Big Data, Large Concerns

The Supreme Court recently affirmed a fundamental right to privacy. The government is now moving to enact a data privacy law, the Justice Shri Krishna Committee has released a comprehensive report and the consultation process is coming to a close. Multiple legal challenges against Aadhaar — many involving citizens' privacy — are being heard before the Supreme Court. The government is keen to assure the judges and the public that there are enough safeguards to keep the programme legal. But the privacy law will impact more than the future of Aadhaar. It will set the terms on which Indians share intimate data about themselves with both the government and a growing number of private companies.

Already, the Committee might be walking a troublesome path. It has suggested that while in the past "it was possible to limit the collection of data to satisfy a particular purpose", in the era of big data "this may no longer hold true". While the Committee does well to endorse the importance of user consent generally, when it comes to big data they suggest, "consent may not be as relevant". Its effort to distinguish big data's privacy modes from other data instead seems to echo an increasingly popular argument in policy circles globally — what Helen Nissenbaum refers to as "big data exceptionalism". It's the belief that regulating the collection of big data is impossible and undesirable. So the focus should be exclusively on preventing harmful uses and outcomes data. In India, prominent tech lawyer Rahul Matthan has argued that India should adopt an "accountability framework" rather than a consent framework. Why not both?

Big data exceptionalism is an attractive position, no doubt. Creating a regulated process to govern data collection can seem impractical, especially when the data is often an unexpected byproduct of everyday interactions — every step we take with our GPS-enabled phone, every post we "like" on Facebook, every purchase we make, every advertisement we watch. Supporters of big data exceptionalism also make the positive "profit" claim that unfettered data collection can unlock innovation. But this is just as likely to create real threats. Some of our research with Kate Crawford (cited by the Shri Krishna Committee) explores the far reaching consequences of big data and its "predictions" on our personal rights, especially when they are used to decide what to sell us, which businesses will interview us for jobs, and even what news we are allowed to see. A more forensic assessment of the threats of big data exceptionalism is needed.

Firstly, unregulated collection of data dramatically increases the risk of breach. If unlimited quantities of data are gathered and stored — even if they are never analysed or applied to any uses — the risk of a single breach grows with each new wave of data scooped up or shared. The frequency and fallout of data breaches becomes more apparent each day, from Aadhaar in India to Equifax in the US. Second, unregulated data collection opens up new modes of surveillance, both government and corporate, that can have an extreme chilling effect on online freedoms. The European Court of Justice noted, the mere collection of data "is likely to generate in the minds of the persons concerned the feeling that their private lives are the subject of constant surveillance".

The concept of big data exceptionalism abandons two globally recognised privacy principles — principles India should be embracing. First is "collection limitation": The principle that there must be a legal basis for any collection of data. User consent is a powerful basis because we can withdraw it when a company abuses our data or its power over us. Even if we accept that consent fatigue is a reality (no one reads privacy policies), there are other bases to regulate data collection. Indeed, most data protection laws around the world already acknowledge that the benefits of big data can be recognised through other means than consent if the circumstances are appropriate, such as when the data collection is in the vital interest of the individual, or fulfils a legitimate interest of the data controller. So concerns over squelching innovation are likely overblown. The second is "data minimisation", the principle that entities must only collect as much

data as is narrowly tailored to the purpose they seek to achieve, and no more. Mandating data minimisation as a design principle compels inquiry into proportional data collection right at the outset — a philosophy often referred to as privacy by design.

For the sake of citizens' privacy, we hope the committee will not abandon these traditional privacy rules under the cloud of big data exceptionalism. It is easy to remove protections; it is hard to put them back in place.

END