UNDERSTANDING HARM FROM PERSONAL DATA PROCESSING ACTIVITIES AND ITS CHALLENGES FOR USER PROTECTION

[*]Srikara Prasad

ABSTRACT

This article seeks to understand the nature of harm emerging from personal data processing activities and its impact on user protection. To this end, the article analyses how personal data-related harms are unique and distinct from other kinds of harms arising from torts, breach of contract, crime or defects in products and services. Then, it explores how these unique features of personal data-related harm could pose practical challenges to existing and upcoming frameworks for user protection and redress in relation to personal data. In doing so, the article analyses how the treatment of 'Harm' under the Personal Data Protection Bill, 2019 risks weakening user protection measures in the Bill. Finally, the article uses the analysis to justify an alternative definition for 'Harm' that would allow future jurisprudence to develop on a case-by-case basis until there is a firmer understanding of how to address the challenges posed by personal data-related harms.

I. INTRODUCTION

The digital economy in India has seen tremendous increase in the user[1] activity over the past few years. A wide bouquet of services, including essential services like finance and social security schemes, are now being delivered digitally. This has made many services accessible and convenient for users. For example, mobile banking is allowing users to make banking transactions from their homes, reducing their reliance on physical bank

[*] Srikara Prasad, Policy Analyst, Future of Finance Initiative at Dvara Research. He was assisted by Ms. Hussanpreet Kaur, II Year Student, Rajiv Gandhi National University of Law, Punjab.

[1] The word 'Consumer' is useful when referring to persons who transact with service providers for commercial purposes. This is different from 'Citizen' who transacts with service providers for social benefits, such as those under various Direct Benefit Transfer schemes. It would be inappropriate to conflate 'Consumers' and 'Citizens' because both have a different set of rights and remedies under the law. This paper will use the word 'User' to refer to both, consumers and citizens who participate in the digital economy for various reasons.

branches in their vicinity. But more importantly, the digital economy has redefined the role of users' personal data in user well-being.

Users generate vast amounts of personal data in the digital economy. This data can help providers understand user preferences better and design products and services better suited to the users' benefit. On the flipside, the mishandling of personal data can be detrimental to users' interests and their right to privacy.[2] The different kinds of risks and harms[3] that arise for users from personal data processing activities and personal data breaches (that this paper addresses as "**personal data-related harm**") can be very difficult to identify and predict.[4] Personal data-related harm, in this sense, is unique and distinct from the general understanding of harm arising from torts, breach of contract, crime or defects in products and services. This unique nature of personal data-related harm could pose difficult practical challenges to user protection under existing and upcoming personal data-related frameworks in India.

Following the Supreme Court's landmark judgment on the right to privacy in *Justice K.S. Puttaswamy v. Union of India*,[5] the Central Government set out to enact a personal data protection law. The ensuing draft of the Personal Data Protection Bill, 2019 ("**the Bill**") seeks to create an elaborate framework to regulate personal data processing[6] activities in India and protect users' interests from harm *ex ante* i.e. preventing harm before it can occur instead of seeking to remedy harm after it occurs.[7] However, one of the most concerning features of this framework is its definition of 'Harm'.[8] The definition prescribes a list of ten outcomes that must be treated as

---

[2] S. Prasad, *Defining 'Harm' in the digital ecosystem*, Dvara Research, available at https://www.dvara.com/blog/2019/05/06/defining-harm-in-the-digital-ecosystem/, last seen on 29/11/2020.

[3] The article uses 'Harm' to refer to harm as defined in the Personal Data Protection Bill, 2019. The word 'Harm' is used for all other purposes.

[4] Supra 2.

[5] Justice K. S. Puttaswamy (Retd.) v. Union of India, (2017) 10 SCC 1.

[6] This article relies on the definition of 'Processing' under the draft Personal Data Protection Bill, 2019 that defines 'Processing' under S. 3 (31) as: *"an operation or set of operations performed on personal data, and may include operations such as collection, recording, organization, structuring, storage, adaptation, alteration, retrieval, use, alignment or combination, indexing, disclosure by transmission, dissemination or otherwise making available, restriction, erasure or destruction."*

[7] Preamble, The Personal Data Protection Bill, 2019 (pending).

[8] S. 3 (20), The Personal Data Protection Bill, 2019 (pending).

harm from personal data processing activities. The definition does not consider the unique nature of personal data-related harm and risks creating practical challenges for user protection and redress in the Bill. These challenges could also arise under other user protection frameworks relating to personal data in India including common law, the Code of Civil Procedure, 1908 ("**the CPC**"), the Information Technology Act, 2000 ("**IT Act**") and the Consumer Protection Act, 2019 ("**the CPA**").

In this context, it is important to understand what makes personal data-related harm unique and what practical challenges they pose for user protection. This article first provides an overview of the role that personal data plays in the digital economy and the benefits and risks that ensue for users (Section 2). Then, it discusses how personal data-related harm is unlike other kinds of harm (Section 3). This is followed by an analysis of the practical challenges that the uniqueness of personal data-related harm creates for user protection under existing and upcoming user protection and redress frameworks (Section 4). Lastly, the article uses this analysis to justify the need for and propose an alternative definition of 'Harm' that would be more suitable for user protection in data processing activities (Section 5).

II. THE ROLE OF PERSONAL DATA IN THE DIGITAL ECONOMY: BENEFITS & RISKS TO USERS

1. Benefits to Users

Users share a lot of personal data when they participate in the digital economy.[9] For example, a user, who creates a profile on a social media platform, shares his name, date of birth, address, E-Mail addresses, preferences, etc. when they sign up. Users further generate personal information when they engage with content on the social media platform signaling what they like, what content they view, who they befriend, etc. Similarly, a user who makes digital financial transactions (like payments)

[9] *The value and role of data in electronic commerce and the digital economy and its implications for inclusive trade and development note by the UNCTAD Secretariat*, U.N. Document TD/B/EDE/3/2, available at https://unctad.org/system/files/official-document/tdb_ede3d2_en.pdf, last seen on 04/12/2020.

has to share his personal information like their Permanent Account Number ("**PAN**") number, Aadhaar number and biometric identifiers to fulfill the financial institution's Know-Your-Customer ("**KYC**") requirements.[10]

Users further generate personal information when they make transactions signaling what they buy, how often they buy etc. Users generate vast amounts of personal data through their activities in the digital economy, some of which (such as health data and financial data) could be highly sensitive in this manner.[11] Developments in data processing techniques and abilities have enabled providers to process these vast amounts of personal data popularly known as big data analytics, and glean valuable insights about users.[12] Providers can use these insights and create products and services that can better match users' preferences. For instance, a food delivery platform could process a user's food order history on the platform to learn their preferences and curate a bespoke list of restaurants for that user.[13] On a larger scale, providers could aggregate users' personal data from different sources to create an aggregated dataset and generate a digital profile that can help in the delivery of services.[14] For example, financial institutions can use such aggregated datasets to create a user's psychometric profile and assess the user's credibility before sanctioning a loan.[15]

---

[10] *Know Your Customer (KYC) Direction, 2016*, RBI Master Direction No. DBR.AML.BC.No.81/14.01.001/2015-16 (25/02/2016), available at https://rbidocs.rbi.org.in/rdocs/notification/PDFs/MD18KYCF6E92C82E1E1419D87323E3869BC9F13.PDF, last seen on 04/10/2020.

[11] Ministry of Electronics and Information Technology, Government of India, *A Free and Fair Digital Economy: Protecting Privacy, Empowering Indians*, available at https://www.meity.gov.in/writereaddata/files/Data_Protection_Committee_Report.pdf, last seen on 30/11/2020.

[12] *Big Data in Action for Government*, World Bank Group, available at https://openknowledge.worldbank.org/bitstream/handle/10986/26391/114011-BRI-3-4-2017-11-49-44-WGSBigDataGovernmentFinal.pdf?sequence=1&isAllowed=y, last seen on 04/12/2020.

[13] E. Knight, *It's the Algorithm, It Decides: An Autoethnographic Exploration of Algorithmic Systems of Management in On-Demand Food Delivery Work in Amsterdam*, available at https://www.google.com/url?sa=t&rct=j&q=&esrc=s&source=web&cd=&ved=2ahUKEwjQx-bR9LTtAhVIzTgGHep_DekQFjAAegQIAxAC&url=https%3A%2F%2Fscripties.uba.uva.nl%2Fdocument%2F674434&usg=AOvVaw1ZpDVb47Ybf_WU7UgFiaio, last seen on 04/12/2020.

[14] *Examples of Data Points Used in Profiling*, Privacy International, available at https://privacyinternational.org/sites/default/files/2018-04/data%20points%20used%20in%20tracking_0.pdf, last seen on 04/12/2020.

[15] Financial institutions rely on a person's credit reports and credit score to assess a person's creditworthiness i.e., their eligibility for a loan. Traditionally, credit reports are

Providers can harness users' personal data towards users' welfare in this manner.

2. Risks and Harm to Users

On the flipside, the use of personal data can expose users to the risk of harm.[16] It is well-established that a variety of harms can emerge when personal data is breached, misused or processed improperly;[17] ranging from mental agony to financial loss, discrimination, surveillance and user manipulation.[18] These harms can be amplified forms of existing and known harms or be newer and inchoate forms of harm (like user manipulation)[19] that emerge with developments in technology and use-cases for personal data.[20] Personal data-related harms could be categorized into two broad categories: primary harms and secondary harms.[21] Primary harms emerge when personal data is breached. It refers to the immediate outcomes of a

based on a person's credit history– previous loans sanctioned, repayment history, defaults etc. However, emerging practices are using different kinds of data including a person's utility bill payments, mobile phone activity etc. to assess the person's creditworthiness. See, A. Singh & S. Prasad, *Artificial Intelligence in Digital Credit in India*, Dvara Research, available at https://www.dvara.com/blog/2020/04/13/artificial-intelligence-in-digital-credit-in-india/, last seen on 22/11/2020. Also see, *High Level Task Force on Public Credit Registry*, Report of the High Level Task Force on Public Credit Registry, available at https://www.rbi.org.in/Scripts/PublicationReportDetails.aspx?UrlPage=&ID=895, last seen on 1/12/2020.

[16] Supra 11.

[17] Supra 2; B. Chugh & N. Kumar, *Harms to Consumers in a Modular Financial System*, Dvara Research, available at https://www.dvara.com/blog/2017/11/07/harms-to-consumers-in-a-modular-financial-system/ , last seen on 24/11/2020; R. Calo, *The Boundaries of Privacy Harm*, 86(113), Indiana Law Journal 1131, 1142-1156 (2011), available at http://ilj.law.indiana.edu/articles/86/86_3_Calo.pdf, last seen on 29/11/20.

[18] Supra 2.

[19] For example, the Cambridge Analytica incident showed how users' personal data could be used to understand and manipulate users' mental models. Insights derived on users seem to have helped in manipulating user preferences and behaviours across the world, including in India. See Cambridge Analytics: The data firm's global influence, BBC, available at https://www.bbc.com/news/world-43476762, last seen on 29/11/2020.

[20] Supra 2.

[21] Ibid.

personal data breach such as privacy infringement,[22] financial loss,[23] harassment,[24] etc.[25] Secondary harms arise when personal data, breached or otherwise, is processed.[26] These harms are a result of processing activity

[22] For example, big data techniques have enabled curation of large, aggregated datasets that can provide a birds-eye view on different groups of users. A breach of any one user's personal data in these datasets' risks infringing the privacy of all other similar users in the dataset collectively. See, B. Chugh & N. Kumar, *Harms to Consumers in a Modular Financial System*, Dvara Research, available at https://www.dvara.com/blog/2017/11/07/harms-to-consumers-in-a-modular-financial-system/, last seen on 24/11/2020. Also see, L. Taylor, L. Floridi & B. van der Sloot, *Safety in Numbers? Group privacy and Big Data Analytics in the Developing World, 24, in Group Privacy: New Challenges of Data Technologies* (Taylor, L. Floridi & B. van der Sloot, eds. 2017). Further, providers' abilities to process user activity and glean insights could also infringe users' privacy. For instance, in 2012, a New York Times report revealed how Target would analyses users' purchases and understand if the user is pregnant. See, C. Duhigg, *How Companies Learn Your Secrets*, New York Times, available at https://www.nytimes.com/2012/02/19/magazine/shopping-habits.html?pagewanted=1&_r=1&hp, last seen on 29/11/2020.

[23] A breach of sensitive datasets, such as those containing account credentials or biometric information, could similarly make users vulnerable to phishing, fraud etc. For example, see *Timeline of Cyber Incidents Involving Financial Institutions,* Carnegie Endowment for International Peace, available at https://carnegieendowment.org/specialprojects/protectingfinancialstability/timeline, last seen on 28/11/2020.

[24] For example, certain digital lending mobile applications access users' personal data on the mobile phone such as contacts, photos etc. as collateral for credit (sometimes without providing due notice to the user). Reports suggest that at the time of loan recovery, the lenders have threatened to leak sensitive images, or have contacted users' social and professional circles to name and shame the user. See, S. Ghosh, *App Lenders Scar Young Debtors*, Livemint, available at https://www.livemint.com/companies/news/app-lenders-scar-young-debtors-11600132261735.html, last seen on 29/11/2020. Also see, P. Mallikarjunan, *How App-Based Lenders are Harassing, Sucking Borrowers Dry*, Moneylife, available at https://www.moneylife.in/article/how-app-based-lenders-are-harassing-sucking-borrowers-dry/60621.html, last seen on 29/11/2020; and D. Sharma, *Shocking Tactics of Lenders Exposed! Loan Recovery Agents Blackmailing Customers After Hacking Phones*, Mid-day, available at https://www.mid-day.com/articles/shocking-tactics-of-lenders-exposed-loan-recovery-agents-blackmailing-customers-after-hacking-phones/22896205, last seen on 29/11/2020.

[25] K. Kemp, *Big Data, Financial Inclusion and Privacy for the Poor*, Dvara Research, available at https://www.dvara.com/blog/2017/08/22/big-data-financial-inclusion-and-privacy-for-the-poor/, last seen on 29/11/2020.

[26] Supra 2.

and can include identity theft,[27] discrimination,[28] exclusion,[29] inaccurate profiling and surveillance.[30]

Personal data-related harms can also be tangible or visceral. Seminal literature on harm related to privacy and personal data by Ryan Calo (2011) classifies harms into two categories: subjective harms and objective harms.[31] Subjective harms refer to harms that are "internal to the person harmed". They refer to "unwelcome mental states" such as anxiety, embarrassment and fear that a user may experience due to unwanted observation. Objective harms refer to harms that are "external to the person harmed". They stem from the "unanticipated or coerced use of information concerning a person against that person" that result in an external effect such as financial loss. It is important to note that these categories are not watertight. Subjective harms could arise due to the

[27] The advent of DeepFakes i.e., highly realistic media synthesized by processing visual or audio data on users using artificial intelligence systems creates a new frontier for identity theft. There have already been instances where DeepFakes have made people vulnerable to extortion and blackmail. See, D.K. Citron & R. Chesney, *Deep Fakes: A Looming Challenge for Privacy, Democracy, and National Security*, 107 California Law Review 1753, 1772 (2019), available at https://scholarship.law.bu.edu/cgi/viewcontent.cgi?article=1640&context=faculty_scholarship, last seen on 24/11/2020.

[28] For example, an investigative report published in April 2016 showed that Amazon, Inc's Prime Free Same-Day deliveries service excluded areas in cities which were predominantly occupied by the African American community. See D. Ingold & S. Soper, *Amazon Doesn't Consider the Race of Its Customers. Should it?* Bloomberg, available at https://www.bloomberg.com/graphics/2016-amazon-same-day/, last seen on 28/11/2020. More recently, Apple, Inc's credit card algorithm sanctioned different amounts of credit for men and women, offering smaller credit amounts to women. See, W. Knight, *The Apple Card Didn't 'See' Gender – and That's the Problem*, Wired, available at https://www.wired.com/story/the-apple-card-didnt-see-genderand-thats-the-problem/, last seen on 28/11/2020.

[29] For example, many public services are increasingly being delivered digitally based on users' Aadhaar numbers. However, data quality issues in the Aadhaar database (such as spelling mistakes in names, biometric mismatches etc.) have led to users being unable to access public services. See, R. Khera, *Aadhaar Failures: A Tragedy of Errors*, EPW Engage, accessible at https://www.epw.in/engage/article/aadhaar-failures-food-services-welfare, last seen on 28/11/2020.

[30] Supra 25; D.J. Solove, *A Taxonomy of Privacy*, 154(3) University of Pennsylvania Law Review 477, 491-553 (2006), available at https://www.law.upenn.edu/journals/lawreview/articles/volume154/issue3/Solove154U.Pa.L.Rev.477(2006).pdf, last seen on 29/11/2020; B. Chugh & N. Kumar, *Harms to Consumers in a Modular Financial System*, Dvara Research, available at https://www.dvara.com/blog/2017/11/07/harms-to-consumers-in-a-modular-financial-system/, last seen on 24/11/2020; L. Taylor, L. Floridi & B. Van Der Sloot, *Safety in Numbers? Group Privacy and Big Data Analytics in the Developing World, 24, in Group Privacy: New Challenges of data technologies* (Taylor, L. Floridi & B. van der Sloot, eds. 2017).

[31] R. Calo, *The Boundaries of Privacy Harm*, 86(113), Indiana Law Journal 1131, 1142-1156 (2011), available at http://ilj.law.indiana.edu/articles/86/86_3_Calo.pdf, last seen on 29/11/20.

perceived risk of objective harms emerging from personal data. However, this distinction is useful in understanding the different dimensions of personal data-related harms: one dimension which is purely internal to a user and another that is external. While objective harms would be visible and objectively verifiable, subjective harms would be visceral and difficult to verify.

Harm is usually understood to be a setback to a person's interests.[32] However, the unique nature of personal data-related harms makes it difficult to identify setbacks to users' interests. The following section analyses how harms relating to personal data are unique and distinct from other kinds of harm.

III. THE UNIQUENESS OF PERSONAL DATA-RELATED HARMS

Although it is clear that mishandling personal data can lead to harm, it is very difficult to understand 'how' and 'when' harm will occur.[33] For instance, once personal data is breached, it is extremely difficult to identify who has access to the personal data, understand what it might be used for or predict when it might be used. The entities breaching personal data could process personal data for themselves or share it with other entities. They could process personal data that they obtained from a breach or they could process personal data after aggregating different datasets obtained from other sources. The entities could process the personal data immediately after the breach or after many weeks, months or years. A similar problem exists in relation to cases of improper processing of personal data. As seen in the case of the Apple Card where the credit decisioning algorithm inconceivably discriminated against female users,[34]

---

[32] S. Perry, *Harm, History, and Counterfactuals*, 40, San Diego Law Review, 1283, 1284-1285 (2003), available at https://www.law.upenn.edu/cf/faculty/sperry/workingpapers/B40SanDiegoLR1283(2003).pdf, last seen on 04/12/2020; D.J. Solove & D. Citron, *Risk and Anxiety: A Theory of Data Breach Harms*, GWU Law School Public Law Research Paper No. 2017-2, 6-7, GWU Law School Public Law Research Paper No. 2017-2 (2017), available at https://scholarship.law.gwu.edu/cgi/viewcontent.cgi?article=2499&context=faculty_publications, last seen on 04/12/2020.

[33] Supra 2.

[34] W. Knight, *The Apple Card Didn't 'See' Gender – and That's the Problem*, Wired, available at https://www.wired.com/story/the-apple-card-didnt-see-genderand-thats-the-problem/, last seen on 28/11/2020.

even *bona fide* processing of personal data could eventually have a *mala fide* impact on the users. In this sense, harms that can arise once personal data is processed or breached can neither be identified nor be predicted with certainty, making them indiscernible and unpredictable.[35] Therefore, a user, provider or a regulator might not even know if a breach of a processing activity has caused a setback to users' interests.

Harms arising from torts, breach of contract, crime, deficiencies in products and services etc., are comparatively a lot simpler to identify and predict. For instance, consumer protection laws have benefited from the reasonable certainty about what is "harmful" for consumers. The Consumer Protection Act, 2019 ("**CPA**") defines harm under S. 2(22):

> 'Harm', in relation to product liability, includes –
> *(i)* damage to any property, other than the product itself;
> *(ii)* personal injury, illness or death;
> *(iii)* mental agony or emotional distress attendant to personal injury or illness or damage to property; or
> *(iv)* any loss of consortium or services or other loss resulting from a harm referred to in sub-clause (i) or sub-clause (ii) or sub-clause (iii), but shall not include any harm caused to a product itself or any damage to the property on account of breach of warranty conditions or any commercial or economic loss, including any direct, incidental or consequential loss thereto;

This definition prescribes what 'Harm' is, based on consumer outcomes emerging from defects[36] and deficiencies[37] in products and services, misleading advertisements,[38] restrictive trade practices[39] and unfair trade practices (which are central to the consumer protection framework in the CPA).[40] Having such a clear and precise definition of 'Harm' allows stakeholders to identify harm with certainty and allows consumers to ask for redress on clear and justifiable grounds. However, the indiscernibility and unpredictability of personal data-related harms makes it impossible to define what harm is, in personal data processing activities. This problem is apparent in the draft Bill, which by attempting to prescribe a specific

[35] Supra 2.
[36] "Defect" is defined under S. 2 (10) of the Consumer Protection Act, 2019.
[37] Ibid, S. 2(11).
[38] Ibid, S. 2(28).
[39] Ibid, S. 2(41).
[40] Ibid, S. 2(47).

definition for 'Harm',[41] creates a definition that is extremely broad, vague and difficult to interpret.

The Bill defines 'Harm' under S. 3(20):

> "'Harm' includes –
> *(i)* bodily or mental injury;
> *(ii)* loss, distortion or theft of identity;
> *(iii)* financial loss or loss of property;
> *(iv)* loss of reputation or humiliation;
> *(v)* loss of employment;
> *(vi)* any discriminatory treatment;
> *(vii)* any subjection to blackmail or extortion;
> *(viii)* any denial or withdrawal of a service, benefit or good resulting from an evaluative decision about the data principle;
> *(ix)* any restriction placed or suffered directly or indirectly on speech, movement or any other action arising out of a fear of being observed or surveilled; or
> *(x)* any observation or surveillance that is not reasonably expected by the data principal;"

This provision provides a list-based definition of 'Harm'. The definition provides ten outcomes that would constitute 'harm' under the Bill's framework. The use of the word 'includes' also suggests that this is an inclusive and non-exhaustive definition that could be expanded further based on the rule of *ejusdem generis* (meaning "of the same kind" in Latin) in statutory interpretation.[42] The *ejusdem generis* rule is helpful in interpreting definitions that contain a set of items. The rule allows expanding a list of items in a definition to include other items if they are of the same kind.[43] For instance, if a provision defines a term as "including, apples, mangoes, pineapples and banana", the provision could be expanded to include other fruits like guava and watermelon. The rule relies on a common thread running through a provision to provide guidance in identifying similar items that can be added to the scope of the provision (and dissimilar items that cannot).

[41] Supra 8.
[42] Ejusdem generis, Legal Information Institute, available at https://www.law.cornell.edu/wex/ejusdem_generis, last seen on 04/12/2020.
[43] Ibid.

Analyzing the Bill's definition of 'Harm' through this lens shows that the definition lacks a common thread passing through the ten outcomes.[44] The definition seems to include (i) all kinds of physical, mental and emotional injury (ii) all kinds of injury to property and reputation (iii) all kinds of interference with constitutional rights and (iv) other distinct outcomes that are not strictly covered by any of the other categories of outcomes. There is no guidance in the provision about how to include and exclude new outcomes into the list in the definition under the rule of *ejusdem generis*. The scope of the definition is extremely wide and could cover almost any outcome and treat it as harm under the Bill.[45] Therefore, in attempting to list all the potential harms that could emanate from the processing of personal data, the Bill creates a very vague and open-ended definition that is very difficult to interpret.[46] As a result, the challenges in identifying harm remain unaddressed.

Further, high amounts of information asymmetry between users and providers (who process personal data) intensifies the problem of identifying personal data-related harms. Millions of users generate many exabytes of personal data that is processed by thousands of entities in a digital economy. However, users rarely have an understanding of the entities which have their personal data and what it is being processed for.[47] This information asymmetry is sharpened by data breaches which cast shadows over how personal data flows in the economy. When harm materializes in this context, the information gaps could make it close to impossible for users to demonstrate a causal link between the harm and a breach or personal data processing activity by a particular entity.[48]

---

[44] S. Prasad, *An Analysis of 'Harm' defined under the draft Personal Data Protection Bill, 2018*, Dvara Research, available at https://www.dvara.com/blog/2019/10/29/an-analysis-of-harm-defined-under-the-draft-personal-data-protection-bill-2018/, last seen on 29/11/2020.

[45] Ibid.

[46] Supra 44.

[47] Supra 11.

[48] *See* D.J. Solove & D. Citron, *Risk and Anxiety: A Theory of Data Breach Harms*, GWU Law School Public Law Research Paper No. 2017-2, 1, GWU Law School Public Law Research Paper No. 2017-2, George Washington University Law School, (2017), available at https://scholarship.law.gwu.edu/cgi/viewcontent.cgi?article=2499&context=faculty_publications, last seen on 04/12/2020.

These challenges with personal data-related harms: indiscernibility, unpredictability and the difficulty in demonstrating causation, make them unique and distinct from other kinds of harm and from the general understanding of harm. The following section discusses the challenges that could arise for user protection and redress, given these characteristics of personal data-related harms.

IV. PERSONAL DATA-RELATED HARM: CHALLENGES TO USER PROTECTION AND REDRESS IN INDIA

The unique characteristics of personal data-related harm create some practical challenges for user protection and redress in India. This section first introduces the existing frameworks for user protection and redress in India in relation to personal data, and then analyses the challenges they face due to personal data-related harms. Then, it introduces the upcoming framework for user protection and redress under the Bill and analyses how personal data-related harm could challenge this framework. In doing so, the section explores how existing and upcoming frameworks could be inadequate or unsuitable for protecting users and providing redress in personal data-related harms.

1. Challenges for User Protection and Redress Under Existing Personal Data Protection Frameworks: India

The prominent legal frameworks in force today that are applicable to user protection in data processing activities include:

i. Torts and common law: Under common law, aggrieved users can approach the court seeking redress when they suffer injury due to an act, omission or negligence of another person or entity.[49]

ii. The IT Act, 2000: Providers (termed as "Body corporates" in the Act) can be held liable for negligence if they do not adopt reasonable security practices and as a result lead to a wrongful loss

[49] Smt. Anguri v. Jiwan Dass, (1988) 4 SCC 189.

to users. Such providers are liable to compensate users for the wrongful loss suffered.[50]

iii. The CPA, 2019: The CPA protects consumers against unfair trade practices[51], which include unlawful disclosure of personal data.[52]

iv. The CPC, 1908: The CPC prescribes procedures for users seeking redress in civil cases before civil courts. Such cases would include breach of contracts, including privacy agreements between providers and users, which define the standard of care that providers owe to users with respect to their personal data.

These frameworks do not regulate personal data processing activities. They provide users redress in cases where users suffer personal data-related harms. Users must fulfil two preconditions before they can obtain redress under these frameworks. It is an established common law principle that aggrieved users (plaintiffs) must be able to prove the cause of action and an injury in order to obtain redress.[53] Injuries must be caused by an act or omission of another person or entity. A "cause of action" establishes a causal link between an act or omission and an injury. A plaintiff's claim is actionable only when they can establish a cause of action. However, the causal link should not be remote. The plaintiff should be able to establish to a reasonable degree that the act or omission led to the injury.[54]

Injury is defined as a legal damage, or an interference with a legal right or damage arising from breach of a legal duty. This would include physical injuries, mental injuries, damage to property, damage to reputation, infringement of other legal rights etc. Other kinds of damage, where there is no interference with rights or breach of legal duties, are not actionable.[55]

[50] S. 43A, The Information Technology Act, 2000.
[51] 'Unfair trade practice' is defined under s.2(47) of the Consumer Protection Act, 2019.
[52] S.2 (47), The Consumer Protection Act, 2019.
[53] Alston v. Marine and Trade Insurance Company Ltd. [1964] 4 SA 112 (Witwatersrand Local Division of High Court of South Africa).
[54] Haynes v. Harwood (1935) 1 KB 146 (1935, Court of Appeal of England and Wales); Overseas Tankship (UK) Ltd v. Morts Dock and Engg. Co. Ltd, (1961) AC 388 (1961, Judicial Committee of the Privy Council).
[55] Bhim Singh v. State of J&K (1985) 4 SCC 677.

Similarly, under the CPC, plaintiffs must be able to establish the facts demonstrating a cause of action and an injury.[56] Under the CPA, complainants[57] must be able to establish defects and deficiencies in goods and services, unfair contractual terms and unfair trade practices that cause an injury when they make a complaint seeking relief.[58] Under the IT Act, plaintiffs must be able to establish that Body Corporates did not adopt reasonable security practices because of which they suffered wrongful loss.[59]

Personal data-related harms pose two broad challenges to user protection and redress under these frameworks.

1.1. Users Could Find it Difficult to Establish a Cause of Action

The first challenge that arises is users being unable to establish a cause of action and prove that injury was a consequence of a breach of processing activity. Admittedly, under common law, plaintiffs can hold defendants liable for negligence (misfeasance) and for not providing (nonfeasance) the standard of care that they owe to plaintiffs.[60] This is also embodied under S. 43A of the IT Act. In the context of personal data processing, users may be able to convincingly argue that providers were negligent and failed to provide the standard of care required in adopting necessary security measures to protect personal data from breaches.

However, it would be difficult for users to make this argument for secondary harms emerging after a data breach. Under established principles of common law, defendants would not be liable for consequences that were too remote from the defendants' acts or omissions i.e. if an injury cannot be reasonably connected to the negligent activity.[61] As mentioned above, information asymmetries that users face in the digital economy

[56] Rule 1, Order VII, The Code of Civil Procedure, 1908.
[57] "Complainant" is defined under S. 2(5) of the Consumer Protection Act, 2019 as (i) a consumer (ii) any voluntary consumer association registered under any law for the time being in force (iii) the Central Government or any State Government (iv) the Central Authority under the Act (v) one or more consumers (vi) legal heir or representative of a deceased consumer or (vii) a parent or guardian of a minor.
[58] Supra 35, S. 2(6).
[59] S. 43A, The Information Technology Act, 2000.
[60] Poonam Verma v. Ashwin Patel, AIR 1996 SC 2111.
[61] Bourhill v. Young (1943) AC 92 (1943, House of Lords).

could make it very difficult for users to demonstrate a causal link between a breach or processing activity and harm.[62] The difficulty could be especially higher in case of secondary harms pursuant to data breach, considering that the harms could have emerged from a breach of the users' personal data from another entity.[63] Similarly, complaints made under the CPA for unlawful disclosure of personal data or through the CPC for breach of privacy agreements may only provide redress for the immediate breach of personal data or failing to provide the standard of care defined under contract, but also for secondary harms.

1.2. Users Could Find it Difficult to Show Injury

The second challenge that arises is users being unable to show injury when personal data is breached or processed improperly. Personal data breaches and improper processing of personal data can make users vulnerable to harm and increase risk of harm. However, harm would only be speculative until it actually materializes. Users whose personal data is breached or processed improperly could speculate that they will be harmed, but they cannot be certain about how and when they will be harmed. Users could find it difficult to obtain redress based on such speculations of harm.

In their scholarly work on personal data-related harms "Risk and Anxiety: A Theory of Data Breach Harms," Daniel Solove and Danielle Citron (2017) present a trend of cases on speculation of harm due to data breaches that were adjudicated by courts in the United States of America ("**USA**").[64] These cases cover three overarching arguments from users seeking redress: "(i) data breaches create a risk of future injury, (ii) plaintiffs take preventative measures to reduce risk of injury, and (iii) plaintiffs experience anxiety as a result of data breaches compromising their personal data."[65] The courts seem to have rejected the first argument (risk of future injury) for being too speculative even in cases where personal data was breached by thieves. The courts seem to have rejected redress when they were unable

[62] Supra 11; See Supra 48.
[63] Supra 2; D.J. Solove, *Privacy and Data Security Violations: What's the Harm?*, TeachPrivacy, available at https://www.linkedin.com/pulse/20140625045136-2259773-privacy-and-data-security-violations-what-s-the-harm/, last seen on 30/11/2020.
[64] Supra 48.
[65] Ibid at 9.

to find immediate harm from the breach or when the period when harm could occur was unascertainable, even if the risk of harm increased greatly.[66] The second argument (preventative measures to reduce risk) seems to have been rejected by the courts on the ground that users could use such measures to 'manufacture' injury.[67] The courts seem to have rejected the third argument (anxiety as a result of a data breach) on grounds that anxiety stemmed from the speculation of harm, and not on certainty of impending harm.[68] The argument seems to have been rejected unless plaintiffs were able to demonstrate impending harm and emotional distress.[69] These judgements are useful in understanding the difficulties that users in India could face in demonstrating injury.

Therefore, although users could hold providers liable for personal data breaches under the existing frameworks, users may not be able to establish a cause of action and demonstrate injury to get redress for secondary harms.

2. Challenges for User Protection and Redress Under the Upcoming Personal Data Protection Bill, 2019

The Bill seeks to create a distinct and elaborate framework to regulate personal data processing[70] activities in India and protect users' interests

---

[66] Ibid at 10; Key v. DSW, Inc., 454 F. Supp.2d 684 (2006, United States District Court, S.D. Ohio, Eastern Division); In re Science Applications International Corp (SAIC) Backup Tape Data Theft Litigation, No. MDL 2360, 2014 U.S. Dist. LEXIS 64125 (D.D.C, May 9, 2014). The courts accepted the argument in cases where the harm was "immediate and very real," such as when unauthorized transactions appeared in users' bank accounts (See Remijas v. Neiman Marcus Group (7th Cir. 2015)) or when new bank accounts were attempted to be opened post a data breach (See Krottner v. Starbucks Corp., 628 F.3d 1139 (2010, United States Court of Appeals, Ninth Circuit)).

[67] Supra 48; Polanco v. Omnicell, Inc (2013, United States District Court of New Jersey).

[68] Ibid; Crisafulli v. Ameritas Life Insurance Co., 2015 WL 1969173 (2015, United States District Court of New Jersey); In re Barnes & Noble Pin Pad Litigation, No. 12-cv-8617, 2013 WL 4759588 (2013, United States District Court for the Northern District of Illinois Eastern Division); Amburgy v. Express Scripts, Inc., 671 F. Supp. 2d 1046 (2009, United States District Court, E.D. Missouri, Eastern Division).

[69] Crisafulli v. Ameritas Life Insurance Co., 2015 WL 1969173 (2015, United States District Court of New Jersey); Maglio v. Advocate Health & Hospitals., 49 N.E. 3d 746, 755 (Ill. App. 2015).

[70] This paper relies on the definition of 'Processing' under the draft Personal Data Protection Bill, 2019 that defines 'Processing' under S. 3 (31) as: "*an operation or set of operations performed on personal data, and may include operations such as collection, recording, organization, structuring, storage, adaptation, alteration, retrieval, use, alignment or combination, indexing, disclosure by transmission, dissemination or otherwise making available, restriction, erasure or destruction*".

from harm *ex ante*.[71] The framework prescribes obligations for data fiduciaries,[72] rights for data principals[73] and powers and functions of the proposed Data Protection Authority (**"DPA"**).[74] However, the Bill predicates twenty-three crucial provisions on the occurrence of 'Harm' as defined in the Bill: It uses harm as a trigger for the enforcement of crucial rights, obligations and regulation.[75] This harms-based approach of the Bill could create practical challenges that significantly weaken its ability to protect users and provide redress.

The problems with the definition of 'Harm' in the Bill have been discussed in S. 3 of this article. The definition is vague, open-ended and very difficult to interpret. In the absence of any guidance in the definition to understand its scope, the task of identifying 'Harm' is left to the subjective interpretation of different stakeholders.[76] This is problematic because what a user interprets as 'Harm' could be interpreted differently by a provider or a regulator.[77] For instance, a user might consider anxiety from personal data breaches as harm, but providers and regulators may be more conservative in their interpretation.[78] This uncertainty could confuse stakeholders about how they should act to mitigate and remedy harm, and compromise the Bill's goal of *ex ante* protection. Users would remain uncertain about whether they have been harmed and whether they can seek redress. Providers and other entities that process personal data would remain uncertain about how to design and conduct data processing activities. Regulators and adjudicators would remain uncertain about when and how

---

[71] Preamble, The Personal Data Protection Bill, 2019 (pending).
[72] 'Data fiduciary' is defined under S. 3(13) of the Bill as "*any person, including the State, a company, any juristic entity or any individual who alone or in conjunction with others determines the purpose and means of processing of personal data*."
[73] 'Data principal' is defined under S. 3 (14) of the Bill as "*the natural person to whom the personal data relates*."
[74] Supra 8, S. 49.
[75] *Our Submission to the Joint Committee of Members of the Indian Parliament on the Personal Data Protection Bill*, dated 25 February 2020, p. 19, Dvara Research, available at https://www.dvara.com/research/wp-content/uploads/2020/03/Dvara-Research-Final-Submission-Comments-to-the-Joint-Parliamentary-Committee-on-PDP-Bill.pdf, last seen on 29/11/2020. Three provision relate to data principals exercising their rights and accessing grievance redress forums. Nine provisions relate to the fulfilment of data protection obligations by data fiduciaries. Eleven provisions relate to the enforcement of the Bill by the Central Government and the DPA.
[76] Supra 44.
[77] Ibid.
[78] Supra 48.

to intervene in data processing activities or provide redress to mitigate and remedy harm.[79]

For example, the provision on 'reporting of personal data breach'[80] is one of the most important user protections safeguards in the Bill. The provision directs data fiduciaries to alert the proposed data protection regulator, the DPA, about personal data breaches so that they may direct necessary actions to mitigate harm from the breach. However, the provision requires data fiduciaries to report personal data breaches only when there is a likelihood of harm. Therefore, when personal data is breached, data fiduciaries can make a subjective assessment of whether the breach can cause harm (as defined in the Bill). Only once it determines that harm could occur does it have to report the breach to the DPA. This risks many sensitive data breaches not being reported to the DPA because data fiduciaries cannot identify or predict harms easily (This is besides the fact that it may be in the interests of data fiduciaries to not report personal data breaches to the DPA).[81] Users would remain vulnerable to secondary harms that could arise when the breached data is processed for *mala fide* purposes.[82]

Similarly, the provision on 'Grievance redress'[83] allows data principals to seek redress when data fiduciaries contravene provisions of the Bill. However, harm or the likelihood of harm are the necessary preconditions for data principals to seek redress. This precondition raises high barriers for seeking redress. It would preclude users from seeking redress when there is (i) a violation of the Bill without corresponding harm or (ii) a harm but without any violation of the Bill.[84] The precondition also levies a disproportionately heavy burden on the users to demonstrate that harm could occur because of the data fiduciary violating a provision in the Bill.[85]

---

[79] Supra 44.
[80] Supra 8, S. 25.
[81] Supra 44.
[82] R.J. Cronk, *Why Privacy-Risk Analysis Must Not be Harm Focused*, IAPP, available at https://iapp.org/news/a/why-privacy-risk-analysis-must-not-be-harm-focused/, last seen on 30/11/2020.
[83] Supra 8, S. 32.
[84] *Our Response to the Draft Personal Data Protection Bill, 2018*, Dvara Research, available at https://www.dvara.com/blog/wp-content/uploads/2018/10/Response-to-draft-Personal-Data-Protection-Bill_DvaraResearch.pdf, last seen on 29/11/2020.
[85] Supra 44.

The precondition assumes that users will have knowledge of all processing activities and their effects and that they will be able to identify and demonstrate harm. However, as mentioned above, information asymmetries that users face could make it impossible for them to do so.[86]

Other key provisions in the Bill that are linked to harm would be impacted by similar challenges. Although harm is an important factor to consider in user protection, it cannot be the yardstick for protecting users against personal data-related harms. Harm can be an outcome that regulation can protect against, but harm should not be the trigger for regulation. Data protection frameworks must be designed to be independent of harm to be effective in protecting users from harm.[87] Defining the rights of users and the responsibilities of providers in a way that imposes an obligation on providers to make reasonable efforts to not cause harm could be a more suitable approach for user protection in data processing activities.[88] Yet, the question about how harm should be defined remains to be answered.

V. DEFINING 'HARM': ENABLING A GRADUAL JURISPRUDENTIAL EVOLUTION OF THE DEFINITION

The article so far has discussed how personal data-related harm is unique. These harms can manifest in many forms, including in visceral and tangible forms, and at any point in time once personal data is breached or processed

---

[86] Supra 11; Also see D.J. Solove& D. Citron, *Risk and Anxiety: A Theory of Data Breach Harms*, GWU Law School Public Law Research Paper No. 2017-2, 1, GWU Law School Public Law Research Paper No. 2017-2, George Washington University Law School, (2017), available at https://scholarship.law.gwu.edu/cgi/viewcontent.cgi?article=2499&context=faculty_publications.

[87] Supra 44; *Our submission to the Joint Committee of Members of the Indian Parliament on the Personal Data Protection Bill*, dated 25 February 2020, p. 19, Dvara Research, available at https://www.dvara.com/research/wp-content/uploads/2020/03/Dvara-Research-Final-Submission-Comments-to-the-Joint-Parliamentary-Committee-on-PDP-Bill.pdf, last seen on 29/11/2020; R.J. Cronk, *Why Privacy-Risk Analysis Must Not be Harm Focused*, IAPP, available at https://iapp.org/news/a/why-privacy-risk-analysis-must-not-be-harm-focused/, last seen on 30/11/2020.

[88] See Supra 84; *Our Response to the White Paper on a Data Protection Framework for India*, Dvara Research, available at https://www.dvara.com/blog/wp-content/uploads/2018/02/Response-to-White-Paper-Public-Consultation-Dvara-Research.pdf, last seen on 30/11/2020; and *The Data Protection Bill, 2018*, p.3, Dvara Research, available at https://www.dvara.com/blog/wp-content/uploads/2018/02/Data-Protection-Bill-Draft-Dvara-Research.pdf, last seen on 30/11/2020; Also see D. Solove, *How Should the Law Handle Privacy and Data Security Harms?*, TeachPrivacy, available at https://teachprivacy.com/law-handle-privacy-data-security-harms/, last seen on 24/12/2020.

improperly. Further, the harms can manifest in new and inchoate forms that could emerge with developments in data processing technology and use-cases. These characteristics of harm make it impractical to create a prescriptive definition of 'Harm' (such as that under S. 3(20) of the Bill). The sheer scope of personal data-related harm makes it impossible to posit all the different kinds of outcomes that must be treated as harm without becoming vague. Not defining harm at all on the other hand could be as problematic as having an expansive definition.

Attempts to define harm could take a middle path in which the definition can broadly indicate what harm means without attempting to be too specific. The definition could guide the gradual development of jurisprudence on personal data-related harm on a case-by-case basis.[89] Simultaneously, the developments could guide providers' data practices in order to mitigate the risk of causing harm to users.[90] A potential definition for this purpose could contain a fundamental or conceptual understanding of harm; such as:

> 'Harm' is actual or potential injury or loss to an individual, whether such injury or loss is economic or non-economic, quantifiable or non-quantifiable.[91]

This definition builds upon existing notions of harm that are already being used by regulators to address harm and the risk of harm when it is difficult to identify. For instance, the Federal Deposit Insurance Corporation ("**FDIC**") relies on this definition for guiding its enforcement activities

---

[89] *Our Response to the White Paper on a Data Protection Framework for India*, p.59-60, Dvara Research, available at https://www.dvara.com/blog/wp-content/uploads/2018/02/Response-to-White-Paper-Public-Consultation-Dvara-Research.pdf, last seen on 02/12/20.

[90] Ibid; *Our Submission to the Joint Committee of Members of the Indian Parliament on the Personal Data Protection Bill*, dated 25 February 2020, p. 20, Dvara Research, available at https://www.dvara.com/research/wp-content/uploads/2020/03/Dvara-Research-Final-Submission-Comments-to-the-Joint-Parliamentary-Committee-on-PDP-Bill.pdf, last seen on 29/11/2020.

[91] S. 2 (m), The Data Protection Bill, 2018, available at https://www.dvara.com/blog/wp-content/uploads/2018/02/Data-Protection-Bill-Draft-Dvara-Research.pdf; FDIC Consumer Compliance Examination Manual- June 2019, p. 2.1., Federal Deposit Insurance Corporation, available at https://www.fdic.gov/resources/supervision-and-examinations/consumer-compliance-examination-manual/documents/2/ii-2-1.pdf, last seen on 02/12/2020; Supra 87; *Our Submission to the Joint Committee of Members of the Indian Parliament on the Personal Data Protection Bill*, dated 25 February 2020, p. 20, Dvara Research, available at https://www.dvara.com/research/wp-content/uploads/2020/03/Dvara-Research-Final-Submission-Comments-to-the-Joint-Parliamentary-Committee-on-PDP-Bill.pdf, last seen on 29/11/2020.

towards identifying, addressing, preventing and remedying consumer harm from financial institutions.[92]

Admittedly, the definition appears similar to that of 'Harm' in the Bill, but with prominent distinctions. The definition under the Bill attempts to list the different kinds of personal data-related harm. The definition was also open-ended without a defined scope within which outcomes could be treated as harm. The definition proposed above does not attempt to list the various kinds of personal data-related harms. It draws broad contours i.e., all actual or potential injuries, economic and non-economic injuries and quantifiable and non-quantifiable injuries, that firmly define its scope within which outcomes should be treated as harm.

These contours in the definition proposed above would guide courts in examining harm in each case and gradually developing jurisprudence through precedent. The contours would also set broad goalposts that can guide stakeholders in identifying the boundaries of legitimate activity. This could allow the stakeholders to take necessary steps: such as designing processing activities, approaching regulators for redress or taking enforcement action towards mitigating and redressing harm. Such a definition would be useful in addressing the elusive issue of identifying personal data-related harm until there is a better understanding about how the unique challenges posed by personal data-related harm can be addressed.

VI. CONCLUSION

India is making strides towards becoming a digital economy. A host of public entities and private entities are able to provide digital services for users' benefit. More users in India are becoming digitally active for these services in turn and are generating large amounts of personal data in the process. These gains and benefits could be undone if users' privacy and interests are not safeguarded in this new digital ecosystem. The law must

[92] *FDIC Consumer Compliance Examination Manual- June 2019*, p. 2.1-2.3, Federal Deposit Insurance Corporation, available at https://www.fdic.gov/resources/supervision-and-examinations/consumer-compliance-examination-manual/documents/2/ii-2-1.pdf, last seen on 02/12/2020.

keep pace with technical developments to make the ecosystem safe for users.

Unfortunately, personal data-related harms pose unique challenges to user protection. It is very difficult to discern what kinds of harm can emerge from personal data processing and to predict when they may emerge once personal data is processed. Together, these factors complicate user protection greatly. One of the biggest complications for the law lies in defining personal data-related harms for the purpose of regulation. For instance, in an attempt to define 'Harm' exhaustively as a list of outcomes, the Bill posits a vague and overbroad definition of 'Harm'. As a result, it risks weakening and diluting several regulatory and user protection provisions in its framework.

Currently, there is inadequate understanding about personal data-related harms among stakeholders involved. Until personal data-related harms are understood better, the law could rely on existing notions of harm that are already being used by regulators to address harm and the risk of harm when it is difficult to identify. The jurisprudence on personal data-related harms and the finer nuances could be developed gradually on a case-by-case basis. Such an approach could be more practicable and helpful in user protection. The law is most effective in safeguarding users' interests when it is practicable.