Why data localisation might lead to unchecked surveillance

Pranesh Prakash

2018-10-16

The Quint

Abstract

Technically, surveillance in India is not a challenge for the NSA.

In recent times, there has been a rash of policies and regulations that propose that the data that Indian entities handle be physically stored on servers in India, in some cases exclusively. In other cases, only a copy needs to be stored.

In April 2018, the Reserve Bank of India put out a circular requiring that all "data relating to payment systems operated by them are stored in a system only in India" within six months.

Lesser requirements have been imposed on all Indian companies' accounting data since 2014 (the back-up of the books of account and other books that are stored electronically must be stored in India, the broadcasting sector under the Foreign Direct Investment policy, must locally store subscriber information, and the telecom sector under the Unified Access licence, may not transfer their subscriber data outside India).

The draft e-commerce policy has a wide-ranging requirement of exclusive local storage for "community data collected by Internet of Things devices in public space" and "data generated by users in India from various sources including e-commerce platforms, social media, search engines, etc", as does the draft e-pharmacy regulations, which stipulate that "the data generated" by e-pharmacy portals be stored only locally.

While companies such as Airtel, Reliance, PhonePe (majority-owned by Walmart) and Alibaba, have spoken up in support the government's data localisation efforts, others like Facebook, Amazon, Microsoft, and Mastercard have led the way in opposing it.

Just this week, two US Senators wrote to the Prime Minister's office arguing that the RBI's data localisation regulations along with the proposals in the draft e-commerce and cloud computing policies are "key trade barriers". In her dissenting note to the Srikrishna Committee's report, Rama Vedashree of the Data Security Council of India notes that, "mandating localisation may potentially become a trade barrier and the key markets for the industry could

mandate similar barriers on data flow to India, which could disrupt the IT-BPM (information technology-business process management) industry."

Justification for data localisation

What are the reasons for these moves towards data localisation?

> Given the opacity of policymaking in India, many of the policies and regulations provide no justification at all. Even the ones that do, don't provide cogent reasoning.

The RBI says it needs "unfettered supervisory access" and hence needs data to be stored in India. However, it fails to state why such unfettered access is not possible for data stored outside of India.

> As long as an entity can be compelled by Indian laws to engage in local data storage, that same entity can also be compelled by that same law to provide access to their non-local data, which would be just as effective.

What if they don't provide such access? Would they be blacklisted from operating in India, just as they would if they didn't engage in local data storage? Is there any investigatory benefit to storing data in India?

As any data forensic expert would note, chain of custody and data integrity are what are most important components of data handling in fraud investigation, and not physical access to hard drives. It would be difficult for the government to say that it will block all Google services if the company doesn't provide all the data that Indian law enforcement agencies request from it.

However, it would be facile for the RBI to bar Google Pay from operating in India if Google doesn't provide it "unfettered supervisory access" to data.

The most exhaustive justification of data localisation in any official Indian policy document is that contained in the Srikrishna Committee's report on data protection. The report argues that there are several benefits to data localisation:

1. Effective enforcement,
2. Avoiding reliance on undersea cables,
3. Avoiding foreign surveillance on data stored outside India,
4. Building an "Artificial Intelligence ecosystem

Of these, the last three reasons are risible.

Not a barrier to surveillance

Requiring mirroring of personal data on Indian servers will not magically give rise to experts skilled in statistics, machine learning, or artificial intelligence, nor will it somehow lead to the development of the infrastructure needed for AI.

The United States and China are both global leaders in AI, yet no one would argue that China's data localisation policies have helped it or that America's lack of data localisation polices have hampered it.

> On the question of foreign surveillance, data mirroring will not have any impact, since the Srikrishna Committee's recommendation would not prevent companies from storing most personal data outside of India.

Even for "sensitive personal data" and for "critical personal data", which may be required to be stored in India alone, such measures are unlikely to prevent agencies like the US National Security Agency or the United Kingdom's Government Communications Headquarters from being able to indulge in extraterritorial surveillance.

In 2013, slides from an NSA presentation that were leaked by Edward Snowden showed that the NSA's "BOUNDLESSINFORMANT" programme collected 12.6 billion instances of telephony and Internet metadata (for instance, which websites you visited and who all you called) from India in just one month, making India one of the top 5 targets.

This shows that technically, surveillance in India is not a challenge for the NSA.

So, forcing data mirroring enhances Indian domestic intelligence agencies' abilities to engage in surveillance, without doing much to diminish the abilities of skilled foreign intelligence agencies.

As I have noted in the past, the technological solution to reducing mass surveillance is to use decentralised and federated services with built-in encryption, using open standards and open source software.

Reducing reliance on undersea cables is, just like reducing foreign surveillance on Indians' data, a laudable goal. However, a mandate of mirroring personal data in India, which is what the draft Data Protection Bill proposes for all non-sensitive personal data, will not help. Data will stay within India if the processing happens within India. However, if the processing happens outside of India, as is often the case, then undersea cables will still need to be relied upon.

The better way to keep data within India is to incentivise the creation of data centres and working towards reducing the cost of internet interconnection by encouraging more peering among Internet connectivity providers.

While data mirroring will not help in improving the enforcement of any data protection or privacy law, it will aid Indian law enforcement agencies in gaining easier access to personal data.

The MLAT route

Currently, many forms of law enforcement agency requests for data have to go through onerous channels called 'mutual legal assistance treaties'. These MLAT requests take time and are ill-suited to the needs of modern criminal investigations. However, the US, recognising this, passed a law called the CLOUD Act in March 2018. While the CLOUD Act compels companies like Google and Amazon, which have data stored in Indian data centres, to provide that data upon receiving legal requests from US law enforcement agencies, it also enables easier access to foreign law enforcement agencies to data stored in the US as long as they fulfill certain procedural and rule-of-law checks.

While the Srikrishna Committee does acknowledge the CLOUD Act in a footnote, it doesn't analyse its impact, doesn't provide suggestions on how India can do this, and only outlines the negative consequences of MLATs.

Further, it is inconceivable that the millions of foreign services that Indians access and provide their personal data to will suddenly find a data centre in India and will start keeping such personal data in India.

Instead, a much likelier outcome, one which the Srikrishna Committee doesn't even examine, is that many smaller web services may find such requirements too onerous and opt to block users from India, similar to the way that Indiatimes and the Los Angeles Times opted to block all readers from the European Union due to the coming into force of the new data protection law.

The government could be spending its political will on finding solutions to the law enforcement agency data access question, and negotiating solutions at the international level, especially with the US government. However it is not doing so.

Given this, the recent spate of data localisation policies and regulation can only be seen as part of an attempt to increase the scope and ease of the Indian government's surveillance activities, while India's privacy laws still remain very weak and offer inadequate legal protection against privacy-violating surveillance. Because of this, we should be wary of such requirements, as well as of the companies that are vocal in embracing data localisation.

(Pranesh Prakash is a Fellow at Centre for Internet and Society, and an Affiliated Fellow at Yale Law School's Information Society Project.)

(This opinion was first published on BloombergQuint and has been republished with permission.)