Internet Democracy Project
Submission in response to the
White Paper of the Committee of Experts on
Data Protection Framework for India

At the outset, the Internet Democracy Project would like to thank the Committee for this opportunity to share our ideas. However, we would also like to express our dissatisfaction at the process through which these ideas are being sought.

To be meaningful, a consultation of this scope should have taken place over a much longer period of time. By leaving such little time for a consultation with over 200 questions, only very few stakeholders will be able to respond in detail to a consultation that may well determine the future for all Indians in the digital age for decades to come. This is not the route a consultation process worth that name should take. The consultation should also have been sequenced, as our answers to some of the questions raised here will be determined by the consensus that emerges around other aspects, such as the definitions of data controller and data processing. In addition, submissions of all should have been made public, and a period for counter-comments should have been incorporated. The speed at which this consultation regarding such a far reaching, important draft law is rushed through is of great concern.

While the severe limitations of this process have made meaningful participation complicated, we would nevertheless like to submit a number of comments, within these limitations, as follows. As this is a coherent set of arguments, we would like to present them together, something that the web form unfortunately made really difficult.

Our submission proceeds from the value of the right to privacy for the realisation of the right to life and dignity of people. Towards this, any framework should center people and account for the most vulnerable of the population.

Privacy is not merely something of value for an individual, but is a larger question of power. A few entities amassing large amounts of data creates a strong power dynamic in favour of those amassing data, and in democracy-bending ways. That privacy has a social value is also acknowledged in the Puttaswamy judgment.

As changes in law and regulation have shown, it is possible to shift what once seemed unchangeable and to return greater control over their data to the people. We hope that this framework will result in a radical shift in the way personal data is dealt with.

Structure

We believe that the Report of the Group of Experts on Privacy, chaired by Justice AP Shah ("AP Shah Report") remains relevant, and so we make our submissions under the high-level principles laid down there as well as under the 'Privacy Principles'. In the below, we propose some additions to these principles as well as add details where relevant under the original set of principles. While we believe the principles laid down in the White Paper are all worthy of inclusion, treating high-level recitals (such as "holistic application") at the same level as the more detailed principles (such as "deterrent penalties") might create room for confusion.

Although we do not explicitly address the high-level principle of multi-dimensional privacy or the privacy principles of access and correction, disclosure of information, security, openness and accountability, all addressed in the AP Shah Report, we believe these remain equally relevant as well, and should be reflected in any data protection legislation in India.

In what follows we will address:

A. Centrality of data subject

The data protection framework has emerged from the right to privacy under the Constitution. In the landmark Puttaswamy judgment[1], the right to privacy was affirmed as being an intrinsic part of the right to life and personal liberty under Article 21. The right was found to be central to human dignity, liberty and autonomy.

[1] Justice K. S. Puttaswamy (Retd.) and Anr. vs Union Of India And Ors. W. P. (CIVIL) No. 494 of 2012.

The data protection framework emerging from this right should therefore center the natural person: a citizen or a resident. Thus, the framework should be applicable to data of natural persons only.

The framework's territorial scope, too, should be framed around this understanding . Entities based outside the country but offering goods and services to Indian residents, or monitoring their behaviour (similar to the GDPR) should also fall within the scope, and appropriate enforcement requirements should be put in place.

Predictably, there will be concerns that the ease-of-doing-business and other such parameters will be affected. But considering that for a lot of entities with Internet-based business models, India is considered a market to be tapped, such parameters cannot be at the cost of violation of rights.

B. Technological neutrality and interoperability with international standards

The first principle of the White Paper should be expanded to include interoperability with international standards. Consistency and harmonisation with international best practices is beneficial for many reasons, and there is an increasing trend towards alignment of data protection regimes.

Technological neutrality, and by extension, having future-proof definitions
Definitions of personal data and sensitive personal data are being challenged due to improvements in re-identification techniques, increasing number of public datasets that can be combined leading to increased risk of re-identification and other factors. In light of these challenges, definitions in the data protection framework should be able to accommodate these changes.

There are many issues of data protection that need attention when it comes to big data and artificial intelligence in particular.[2] The privacy concerns (apart from bias and associated harms) arising due to pattern recognition already require articulation of a new generation of rights. Consider, for example, that it is possible to 'derive the intimate from the available'[3]. The implication of this is that new rights will have to be articulated and existing assumptions questioned: examples include explicit inclusion of metadata within the definition of 'personal data'; no assumption that there is no expectation of privacy in public (already acknowledged in Puttaswamy judgment); right to object to profiling; and rights against direct marketing on the basis of one's personal data.

---

[2] Calo, 2017. Artificial Intelligence Policy: A Primer and Roadmap. (August 2017). https://ssrn.com/abstract=3015350.
[3] Ibid.

C. Horizontal applicability

The data protection law should be applicable to both public and private entities.

As the White Paper recognises, where the public sector is exempted, it should not be a blanket exemption from all principles, and such an exemption should not apply to all operations or functions. This also means that a blanket exemption for law enforcement and intelligence agencies should not apply. According to the Puttaswamy judgment[4], for a derogation from the right to privacy, there should be a legitimate State aim, a law and a proportionate limitation for the derogation to be justified. Any exemptions should reflect this.

The example of data collected for the purpose of delivery of welfare benefits under Aadhaar have famously been used for purposes far more ambitious than that. Egregiously, it has been used for creating parallel surveillance databases called State Resident Data Hubs that are outside the scope of the Aadhaar Act.[5] The Aadhaar program should serve as a reminder to put in place a data protection framework that meaningfully holds both public and private entities accountable. A 'public interest' or equivalent ground for data processing should not obviate the need for public entities to adhere to all proposed data protection principles like Privacy-by-Design, purpose limitation.

D. Conformity with Privacy Principles

1. Notice

'Notice' is the first principle in the AP Shah Committee report, which gave a 'distillation of global best practices'. The list of heads for which notice is required, and its manner, remains relevant, including the requirement to notify individuals 'of any legal access to their personal information after the purposes of the access have been met'. Notice in cases of profiling should be added to this list.

2. Choice and consent

Consent has an important but limited role to play in the data protection framework. It should be one of the different grounds for for processing. Consent should be (i) unbundled (ii)

---

[4] Supra n 1.

[5] Unique Identity Authority of India, 2012. State Resident Data Hub Institutional Document. April 2012. https://ia800806.us.archive.org/14/items/UIDAISRDHStateAdoptionStrategy/UIDAI%20-%20SRDH%20-%20Institutional%20Framework.pdf

unambiguous (iii) uncoerced and (iii) express and (iv) obligations to comply with proposed principles like privacy-by-design, purpose limitation should still apply.

Further, details like the burden of proof for consent should be decided with the intention of restoring power with data subjects.

3. Collection limitation

While it might be argued that collection limitation is irrelevant in the age of big data, where all actions are thought of as data waiting to be captured ("data is the new oil"), the committee should not be swayed by the rhetoric of 'data driven innovation'[6]. This rhetoric is employed by public entities as well. The Task Force for Artificial Intelligence set up by the Ministry of Commerce & Industry[7], which speaks in grand terms like 'India's Economic Transformation' does not seem to have considerations of ethics or privacy. More than ever, regulating collection is relevant[8] in the age of big data, and should be looked into.

4. Purpose limitation

Purpose limitation is another area where there is likely to be much resistance. However, this principle has a crucial role to play in curbing excesses of data monetisation (for money or insight) at the cost of contextual integrity by public entities as well as private entities.

To move away from the wealth of lessons that Aadhaar has to offer about how not to treat personal data, take the example of the UK government. Sensitive personal data (on mental health, gender, nationality) collected on homeless people with the purpose of understanding their needs, was used on a large scale to identify individuals vulnerable for deportation.[9] Absence of purpose limitation is standard practice in the private sector. The Register reports

---

[6] Bajaj, 2018. Creating Data Protection Regime - encouraging innovation while enhancing privacy. *Economic Times*. January 2018. https://tech.economictimes.indiatimes.com/news/internet/creating-data-protection-regime-encouraging-innovation-while-enhancing-privacy/62389379

[7] Press Information Bureau, Government of India, Ministry of Commerce & Industry. Commerce and Industry Minister Sets up Task Force on Artificial Intelligence for Economic Transformation. August 2017. http://pib.nic.in/newsite/PrintRelease.aspx?relid=170231

[8] Nissembaum, 2017. Deregulating Collection: Must Privacy Give Way to Use Regulation?. May 2017. https://ssrn.com/abstract=3092282

[9] Townsend, 2017. Home Office used charity data map to deport rough sleepers. *The Guardian*. August 2017. https://www.theguardian.com/uk-news/2017/aug/19/home-office-secret-emails-data-homeless-eu-nationals

that Chief Data Officers are 'increasingly asked to help monetise the data companies hold, rather than purely managing and protecting that data.'[10]

5. Control

Towards ensuring greater control over one's personal data, and with regard to new challenges arising from widespread profiling, among other developments in privacy violations, the following should be included:

- A right to object to profiling (which has already been recognised in the GDPR, for example).

- A prohibition on fully-automated decision making (similarly, the GDPR provides a right to object to evaluative decisions based on automated decision making as well as access to 'at least right of human intervention').

- A limited right to be forgotten, in the form of a right to erasure, is an important tool through which to shift the balance of power from business entities and public and government organisations back to data subjects. To the extent that it corrects this current power imbalance, it should be included in India's data protection law. It should include clear exemptions (including for journalistic purposes and where data concerns public figures and public authorities in the course of their public duties) and a wide range of substantive and procedural safeguards to protect the right to freedom of expression and information as guaranteed under India's Constitution and further detailed in laws such as the Right to Information Act. Requests for de-listing should not be entertained under the right to erasure in a data protection law.

- Data portability should be part of the rights available within the right to data protection so as to ensure meaningful control by users over their data, as well as incidentally resulting in increasing competition in respective markets. Such a right would be able to tackle some of the issues arising from Internet-era monopolies, described earlier. TRAI had released a statement in 2017 stating that since the introduction of mobile number portability in 2015, 29 crore subscribers had availed the facility.[11]

---

[10] Hill, 2017. Sucks to be a... chief data officer, when they're being told: Boost revenues. *The Register*. December 2017. https://www.theregister.co.uk/2017/12/06/chief_data_officers_increasingly_told_to_monetise_data/

[11] The Hindu Business Line Bureau, 2017. In two years, requests for mobile number portability touched 29 cr. *The Hindu Business Line*. August 2017. http://www.thehindubusinessline.com/info-tech/in-two-years-requests-for-mobile-number-portability-touched-29-cr/article9818181.ece

6. Privacy by design and default

Privacy-by-design[12] is one of the global best practices that have remained relevant over decades since it was conceptualised, and the White Paper is right in recognising it as central to the principle of accountability. It should be one of the main principles in the data protection framework and should apply to public and private entities alike. Use of Privacy Enhancing Technologies (PETs), most prominently strong encryption, should be promoted.

E. Balance with other rights and interests

Exemptions for household, journalistic, literary, research, artistic, statistical purposes should be included, to balance the competing right to freedom of speech and expression. In addition, an explicit exemption for security researchers is also extremely important. Documentation of the determination that an exemption applies may be required as one of the safeguards.

Where such exemptions apply, specific principles may be allowed derogation from, not all. For example, where a research exemption applies, a strict limitation on sharing of datasets can continue to apply.

Exemptions for investigation and detection of crime and national security should be included, but herein lies the true test of many of the other principles like horizontal applicability and accountability: these exceptions have to be narrow and accompanied with checks and balances. Lessons should be drawn from checks and balances (like prior judicial approval) that have failed to uphold rights in other jurisdictions.[13]

A cautious framing should be done when it comes to the exemption for investigation and detection of crimes, as mass surveillance combined with artificial intelligence can have disastrous consequences for the right to privacy. A concrete example would be to have transparency requirements around 'predictive policing' algorithms.[14]

The contours of the right to erasure, which is part of the menu of rights proposed for data subjects should be carefully determined.

---

[12] Cavoukian, 2009. Privacy by Design. *Office of Information and Privacy Commissioner of Ontario*. August 2009. https://www.ipc.on.ca/wp-content/uploads/Resources/7foundationalprinciples.pdf

[13] See FISA example: Greenwald, 2013. Fisa court oversight: a look inside a secret and empty process. *The Guardian.* https://www.theguardian.com/commentisfree/2013/jun/19/fisa-court-oversight-process-secrecy

[14] See Winston, 2018. Transparency Advocates Win Release of NYPD "Predictive Policing" Documents. *The Intercept*. January 2018. https://theintercept.com/2018/01/27/nypd-predictive-policing-documents-lawsuit-crime-forecasting-brennan/

F. Co-regulatory regime

A co-regulatory model of regulation should be put in place. However, it is important to learn from co-regulatory models that have failed. The Australian case example of the Office of the Information Commissioner is instructive as it is an example where such a model failed due to underfunding, lack of staff and other operational issues.[15] While this is the best model available, its success depends on ensuring that the regulatory body is equipped suitably and independent.

The statute for data protection should delineate rights of data subjects, set up an independent institution with supervisory, investigative and enforcement powers and equip it with soft and hard regulatory tools and resources.

ADDITIONAL ISSUES

Retrospective application:

As the data subject's rights are affected even where data collection and processing has already been done, the law should apply retrospectively where the scale of data collection and processing exceeds a certain threshold.

There is need for especial focus on the Aadhaar program when it comes to retrospective protection. This program has been recorded to derogate on a large scale from consent explicitly withheld for sharing of data. Further, it has been a patently coercive exercise in linking the number to various databases to avail various services. The Aadhaar Act retrospectively gave legitimacy to large scale data collection and use, therefore this is a program that has truly run roughshod with the data of Indian residents. The data protection framework should apply retrospectively, and Aadhaar should not be exempt from its sweep.

A transition period should be there where entities that have already collected data can comply with requirements.

Data localisation

The White Paper lists the following issues in favour of data localisation: (i) protecting right of data subjects, (ii) preventing foreign surveillance and (iii) easy access of data in support of law enforcement and national security.

---

[15] Farrell, 2015. 'Governments do not like freedom of information': the war on Australia's privacy and information watchdog. The Guardian. October 2015. https://www.theguardian.com/australia-news/2015/oct/01/governments-do-not-like-freedom-of-information-the-war-on-australias-privacy-and-information-watchdog

To illustrate the case for the first point, the White Paper gives the example of Microsoft Corporation v. United States of America. The White Paper wrongly notes this case as one about protection of rights of data subjects, while it is actually a case about a warrant for emails and other information associated with a Microsoft Networks email address.

The rights of data subjects will not be protected simply by virtue of having their data locally stored, even as businesses mount the case for 'data to be the next oil'. Protection of rights of data subjects will be dependant on a strong set of such rights being elucidated, a good enforcement mechanism and a number of other factors. Besides, there are ways to protect rights of data subjects through other enforcement mechanisms without having to resort to data localisation.

Second, the need for data localisation due to foreign surveillance. There are many factors that ultimately affect whether or not data is accessible by governments: *Whether data are accessible by governments is often determined by the encryption and privacy design choices companies make, by their data collection and retention policies, and their decisions to agree to government surveillance terms and to comply with law enforcement requests* (DeNardis & Hackl, 2015; Soghoian, 2010) quoted in Data Localization and the Role of Infrastructure for Surveillance, Privacy, and Security[16]. Efforts to ensure there is privacy by design, strong encryption and anonymisation requirements etc. go a long way in ensuring that foreign governments are not able to access data of Indian residents.

Only the last of these points can be ensured through data localisation.

Beyond the cost for foreign companies that cater to an Indian customer base, smaller companies will be hugely impacted if they cannot use the services of many cloud service providers because these offerings don't have a local data center. This has been pointed out in the White Paper as well.

[16] Sargsyan, 2016. Data Localization and the Role of Infrastructure for Surveillance, Privacy, and Security. International Journal of Communication. http://ijoc.org/index.php/ijoc/article/viewFile/3854/1648