COMMENTS ON THE:

PERSONAL DATA PROTECTION BILL (PDP BILL), 2019

February 2020 | *Issue No. 102*

ESYA centre

ABOUT THE ESYA CENTRE

The Esya Centre's mission is to generate empirical research and inform thought leadership to catalyse new policy constructs for the future. It simultaneously aims to build institutional capacities for generating ideas which enjoin the triad of people, innovation and value, consequently helping reimagine the public policy discourse in India and building decision-making capacities within government.

Esya invests in ideas and encourages thought leadership through collaboration. This involves curation of niche and cutting-edge research, and partnerships with people, networks and platforms. Moreover, it prioritises multi-disciplinary research to engender "research clusters", through which practitioners and researchers collaborate.

DEFINITION OF PERSONAL DATA

At its heart, the Bill lays down standards for protection of the privacy of individuals, relating to their personal data. It specifies guidelines for the flow and use of personal data. The term 'personal data' is defined under Section 3(28) of the Bill to include *'data about or relating to a natural person who is directly or indirectly identifiable, having regard to any characteristic, trait, attribute or any other feature of the identity of such natural person, whether online or offline, or any combination of such features with any other information, and shall include any inference drawn from such data for the purpose of profiling.*

The Bill expands the scope of 'personal data' envisioned originally, and now includes "any inference drawn from such data". Extending the application of the PDP Bill to inferences derived from any indirectly identifiable personal data including those that are aggregate in nature, goes beyond the Preamble of the PDP Bill, 2019.
Further, to avoid conflicts with the intellectual property rights (IPRs) afforded to market participants, the standard for determining whether data is personal is whether such data is related to an identified or identifiable individual. Jurisdictions such as the European Union and Singapore employ some version of this formulation. For instance, Article 4 of the General Data Protection Regime (GDPR) defines personal data' to mean *any information relating to an identified or identifiable natural person ('data subject'); an identifiable natural person is one who can be identified, directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, an online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person*[1].

The Committee of Experts under the chairmanship of Justice (Retd.) B.N. Srikrishna recognised that there is no alternative (to the identifiability standard) which provides a workable standard for demarcating data that must be protected under the law[2]. Consequently, the scope of personal data under the 2018 draft was limited to 'data about or relating to a natural person who is directly or indirectly identifiable'[3].

Inclusion of inferred data or derived data however creates direct conflict with IPR. It is worth pointing out here that when raw-data is compiled, arranged, processed and analysed, such databases acquire a proprietary nature due to the effort and innovation put in by a data fiduciary.

The copyrightability of databases has been settled by the Supreme Court in Eastern Book Company v. DB Modak[4]. Here, the question was whether the petitioner, a company which created databases of Supreme Court cases (which are in the public domain) could claim copyright protection for their databases. It was held by the Court that the petitioner's input of independent skill, labour and capital, in editing and arranging the information as well as adding inferences from it in the form of headnotes, resulted in the database being a copyrightable work.

This rationale would also apply to databases created and used by data processors as these involve skill, labour and capital investment in arrangement of the data, as well as drawing inferences from the data through processing.
We recognise the risks of sole reliance on the identifiability standard, particularly from the failure of methods of de-identification. However, an alternative approach is to adopt a definition which is applied to various contexts in which the data of a person may be processed. That said, flexibility in the definition should not be achieved at the cost of certainty.
As noted in B.N. Srikrishna Committee's Report, the Data Protection Authority will have to offer guidance, explaining the standards in the definition as applied to different categories of data in various contexts, especially with regard to newer categories of data developed as a result of advances in technology. **We therefore submit that the definition proposed under the PDP Bill of 2018 should be restored.**

1 Article 4 of General Data Protection Regime; available at https://gdpr-info.eu/art-4-gdpr/
2 Committee of Experts, A Free and Fair Digital Economy: Protecting Privacy, Empowering Indians; available at https://MEITy.gov.in/writereaddata/files/Data_Protection_Committee_Report.pdf
3 Personal Data Protection Bill, 2018; available at https://MEITy.gov.in/writereaddata/files/Personal_Data_Protection_Bill,2018.pdf
4 2008ISCC1

SENSITIVE PERSONAL DATA

In addition to the list given under Section 3(36) of the Bill, we submit that the law should provide a clear and unambiguous guideline to determine data which can be categorised as sensitive personal data. We suggest that the test of identifiability applicable to 'personal data' should be used for sensitive personal data as well.

Therefore, we recommend that the phrase *'be related to'* is deleted, and the definition should be limited to data which 'reveals' or 'constitutes' sensitive personal data.

CHILDREN'S DATA

Inclusion of single, flat threshold that inhibits children's ability to participate in the digital ecosystem may be counterproductive. It is important to note that digital platforms allow children to access to a wide variety of information, provide opportunities for learning, and also economic opportunities (UNICEF, 2017).

India has also witnessed introduction to innovation and interactive platforms aimed at promoting digital learning. While digital safeguards for child protection should be encouraged, the age threshold and related obligations should be nuanced and graded.

The rationale behind adopting the threshold in the PDP Bill stems from the age of majority provided under the Indian Contract Act. Despite noting that that from the perspective of the full, autonomous development of the child, the age of 18 is too high, the B.N. Srikrishna Committee in its Report recommended this threshold to ensure consistency with the existing legal framework. However, this logic appears tenuous.

Besides diverging from global practices, the threshold of 18 years envisaged in the Bill is in itself not consistent across Indian legislation. There is precedence in other laws prescribing different age limits. For instance, the Reserve Bank of India allows minors above the age of 10 years to independently operate savings bank account[5].

Globally, the age threshold for Data Principals requiring parental or guardian consent is recognized as 13. In its present form, the PDP Bill does not recognize the varying maturity levels of children at various age groups. Parental/ guardian approval should be required in relation to collection of personal data from children below the age of 13. However, children between the ages of 13 and 18 years must be permitted and empowered to make decisions about their data in relation to activities in ordinary course.

Article 8 of EU's GDPR allows collection and processing of any person below the age of 16 with parental consent and this age threshold can be brought down by member states to 13 years under their domestic law.

Moreover, Recital 38 notes that the use of child data in marketing, or for profiling purposes or in connection with the supply of services to children are areas of concern requiring specific protection under the GDPR.
Even per the US Children's Online Privacy Protection Act (COPPA), the threshold for collecting and processing of personal data without parental consent is 13 years. Similarly, The Australian Privacy Principles Guidelines allows an entity to presume that "an individual aged 15 or over has the capacity to consent, unless there is something to suggest otherwise"[6].

Therefore, in line with the global standards, India should adopt a more flexible and graded approach.

5 Opening of Bank Accounts in the Name of Minors, Reserve Bank of India; available at https://www.rbi.org.in/Scripts/NotificationUser.aspx?Id=9227&Mode=0
6 Office of Australian Information Commissioner; available at https://www.oaic.gov.au/privacy/your-privacy- rights/children-and-young-people/

COMPLIANCES

The PDP Bill lists several compliances to be maintained by a data fiduciary. The compliances of the Bill will require fiduciaries, particularly small and businesses, to allocate significant resources. Therefore, it is important that sufficient time is afforded to comply.

In line with the global best practices, we suggest that businesses be provided a minimum time frame of 24 months to comply with the Bill. This period should commence from the date of notification of the Act or from the data the regulations are prescribed (if they are not specified in the act itself), whichever is later.

IDENTIFYING HARM

Section 3(20) of the Bill providing a list of different types of harm, which includes

(i) bodily or mental injury;
(ii) loss, distortion or theft of identity;
(iii) financial loss or loss of property;
(iv) loss of reputation or humiliation;
(v) loss of employment;
(vi) any discriminatory treatment;
(vii) any subjection to blackmail or extortion;
(viii) any denial or withdrawal of a service, benefit or good resulting from an evaluative decision about the data principal;
(ix) any restriction placed or suffered directly or indirectly on speech, movement or any other action arising out of a fear of being observed or surveilled; or
(x) any observation or surveillance that is not reasonably expected by the data principal'.

The definition provided in the Bill fails to link the notion of harm to the compromise of a data principal's personal data. Instead, the harms listed could be caused by factors unrelated to the misuse of personal data of an individual. For example, under the current definition for the term, *any discriminatory treatment"; "any denial or withdrawal of a service, benefit or good resulting from an evaluative decision about a data principal"; and "any observation or surveillance that is not reasonably expected by the data principal* can be considered harmful. Moreover, given that the PDP Bill accounts for several disclosure requirements, the broad terms adopted for identifying harm can be construed subjectively, giving wide discretion to the Data Principal. This may lead to significant amount of frivolous complaints and misuse of the Data Principal's rights.

The definition of harm proposed under the Bill doesn't spell out the underpinning principle on the basis of which existing harms or newer forms of harm may be included in the list. Further, the use of the doctrine of reasonable expectations in the context of privacy is problematic. The test of reasonable expectations is "inherently uncertain because reasonable expectations of privacy vary across social groups, time and social culture. the boundaries of what amounts to a reasonable expectation of privacy shift over time"[7].

This uncertainty in varied reasonable expectations introduces a subjective element which could be abused in this context.

Therefore, the definition of harm should be restricted to

(a) any discriminatory treatment affecting the fundamental rights of the data principal;
(b) any denial or withdrawal of a service, benefit or good resulting from an evaluative decision about a data principal which a data principal is entitled to receive as a fundamental right; and
(c) any observation or surveillance that is not consented to by the data principal.

7 Barocas and Selbst, 'Big Data's Disparate Impact', California Law Review, Vol. 104 issue 3 (2016); available at https://lawcat.berkeley.edu/record/1127463

NON-PERSONAL DATA ACCESS TO DATA

Section 91 of the Bill empowers the Central Government to mandate any data fiduciary or processor to disclose such information that constitutes non-personal Data or anonymized data. Notably, non-personal/anonymized data has wide implications and could include proprietary information, insights, trade secrets, algorithms, source codes etc. Given anonymized data sets and non-personal data in the form of inferences are proprietary to businesses, asking them to share this data, even to achieve non- commercial public policy objectives, may undermine the competitiveness of such businesses.

As noted above, it is a well settled principle of law that processed and analysed, such databases acquire a proprietary nature due to the effort and innovation put in by the data fiduciary. This provision has been framed under an incorrect assumption that that there is no element of ownership of private entities in these datasets, consequently, summarily dismissing attempts to explore other mechanism such as the use of Fair, Reasonable and Non-Discriminatory (FRAND) terms for sharing non- personal data. Extrapolating FRAND terms, which is generally used in the domain of standard essential patents (SEPs) could enable greater access to data while respecting IPRs of entities holding such data.

Moreover, it is often difficult or impossible to distinguish between personal and non- personal data. For example, it is difficult to distinguish between personal and non-personal data generated by consumer devices (e.g. from connected vehicles, smart appliances and smart meters). In certain situations, non-personal and personal data may be 'inextricably linked', consequently giving rise of new challenges.

We suggest that requests by Central Government to share non-personal data or anonymized should be subject to judicial scrutiny to ensure that conflicts with legitimate business interests are avoided.

This document has been prepared by the Fellows at the Esya Centre.
For any further contact, please get in touch with us at:

www.esyacentre.org