NOTES AND COMMENTS

Regulating Artificial Intelligence under Data Protection Law: Challenges and Solutions for India

Paarth Naithani[1]

Abstract

As India moves toward enacting a comprehensive data protection legislation, it becomes essential to examine the possible application of India's proposed data protection law to the use of Artificial Intelligence (AI). The various challenges posed by AI to data protection principles and data principals' rights need to be examined. The need for data maximisation in the use of AI challenges the principle of collection limitation. The difficulty in anticipating the processing purposes of AI challenges the principle of purpose limitation. With a brief introduction to AI and data protection law in India, the paper examines the compatibility of various data protection provisions under India's Digital Personal Data Protection Act, 2023 with AI. The paper also provides recommendations for data protection regulation of AI. The paper proposes the need to hold data fiduciaries accountable using Data Protection Impact Assessments, Codes of Practice and Security Measures. Besides, there is a need to define the fiduciary duty of care between the data principal and data fiduciary. There is a need recognize data protection by design and default and the Right against automated decision making. Technical solutions need to be explored, but at the same time, AI must not be over-regulated. Lastly, there is a need for flexibly interpreting the provisions of the proposed data protection law.

Keywords - Artificial Intelligence, Data Protection Law, Data Protection Act, 2021, India, Regulation, Rights, Principles

[1] Assistant Lecturer and Research Fellow with the Jean Monnet Chair in Multi-dimensional Approaches to the Understanding of the EU Data Protection Framework at O. P. Jindal Global University.

I. Introduction

As the name suggests, Artificial "Intelligence" (AI) is a technology that is "intelligent". AI is considered "intelligent" because it performs tasks which require intelligence such as perception and decision making.[2] AI analyzes data through algorithms to detect patterns and make predictions.[3] AI can be used for providing better services and also for profiling, tracking, and targeting individuals.

Today, AI technology has applications in various sectors.[4] Individuals use products and services which are powered by AI.[5] Personal assistants such as Siri, Alexa, and Google Assistant use the AI technique of voice recognition and natural language processing.[6] Predictive text used in products such as Gmail and Google Search works on the AI technique of machine learning.[7] Facial recognition used to identify persons in photos on Social Media uses machine vision.[8] Product recommendations and personalised advertising on Facebook and Amazon make use of AI to find patterns and profile individuals.[9]

---

[2] Defense Science Board, *Report of the Defense Science Board Summer Study on Autonomy* (Jun. 2016), https://www.hsdl.org/?view&did=794641.

[3] Paul Scharre, Michael C. Horowitz, and Robert O. Work, *What Is Artificial Intelligence,* JSTOR (Jun. 1 2018) http://www.jstor.org/stable/resrep20447.5.

[4] NITI Aayog, *National Strategy for Artificial Intelligence*, INDIAai (Jun. 13, 2019) https://indiaai.gov.in/research-reports/national-strategy-for-artificial-intelligence.

[5] Bernard Marr, *The 10 Best Examples Of How Companies Use Artificial Intelligence In Practice*, Forbes (Dec. 9, 2019) https://www.forbes.com/sites/bernardmarr/2019/12/09/the-10-best-examples-of-how-companies-use-artificial-intelligence-in-practice/?sh=497bbed77978.

[6] Sakshi Gupta, *Natural Language Processing Use Case – How Do Personal Assistant Apps Work?*, Springboard Blog (Jun. 10, 2020) https://www.springboard.com/blog/data-science/nlp-use-cases/.

[7] Yonghui Wu, *Smart Compose: Using Neural Networks To Help Write Emails*, Google AI Blog (May 16, 2018) https://ai.googleblog.com/2018/05/smart-compose-using-neural-networks-to.html.

[8] *Machine Learning and Facial Recognition*, PXL Vision (Jan. 21, 2021) https://www.pxl-vision.com/en/blog/machine-learning-and-how-it-applies-to-facial-recognition-technology.

[9] Mike Kaput, *AI in Advertising: Everything You Need to Know*, Marketing Artificial Intelligence Institute (Mar. 10, 2022) https://www.marketingaiinstitute.com/blog/ai-in-advertising.

AI is also used to draw inferences and interpret Big Data. Big Data is high-volume-velocity-variety information that is obtained real time and is processed using Machine Learning.[10] The processing of Big Data using AI is referred to as Big Data Analytics.[11]

Data protection is an area of law which aims to regulate the processing of personal data. It aims to protect informational privacy, which is a part of the right to privacy[12] recognised as a fundamental right under the Constitution of India.[13] Data protection is gaining significance and India has enacted a comprehensive legislation on data protection.[14]

Earlier, a proposal for comprehensive data protection framework was made by the Srikrishna Committee.[15] The proposal was accompanied by a draft data protection legislation, the Personal Data Protection Bill, 2018 (PDP Bill, 2018).[16] The PDP Bill, 2018 was revised and tabled in the Parliament with considerable changes as Personal Data Protection Bill, 2019 (PDP Bill, 2019).[17] The PDP Bill, 2019 was sent to a Joint Parliamentary Committee (JPC) for its recommendations. The JPC then released its report and proposed the Data Protection Bill, 2021 by amending the PDP Bill, 2019. Next, a new draft data protection framework, the Digital Personal Data Protection Bill, 2022, was released for public consultation. Finally, an updated version of the Digital Personal Data Protection Bill, 2022 has become law in India and is titled the Digital Personal Data Protection Act, 2023 (hereinafter DPDP).

---

[10] *Big data, artificial intelligence, machine learning and data protection*, Information Commissioner's Office https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf.
[11] *Id.*
[12] Preamble, Data Protection Bill, 2021.
[13] Justice K.S. Puttaswamy and another v. Union of India, AIR 2017 SC 4161.
[14] The Digital Personal Data Protection Act, 2023.
[15] Committee of Experts under the Chairmanship of Justice B.N. Srikrishna, *A Free and Fair Digital Economy Protecting Privacy, Empowering Indians*, Ministry of Electronics and Information Technology, https://www.meity.gov.in/writereaddata/files/Data_Protection_Committee_Report.pdf
[16] The Personal Data Protection Bill, 2018 (India).
[17] Anurag Vaishnav, *The Personal Data Protection Bill, 2019: How it differs from the draft Bill*, The PRS Legislative Research Blog (Dec. 27, 2019) https://prsindia.org/theprsblog/personal-data-protection-bill-2019-how-it-differs-draft-bill.

In the background of the increasing significance of both AI and data protection law, it is important to examine the possible application of the DPDP to AI.

II. AI and Data Protection

The discussions on AI and data protection in India find mention in the NITI Aayog Strategy Paper,[18] NITI Aayog Approach Paper[19] and the Srikrishna Committee Report[20]. These discussions have identified the conflict between data protection and AI. For instance, there is the issue of AI causing discrimination and harm to data subjects.[21] There is the possibility of emotional and economic harm when sensitive personal data is used with AI.[22] There is a need for explainability of AI.[23] India can learn from the global standard on data protection which is the EU General Data Protection Regulation[24] (GDPR). The following sub-sections examine the possible issues in the application of the DPDP to AI.

There is discussion on data protection concepts such as 'notice and consent' and personal data, data protection principles such as transparency, collection

[18] NITI Aayog, *National Strategy for Artificial Intelligence*, INDIAai (Jun. 13, 2019) https://indiaai.gov.in/research-reports/national-strategy-for-artificial-intelligence.

[19] NITI Aayog, *Appraoch Document for India Part 1- Principles for Responsible AI*, INDIAai (Feb. 24, 2021) http://indiaai.gov.in/research-reports/responsible-ai-part-1-principles-for-responsible-ai.

[20] Committee of Experts under the Chairmanship of Justice B.N. Srikrishna, *A Free and Fair Digital Economy Protecting Privacy, Empowering Indians*, Ministry of Electronics and Information Technology, https://www.meity.gov.in/writereaddata/files/Data_Protection_Committee_Report.pdf

[21] Amber Sinha and Elonnai Hickok, 'The Srikrishna Committee Data Protection Bill and Artificial Intelligence in India'<https://cis-india.org/internet-governance/blog/the-srikrishna-committee-data-protection-bill-and-artificial-intelligence-in-india>19 Jun. 2021.

[22] NITI Aayog, *Appraoch Document for India Part 1- Principles for Responsible AI*, INDIAai (Feb. 24, 2021) http://indiaai.gov.in/research-reports/responsible-ai-part-1-principles-for-responsible-ai.

[23] NITI Aayog, *National Strategy for Artificial Intelligence*, INDIAai (Jun. 13, 2019) https://indiaai.gov.in/research-reports/national-strategy-for-artificial-intelligence

[24] General Data Protection Regulation (EU).

limitation, purpose limitation, data quality and retention limitation, and data protection rights such as the Right to be forgotten and the Right to data portability.

A. AI and Personal Data

The DPDP applies to the processing of digital personal data.[25] The DPDP defines personal data as "any data about an individual who is identifiable by or in relation to such data."[26] Unlike the Data Protection Bill, 2021, the DPDP does not define profiling, sensitive personal data and non-personal data. The Data Protection Bill, 2021 (hereinafter Bill) had clarified that personal data "*shall include any inference drawn from such data for the purpose of profiling*."[27] Profiling had been defined as the analysis or prediction of behaviour, attributes or interests through the processing of personal data of data principals.[28] The Bill had defined sensitive personal data to include health data, biometric data, genetic data and religious or political beliefs.[29] Sensitive personal data is a special category of personal data which requires a higher level of protection and has been recognized as a separate category of personal data under the EU GDPR. The Bill had also distinguished personal data from non-personal data.[30] Non-personal data is "*data other than personal data*".[31] Non-personal data includes anonymized data, which is data that has been put through technical processes which make it difficult to identify a person.[32]

Although there is an absence of clarification by the DPDP that personal data includes inferences drawn for profiling, the wording of the DPDP definition of personal data suggests that it does. Personal data has been defined an "any data" from which the individual is identifiable, and inferences drawn from profiling are data from which the individual can be identifiable. AI is used to make inferences from existing data which are used for profiling.[33] The data inferred through

[25] Section 3, DPDP
[26] Section 2(t), DPDP
[27] *Id.*
[28] Clause 3(37), Data Protection Bill, 2021.
[29] Clause 3(41), Data Protection Bill, 2021.
[30] Clause 3(28), Data Protection Bill, 2021.
[31] Clause 3(28), Data Protection Bill, 2021.
[32] Clause 3(2)-(3), Data Protection Bill, 2021.
[33] Panel for the Future of Science and Technology, *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence*, European Parliament (Jun. 2020)

existing personal data using AI would constitute personal data as per the definition of personal data. Thus, even data inferred by AI must be regulated under the provisions of the DPDP as it applies to the processing of personal data.[34]

Although the DPDP does not define non-personal data and therefore does not make a distinction between personal data and non-personal data, the distinction between personal data and non-personal data is essential as the two categories of data would be regulated under different frameworks. AI challenges the distinction between personal data and non-personal data. AI challenges the distinction because AI makes it possible to identify individuals even from anonymised data sets.[35] AI can link datasets and recognise patterns in data leading to persons becoming identifiable from the data.[36]

Although the DPDP does not make a distinction between personal data and sensitive personal data, this distinction is important as sensitive personal data has a higher risk of harm to privacy. AI blurs the line between personal data and sensitive personal data. AI can be used to make sensitive inferences about a person. For instance, even health data can be inferred from data sets on shopping databases

B. AI and Notice and Consent

The DPDP requires that the data fiduciary must give the data principal notice about the personal data processed and the purposes of processing.[37] The notice must be given before or at the time of collecting personal data. [38] The Bill also provides for consent as a legal ground of processing.[39] The Bill requires consent

https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU(2020)641530_EN.pdf

[34] Section 3, DPDP.

[35] Rekha Jain and Viswanath Pingali, *India's non-personal data framework: a critique*, 9 *CSIT* 171 (2021).

[36] Robert Walters and Matthew Coghlan, *Data Protection and Artificial Intelligence Law: Europe Australia Singapore - An Actual or Perceived Dichotomy?*, SSRN (Feb. 18 2020) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3503392

[37] Section 5, DPDP

[38] *Id.*

[39] Section 4, DPDP

to be free, informed, specific, clear, unconditional, unambiguous and capable of being withdrawn.[40]

First, notice and consent are not practical when it comes to Big Data analytics. Big Data analytics is used to make correlations such as between people's lifestyle and credit worthiness.[41] Notice about the purpose of processing cannot be provided at the time of collecting personal data because of unforeseeable correlations.

Second, consent cannot be valid when the nature of the analysis done by AI is opaque.[42] In such cases, consent cannot be informed as the purpose of processing is unknown and the scope of processing is indeterminable.

Third, even when a person explicitly denies consent to the processing of his personal data, it is possible to make inferences about the person by drawing extrapolations from connected and related persons.[43] Thus, in an era of machine learning where group profiling is possible, it is difficult to opt out[44] or withdraw consent.

C. AI and Transparency

The provision on notice under DPDP requires information to be made available to the data principal.[45] Unlike the Bill, the DPDP does not require information to be provided about the "*fairness of algorithm or method used for processing personal data*".[46] An absence of such a provision in the DPDP is concerning as it should to be transparent to the data principal how AI is used to process their personal data.

The problem is exacerbated as transparency is a challenge with AI . It is difficult to look into the black box of AI. The black box effect is the inevitable opacity

[40] Section 6, DPDP.
[41] *Big data, artificial intelligence, machine learning and data protection*, Information Commissioner's Office https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf
[42] *Id.*
[43] Matt Bartlett, *Beyond Privacy: Protecting Data Interests in the Age of Artificial* Intelligence, 3 Law, Tech & Hum 96 (2021).
[44] *Id.*
[45] Section 5, DPDP.
[46] *Id.*

which makes it unlikely to understand and explain AI's working.[47] Besides, the logic of machine reasoning is difficult to explain in human terms.[48] It is also difficult to trace the outcome of the AI.[49] When it comes to unsupervised learning (a kind of AI), it is difficult to explain its working as there is a lack of data labels and relationships which can help explain the processes behind AI.[50]

D. AI and Purpose Limitation

While the DPDP does not explicitly recognize the purpose limitation principle, the DPDP states in section 6 that consent must be given for a specified purpose. The DPDP further states in section 6 that consent shall be "limited to such personal data as is necessary for such specified purpose." Previously, the Bill had recognized the purpose limitation principle which requires that personal data must be processed for consented purposes or purposes incidental or connected to the consented purposes.[51] The data principal's reasonable expectations regarding the use of the data need to be considered.[52] Personal data must also be processed in a fair and reasonable manner while ensuring privacy.[53] Purpose limitation implies that voice recordings used by Siri and Alexa should not be used to extract biometric findings.[54] Purpose limitation also implies that fitness trackers must not become pharmacy shops.[55]

AI challenges the purpose limitation principle. First, when AI is used to process data, the purpose cannot always be specified. The purpose may be unknown or

---

[47] *Big data, artificial intelligence, machine learning and data protection*, Information Commissioner's Office https://ico.org.uk/media/for-organisations/documents/2013559/big-data-ai-ml-and-data-protection.pdf

[48] Lilian Mitrou, *Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof'?*, SSRN (Jun. 3, 2019) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3386914

[49] *Id.*

[50] Matthew Humerick, *Taking AI Personally: How the E.U. Must Learn to Balance the Interests of Personal Data Privacy & Artificial Intelligence*, 34 Santa Clara High Tech. L.J. 393 (2018).

[51] Clause 5, Data Protection Bill, 2021.

[52] *Id.*

[53] *Id.*

[54] Robert Walters and Matthew Coghlan, *Data Protection and Artificial Intelligence Law: Europe Australia Singapore - An Actual or Perceived Dichotomy?*, SSRN (Feb. 18 2020) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3503392

[55] *Id.*

undecided at the time of processing. Second, AI can be used to process data for multiple purposes. AI makes it possible to re-purpose and multipurpose data for initially unknown and wide-ranging purposes.[56] Thus, AI challenges the notion of purpose being limited to the specified purpose or incidental purposes.

E. AI and Collection Limitation

While the DPDP does not explicitly recognize the collection limitation principle, it states in Section 6 that consent must be given for a specified purpose and consent must be limited to data necessary for specified purpose. As per the Bill which had recognized the principle of collection limitation, only that data must be collected which is necessary for processing purposes.[57] First, AI challenges collection limitation because it requires a massive amount of data to make accurate inferences. "*With few exceptions, more data is better than less, and there is almost never enough.*"[58] Second, collection limitation is challenged by AI because it is usually not possible to predict what data would be relevant for the AI.[59] This makes it difficult to limit the amount of data collected for training the AI

F. AI and Data Quality

The DPDP provides in Section 12 the right to correction and erasure, which allows the correction of inaccurate or misleading personal data, the completion of incomplete data and the updation of personal data. Previously, the Bill required that the quality of personal data must be maintained by ensuring completeness, accuracy, up-datedness and non-misleading nature of the data.[60]

Data quality is essential for maintaining the accuracy of the output of AI. If inaccurate data is input into the AI, it could lead to inaccurate inferences and

---

[56] Lilian Mitrou, *Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof'?*, SSRN (Jun. 3, 2019) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3386914

[57] Clause 6, Data Protection Bill, 2021.

[58] Christopher Kuner, Fred H Cate, Orla Lynskey, Christopher Millard, Nora Ni Loideain, and Dan Jerker B Svantesson, *Expanding the artificial intelligence-data protection debate*, 8 International Data Privacy Law 289 (2018).

[59] Lilian Mitrou, *Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof'?*, SSRN (Jun. 3, 2019) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3386914.

[60] Clause 8, Data Protection Bill, 2021.

biased decisions about a person.[61] Bias can result because of incomplete or outdated data being input into the AI. Bias in AI can also be a result of unrepresentative training data.[62] Bias in AI can be a result of using attributes such as gender without tracking how they are being considered by the AI.[63] Thus, it is essential to maintain data quality as the output of the AI depends on the data quality of training data and data input into the AI

G. AI and Retention Limitation

The Bill provided the principle of retention limitation which requires that data must not be retained beyond the period necessary to satisfy processing purposes, after which the data must be deleted.[64] The DPDP provides in Section 8 for an obligation on the data fiduciary to erase personal data when the data principal withdraws consent or when it is reasonable to assume that specified purpose is not being served anymore.

AI challenges the principle of retention limitation because it is not feasible to delete the data once it has been processed for the purposes for which it was collected.[65] Organisations may want to use the data for development and deployment of AI which may carry potential benefits.[66] One does not know when data would become relevant for processing by AI. Data fiduciaries would have an interest in storing data for a longer time so that it can be used when it becomes relevant.

H. AI and the Right against Automated Decision Making

The DPDP lacks a provision on the Right against automated decision making including profiling. Such a provision is present in the GDPR. The GDPR provides data subjects (data subjects is the term in the EU and data principal is the term

---

[61] KOAN Advisory and Digital India Foundation, *Handbook on Data Protection and Privacy for Developers of Artificial Intelligence (AI) in India: Practical Guidelines for Responsible Development of AI*, DSCI (Jul. 2021) https://www.dsci.in/sites/default/files/documents/resource_centre/AI%20Handbook.pdf
[62] *Id.*
[63] *Id.*
[64] Clause 9, Data Protection Bill, 2021.
[65] Christopher Kuner, Fred H Cate, Orla Lynskey, Christopher Millard, Nora Ni Loideain, and Dan Jerker B Svantesson, *Expanding the artificial intelligence-data protection debate*, 8 International Data Privacy Law 289 (2018).
[66] *Id.*

used in India) the right not to be subject to decisions taken solely on the basis of automated processing including profiling, which have a legal effect or other significant effects on the data subject.[67] An example is e-recruiting practices which use AI to shortlist applications. In such cases, data subjects also have the right to be informed of the existence of automated decision making, meaningful information about the logic of the automated decision making, and the significance and envisaged consequences.[68] For instance, in e-recruiting through the use of AI, the data subject should be informed of the kind of AI used to make a decision, the data input into the AI, and the possible consequence of the use of AI. The data subject should be informed, for example, that Machine Learning was used to input his personal details such as name, age, gender, marks, and experience into AI and the envisaged consequence could be that the job application may be rejected by the AI. Data subjects in the EU also have a right to obtain human intervention and a right to contest the decision made using AI.[69] In e-recruiting through the use of AI, the data subject has a right to have a human review the decision taken by the AI.

India's Bill merely required that information must be provided to the data principal about the "*fairness of algorithm or method used for processing of personal data*".[70] This clause has been removed from the DPDP. In fact, India's DPDP lacks a right to contest decisions taken by AI, right to obtain human intervention when AI makes decisions and the right not to be subject to automated decisions including profiling.

I. AI and the Right to Data Portability

The DPDP does not provide for a right to data portability. The Bill had provided the right to data portability which is the right to have personal data accessed and transferred to another data fiduciary in a structured, machine readable, and commonly used manner.[71] The data includes data that has been provided by the

[67] Article 22, GDPR.
[68] Article 13, 14 GDPR.
[69] *Id.*
[70] Clause 23, Data Protection Bill, 2021.
[71] Clause 19, Data Protection Bill, 2021.

data principal to the data fiduciary, data relating to any profile on the data principal and data generated while providing services or goods.[72]

AI is used for profiling and also for generating data in the course of providing goods or services. Under the right to data portability, such data needs to be shared with the data principal and other data fiduciaries. The requirement of sharing profiling data and data generated while providing goods or services may come in conflict with the data fiduciaries' trade secrets and intellectual property.

J. AI and the Right to be Forgotten

As per the Bill, the Right to be forgotten requires that personal data must be restricted in processing and disclosure when a person withdraws consent to the processing of personal data.[73] The right can come in conflict with the working of AI. Theoretically, if a person withdraws consent and the AI still continues to function through its learnings from previously learned behaviours, the data protection law would be violated.[74]

Under the DPDP, there is no explicit mention of the right to be forgotten. There is the right to erasure under Section 12 that allows the data principal to make an erasure request. Section 12 requires that the data fiduciary must erase the data "unless retention of the same is necessary for the specified purpose" or for a lawful purpose. But if AI is made to forget the data and the learning it has done from the data by erasure of the data, the functioning of AI would be affected.[75] This would make it difficult for the AI to function optimally.[76]

---

[72] *Id.*
[73] Clause 20, Data Protection Bill, 2021.
[74] Matthew Humerick, *Taking AI Personally: How the E.U. Must Learn to Balance the Interests of Personal Data Privacy & Artificial Intelligence*, 34 Santa Clara High Tech. L.J. 393 (2018).
[75] Lilian Mitrou, *Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof'?*, SSRN (Jun. 3, 2019) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3386914
[76] *Id.*

III. Possible Solutions

One possible solution is to hold data fiduciaries accountable.[77] As per the Bill, significant data fiduciaries intending to use new technologies or carry out processing having risk of significant harm should undertake a Data Protection Impact Assessment (DPIA) prior to the processing.[78] But such a provision that requires carrying a DPIA for using new technologies is missing in the DPDP.

Significant data fiduciaries are data fiduciaries that may be notified by the Central Government by assessing relevant factors including risk to rights and volume and sensitivity of data being processed.[79] A DPIA should be required when AI is used[80] because the use of AI systems carries a risk to rights of the data principal. Studies suggest that the use of AI may lead to discrimination.[81] AI can also be used to make evaluative decisions about an individual which could lead to denial of a benefit.[82]

As per the Bill, a DPIA includes an assessment of potential harm to data principals and measures for mitigating, minimising and managing risks.[83] A DPIA would include a systematic description of the processing, identifying risks to individuals and measures to reduce risk.[84] The description of processing would include describing data flows, stages of processing by AI, and effects on individuals.[85] The risks should be identified and could be categorised according to the likelihood of occurrence and severity of impact on data principals.[86] These risks could

[77] Clause 10, Data Protection Bill, 2021.
[78] Clause 27, Data Protection Bill, 2021.
[79] Section 10, DPDP
[80] Robert Walters and Matthew Coghlan, *Data Protection and Artificial Intelligence Law: Europe Australia Singapore - An Actual or Perceived Dichotomy?*, SSRN (Feb. 18 2020) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3503392
[81] Frederik Zuiderveen Borgesius, *Discrimination, Artificial Intelligence and Algorithmic Decision-Making* , Council of Europe (2018) https://rm.coe.int/discrimination-artificial-intelligence-and-algorithmic-decision-making/1680925d73
[82] Clause 3(23), Data Protection Bill, 2021.
[83] Clause 27, Data Protection Bill, 2021.
[84] Simon Reader, *Data Protection Impact Assessments and AI* , Information Commissioner's Office (Oct. 23 2019) https://ico.org.uk/about-the-ico/media-centre/ai-blog-data-protection-impact-assessments-and-ai/
[85] *Id.*
[86] *Id.*

include the risk of discrimination and impact on fundamental rights.[87] The mitigation of such risks must be planned early in the AI lifecycle.[88] The DPIA must be a live document which is regularly reviewed and re-assessed.[89]

Other than DPIA, the Bill had recognized Codes of Practice to promote data protection good practices and facilitate compliance.[90] The Codes of Practice may include various matters such as measures of ensuring data quality, the exercise of rights by data principal, standards of security safeguards and manner of carrying out DPIA.[91] These Codes of Practice can also cater to a specific sector. For example, a Code of Practice on the use of AI by the healthcare sector could include guidelines on demonstrating that data is collected and processed in a fair and lawful manner.[92] But the DPDP does not recognize Codes of Practice.

An obligation on the data fiduciary under the DPDP is to implement necessary security safeguards.[93] As per the Bill, these measures include taking necessary steps to prevent data misuse, and unauthorised access, disclosure and modification.[94] The data fiduciaries must implement necessary security safeguards when they use AI.

Another solution is to interpret the proposed relationship between data fiduciary and data principal. Presently, the DPDP does not define the fiduciary nature of the relationship or what it would entail. It is unaddressed whether the fiduciary duty means the entire set of obligations contained in the DPDP.[95] It is also unaddressed whether there is an additional duty of care to be undertaken by the

[87] *Id.*
[88] *Id.*
[89] *Id.*
[90] Clause 50, Data Protection Bill, 2021.
[91] *Id.*
[92] National Health Service UK, *A guide to good practice for digital and data-driven health technologies*, UK Government (Jan. 19 2021) https://www.gov.uk/government/publications/code-of-conduct-for-data-driven-health-and-care-technology/initial-code-of-conduct-for-data-driven-health-and-care-technology.
[93] Section 8, DPDP.
[94] Clause 24, Data Protection Bill, 2021.
[95] Smitha Krishna Prasad, *Information Fiduciaries and India's Data Protection Law*, Data Catalyst (Sept. 2019) https://datacatalyst.org/wp-content/uploads/2020/06/Information-Fiduciaries-and-Indias-Data-Protection-Law.pdf.

data fiduciaries.[96] Thus, a duty of care to protect privacy[97] can be provided especially when AI is used. This is important because the individual may not be in a position to understand complicated algorithms or the consequences of their use.[98]

Another solution is data protection by design and default. The underlying ideas of data protection by design and default are - privacy as a default setting, privacy embedded into the design, end-to-end security to ensure protection during the full lifecycle, respect for user privacy and transparency.[99] The products which incorporate AI and processing operations must be designed in such a manner that privacy protections are considered right at the beginning.[100] By default, the highest privacy protections must be ensured in the use of AI.[101] But, the DPDP does not explicitly recognize data protection by design and default.

Other technical solutions must also be explored. For example, the model of explainable AI can be used to make AI transparent. The solution is an easily explainable model of the decision-making process, and a way of ascertaining the attributes and weightage given to each attribute by the AI.[102] The outcomes of the AI must be measured for different attributes to assess whether there is bias against any given attribute.[103] To maintain data quality and avoid bias in the use of AI, the representativeness of the data input into the AI must be ensured.[104]

---

[96] *Id.*

[97] Matt Bartlett, *Beyond Privacy: Protecting Data Interests in the Age of Artificial* Intelligence, 3 Law, Tech & Hum 96 (2021).

[98] *Id.*

[99] Information Commissioner's Office, *Data protection by design and default,* https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/data-protection-by-design-and-default/

[100] European Commission, *What does data protection 'by design' and 'by default' mean?* , https://ec.europa.eu/info/law/law-topic/data-protection/reform/rules-business-and-organisations/obligations/what-does-data-protection-design-and-default-mean_en

[101] *Id.*

[102] KOAN Advisory and Digital India Foundation, *Handbook on Data Protection and Privacy for Developers of Artificial Intelligence (AI) in India: Practical Guidelines for Responsible Development of AI*, DSCI (Jul. 2021) https://www.dsci.in/sites/default/files/documents/resource_centre/AI%20Handbook.pdf

[103] *Id.*

[104] *Id.*

While the above solutions for regulating AI are essential, over-regulating AI may not be the appropriate solution. Over-regulating AI could limit and stagnate AI research and use of AI for beneficial purposes.[105] AI technology must be allowed to flourish and some flexibilities are essential. First, the DPDP provides exemptions for research or statical purposes.[106] Thus, there is a possibility that the use of AI for research purposes may be exempted from provisions of the Bill. At the same time, there is a need for certain safeguards. These safeguards were recognised by the Bill and include the principle of necessity, avoiding the risk of significant harm, avoiding specific decisions or directed actions and the requirement of de-identification as per Codes of Practice. [107]

Second, the Data Protection Board of India could create a Sandbox for encouraging innovation in AI and machine learning or emerging technology, a possibility recognised by the Bill.[108] Sandbox has been defined as live testing in controlled or test regulatory environments of new products or services.[109] Sandbox implies that regulatory relaxations may be provided for a specified time for limited testing purposes.[110] The relaxations could be from data protection principles and data protection obligations.[111] Sandbox addresses the fear that data protection requirements may impede the development of AI technologies.

While the PDP Bill, 2019 stated that the Authority "shall" create a Sandbox, the DP Bill, 2021 has replaced the word "shall" with the word "may".[112] This indicated that Sandbox may be given at the discretion of the Authority. Now, the DPDP does not have a Sandbox provision. The Board should consider Sandbox

---

[105] Matthew Humerick, *Taking AI Personally: How the E.U. Must Learn to Balance the Interests of Personal Data Privacy & Artificial Intelligence*, 34 Santa Clara High Tech. L.J. 393 (2018).
[106] Section 17, DPDP.
[107] Clause 38, Data Protection Bill, 2021.
[108] Clause 40, Data Protection Bill, 2021.
[109] *Id.*
[110] *Id.*
[111] *Id.*
[112] Joint Committee on The Personal Data Protection Bill, 2019, *Joint Committee on the Personal Data Protection Bill, Report of the Joint Committee on the Personal Data Protection Bill,2019*, http://164.100.47.193/lsscommittee/Joint%20Committee%20on%20the%20Personal%20Data%20Protection%20Bill,%202019/17_Joint_Committee_on_the_Personal_Data_Protection_Bill_2019_1.pdf

especially for small and medium enterprises that may benefit from increased innovation.[113]

Third, there is a need to rethink the fundamental principles of data protection. These principles may be inadequate to regulate AI and may also restrict AI development if they are given a strict interpretation.[114] The challenge posed by AI to personal data can be addressed by considering the risks of the re-identification of anonymized data. Once re-identified, anonymized data must be governed by all the data protection provisions.[115] The challenge posed by AI to collection limitation principle can be addressed by incorporating the safeguards of pseudonymisation and masking techniques without a reduction in the data.[116] The challenge posed to purpose limitation principle by AI can be addressed by having a flexible idea of processing for incidental purposes.[117] A safeguard can be provided that risks of significant harm must be avoided while repurposing and explicit consent must be taken for action directed to individuals.[118] To reconcile the Right to be forgotten with AI, the possible solution is to isolate or delete strands of AI's learning.[119] But, isolation of learning is not possible in the case of

[113] Lilian Mitrou, *Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof'?*, SSRN (Jun. 3, 2019) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3386914

[114] Christopher Kuner, Fred H Cate, Orla Lynskey, Christopher Millard, Nora Ni Loideain, and Dan Jerker B Svantesson, *Expanding the artificial intelligence-data protection debate*, 8 International Data Privacy Law 289 (2018).

[115] Panel for the Future of Science and Technology, *The impact of the General Data Protection Regulation (GDPR) on artificial intelligence*, European Parliament (Jun. 2020) https://www.europarl.europa.eu/RegData/etudes/STUD/2020/641530/EPRS_STU(2020)641530_EN.pdf

[116] *Id.*

[117] *Id.*

[118] Committee of Experts under the Chairmanship of Justice B.N. Srikrishna, *A Free and Fair Digital Economy Protecting Privacy, Empowering Indians*, Ministry of Electronics and Information Technology, https://www.meity.gov.in/writereaddata/files/Data_Protection_Committee_Report.pdf

[119] Matthew Humerick, *Taking AI Personally: How the E.U. Must Learn to Balance the Interests of Personal Data Privacy & Artificial Intelligence*, 34 Santa Clara High Tech. L.J. 393 (2018).

Robert Walters and Matthew Coghlan, *Data Protection and Artificial Intelligence Law: Europe Australia Singapore - An Actual or Perceived Dichotomy?*, SSRN (Feb. 18 2020) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3503392

AI such as neural networks.[120] Thus, the possible solution is for the Right to be forgotten to allow retention of information up to the point the Right has been requested.[121]

IV. Conclusion

As India has now enacted a data protection legislation, the potential challenges presented by AI need to be considered. The proposed solutions are that India could provide a fiduciary duty of care on the data fiduciary towards the data principal. The Data Protection Board of India could recognize data protection by design and default. Technical solutions must be explored such as designing AI in a manner that rights such as the Right to correction and the Right to be erasure are provided from the beginning and irrespective of the kind of AI.[122] Codes of Practise must be used to define data protection standards for use of AI in specific sectors. India also needs protect data principals from automated decision-making including profiling. Lessons can be learnt from the EU which provides the right to contest decisions made by AI, the right to obtain human intervention and the right not to be subject to automated decision-making affecting the individual.[123]

Data protection rights must be protected throughout the processing life-cycle of AI - both at the time of development of AI and also while employing AI for making decisions.[124] At various stages of processing, there is also a need for qualified human oversight to ensure that rights are respected and negative effects for individuals are avoided.[125] There is also a need for transparency by providing

---

[120] Matthew Humerick, *Taking AI Personally: How the E.U. Must Learn to Balance the Interests of Personal Data Privacy & Artificial Intelligence*, 34 Santa Clara High Tech. L.J. 393 (2018).
[121] *Id.*
[122] European Data Protection Board and European Data Protection Supervisor, *Joint Opinion 5/2021on the proposal for a Regulation of the European Parliament and of the Council laying down harmonised rules on artificial intelligence (Artificial Intelligence Act)*, EDPB (Jun. 18 2021) https://edpb.europa.eu/system/files/2021-06/edpb-edps_joint_opinion_ai_regulation_en.pdf
[123] Article 22, GDPR
[124] Lilian Mitrou, *Data Protection, Artificial Intelligence and Cognitive Services: Is the General Data Protection Regulation (GDPR) 'Artificial Intelligence-Proof'?*, SSRN (Jun. 3, 2019) https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3386914
[125] *Id.*

information to the data principal about the logic of the AI, the scope of processing, and legal basis for processing at various stages of processing.[126]

While data protection is essential, there is also a need to ensure that the development of AI technology is not hindered. AI technology has potential benefits which can lead to the progress of society. Thus, a delicate balance needs to be established between protecting privacy and data protection while allowing AI technology to develop.

[126] *Id.*