highlight the key discussion points and the takeaways from the event.

4. Key Discussion Points

4.1 Definitions and Key roles in the NPD ecosystem

The data subject, in certain cases could be a community. Therefore, the definition of a community was important to establish the key roles within the community that would allow for proper management of the community NPD and would establish corresponding rights and duties. The definition of "community NPD" that has been put forward in the framework is a fairly amorphous concept. A literal interpretation of this definition may lead to several competing claims over the same data. The key roles that were identified were those of the data principal, data custodian and the data trustee. An important difference that needs to be established is between a data custodian and a data processor. Some of these companies only store and process data on behalf of their clients. More often than not, confidentiality agreements are entered into to ensure that the processors do not have any viewing access. Additionally, the role of the government within these key roles was not appropriately identified. What would happen if the data custodian has a conflict of interest with the government, or what if the government is both the custodian as well as the trustee?

Each trust is established for a particular purpose where there is a beneficiary group/person/entity. In this case, the purposes for creation of data trusts are broad and ambiguous, while the beneficiary can not be clearly mapped out. As a result of such a formulation, it follows that enforcement of rights and obligations from these structures will be hard to enforce.

Questions abound on grievance redressal, oversight mechanisms etc. Right from the definition stage, it is important to ensure the scope is narrow to prevent issues pertaining to misinterpretation. Definitions, in general, play an important role in avoiding ambiguity and determining the enforceability of the concerned law. In addition to this, definitions must also be reflective of the inclusive nature of the policy. For example, with regards to this framework, India's poor digital literacy is of immense importance.

There are a disproportionate number of women that engage with the digital space. Without the definitions exhibiting inclusivity, the discussion and debate will take place only within a specific sphere of the community.

4.2 Rights over data and Data Ownership

Every community from which data is taken from was to be considered as data commons, where the communities can exercise certain rights and privileges over. However, it was opined that articulation of the contours of the rights that communities will accrue will take a long time to take shape. From a legal standpoint, unless there is absolute clarity over data ownership, it will be extremely difficult to allocate any rights and/or duties. Given the fact that community is an amorphous entity, it is harder to ascertain specific boundaries. For example, a community under the framework, does not take into account the various subsets under each community; many of which overlap with one another. Additionally, it's important to determine the criteria based on which such definitions are employed. In case of a geographical criteria, the consequence of someone moving from one place to another has not been iterated.

Another aspect of data ownership is its conflict with intellectual property rights. As it stands today, data falls under "literary work" under the Copyright Act, 1957. There is a lot of time, work and investment that goes into collecting datasets but more importantly, in making said datasets usable. Datasets tend to satisfy the tests of 'modicum of creativity' and the sweat of the brow' theories. With the key roles in place, it is unclear if there will be a clash with ownership of intellectual property rights over datasets. Having said this, it is possible that there will be considerable litigation when it comes to claims over data, especially when mandatory sharing is envisaged. It is prudent to revisit clauses that specify sharing of raw data as they might be protected under copyright law if they pass the test of "modicum of creativity".

Moreover, if data sharing is made mandatory, including private proprietary data- India could be violating its obligations under international agreements and treaties such as TRIPS. While the report mentions the possibility of compensation that could be provided to the entity sharing such data, establishing a data market is a scary precedent. If at all, a well regulated market can be created for this purpose, it is left to be seen how valuation of data would be done. It was pointed out that the value of datasets for businesses is extremely contextual.

4.3 Mandatory Data sharing by Data businesses

Including data businesses within the framework was important because, as was argued throughout the discussion, data is currently stored in limited vertical silos. It was pointed out that data serves as an infrastructural element in the development of a digital economy and the sharing must facilitate growth. Data is to the digital economy, what banks and roads were to the industrial economy. However, there were several ambiguities that were pointed out by the panellists, such as the relation between data businesses and the regulators.

The speakers agreed that data is an integral part of doing business. There are very few powerful players in the market which exercise almost all control over the availability of datasets. The Committee has attempted to create an enabling framework which will ensure that data is more readily available for smaller businesses as well. One of the proposals was that of creating a market that is governed by the forces of demand and supply. However, a key issue with the same is that the market would be largely monopolistic. It was opined that if data was made accessible to all, companies will have to invent new ways to establish their competitive advantage over the others in the market. There was consensus among the speakers regarding the importance of a competitive market, not only to encourage innovation but to also ensure accessibility and affordability.

Mandated data sharing creates extremely high compliance costs for those categorised as data businesses. However, the framework doesn't provide for a principled or guided approach towards the same. There seemed to be unanimous support for an 'incentive-based' approach. Additionally, the monetisation of data was a concept that was explored throughout the discussion. If sellers and buyers can come together in order to engage in price discovery, without any violation of privacy, a market based model could be effective.

It would be wise to caution against any kind of horizontal legal obligations or rules that are overly prescriptive in nature, especially those regulating private enterprises. For one, this would create significant entry barriers. Secondly, different sectors and different companies within sectors, have singular needs. The manner of engagement with data will differ on a case-to-case basis. Horizontal application of law bears the risk of over-regulation and creating easily accessible loopholes that may lead to misuse of data.

It is also important to identify the ends the regulation seeks to meet. On one hand, it aims to promote economic activity, which could get affected due to an uncertain regulatory landscape. On the other, it aims to promote social welfare and ends by contributing to "public good". Even if this could be justified, the panelists opined that the "public good" must be defined carefully to prevent misuse.

Moreover, in cases where mandatory data sharing is allowed, it must be noted that some level of agency is provided to the businesses as well, as opposed to a regulator having the final word.

4.4 Surveillance concerns and PDP Bill

It was pointed out that increasingly, every technology by itself is becoming a surveillance technology. Most products or services that we avail rapidly and constantly generate data that forms the basis for future development. In such a context, the policy approach and design must set up a mechanism that can guide the market, and prevent the citizens from unwarranted harm. However,

the report seems to raise alarm bells with regard to unwarranted surveillance. The need for a surveillance law can not be understated lest we move towards a police state like China.

There is a fundamental shift from the PDP bill which starts with the protection of individual rights and is mostly a protection framework. It was observed that the NPD report hardly engages with the individual or their rights. It makes a swift dive into government access to data and industry access to data. This is premised on the assumption of marginal good to communities and the public as a whole. It was pointed out that the intelligence agencies have not been created by an Act of Parliament, thus weak in any oversight mechanism. This creates considerable fears and apprehension on the usage of data, purpose of such use etc. It was also pointed out that the enforcement of law and development of law in this regard is extremely minimalistic and is far away from global standards and best practices. During the session, a panelist pointed out the fallacy of dissociating NPD and surveillance. Under certain intelligence operations around the world, they start with deanonymising data by combining various data points. The resultant data, which could be extremely sensitive and personal, can then be used to target individuals.

It was pointed out that the Government was one of the largest custodians of data. Due to the wide corpus it has, in its possession, there is information asymmetry between the state and its subjects. Thus creating fears on the resulting power dynamics and trust deficit. Similar arguments were brought to fore during the discussions surrounding the PDP bill, particularly on section 35 of the bill that grants widespread exemptions for the Government. It was opined that, in both cases effective checks and balances on the government control over individual and private lives is a crucial facet enforcing citizen trust in the system. The report leaves a blank slate on this front, and the need for a surveillance law is more imminent than ever.

4.5 Regulatory Challenges and harmonisation

The panelists pointed out that very often the mandates and set up of many regulators are unclear leading to more problems than solutions. In this context, for instance, there is the need for democratic processes and accountability mechanisms to be implemented, even by statutory entities and independent regulators. There needs to be considerable thought put into the design and processes of regulation and it's not sufficient to establish a toothless yet “independent” regulator. The report provides insufficient detail about the regulatory setup and accountability grievance mechanisms that need to be articulated

It was agreed that since regulation of sharing non-personal data was a relatively novel concept, it would be wiser to enforce a far more relaxed regime that allows data sharing on a voluntary basis where the data requests are analysed and are then accepted or rejected by the Non Personal Data Protection Authority (NPDA). The companies must be given a choice to refuse the request for data sharing. Once the implications of such regulation can be effectively analysed, a more well-rounded policy can be introduced. Frameworks and standards that bridges the trust deficit are crucial for the future of data governance. An impartial regulator with all stakeholders can go to for their grievances is very important to achieve this end.

The legal regimes created by the NPDA regulating non-personal data and the Data Protection Authority (DPA) regulating personal data are conflicting in nature. It is impossible to create a binary distinction between two overlapping concepts such as these. Therefore, enforcement is going to be fairly problematic. Additionally, such a regulatory body must be formed in the most transparent manner possible in order to minimise misuse of power. For example, start-ups and small businesses, the intended beneficiaries of this framework might end up facing the overlapping requirements of compliance that could affect the growth much more than the larger organisations in the ecosystem.

It was also pointed out that creating a regulator involves considerable effort and care must be taken to prevent regulatory arbitrage. It was opined that the reasons for creating a regulator must be narrow, and there was a need for harmonising the internal laws. It was pointed out that the framework tries to cover three regulatory regimes broadly- competition, data governance, and IP. So while deciding on how the powers will be divided, and harmonised each of those must be tested against the problem it tries to solve. For example, if it is a question of breaking up monopolies, the answer lies in Competition law.

If the issue is about incentivizing good data sharing practices or sharing proprietary knowledge and that's a conversation in the IP regime, and so on. It was pointed out that there was barely any mention of individual autonomy throughout the report. Additionally, any redressal mechanism in place must take into consideration the low rates of digital literacy in India. The lack of inclusivity within this sphere will only end up increasing the digital divide in the country.