VOLUME 16 ISSN 2456-9704

REGULATION OF NON-PERSONAL DATA IN INDIA

By Apoorva Singh
From Symbiosis Law School, Pune

The Personal Data Protection Bill, when enacted, seeks to replace section 43A of the IT Act, 2000 which is currently responsible for data protection in the nation. It was introduced to the Lok Sabha on December 11, 2019 by the MeitY. As the name suggests the bill is application to the processing of personal data, this personal data has been further classified into sensitive and critical personal data. The term "Data Principal" in the bill has been defined as any natural person who alone or in tandem with another determines the reason and way to process personal data.[1] Non-Personal Data is any data that does not come under the definition of Personal Data i.e. any data that is not personal.[2] Users of online businesses often leave behind digital data trail, the PDP bill seeks to protect any personal information that gets left behind in that trail, however, information such as weather trends collected by apps or data on which locations a site is getting traffic from cannot fall into the ambit of personal data as it is anonymized, hence to regulate this will be more of a challenge than personal data. In this article, I will be dealing with the developments of non-personal data in India and its interplay with the PDP bill.

1. Developments before the Draft Personal Data Protection Bill, 2018

1.1 TRAI Consultation Paper

The Telecom Regulatory Authority of India on 9th August 2017 released a consultation paper titles "Privacy, Security, and Ownership of the Data in the Telecom Sector"[3]. This paper discussed the economic value of data. The paper by TRAI described the protection of data in the form of legal control over access and usage of data that has been stored in a digital format, it also includes the ability of a person to understand and control the way in which information that pertains to them can be used and accessed by other individuals.

TRAI stated that ownership over ones data was important and if it could be ascertained who held ownership of a given data and the ownership of this data could be of some benefit to the owner's life. This would make data an empowering commodity.

The TRAI paper suggested that the government should enable the data ownership sector grow by way of creating new services. They emphasized on 2 aspects, firstly, data portability, which is the ability to take user data from one service and share it with another, and secondly, the creation of anonymized public datasets also referred to as a data sandbox, this is a mechanism by which entities can contribute data that has been anonymized that could help in the development of new products. This public data set could be used as a test bed by new service providers.

1.2 NITI Aayog Discussion Paper

In June 2018, a discussion paper was released by the NITI Aayog named "National strategy

[1] Section 3(14), Personal Data Protection Bill, 2019
[2] Explanation to Section 91(2), Personal Data Protection Bill, 2019
[3] Consultation No:09/2017, Telecom Regulatory Authority of India

for Artificial Intelligence"[4]. In this paper it was discussed that to incentivize more players to supply AI training data sets and services, it was important to have data available. There should also be an audit mechanism to curtail the reselling of the same data, there should also be means to address privacy and security concerns. A solution given in the discussion paper was that all the data could be centralized and hosted by a trusted party on behalf of the data providers. Obvious challenges arose with this idea, the largest being that data is a replicable resource. The technology was recommended as to form a decentralized data marketplace. The process of data exchange would come with provisions of anonymization, so as to remove personal information to give the data a market determined worth and move towards a formal data market economy.

In this paper however, it was suggested that the concentration of this data in the hands of a few persons would create an entry barrier for entrepreneurs, hence suggested the market to share data. It was also stated that data must be shared for the purpose of good governance and creating public policies. It was also suggested that companies might be required to share data that they have aggregated for the good of the public.

2. Developments after the Draft Personal Data Protection Bill, 2018

The PDP bill has not been able to take Anonymized data in its ambit, even though it has not been expressly provided, the anonymization of large datasets is not an impossible task under this legislation. The Justice Srikrishna Committee Report discussed the degree of anonymity of anonymous data, data in modern times does not only exist in identifiable or un-identifiable states, identifiability can depend on many different factors as persons in processing the data can have additional data that can aid them in recognizing an individual based on data that was supposed to be anonymous. This will only increase as technology advances. Anonymization takes place when personal identifiers are removed so that the data principal cannot be identified. Several countries such as South Africa and EU do not include anonymized data into its ambit, the EU even has an entire system of pseudonyms where identifiers get supplanted by pseudonyms.

The Sarkrishna Committee recommended to the government to provide standards for the anonymization and de-identification of data, all data that has gone through the process of de-identification will come under the purview of law meanwhile anonymized data that meets the given standard should be exempt from the law.[5]

Section 105 of the Draft Bill, 2018 gave government the powers to form policies for a digital economy, this included methods to grow, secure, and, define non-personal data give a scope to its utilization.

3. Developments under the Personal Data Protection Act, 2019

Under the 2019 bill, Personal data can be defined as any data that can identify a natural person in any manner, it can be their characteristics, traits, or any feature both online and offline.[6] Information that can reveal a person's financial, health, biometric, genetic, or sex life data among others is

[4] Discussion Paper, National Strategy for Artificial Intelligence, Anna Roy Advisor (Industry), NITI Aayog

[5] Recommendation for Sections 3(3), 3(16), and 61(6)(m) of the Draft Bill.

[6] Section 3(28), Personal Data Protection Bill, 2019

considered to be "Sensitive Personal Data"[7], the government can however declare certain data to be Critical Personal Data but no clarification as to what can be considered such has been provided in the Bill.[8] It should be noticed that the definition of Personal Data is very wide in comparison to the present law under the IT act, on the other hand, non-personal data is any data that is not personal.

Under Section 91 non-personal data is defined and gives the government the power to direct a data fiduciary to provide de-identified/anonymized data to provide them with any anonymized data, the real given for this is to enable the government to form more "evidence based" policies.[9] The bill also provides that it would not apply to anonymized data other than what is provided in section 91 itself.[10] This also lets the government ask the data fiduciaries to provide the government with anonymized data that is actually de-identified personal data.

MeitY in September 2019 formed a committee to deliberate over regulations regarding non-personal data known as the Gopalakrishnan Committee for Non-Personal Data. This committee was formed to deal with issues related to non-personal data regulation, the government stated that the committee needed to look at the 'economic dimension' to non-personal.[11] In a circular given by MeitY it was stated the benefits of using "privately collected data" for the greater public good. The committee has not published a report as of yet and hence there currently exists no solid framework to regulate Non-Personal Data.

4. Government Involvement in Non-Personal Data

The government through section 91 of the PDP bill has the sovereign power to control non-personal data in the country. The government's rationale in this situation is that according to them large companies such as Amazon and Google have a very large inventory of datasets which are like a natural resource towards development and should be utilized for the good of the public.[12]

A flipside to this argument can be the fact that businesses invest a very large amount of resources to collect data, this collection can require building of new products and using several efficient data collection methodologies that helps them effectively aggregate data. This data can also put certain companies at advantageous position as compared to their competitors and can also be essential for their growth as a business. Sharing privately built assets with the government in which the government becomes the custodian of such assets is not fair to the companies.[13]

[7] Section 3(36), Personal Data Protection Bill, 2019
[8] Explanation to Section 33(2), Personal Data Protection Bill, 2019
[9] Explanation to Section 91(2), Personal Data Protection Bill, 2019
[10] Section 2(B), Personal Data Protection Bill, 2019
[11] MeitY, Constitution of a Committee of Experts to deliberate on Data Governance Framework, 13 September 2019,
[12] Agarwal, S. (2020). Community social media platform, LocalCircles, highlights data misuse worries. [online] The Economic Times. Available at: https://economictimes.indiatimes.com/small-biz/startups/newsbuzz/community-social-media-platform-localcircles-highlights-data-misuse-worries/articleshow/71041420.cms?from=mdr [Accessed 29 Jan. 2020].
[13] Tech2. (2020). Exclusive: Govt wants unrestricted access to non-personal data of citizens from e-commerce, social media companies- Technology News, Firstpost. [online] Available at: https://www.firstpost.com/tech/news-

An argument for government intervention can be three pronged. Firstly, there is a blatant information disproportionateness between a consumer and the data user, this asymmetry is caused because a customer often under-estimates the value of their personal data and is ignorant about the massive scale at which their data is being collected and used. Data collectors can also unilaterally change their data policies hence making the system more disproportionate.[14] Secondly, customers often lose sight of the long term consequences of their actions, when they share their personal information to avail specific services, this is a bounded rationality that favors the data user. Thirdly, since certain service providers hold the data, it gives them an advantage, this can end up giving these users a data monopoly. The results of this can be harmful to the market.[15]

5. Probable Objectives for the Governance of Non-Personal Data

5.1 Ensuring digital Competition

Barriers to entry can be created primarily by economies of scale as Data-intensive businesses profit when they serve more consumers instead of a few[16], which saw the rise in zero-price services as more the users more the company's profit. Secondarily, network affects which are the effects the users of a good can have on the value of the same good to other customers or even potential customers.[17] Thirdly, economies of scale with network effects can lead to the creation of more data which can give the data user control over more data. Lastly, the data user after having access to various datasets for a long tune can enter other markets easily and hinder the development of some secondary markets.[18] All the above factors can asymmetrically favor the service provider which can in the future be detrimental to the consumer.[19]

5.2 Development of International Trade

The development of International Trade is very important to the Indian Government and hence data-driven economies will have to deal with the commodity that is non-personal data. International rules matter when it comes to non-personal data as they affect cross border data flow by regulating the trade of goods and services and protect intellectual property, they also give rules that can warrant a change in national laws and limit the space in which national governments can make policies.[20]

5.3 National Security

The Government of India has many data sharing initiatives that can collect as well as process data for the enhancement of public security. Systems collect information which can be personal and non-personal such as

analysis/exclusive-govt-wants-unrestricted-access-to-non-personal-data-of-citizens-from-e-commerce-social-media-companies-7710181.html [Accessed 29 Jan. 2020].

[14] Indian Television Dot Com. (2020). TRAI begins work on data protection and government's role. [online] Available at: https://www.indiantelevision.com/regulators/trai/trai-begins-work-on-data-protection-and-governments-role-170809 [Accessed 29 Jan. 2020].

[15] Id.

[16] OECD Annual Report, 2002

[17] UNCTAD, Trade and Development Report 2019

[18] Id.

[19] Unlocking digital competition, Report of the Digital Competition Expert Panel, 2019

[20] The Governance of Data and Data Flows in Trade Agreements: The Pitfalls of Legal Adaptation, UC Davis Law Review, Vol. 51, 2017, pp. 65-133

satellite data, traffic, financial information and etc. Laws also give the state the authority to collect documentation. Any policy that comes into place with respect to Non-Personal data will need to account for existing systems that give access to information data sets for the purpose of national security.[21]

5.4 Privacy

As discussed previously, anonymization does not guarantee privacy. As data can be de-identified, there are risks of it being re-identified as the anonymization reverses. Post the re-identification, the data could be used for malicious purposes. Even though the PDP bill 2019 gives for an irreversible de-identification, a combination of de-identified data can possibly identify a data principal, hence a very rigorous standard needs to be observed so that the anonymized data cannot be re-identified.[22]

6. CONCLUSION

It is abundantly clear at this point that the Section 91 of the PDP Bill, 2019 will not be suitable as the sole regulator for the different objectives one needs for the regulation of Non-Personal Data. Given how different the considerations are for the two, a blanket framework or a one size fits all governance model will not be adequate. While there are benefits to the free flow of data, there are also concerns and policies should not be formed before addressing the concerns first.

*****

[21] Xynou, M., Hickok, Security, Surveillance and Data Sharing Schemes and Bodies in India. Retrieved from Centre for Internet and Society [ONLINE] Available at https://cis-india.org/internet-governance/blog/security-surveillance-and-data-sharing.pdf [Accessed on 29 January 2020]

[22] Wes, M., Looking to comply with GDPR? Here's a primer on anonymisation and pseudonymisation. [ONLINE] Available at https://iapp.org/news/a/looking-to-comply-with-gdpr-heres-a-primer-on-anonymization-and-pseudonymization/ [Accessed on 29 January 2020]