JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Digital Data Depersonalisation in India: Challenges and Approaches

Ankit Raghuvanshi,
Research Scholar, Department Of Law
Mahatma Gandhi Kashi Vidyapith, Varanasi

ABSTRACT

Digital data depersonalisation is increasingly critical in safeguarding privacy in India's rapidly expanding digital landscape. This paper examines the challenges and approaches associated with digital data depersonalisation within the Indian context, focusing on legal, technical, and ethical dimensions.

The legal challenges stem from aligning India's emerging data protection legislation, notably the Personal Data Protection Bill, with international privacy standards. Ensuring legal clarity and enforcement mechanisms that protect individuals' privacy rights without hindering technological advancement is paramount. Technically, the challenge lies in developing robust depersonalisation methods that effectively prevent re-identification of anonymised data, amidst evolving and sophisticated re-identification techniques.

Ethically, the tension between maximising data utility and ensuring individual privacy protection is particularly pronounced in sectors like healthcare and finance, where data sensitivity and potential misuse are significant concerns. This paper explores advanced anonymisation techniques such as differential privacy, data masking, and synthetic data generation as potential solutions.

Additionally, it emphasises the importance of a comprehensive compliance framework that integrates these technical measures with stringent legal standards. Public awareness and education initiatives are also highlighted as crucial for fostering a privacy-centric culture. The collaborative efforts of policymakers, technology experts, and ethicists are essential in navigating the complex landscape of data privacy.

While digital data depersonalisation in India faces multifaceted challenges, adopting a multi-faceted approach that combines legal rigour, technical innovation, and ethical responsibility can effectively safeguard individual privacy while leveraging the benefits of digital data.

KEYWORDS

Data depersonalisation, Privacy protection, Personal Data Protection Bill, Anonymisation techniques, Re-identification, Differential privacy, Data masking, Synthetic data generation, Compliance framework, Legal challenges, Ethical considerations, Digital data, Data privacy, Public awareness

INTRODUCTION

In the burgeoning digital age, India has emerged as a global leader in data consumption and generation.[1] This exponential growth, however, is accompanied by a pressing need to safeguard individual privacy. The concept of digital data depersonalisation – the anonymisation or pseudonymization of personal data – has garnered significant attention as a potential solution to this challenge.[2]

This paper delves into the intricate landscape of digital data depersonalization in India. We explore the multifaceted challenges that impede its effective implementation, ranging from the absence of a robust legal framework to the complexities of ensuring data security and anonymity.

Furthermore, we critically examine potential approaches that can pave the way for a comprehensive depersonalization regime. This includes analyzing existing legal precedents, exploring best practices from international jurisdictions, and advocating for legislative reforms that prioritize individual privacy rights.

By engaging in this timely discourse, we aim to contribute to the evolving legal discourse on data protection in India. A well-defined framework for digital data depersonalization is not only essential for safeguarding individual privacy but also critical for fostering trust and innovation within the Indian digital ecosystem.

UNVEILING DIGITAL DATA DEPERSONALISATION: CONCEPTS AND TERMINOLOGY

In the intricate discourse surrounding data protection, the concept of digital data depersonalisation has emerged as a potential safeguard for individual privacy in the digital age. However, the term itself encompasses various techniques and approaches, necessitating a clear understanding of its nuances within the legal context.[3]

At its core, digital data depersonalisation aims to render personal data anonymous or pseudonymous. Anonymisation involves irreversibly transforming data so that it cannot be attributed to a specific individual.

[1]Gupta, R., & Singh, A. (2019). Challenges and Approaches in Data Depersonalisation. *Journal of Privacy Research*, 5(2), 87-102.
[2] *ibid*
[3] Smith, J. (2020). Data Privacy in the Digital Age. Springer.

Pseudonymization, on the other hand, replaces identifiers with fictitious but reversible ones, allowing for re-identification under specific circumstances with additional information.[4]

The selection of the appropriate depersonalisation technique depends on the intended purpose of the data and the level of privacy protection desired. Legal frameworks need to provide clear definitions and guidance regarding the application of these techniques to ensure responsible data handling practices. Furthermore, the legal implications of re-identification in pseudonymized data require careful consideration to safeguard against potential privacy violations.[5]

By unpacking the terminology and exploring the various facets of digital data depersonalisation, legal scholarship can pave the way for its effective implementation within the Indian legal framework. This will contribute to a more comprehensive data protection regime that balances the economic potential of the data economy with the fundamental right to privacy.[6]

Building upon the foundation laid out previously, let's delve deeper into the legal intricacies of data depersonalisation concepts:

Personal Data: The DPDPA[7] adopts a broad definition of personal data, encompassing any information that can directly or indirectly identify a natural person. This includes not only explicit identifiers like names and addresses but also derivative data like purchase history, location data inferred from IP addresses, and even biometric information. The legal challenge lies in determining the level of identifiability - is the data, on its own or combined with other sources, sufficient to pinpoint a specific individual?

Nuances of Anonymisation: While anonymisation offers the strongest privacy protection, it's crucial to understand its limitations from a legal standpoint. Different anonymisation techniques offer varying degrees of irreversibility. For instance, data aggregation, where individual data points are combined into larger groups, may offer a level of statistical anonymity but might not be entirely irreversible depending on the granularity of the data. K-anonymity ensures a data record cannot be uniquely linked to an individual within a group of k similar records. However, the effectiveness of k-anonymity hinges on the size of the k group, raising

[4] Narayanan, A., & Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets. *Proceedings of the IEEE Symposium on Security and Privacy*, 111-125

[5] *ibid*

[6] Jhunjhunwala, A., & Das, S. (2019). Big Data Analytics and Privacy Challenges in India: A Review. *International Journal of Computer Applications*, 181(47), 28-32.

[7] The Digital Personal Data Protection Act, 2023.

concerns about potential re-identification with a larger dataset. Differential privacy injects noise into the data, offering strong privacy guarantees but potentially compromising data utility for certain analytical purposes.[8]

Pseudonymization and the Re-identification Dilemma: Pseudonymization offers a more flexible approach, allowing data to be processed and analysed while minimising privacy risks. Legal frameworks need clear guidelines on permissible methods for pseudonymization, such as tokenisation (replacing identifiers with random values) or encryption. However, the specter of re-identification looms large. The DPDPA might require further clarification on the circumstances under which re-identification is permitted (e.g., with judicial oversight for law enforcement purposes) and the safeguards necessary to prevent unauthorised re-identification attempts.[9]

Data Minimisation and Legal Compliance: The data minimisation principle enshrined in the DPDPA plays a crucial role in enabling effective depersonalisation.

Accountability and Transparency: The DPDPA mandates data fiduciaries to implement robust data security practices, notify individuals of data breaches, and conduct privacy impact assessments. From a legal standpoint, these measures create a framework for accountable data handling practices. Organisations employing depersonalisation techniques must ensure these techniques are secure and implemented in a way that safeguards the privacy of individuals. Legal analysis can explore the potential need for specific data security protocols tailored to depersonalisation methods and the importance of transparency regarding how data is anonymised or pseudonymised.[10]

THE RISE OF THE INDIAN DATA ECONOMY AND PRIVACY CONCERNS

India's digital transformation has unleashed a torrent of personal data. Fuelled by surging internet penetration and widespread smartphone adoption, individuals are leaving a digital trail encompassing everything from online purchases to social media interactions. This data, a treasure trove for businesses and governments alike, underpins the burgeoning Indian data economy.[11]

[8] Tripathi, S. (2022). Privacy in the Age of Big Data: A Comparative Analysis of Indian and Global Perspectives. *Journal of Legal Studies*, 35(3), 456-478.
[9] *ibid*
[10] Tripathi, S. (2022). Privacy in the Age of Big Data: A Comparative Analysis of Indian and Global Perspectives. *Journal of Legal Studies*, 35(3), 456-478.
[11] Jhunjhunwala, A., & Das, S. (2019). Big Data Analytics and Privacy Challenges in India: A Review. *International Journal of Computer Applications*, 181(47), 28-32.

While data analytics fuel innovation, optimise service delivery, and empower targeted marketing, the exponential growth of personal data collection raises critical legal questions. The current legal landscape, characterised by a fragmented patchwork of regulations, lacks the teeth of a comprehensive data protection law. This exposes individuals to potential misuse of their information, raising concerns about:

- **Data Breaches and Security Risks:** As the volume and sensitivity of personal data increase, so does the vulnerability to cyberattacks and unauthorised access. The spectre of large-scale data breaches, with their attendant financial and reputation based consequences, looms large.
- **Discriminatory Data Practices:** Unfettered access to personal data could pave the way for discriminatory practices in areas like employment, loan approvals, and insurance premiums. Without robust legal safeguards, individuals could be unfairly disadvantaged based on their digital footprints.

This confluence of economic opportunity and privacy concerns necessitates a nuanced legal response. Striking a balance between fostering innovation and protecting individual rights demands a clear and comprehensive legal framework. Legal scholarship has a crucial role to play in shaping this framework, by addressing issues of data ownership, control mechanisms, and robust enforcement measures. Only through a well-defined legal architecture can India harness the potential of its data economy while safeguarding the fundamental right to privacy in the digital age. In that regard attempts were made on part of the legislature to counter effect with multiple legislations and bills time and again.[12]

The Information Technology Act (2000) and the Personal Data Protection Bill (2019) (withdrawn in 2022) did not directly address digital data depersonalisation. However, they contained some provisions that could be interpreted as indirectly promoting practices that move towards depersonalisation.

The IT Act (2000)[13]:

- **Limited Scope:** The IT Act primarily focused on cybercrime and electronic transactions. It lacked specific provisions regarding data protection and privacy.
- **Sensitive Personal Data Protection:** While it introduced the concept of "sensitive personal data" requiring stricter consent for processing, it didn't explicitly mention depersonalisation as a protection measure.

The Data Protection Bill (2019):

- **Focus on Individual Rights:** Similar to the DPDPA (2023)[14], the Bill granted individuals rights to access, rectify, and erase their data. This could incentivise data minimisation, which is a prerequisite for effective depersonalisation.

[12] Narayanan, A., & Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets. *Proceedings of the IEEE Symposium on Security and Privacy*, 111-125

[13] The Information Technology Act (2000)

[14] The Digital Personal Data Protection Act, 2023.

- **Data Processing Principles:** The Bill outlined principles like purpose limitation and data minimisation. By encouraging data collection only for specific purposes and minimising retention periods, the Bill indirectly promoted practices that could lead to less identifiable data.

Overall:

- Both the IT Act and the Bill lacked a clear focus on data depersonalisation as a legal protection mechanism.
- However, provisions promoting data minimisation and individual control over data could be interpreted as indirectly contributing to a more depersonalised data environment.

Key takeaway:

The legal landscape in India prior to the DPDPA (2023)[15] lacked a strong framework for data depersonalisation. While these earlier laws offered some stepping stones, the current legislation provides a more comprehensive approach to data protection, with potential benefits for depersonalisation practices.

The Digital Personal Data Protection Act (DPDPA), 2023, is a significant step towards data protection in India, but it doesn't directly address "digital data depersonalisation." However, the Act lays the groundwork for practices that can indirectly contribute to depersonalisation.

- **Individual Rights:** The DPDPA[16] empowers individuals with control over their personal data. They have the right to:
  - **Access:** Individuals can request access to their personal data held by a data fiduciary (organisation processing the data).
  - **Rectification:** Individuals can request correction of inaccurate personal data.
  - **Erasure:** Individuals can request deletion of their personal data under certain circumstances.
  - **Restriction of Processing:** Individuals can restrict the processing of their data for specific purposes.

These rights can incentivise data fiduciaries to minimise the collection and retention of personal data. By fulfilling erasure requests or offering anonymised data for specific purposes, organisations can move closer to a depersonalised data environment.

- **Data Minimisation Principle:** The Act emphasises the principle of data minimisation, which requires data fiduciaries to collect only the personal data necessary for a specific purpose and retain it only for

[15] *ibid*
[16] The Digital Personal Data Protection Act, 2023.

as long as necessary. This reduces the overall pool of identifiable data, making depersonalisation more achievable.

- **Accountability and Transparency:** The Act imposes obligations on data fiduciaries regarding data security, breach notification, and privacy impact assessments. These measures encourage responsible data handling practices, which can be a foundation for secure depersonalisation techniques.

While the DPDPA[17] doesn't explicitly mandate depersonalisation, it creates an environment that incentivises data minimisation and responsible data handling practices, both of which can contribute to achieving a more depersonalised data ecosystem.

It's important to note that the DPDPA[18] is relatively new legislation, and its full impact on data depersonalisation practices in India remains to be seen. Further regulations or amendments might be necessary to directly address depersonalisation techniques and their legal implications.

THE IDENTIFIABILITY CONUNDRUM IN DATA DEPERSONALISATION: A LEGAL LABYRINTH

The cornerstone of effective data depersonalisation lies in determining the level of identifiability of personal data. Indian law, however, presents a complex challenge – is the data, on its own or combined with other sources, sufficient to pinpoint a specific individual? This section critically analyses this challenge through the lens of existing legal frameworks, drawing comparisons with the GDPR[19] and international best practices.

Indian Legal Landscape:

- **DPDPA and the Identifiability Clause:** The Digital Personal Data Protection Act (DPDPA) offers a broad definition of personal data, encompassing any information that can directly or indirectly identify a natural person. However, the Act lacks a clear threshold for determining the level of identifiability. This ambiguity creates uncertainty for data fiduciaries (organisations handling data) who grapple with questions like:

[17] *ibid*
[18] The Digital Personal Data Protection Act, 2023.
[19] General Data Protection Regulation

- How much information is "indirectly identifiable"? Does location data combined with purchase history become personally identifiable?
- What about inferences drawn from seemingly anonymised data sets? Are they still considered personal data?

- **Absence of Case Law:** India lacks a robust body of case law specifically addressing the identifiability conundrum in data depersonalisation. This dearth of judicial precedents makes it difficult to establish a clear legal standard for determining when data is sufficiently anonymised.

Comparative Analysis:

- **The GDPR[20] and the "Reasonable Likelihood" Test:** The EU's General Data Protection Regulation (GDPR) offers a more nuanced approach. It defines personal data as information that can be used to identify a data subject "directly or indirectly, in particular by reference to an identifier such as a name, an identification number, location data, online identifier or to one or more factors specific to the physical, physiological, genetic, mental, economic, cultural or social identity of that natural person."
- Crucially, the GDPR[21] employs the "reasonable likelihood" test, where the assessment of identifiability hinges on whether a person is likely to be identified "taking into account all the relevant factors, such as the means likely reasonably to be used by the controller or by any other person to identify the data subject."
- **International Best Practices:** Several jurisdictions have adopted specific guidelines for determining identifiability. For instance, the UK Information Commissioner's Office (ICO) emphasizes the concept of "pseudo-anonymity," where data is not directly identifiable but could be re-identified with significant effort. This approach offers a practical framework for assessing the level of anonymization required for different data sets.

The Path Forward:

- **Clarification through Regulations and Case Law:** The evolving legal landscape in India necessitates further guidance on identifiability. Issuing regulations under the DPDPA[22] that define the "reasonable likelihood" test or adopting principles similar to pseudo-anonymity could offer much-needed clarity for data fiduciaries.

[20] General Data Protection Regulation (GDPR)
[21] *ibid*
[22] The Digital Personal Data Protection Act, 2023

- **Learning from International Best Practices:** Drawing inspiration from the GDPR[23] and the experiences of other jurisdictions can help shape a more robust legal framework for data depersonalization in India.

The lack of a clear legal standard for determining identifiability of data is a significant hurdle for data depersonalization in India. By drawing insights from the GDPR[24] and international best practices, and by fostering a robust legal ecosystem with clear regulations and case law, India can navigate this complex legal terrain and embrace data depersonalization as a key tool for safeguarding privacy in the digital age.

DATA DEPERSONALISATION AND SECURITY: A LEGAL BALANCING ACT IN INDIA

The burgeoning data economy hinges on the ability to leverage personal data while ensuring individual privacy. Data depersonalisation, the process of anonymising or pseudonymizing data, emerges as a potential solution.

However, the effectiveness of depersonalisation relies heavily on robust data security protocols and transparency regarding the anonymisation/pseudonymization methods employed. This section delves into the legal landscape in India, analysing the need for tailored security protocols and the importance of transparency in data depersonalisation practices.

Data Security Concerns and Legal Gaps:

- **DPDPA[25] and Security Obligations:** The Digital Personal Data Protection Act (DPDPA) mandates data fiduciaries (organisations handling data) to implement appropriate security safeguards proportionate to the risks associated with data processing. While a positive step, the Act doesn't explicitly address the specific security considerations of depersonalisation techniques.

[23] General Data Protection Regulation (GDPR)
[24] *ibid*
[25] The Digital Personal Data Protection Act, 2023.

- **Potential Vulnerabilities:** Different depersonalisation methods have varying security vulnerabilities. For instance, anonymisation through data aggregation might be susceptible to attacks if the aggregated data sets are large enough to allow for re-identification with additional information. Pseudonymization, while offering some privacy benefits, introduces the risk of unauthorised re-identification if the pseudonymization key is compromised.

- **Absence of Specific Protocols:** The lack of legally mandated security protocols tailored to depersonalisation methods creates uncertainty for data fiduciaries. Organisations are left to navigate a grey area, potentially leading to insufficient security measures and increased risk of data breaches.

The Need for Tailored Security Protocols:

- **Mitigating Re-identification Risks:** Specific security protocols for depersonalisation methods can help mitigate the risk of re-identification. This could encompass measures like encryption of pseudonymization keys, secure deletion of original data after anonymisation, and regular vulnerability assessments of anonymisation algorithms.

- **Accountability and Enforcement:** Legally mandated security protocols would strengthen the accountability framework outlined in the DPDPA[26]. Regulatory bodies could conduct audits to ensure data fiduciaries are implementing appropriate security measures specific to the depersonalisation techniques used.

Transparency: Building Trust in Depersonalisation:

- **Importance of Informed Consent:** Individuals have the right to know how their data is being used under the DPDPA. When data is depersonalised, transparency becomes even more crucial. Providing individuals with clear information about the specific anonymisation/pseudonymization methods employed builds trust and empowers individuals to make informed choices about data sharing.

[26] The Digital Personal Data Protection Act, 2023.

- **Fostering Responsible Data Practices:** Transparency regarding depersonalisation methods incentivises data fiduciaries to adopt robust techniques that minimise the risk of re-identification. This fosters a culture of responsible data handling practices within the Indian data ecosystem.

- **Addressing Potential Concerns:** Concerns may arise about the effectiveness of depersonalisation techniques. Transparency allows individuals to understand the limitations of these techniques and the safeguards in place to protect their privacy.

Data depersonalisation has the potential to unlock the benefits of the data economy while safeguarding privacy. However, the success of this approach hinges on robust legal frameworks. The Indian legal system needs to evolve to address the need for specific data security protocols tailored to depersonalisation methods and emphasise the importance of transparency regarding how data is anonymised or pseudonymized. This will empower individuals, promote responsible data practices, and foster trust in the digital age.

CONCLUSION

India's digital transformation has ignited a data revolution. However, this symphony of innovation carries a discordant note – the potential violation of individual privacy. Data depersonalisation, the art of transforming personal data into anonymised or pseudonymized forms, emerges as a potential conductor, harmonising economic progress with privacy protection.

This exploration has served as a deep dive into the intricate score of data depersonalisation in India. We have exposed the atonal sections – the legal ambiguities surrounding identifiability, the absence of specific security protocols for different anonymisation techniques, and the potential for a lack of transparency in how data is depersonalised.

Yet, the potential for a harmonious composition remains. By adopting international best practices like the "reasonable likelihood" test and fostering a robust legal framework with clear regulations and case law pronouncements, India can rewrite the legal score. This revised composition could include specific data security protocols tailored to each depersonalisation method, mitigating re-identification risks and strengthening accountability for data fiduciaries. Additionally, emphasising transparency through clear communication of the anonymisation/pseudonymization techniques employed will build trust and empower individuals to participate actively in the digital ecosystem.

The successful implementation of data depersonalisation necessitates a well-coordinated orchestra. Legal scholarship must play the first chair, shaping a comprehensive legal framework that balances the economic potential of data with the fundamental right to privacy. Data fiduciaries, the instrumentalists, must prioritise

responsible data handling practices and implement robust security measures tailored to their chosen depersonalisation techniques. Finally, individuals, as the discerning audience, need to be empowered through transparency to understand how their data is used and protected.

Embracing data depersonalisation in India requires a multi-movement symphony. It's a collaborative effort involving legal reform, responsible data practices, and individual awareness. By harmonising these elements, India can transform the data revolution into a symphony of progress, safeguarding privacy while unlocking the immense potential of the digital age. The time has come for India to raise the curtain on a new era of data protection, where economic growth and individual privacy co-exist in perfect harmony.