CAFRAL Virtual Conference on Data Protection, Data Privacy and Data Localization

December 27, 2021, Online

Data Protection, Data Privacy and Data Localization –

Current Issues Keynote Address at CAFRAL on December 27, 2021

By Jayant Kumar Dash, Executive Director, Reserve Bank of India

Exactly after one month from today that is on 28th January India will be joining many other nations in celebrating what is called Data Protection Day in Europe and Data Privacy Day in the USA. I am saying this to hint that while the subject matter of today's conference may have three different adjuncts to data, a close look would reveal that they are joined at the hip. India's journey for a data protection framework has genesis in the Supreme Court judgment on the Right to Privacy delivered in August 2017. The judgment articulated the judicial intent of protecting personal information of individuals with the greater objective of protecting civil liberties. India is in the cusp of giving a final form to the Data Protection Bill, 2021, which has been in the works for over three years since Justice Srikrishna Committee report was submitted in July 2018.

2. When the first digital offset printer was invented in 1993, its Israeli inventor Benny Landa had famously said "Everything that can be digital, will be". There has been no stopping to the digitalization, first through evolution and then through revolution. The raw materials, the fuels, the final products, and by-products of this digitalization process has been a word always addressed in plural terms - 'data'. Artificial intelligence, advanced analytics, cloud computing, the Internet of Things are examples of emerging technologies that either rely on or produce data. The type of data produced and consumed in providing financial services today were beyond imagination even a couple of years ago. The management saying "where there is data smoke, there is business fire" will have an entirely different meaning as the regulations around data governance take shape and effect. Banks and other financial service providers and even their regulators / supervisors would need to re-imagine their strategies and re-engineer operations to be on the right side of such regulation. To come back to the subject of the day and to simplify the context, I hope that the main takeaway from this conference could be around how to count the benefits and costs of the impending regime, how it would affect banking business and how to prepare for it. I will try to give conceptual context to some of the current issues surrounding the subject as the detailed deliberations on each component are lined up for discussion in following sessions.

A. Data Localization

3. It may be worthwhile to parse the last piece of the subject first i.e Data localization, being a relatively tangible concept. It refers to hard or soft restrictions placed on the ability of banks to move, store, process or otherwise handle their users' data across the physical borders of a nation. The term 'data sovereignty' is

often used interchangeably with 'data localization', though there are shades of differences. The data sovereignty conceptually refers to control over personal data that was either created or collected in that country. Unlike data localization, data sovereignty doesn't dictate where data must be stored. Instead, it determines who governs and who can access the data once created, stored, processed or collected within a certain region. For example, the US Patriot Act enables US authorities access any information physically found within a server in the United States regardless of where it had originated. The USA also passed the Clarifying Lawful Overseas Use of Data Act (CLOUD Act), in 2018, which allowed US law enforcement agencies to access data stored by cloud providers even if held outside the US.

4. The significance of this policy issue is traditionally argued on the strength of increased security and economic benefits. More granularly, (i) better data access for law enforcement (ii) better enforcement of data protection law (iii) preventing foreign surveillance from national security perspective (e.g data manipulation by third parties like Cambridge Analytical episode) (iv) advancing national economic interest and competitiveness and (v) leveling cross border regulatory grounds around data governance. The case for economic interest is also increasingly gaining weight as data becomes more commodified. Data localisation laws are presented as a means of ringfencing the value in and profits derived from relevant data in the domestic market.

The global big data and business analytics market size[1] was valued at $198.08 billion in 2020, and is projected to reach $684.12 billion by 2030, growing at a CAGR of 13.5% from 2021 to 2030. The share of IT/BPM sector in the GDP of India in FY 2020 was 7.7 percent which was lower than highest of 9.5 percent achieved in FY 2015[2]. No wonder, the number of data-localization measures in force around the world[3] has more than doubled in four years. In 2017, 35 countries had implemented 67 such barriers. Now, 62 countries have imposed 144 restrictions and still counting.

5. In India, sector- and activity-specific data localization measures, such as that for telecom sector or for payment service operators or Video KYC data, have already been implemented. The report of Working Group on Digital Lending, which submitted its report to RBI recently, has also recommended localization of data relating to digital lending apps or platforms. The Personal Data Protection Bill, 2019 introduced India's first

[1] Source: Biga Data and Business Analytics Report by Allied Market Research, September 2021
[2] Source: Statista website
[3] Source: How Barriers to Cross-Border Data Flows Are Spreading Globally, What They Cost, and How to Address Them" by NIGEL CORY AND LUKE DASCOLI | JULY 2021 – Information Technology and Innovation Foundation

economy-wide data localization making it applicable to sensitive personal data and critical personal data with conditions around its transfer outside India. It has now been reported in media of Data Protection Bill, 2021 that the Joint Committee of Parliament has recommended for Government's mandating a mirror copy of all such data available with foreign entities be brough to India in a time bound manner. This will be in addition to developing a comprehensive policy on data localization in consultation with sectoral regulators.

5. The data localization mandate would have several considerations for the banks operating in India, depending on the final provisions of law.

(a) The cost aspects: The banks that have not yet set up storage infrastructure in India, would need to invest to set up local servers. This may have more relevance for existing foreign banks where the scale of operations and future growth strategy may have a bearing on this investment. Some of the foreign banks who are awaiting on their wings for an India entry or expand their presence may have this to factor in. In general terms, one business cost that strict localisation might impose is inability to take advantage of cheaper storage solutions off shore. For example, options like "data sharding", which splits a piece of data into smaller "shards" and stored across multiple systems to minimize costs or improve redundancy mechanisms may vanish from data strategies. Large investments may need to be made by some banks collecting Sensitive Personal Data or Information (SPDI) of data principals within India and banks that have been storing and processing Indian data at facilities outside India may have to either establish or contract for local infrastructure to store such data. Banks may also need to have their offshoring agreements vetted by the DPA before the transfer of data. The dusts need to settle on this front as the PDP, 2019 did not contemplate the use of pre-approved standard contractual clauses.

(b) Global interoperability: Data localisation regimes in different countries following different protocols can potentially create complex compliance problems for global banks. Some quote the inclusion of non-personal data and data flow restrictions in the Data Protection Bill in India as example. The costs of navigating data in motion for mid-air change of protocols for multinational banks or Indian banks with international presence need to be understood, not only from their own perspective but also from that of their service providers as popularity of 'banking as a service' (BaaS) has been gaining currency.

(c) Data localisation laws, by simply locating data in a "safe" home jurisdiction may not afford progressive protection if appropriate technical and governance measures are not implemented by the banks in tandem with the digital technology frontiers. Some also argue that with the realized importance of data, a country with deficient data security infrastructure may run the risk of being a honeypot for data breach attacks. Bank

CISOs may have a role cut out to play here.

(d) As the internet and technology are emerging as the default delivery mode for most of the financial services, emerging data localisation and data sovereignty regimes may create complex problems for banks, particularly in the cloud computing environment and its customers. With a lot of data scattered in the cloud, the challenges that bank CISO may have to meet have to be understood in advance.

(e) In order to mitigate compliance risks, banks may need to have clearly mapped ongoing data streams with up-to-date understanding of data in use, its provenance and its storage location. Data maps may not be permanent records of a bank's information flows; every time a decision-flow process or organisational structure changes, or some data is added in a new format, data maps may continuously need updating to reflect those changes.

(f) Segregation of different types of data for compliance with a new statute - Segregating large volumes of data into SPDI and other personal data, for example, may be practically dufficult and may lead to situations where banks decide that all data be mirrored. The fact that critical personal data cannot be transferred outside India will need further segregation before any data is transferred. For introducing a risk-based model to compliance with a data regime, it is suggested in some quarter that data can be broadly grouped into two categories viz. (i) those associated with corporate operations of the bank such as HR & finance and (ii) customer personal data to be dealt with in accordance with the provisions of the law to simplify the matter.

(g) Compliance with data localisation requirement under Distributed Ledger Technology, with underlying architecture of decentralized data storage, is another area for which some clear understanding will be required.

B. Data Protection and Data Privacy

6. Coming back to the broader theme of today's conference, data protection or data privacy is a branch of data security dealing with proper handling of data mainly covering consent, notice, and regulatory obligations. The privacy landscape has transformed dramatically in the last few years, so much so that privacy has become core to banking business rather than an auxiliary consideration. As the world transitions to a data-driven society, privacy seems set to play an even more important role than it already does. Data Protection laws are complex and are based on broad principles. It may be illuminating to understand the

seven key principles[4] related to the privacy of an individual under the PDP Bill, 2019:

i. Technology agnosticism i.e. the law must be flexible to take into account changing technologies and standards
ii. Holistic application i.e. the law must apply to both private sector entities and government. However, differential obligations may be carved out in the law for certain legitimate State aims.
iii. Informed consent i.e. consent is an expression of human autonomy. For such expression to be genuine, it must be informed and meaningful.
iv. Data minimization i.e. data that is processed ought to be minimal and necessary for the purposes for which such data is sought and other compatible purposes.
v. Controller accountability i.e. the data fiduciary shall be held accountable for any processing of data, whether by itself or entities with whom it may have shared the data
vi. Structured enforcement i.e. enforcement must be by a high-powered statutory authority with sufficient capacity. This must coexist with appropriately decentralized enforcement mechanisms.
vii. Deterrent penalties i.e. penalties on wrongful processing must be adequate to ensure deterrence.

To further stylize the difference between the 'before' and 'after' scenario, borrowing two key General Data Protection Regulation (GDPR) principles may be helpful – 'privacy by design' and 'privacy by default'. Privacy by design means banks design things with privacy at heart of the systems, while privacy by default means that if a customer doesn't elect to do something positive, the default position is that their data will be protected. Privacy by default was not a previous requirement. On the face of a new data regime, banks can be expected to be better placed compared to entities in other segments as data security has been an established part of the extant regulations with RBI taking keen interest progressively. Nonetheless, meeting some of the diverse technical, legal and commercial strictures of any new regime would definitely bring some challenges in plain sight. Retrofitting 'privacy by default' into IT systems of banks have been expensive for some of the European banks.

7. Challenges: There have been extensive discussion on various features and provisions of the proposed Data

[4] https://www.digitalindia.gov.in/writereaddata/files/6.Data%20Protection%20in%20India.pdf

Protection Act in India and I am sure that the participants of this conference would be generally familiar with the discourses. I would rather limit myself to flag certain aspects of a data protection regime that may be relevant for bankers to prime for the day the regime may be upon them.

(a) Legacy Data: Access to data for compliance with new regime may have different stumbling blocks for different banks. For larger and older banks, data amassed over a long period of time still in various file types or held on different legacy IT systems 'not talking to each other', accessing the relevant data for compliance purpose may be complex.

(b) Meeting the Timeline: The PDP Bill, 2019 did not have provisions for phased roll out. The data fiduciaries and data processors would need lead time to make the necessary changes to their policies, infrastructure, processes etc. It has been reported in the media that the Joint Committee of Parliament has recommended an approximate period of 24 months for implementation of any and all the provisions of the Act with a time line of phased roll-out. Once the contour of the Data Protection Bill 2021 is clear, it will be for banks to discover/understand the technical/operational and managerial requirements for compliance of the provisions of the Bill. It is expected that initial clarity on various aspects of the new law may take a while to settle down.

(c) With emergence of every new technology, new ways of collecting and processing data emerge. A smart data privacy laws approach, as hinted earlier, is characterised by rules that are risk-based, technology and sector-neutral and promote the concept of 'Accountability'. Under the principle of Accountability, institutions are encouraged to not only comply, but also be able to exhibit how they comply through effective data governance policies and processes, for example, to conduct data privacy impact assessments, maintain transparency and to avoid/mitigate the risk of harm to individuals through good 'Privacy-by-Design' practices. Some of the emerging technologies for banks that may need attention of banks while buying-in and I would like to draw attention to the following oft-discussed aspects:

(i) Block Chain: In theory, blockchain is meant to decentralise computer networks, allowing equal access and power to the same set of electronic files and software. As the created chains are operated and maintained in a decentralised network, the nodes forming that network may be located in different jurisdictions and can thus be subject to various data protection regulations. This decentralised situation results in a significant burden of verifying compliance of the blockchain-based solution as there is not only one particular data protection regulation to abide by, but potentially many other ones to follow.

- One of the key USP of a blockchain is inalterability of stored data. This clearly has implications for data privacy, particularly where the relevant data is personal data or metadata sufficient to reveal someone's personal details. Data protection regulation may require that personal data be kept up-to-date and accurate or deleted at the discretion of the individual, and the immutability of a blockchain system may pose some challenges for such requirements.

- Decentralised autonomous organisations (DAOs) are usually online/digital organisations operating through the implementation of pre-coded rules maintained on a blockchain platform. The decentralised nature of DAOs presents questions that did arise previously for centralised entities with cognisable legal structure and form.

- Exit assistance by blockchain vendors, if required, might in large part be determined by the specific solution and the extent to which the vendor holds the customer's data and manner of its storage on the blockchain. If the customer does not have its own copy of the data, obligatory data migration assistance may be required to hand over all such data on expiry or termination.

- Insofar as the draft bill in India is concerned, there are a few known challenges that have been publicly in the context of adopting block chain technology.

Firstly, with respect to the Storage Limitation principle, the immutable nature of the technology prevents the data from being deleted once the purpose has been fulfilled.

Secondly, under the decentralised blockchain, determining the exact data collection purpose over a widespread network and keeping check on the data being used only for predefined purposes may be challenging.

Thirdly, implied consent for data sharing may not meet the requirements of consent being clear, through an affirmative action. This gives birth to concomitant regulatory issues over a decentralised system as to who shall oblige with these compliances under the law and who should be made responsible / liable for any lapses in compliance.

(ii) Internet of Things (IoT):

RBI's allowing wearables for payment services is not the only example of IoTs in banking space. Although we are shifting away from bricks and mortars of the past, physical branches or customer service points will still be present in the future and use of IoT technologies for better customer experience may be normal. As the

current Bill haf no provision to keep a check on hardware manufacturers that collect the data through digital devices, the committee has suggested insertion of a new subclause 49(2)(o) enabling DPA to frame regulations to regulate hardware manufacturers and related entities.

(iii) Fog / Edge Computing: In the banking industry, edge computing is becoming a persuasive ingredient for many mission-critical infrastructure. Edge computing is a distributed information technology (IT) architecture, similar to cloud computing, in which client data is processed at the periphery of the network, as close to the originating source as possible. Combined with AI, cloud, and 5G, the potential of edge computing in finance is nigh endless. Some of the edge computing and AI use cases in banking as early adopters of technologies include[5]:

- Hyper-personalization – Bots using natural language processing to interpret and comply with customer information requests as well as perceive basic human emotions and adjust behavior accordingly.
- Retail banking – features like SmiletoPay, whereby a user just smiles to a camera, the AI captures your features and can complete your transactions. This can be applied to retail stores, etc. Edge allows seamless integration with non-banking apps, facial recognition for frictionless payments, and more.
- Corporate banking – customized lending solutions for loans based on microexpression analysis to review loan applications. The entire process is service by an AI-powered virtual adviser.
- Banking security – edge computing delivers low latency analytics that guarantees data sovereignty and security.
- Cybersecurity – real-time geo-location tagging, digital footprint analyzer, suspicious beneficiary detection, microexpression analysis for facial expressions are just some of the use cases banks can consider.

(iv) Data Protection and 5G technology: As the volume and granularity of traffic and location data generated during 5G communications may increase, banks will tend to tailor their virtual network requirements for specific use-cases and more data-driven applications that leverage 5G could lead to a greater volume and

[5] https://techwireasia.com/2021/07/edge-computing-with-ai-brings-real-time-insights-to-banking/

variety of personal data usage. The features of 5G that may have data protection implications include high frequencies with smaller cells, multiple inputs multiple outputs, network slice configuration and mobile edge computing requirements.

(v) The Joint Committee of Parliament have observed that data protection in the financial sector is a matter of genuine concern worldwide, particularly when through the SWIFT network, privacy has been compromised widely. As Indian citizens are engaged in huge cross border payments using the SWIFT, the Committee have recommended that an alternative indigenous financial system should be developed on the lines of similar systems elsewhere such as Ripple (USA), INSTEX (EU), etc. which would not only ensure privacy but also give a boost to the digital economy.

(d) New FinTech players supporting banking business in a growing pace such as Account Aggregators, third party payment service providers or increasing use of APIs for data sharing will have to be analyzed in the context of the upcoming data regime in India. It will not only be ensuring compatibility with other jurisdictions where banks have presence, at the same time compliance with home jurisdiction will assume more importance. The banks will have one more regulator, namely Data Protection Authority (DPA) to deal with. Information sharing under Rights to Information Act and even regulatory / supervisory submission of information may undergo a paradigm change.

8. Costs of Compliance: This may be premature to estimate the costs of compliance of the proposed DPA in India. However, a look at GDPR, may afford some glimpses about the shape of things. The cost of GDPR compliance is expected to be around $8 billion per year for Fortune 500 organizations[6]. GDPR implementation costs for UK banks run to an average of £66 million ($ 88 million), the highest spend of any sector[7]. Certain costs are likely to recurring in long term as the ongoing need to demonstrate compliance will add to the administrative costs by the need to keep a record of processing activities and a log of all data breaches, however minor.

9. Penalties for non-compliance: While the compliance may be costly, non-compliance may be still costlier. Under the GDPR, the EU's data protection authorities can impose fines of up to up to €20 million (roughly $20,372,000), or 4 percent of worldwide turnover for the preceding financial year—whichever is higher. Since the GDPR took effect in May 2018, we've seen over 800 fines issued across the European Economic Area

[6] https://www.tessian.com/blog/biggest-gdpr-fines-2020/
[7] https://www.financialdirector.co.uk/2018/06/21/gdpr-how-is-it-affecting-banks/

(EEA) and the U.K. GDPR fines have ramped up significantly in recent months. The sum total of GDPR fines levied in Q3 2021 hit nearly €1 billion—20 times greater than the totals for Q1 and Q2 2021 combined[8]. Mercifully, the penalties paid by banking companies is minuscular in the whole gamut. Among top 22 cases of top penalties of all time amounting to €1229 million ($ 1647 million), there were only two cases of banks accounting for €11 million ($15 million) only.

C. Conclusion

10. The imminent Data Protection regime in India may not be exactly a revolution for banks as it might be for non-banks. One should not paint this with a single colour of an ever-increasing and stricter regulation going forward. In medium term it may induce some of the old banks to behave more like new age banks replacing their legacy systems with a more connected digital offering. Any new data project would also be designed with a 'privacy by default' architecture. Compliance with the new data regime will add to the brand equity of the early adopter banks. It can actually lead to increased level of customer engagement and trust. The challenge before the banking industry in India would be to keep ahead of technological advances from data privacy perspective, while keeping the customer at the heart of it.

Thank you and wish the conference a very productive day.

[8] https://www.tessian.com/blog/biggest-gdpr-fines-2020/