India's new data bill is a mixed bag for privacy

India's parliament has released its Digital Personal Data Protection Bill, the second pass at a comprehensive data privacy law after the government withdrew its Personal Data Protection Bill earlier this year. The current draft is shorter than the other bill, though it has many of the same components.

Indian policymakers have also circulated comments on the bill that have not been made publicly available, but which I discuss here. Importantly, the new bill offers a mixed bag for privacy—with some requirements for companies to receive individual "consent," correct inaccurate personal data, and protect data rights alongside concerning provisions for government data access. This analysis is not comprehensive, but highlights some key points of note in the legislation which would constitute a major development in global data regulation.

The idea of 'consent'

As with the old bill, the Digital Personal Data Protection Bill requires organizations processing data (which it calls "data fiduciaries") to obtain "consent" from individuals about whom they are processing data. When an individual gives their "consent," the bill says, the organization must provide that person with "an itemized notice in clear and plain language" that describes the data collected and the purpose for which it is being processed, "as soon as it is reasonably practicable." The bill also proposes that individuals be able to withdraw their "consent" for organizations to process their data, at which point the organization in question must stop doing so.

This notion of consent is broken, and it is not specific to India, either. US and European companies and policymakers push the same idea. People do not read terms of service agreements, and companies that flash a lengthy document with inaccessible legalese—waiting for individuals to just click "accept"—know that users neither read nor understand the document.

This defies the very notion of consent.

Likewise, people do not read privacy policies, yet many websites and applications will literally assert that viewing the website or opening the application in and of itself constitutes agreement with its privacy policy. Furthermore, consent is not fully and freely given in a world in which citizens—including Indian citizens—cannot access basic services without subjecting themselves to data collection. Nonetheless, the new bill's language on consent puts India relatively in the same direction as "consent" provisions in US state privacy laws and the General Data Protection Regulation in the European Union.

Government data collection carve outs

The bill weakens this notion of consent even further by specifying numerous exceptions, including many exceptions for the Indian government. While some are arguably more reasonable, others create privacy risks to Indian citizens and generate complex legal questions for companies and organizations operating in India.

Individuals, in the most recent public bill draft, are deemed to have given consent if the data processing "is necessary" for "the performance of any function under law," "for compliance with any judgment or order issued under any law," "for responding to a medical emergency involving a threat to the life or immediate threat to the health of the [individual] or any other individual," and "for taking measures to ensure safety of, or provide assistance or services to any individual during any disaster, or any breakdown of public order," among others. It would also exempt situations in which it

is “reasonably expected” that someone would provide their personal data to an organization voluntarily, the processing of “publicly available personal data,” and credit scoring.

While some of these exceptions may appear reasonable on their face (such as credit scoring), many are highly concerning from a privacy and civil rights perspective. The government of Indian Prime Minister Narendra Modi has frequently made bogus claims of threats to public safety and order to crack down on protest and dissent. Authorities have likewise drawn on similarly dubious, but public order-framed, claims to legally and rhetorically justify shutting down the internet more times than any other country on earth. In similar form, the bill’s current language would allow for data gathering based on the “public interest,” defined extremely broadly to include the interest of India’s sovereignty and integrity, state security, friendly relations with foreign states, maintenance of public order, preventing incitement of any of the aforementioned activities (like undermining public order), and preventing the dissemination of “false statements of fact.”

The Modi government is hardly the only government in a nominally democratic country engaged in outright undemocratic practices. Yet, the proposal for incredibly broad government data carve outs in the bill would empower state surveillance at the expense of Indian citizens’ privacy and civil rights. It would also force companies and organizations operating in India to constantly grapple with an even more complex set of legal questions around expanded government access to data.

Concerningly, this is just one component of the Modi government’s broader push to undermine encryption and increase its ability to coerce technology companies.

Data localization

The most contentious element of the old Personal Data Protection Bill was probably its data localization requirements. These stipulations would have required organizations with personal data on Indian citizens to keep that information stored in the country, in some cases merely a copy and in other cases preventing outbound transfer entirely. Different Indian policymakers wanted these requirements in place for a range of reasons, including to impose costs on foreign companies, boost India’s data storage industry, increase Indian government oversight over the storage of data related to Indian citizens, and, as some saw it, enable better Indian law enforcement access to crime-relevant data held by US companies, which is currently less than accessible through a broken mutual legal assistance treaty process.

The new bill steps away from that emphasis on localization. Instead of requiring local storage of data per se, it proposes to allow the Indian government to evaluate foreign countries’ data protection regimes and then certify those as sufficient to provide destinations for Indian citizens’ data. Specifically, it states, “The Central Government may, after an assessment of such factors as it may consider necessary, notify such countries or territories outside India to which a Data Fiduciary may transfer personal data, in accordance with such terms and conditions as may be specified.”

Foreign businesses will certainly be happy about this change. Most often, their complaints about data localization came back to costs—not wanting to pay for the technical changes and technical infrastructure to store data locally in India, as well as other legal and organizational costs. But there are also other concerns generated by data localization, including climate emissions costs for duplicate data storage infrastructure and the cybersecurity risks of storing another copy of information when there was previously one fewer.

These provide at least several reasons for the Indian government to move away from a highly controlled data localization regime.

Other ideas undergoing debate

The current version of the bill on the website for the Indian Ministry of Electronics and Information Technology does not apply to "non-automated processing of personal data," "offline personal data," "personal data processed by an individual for any personal or domestic purpose," and "personal data about an individual that is contained in a record that has been in existence for at least 100 years."

Interestingly, one marked-up version of the bill I reviewed—whose changes are not reflected in the current draft online—contained a suggestion to also exempt "anonymized personal data" that is "owned by the central or state government, depending on the subject matter to which the data pertains" from the bill. It also contained an edit to explicitly use the phrase "data free flow with trust" to describe the provisions on the Indian government's approval for foreign data transfers and storage.

The Economic Times reports, in this vein, that the next iteration of the bill will have the Indian government define "trusted" geographies to which data fiduciaries can transfer and store data related to Indian citizens.

The former is (here, was) a bad suggestion. In the United States, federal and state legislators continue to exempt "anonymized" or "deidentified" data from privacy laws and bills under the false belief that such terms are technically meaningful. Numerous studies have shown how easy it is to link supposedly "deidentified" or "anonymized" data to real people, given advancements in statistical data analysis, the sheer volume of data in the world today, and companies' access to data points that are remarkably unique to individuals, such as geolocation history or a device's Wi-Fi connection patterns.

Yet, legislators continue to include these carve outs in laws, including because companies push the bogus line that "anonymization" is real. Hopefully, India's parliament does not fall into the same trap. New techniques to better protect the confidentiality of data while still enabling organizations to process it, such as differential privacy, are valuable. The better path forward is to recognize that there is a spectrum of capabilities to link data to people by name or by another clear individual identifier—and that total "anonymization" is a myth.

The second suggestion in the unpublished bill markup I reviewed—the phrase "data free flow with trust"—is a reference to the Data Free Flow with Trust initiative spearheaded by the Japanese government at the 2019 G-20 in Osaka. At the time, it described a general belief that countries have an interest in allowing the free flow of data between one another, but with some safeguards in place. New Delhi refused to sign onto the initial Data Free Flow with Trust agreement in 2020—a vague proclamation to pursue cooperation for a "data free flow with trust" framework—because it saw the effort as too driven by high-resource countries. India's minister of commerce and industry had then said that "in view of the huge digital divide among countries, there is a need for policy space for developing countries who still have to finalize laws around digital trade and data. Data is a potent tool for development and equitable access of data is a critical aspect for us."

As governments expand their data flow regulations, India's shift away from such an emphasis on localization and toward a focus on trusted geographies aligns India with some of these efforts—while still leaving space for policymakers to carve out a so-called fourth way of data governance aimed at global south countries. Significantly, India is also taking over the G-20 presidency, which means a data free flow with trust type approach could put the country in an influential position to drive global data conversations in the next year.

Conclusion

There was intense debate about the last attempt at comprehensive data regulation in India, including among Indian policymakers, US policymakers, US tech companies, and civil society stakeholders in India. Undoubtedly, those kinds of debates will continue happening around the new legislation.

There are also numerous other issues and questions raised by the proposal that merit deeper analysis and discussion—far more than is covered in this article. For now, though, it's clear that India is intent on planting its flag on data regulation, and that where "trusted geographies" are concerned, there is plenty of space for the US government to productively engage.