Data Privacy Considerations for Artificial Intelligence Systems Use in Nigeria: The Nigeria Data Protection Act (2023) in Focus

Author: Samuel Uzoigwe and Anastasia Edward[1]

1. Introduction:

Artificial Intelligence (AI) is rapidly altering the mode and manner of human and business interactions in the world, with its applications present in virtually every industry and sector ranging from financial to health, real estate, human resources, cloud computing and storage, telecommunications, entertainment/content creation etc. The common denominator in the application of AI in these industries is that data is processed on a massive and continuous scale throughout the lifecycle of AI systems, and much of this data is personal data. The interplay between AI and privacy is evolving and because developing AI systems is iterative coupled with the presence of personal data processing after deployment, AI utilization raises new and complex privacy concerns which pose risks to the rights and freedoms of humans.

Pre-deployment, AI systems are often trained on massive datasets that contain personal data, which could include sensitive personal data. When AI systems are trained using personal data, they acquire the capacity to make inferences and identify objects, patterns and relationships that can be used to make predictions about human behaviour and preferences. This could be useful in many cases, but it could also pose risks to the rights and freedoms of data subjects and individuals at large. The personal data in training datasets could also be deployed for purposes beyond that for which it was initially processed, and privacy breaches could expose data subjects to numerous risks in the hands of malicious threat actors.

This work examines the delicate and intricate nexus between AI development and deployment and data privacy and protection. It also outlines key privacy considerations that AI developers or organizations should consider to enable compliance with the Nigeria Data Protection Act (NDPA) 2023. The work is structured into three (3) major sections comprising of the examination of AI systems use in a Nigerian context; the nexus between AI and the NDPA 2023, as well as data privacy considerations for AI systems use in Nigeria. There will be references to foreign guides for context where the NDPA does not provide a sufficient description of a subject or concept.

[1] Samuel Uzoigwe is an Executive Associate at Alliance Law Firm, Lagos, Nigeria.

2. Artificial Intelligence: Uses and Lifecycle

In the Nigerian context, artificial intelligence (AI) is increasingly deployed in diverse applications, including creditworthiness assessment by financial institutions, talent acquisition processes, document authentication, biometric recognition in consumer electronic devices and smart home systems, plagiarism detection, data analysis, chatbots for consumer related services, and generative AI. All the aforementioned applications involve the processing of personal data at some point, and in various capacities and volumes. Some of the industries in Nigeria where this utilization of AI systems is common include legal, finance, healthcare, security, telecommunications, insurance, real estate, etc. Artificial Intelligence has subsets such as deep learning (DL) – the recognition of complex patterns in pictures, text, sounds, and other data to produce accurate insights and predictions e.g virtual assistants, facial recognition and language translation;[2] machine learning (ML) – the focus on the use of data and algorithms to imitate the way and manner in which humans learn, gradually improving the accuracy of the system e.g customer service chatbots;[3] and natural language processing (NLP - the branch of AI) that enables computers and machines to comprehend, generate, and manipulate human language e.g analysis of large documents and also chatbots. An AI system could sometimes rely on third-party frameworks and codes, which creates increased complexity of relationships involving personal data processing.

Artificial Intelligence systems typically transit from the design and development phase to the deployment phase. The design and development phase includes building, testing, and validation of the models by AI developers in collaboration with software engineers, analysts, enterprises etc., to ensure that it meets performance metrics, baselines and is scalable before deployment at production levels. After an AI model has gone through the iterations of the development phase, it is deployed into production.[4] As summed by Rybalko, "deployment is the process of configuring an analytic asset for integration with other applications or access by business users to serve production workload at scale."[5] Each phase within the outlined life cycle of AI systems exhibits distinct activities, contexts, and goals, entailing a spectrum of varying privacy risks and considerations. This is further accentuated by the pervasive practice of collecting personal data without data subject participation, as exemplified by web scraping and facial recognition technologies.

---

[2] Amazon Web Services, "What is deep learning," https://aws.amazon.com/what-is/deep-learning/#:~:text=Deep%20learning%20is%20a%20method,produce%20accurate%20insights%20and%20predictions.

[3] IBM, "What is Machine Learning," https://www.ibm.com/topics/machine-learning

[4] Dmitriy Rybalko, AI model lifecycle management: Deploy Phase, IBM, https://www.ibm.com/blog/ai-model-lifecycle-management-deploy-phase/, accessed January 15, 2024.

[5] Ibid.

3. Nexus Between Artificial Intelligence Systems Use and the Nigeria Data Protection Act 2023:

Section 2(1) of the Nigeria Data Protection Act 2023 ("the NDPA" or "the Act"), provides that the Act shall apply to all forms of processing of personal data, whether by automated means or not. The processing of personal data in AI systems, whether at the development stage or after deployment, falls within the ambit of the NDPA, and the data controller or processor is expected to carefully ensure the lawfulness of such processing, among other key privacy considerations under the Act.

Data Preparation/Data Pre-Processing

Data Preparation is a critical part of an AI system development process. It is the process of gathering, combining, structuring and organizing data so that it can be used in business intelligence (BI), analytics and data visualization applications.[6] Data preprocessing on the other hand, a component of data preparation, describes any type of processing performed on raw data to prepare it for another data processing procedure.[7]

It is usually a misconception that data privacy obligations do not apply to these phases of an AI model development. Under the NDPA, once any operation or set of operations is performed on any information relating to an identified or identifiable individual during these stages, personal data processing is deemed to have occurred.[8] This includes during data collection, discovery and profiling, cleaning, structuring, transformation, validation, etc. Therefore, the data preparation and pre-processing stages fall within the purview of the NDPA irrespective of how they are designated once an individual can be identified directly or indirectly through the concerned data, until such personal data is deidentified.

4. Data Privacy Considerations for Artificial Intelligence Systems Use;

By virtue of the processing of personal data both at the development and deployment of artificial intelligence models/systems, several privacy considerations must be taken into account by AI system developers and users that operate within the scope and jurisdiction of the NDPA. The factors below are not exhaustive, and new regulations guides, and global best pest practices must be considered at all times to ensure robust compliance with the NDPA. It should also be

[6] Craig Stedman, industry editor | Ed Burns, executive editor | Mary K. Pratt, "What is Data Preparation? An in-depth Guide to Data Prep", TechTarget, https://media.techtarget.com/digitalguide/images/Misc/EA-Marketing/Eguides/What_is_Data_Preparation_An_In-Depth_Guide_to_Data_Prep.pdf, accessed January 15, 2024.
[7] George Lawton, "Data Preprocessing", Techtarget, https://www.techtarget.com/searchdatamanagement/definition/data-preprocessing, accessed January 15, 2024.
[8] Section 65, of the NDPA.

noted that the considerations below should not be confused with pure principles of responsible AI use.

i. Lawfulness, Fairness and Transparency of Processing:

This is one of the key principles of data privacy and is mandated by section 25(1)(a) of the NDPA. This principle has three arms and demands that all personal data processing must be lawful fair and transparent. The NDPA does not elucidate on this principle but the GDPR and the UK Information Commissioner's Office (UK ICO) Guide provide some clarity. The UK ICO, in its Guide on AI and Data Protection, recommends that to achieve adequate compliance under the lawfulness of the processing arm, each distinct operation during the development phase and the deployment phase must be broken down and an appropriate lawful basis for processing identified.[9] An apt illustration of the appropriateness of this recommendation can be found in the difference between the purposes for designing and training an AI model by a software engineer and data scientists, and the purpose for the purchase and use of such system by a financial institution for credit scoring purposes. The software engineer will primarily focus on designing and training models for optimum performance and prediction accuracy, while the financial institutions deploy it to determine credit worthiness of individuals. The Financial institutions and the AI systems designers will typically encounter disparate legal basis for processing personal data in the varying circumstances, and must correctly link each basis to the purpose of each processing.

The Fairness principle requires that personal data should not be processed in a way that is unjustifiably detrimental, unlawfully discriminatory, unexpected or misleading but rather in a way that is reasonably expected by the data subject.[10] It is a key tool to check deception as to the nature and purposes of processing.[11] As enjoined by the UK ICO Guide, when an AI system is developed and deployed to infer data about people, it is expedient that the system is sufficiently statistically accurate and avoids discrimination. Under the NDPA, it is a data subject's reasonable expectation that the training and deployment of an AI system's architecture must be meticulously done to prevent the amplification of societal biases which could lead to detrimental and discriminatory outcomes against such data subject or any class of data subjects. As equally explained by the UK ICO, the fairness principle means that your AI system needs to be

---

[9] UK ICO Guide on AI and Data Protection, https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/how-do-we-ensure-lawfulness-in-ai/, accessed January 15, 2024.

[10] EDPB Guidelines 4/2019 on Article 25 Data Protection by Design, 2020, https://edpb.europa.eu/sites/default/files/files/file1/edpb_guidelines_201904_dataprotection_by_design_and_by_default_v2.0_en.pdf, accessed January 15, 2024.

[11] Damian Clifford & Jef Ausloos, "Data Protection and the Role of Fairness," CiTiP Working Paper 29/2017, KU Leuven Centre for IT & IP Law, available at SSRN ID=1781425

sufficiently statistically accurate for your purposes. AI systems should also be developed and deployed in a way that makes it easy for data subjects to exercise their rights under the NDPA.

The transparency principle as espoused by the General Data Protection Regulation (GDPR) 2018, requires that any information and communication relating to the processing of personal data should be easily accessible and easy to understand, in clear and plain language.[12] In an AI use context, the above principles form an intricate part of the other and is the foundation on which many data privacy and protection rights are built.

ii. Accountability:

Section 24(3) of the NDPA imposes on data controllers and data processors the duty of care and accountability in respect of personal data processing. Ensuring compliance with these duties lie with the controller who shall ensure adherence of employees, processors, or vendors/contractors to this obligation unless where the other party is also a controller. One of the mechanisms devised to ensure accountability in Nigeria is the filing of annual data protection audits with the regulatory authority, although this is not conclusive proof of compliance with the NDPA and should not be approached in box-ticking manner.

iii. Accuracy of Personal Data:

This is a key data privacy obligation under the NDPA which requires that only personal data that is accurate, complete and not misleading is processed.[13] Considering the fact that personal data could be scraped from the internet for use in training AI systems, some of these personal data could lose contextual accuracy leading to a misleading output which could affect the rights of such data subject. For instance, a bank's facial recognition technology. Individuals with good creditworthiness could be denied loans due to inaccurate predictions arising from incomplete personal data processing, restricting access to financial resources and hindering economic opportunities. We also have the possibilities of unauthorized third parties having access to a data subject's mobile device as a result of incorrect processing of a third-party's facial features by the device's facial recognition technology to unlock the data subject's device. These will breach the accuracy of processing obligation under the NDPA. It is without peradventure that demographic data changes creates a challenge for the accuracy of personal data, and where an AI system has already been trained using personal data, updating such personal data could create a challenge.

iv. Data Protection Impact Assessment:

[12] Recital 39, Regulation (EU) 2016/679 (General Data Protection Regulation)
[13] Section 24(1)(e)

Considering the use of artificial intelligence systems especially to make inferences and assessment of individuals, the development and deployment of AI systems creates high privacy risks to the rights of data subjects in several instances. There is also the case of third-party dependencies and relationships in the development and deployment of AI systems, further exacerbating the privacy risks inherent in an AI system. Where this is the case, a data privacy impact assessment (DPIA) is mandated under the NDPA.[14] As aptly noted by the U.S National Institute of Standards and Technology, while pre-deployment AI risk assessment in a testing, staging or controlled environment may yield important insights, the true spectrum of risks will only emerge in operational, real-world settings where there may be integration with third- party tools and systems, and these may be in stark contrast to earlier envisaged risks.[15]

The above situation implies that a DPIA may need to be conducted or revisited on several occasions as new privacy risks to data subjects arise and measures to minimize risks are implemented. Where identified high risks to data subjects persist, notwithstanding measures deployed to mitigate such risks, the National Data Protection Commission shall be consulted. The NDPA does not state instances or any baseline that could aid the determination of high risks to data subjects in a privacy context, but the GDPR provides some valuable insight and provides that a DPIA will typically be required in situations where there is:

1. a systematic and extensive evaluation of personal aspects relating to natural persons which is based on automated processing, including profiling, and on which decisions are based that produce legal effects concerning the natural person or similarly significantly affect the natural person;

2. processing on a large scale of special categories of data, or of personal data relating to criminal convictions and offences; or

3. a systematic monitoring of a publicly accessible area on a large scale.[16]

Conducting a DPIA is an intensive process requiring high-level collaboration between privacy and non-privacy professionals. Some key issues to be considered during a DPIA on AI systems include limitation of data subjects exercise of their rights, possibility of identification and exposure of personal data, inability of individuals to access services or opportunities by virtue of automated profiling, discrimination, etc.[17] The UK ICO also recommends the consideration of allocative and representational harms that processing may have on individuals during a DPIA. Allocative harms arise from the decision to allocate goods and opportunities among a group. This leads to disparities in access to financial resources, livelihood, liberty, and even survival in

[14] Section 28 of the NDPA
[15] NIST Artificial Intelligence Risk Management Framework https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf at page 6, accessed January 14, 2024.
[16] Article 35(3) GDPR.
[17] UK ICO Guidelines, Data Protection Impact Assessment, https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/accountability-and-governance/data-protection-impact-assessments-dpias/how-do-we-do-a-dpia/#how5

extreme cases.[18] Representational harm arises when algorithmic systems perpetuate social hierarchies and marginalization by reinforcing negative stereotypes, underrepresenting certain groups, and diminishing their dignity through denigration.[19] These are harms that affect individuals and reduce their capacity to access necessary services.

v. Purpose Limitation:

The purpose limitation principle under the NDPA states that all personal data processing shall only be collected for specified, explicit and legitimate purposes, and not to be processed in a way incompatible with those purposes.[20] This is a delicate principle to navigate that is linked to other principles of processing and will also be dependent on the mode of collection of personal data. For instance, the application of the purpose limitation principle to personal data collected directly from a data subject will vary from that of personal data scrapped off a website or based on other legal bases for processing. This principle dictates that personal data collected for other purposes should not be redeployed to train an AI system unless such training is compatible with the purposes for which it was initially collected. This prohibition can be only circumvented by seeking the consent of the data subject or relying on other appropriate legal bases. To ascertain the compatibility of further processing to the original purpose, regard should be had to the relationship between the original purpose and the purpose for further processing, as well as the consequences of the further processing, among others.[21]

vi. Data Minimization:

The important question that encapsulates this principle is, "Why do we need this data?" Section 24(1)(c) of the NDPA provides that only personal data that is adequate, relevant, and limited to the minimum necessary for the purposes for which it is sought to be processed should be collected. AI systems could be built in-house by organizations or procured and installed as stand-alone applications or as part of an organization's infrastructure with customization capacity. This makes the application of this principle vary in context. For example, at the development stage of an AI chatbot being designed to assist customers of a financial institution, is it relevant to collect information regarding the genotype of individuals to form part of the training data of such a

[18] The UK ICO gives an example of allocative harm in an instance where an organisation may use an AI system in recruitment that disproportionally classifies applications from male candidates as suitable compared to women. The use of this system has implications for the allocation of job opportunities to female candidates and the relevant economic results

[19] The UK ICO gives an example in the instance when an individual belonging to an ethnic minority group uploads pictures of their holiday photos on an internet platform. The image recognition system operated by the platform assigns labels to their 'selfie' photos that are denigrating reflecting racist tropes.

[20] Section 24(1)(b) NDPA 2023.

[21] Section 24(4) NDPA.

model? Upon deployment, is it relevant and necessary for such a virtual assistant to demand access to the phonebook of a customer to be able to verify the identity of a customer? Answers to these questions are not always in black and white and will always depend on the reason why such personal data is needed.

Irrespective of the fact that AI systems typically require vast amount of training data for design or improvement after deployment, the NDPA mandates compliance with this obligation. The UK ICO Guide opines that the key to compliance with this rule is that you only process the personal data you need for your purpose, and does not mean either 'process no personal data' or 'if we process more, we're going to break the law.'[22] This requires high-level collaboration between the privacy team/consultants/professionals and the AI developers/data scientists, alongside buy-in from management. The application of this principle commences at the design stage or as part of the procurement process due diligence when an organization purchases AI systems or implements AI systems operated by third parties.[23] The UK ICO recommends perturbation or adding 'noise,' synthetic data federated learning, converting personal data into less 'human-readable' formats, making inferences locally, and privacy-preserving query approaches, anonymization and pseudonymization as techniques that could be applied to the personal data processing to minimize the personal data utilized in AI systems.

At the training stage, striking a balance between data minimization and having sufficient training data for AI models is also crucial. This necessitates the identification of essential features within training datasets, ensuring models' statistical accuracy and non-discriminatory functionality in line with the principle of fairness.[24] This is also in line with the adequacy component of the data minimization principle, which **dictates that** personal data should not be processed if it is insufficient for its intended purpose.

vii. Security of Personal Data:

As aptly noted by the NIST, the deployment and utilization of AI systems presents unique risks to society at large due to the complex interplay of technical aspects of these systems combined with societal factors related to how a system is designed and deployed. These factors include its interactions with other AI systems or the integration of third-party tools, the decision-making

[22] UK ICO Guide, https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/artificial-intelligence/guidance-on-ai-and-data-protection/how-should-we-assess-security-and-data-minimisation-in-ai/, last accessed January 15, 2024.
[23] Ibid.
[24] Ibid.

process of its operators, and the social context in which it is deployed also exacerbate these risks.[25] Due to the above, privacy risks abound and could pose high risks to data subjects.

Sections 24(1)(f), 21(2) and 39 of the NDPA mandates the processing of personal data using appropriate technical and organizational measures that ensures the security of personal data. The NDPA recommends the implementation of measures such as the deidentification of personal data, encryption, regular assessment of the effectiveness of measures, etc. The privacy of data subjects whose personal data is used to interact with AI systems can be threatened through threat actors employing various attack vectors such as model inversion attacks, membership inference attacks, black box attacks and white box attacks, which could also enable model inversion attacks. Security challenges may also arise in a bid to make AI systems explainable to users and stakeholders in line with the AI explainability principle.

The adoption of appropriate security measures will be dependent on the peculiar circumstances. Proper application of data minimization principles under section 24(1)(c) of the NDPA also helps limit the size of personal data that may potentially be exposed to security risks. This principle can be complied with by having and implementing a privacy policy with state-of-the art security measures, and data breach management protocols in the event of a personal data breach. Considering the evolving nature of security attacks and risks to personal data in AI systems, it is paramount to have periodic training of personnel and keeping tabs with advancements within the field.

viii. Designation of a Data Protection/AI Governance Officer:

The NDPA mandates data controllers of major importance to designate data protection officers with expert knowledge of data protection laws and practices to ensure compliance with the NDPA and other subsidiary legislation.[26] The NDPC recently issued its Guidance Notice on the Registration of Data Controllers and Processors of Major Importance (DCMI/DPMI). By virtue of paragraph 1(1) of the Notice, Data controllers are deemed to be of Major Importance if they **maintain a filing system (analog or digital) for processing personal data, AND** Process the personal data of more than 200 individuals within six months, OR Provide commercial Information and Communication Technology (ICT) services on digital devices with storage capacity belonging to others, OR Process personal data as an organization or service provider in any of these sectors: finance, communication, health, education, insurance, import/export, aviation, tourism, oil and gas, or electric power. According to the Guide, it is certain data controllers or processors that design or utilize AI systems where personal data processing is

[25] NIST Artificial Intelligence Risk Management Framework available at https://nvlpubs.nist.gov/nistpubs/ai/nist.ai.100-1.pdf at page 1 accessed January 14, 2024.
[26] Section 32 of the NDPA.

carried out in any form will be classified as DCMI if they meet the designated criteria above at a minimum. It should be noted that there are other risks that arise from the use of AI that are not privacy based and as such an AI Governance Officer could become a necessity in the near future in Nigeria.

ix. Storage Limitation:

This principle under the NDPA dictates that personal data should not be stored beyond the period necessary to achieve the purpose for which it was processed.[27] This principle is particularly tricky to navigate considering the nature of AI systems, which may need retraining to reflect demographic changes or update the model in line with new information or approaches. There is also the importance of maintaining a balance between the adequate functionality of AI systems and this obligation. Observing data minimization principles and employing personal data deidentification techniques aid in observing this principle. It is important to have data retention policies and consistently adapt these policies in line with the best data retention standards.

x. Exercise of Data Subjects Rights:

The NDPA bestows several rights on data subjects which can only be derogated from in limited circumstances.[28] The most applicable rights within an AI use context include the right to information, right of access to personal data,[29] right to correction,[30] right to erasure,[31] right to restriction of processing,[32] right to objection,[33] right to withdraw consent,[34] and the right not to be subjected to automated decision making.[35] Subsumed under the right not to be subjected to automated decision-making is the right of a data subject to contest an automated decision and obtain human intervention on the part of the controller if dissatisfied.[36] For example, if through an automated process, a system concludes that a loan applicant is ineligible to take a loan, such applicant should be able to request a human review of his application if dissatisfied. Irrespective of the challenging nature of the obligation to enable an exercise of these rights without constraints or unreasonable delay, organizations that utilize these systems must have procedures in place to enable seamless exercise of these rights by data subjects. The NDPA does not prescribe a timeframe within which to respond to data subject rights requests. However, it is

[27] Sections 24(1)(d) and 34(2) NDPA.
[28] Sections 3(2), 36(2), 37(2) NDPA
[29] Section 34(1)(b) NDPA.
[30] Section 34(1)(c) NDPA.
[31] Section 34(1)(d) NDPA.
[32] Section 34(1)(e) NDPA.
[33] Section 36 NDPA.
[34] Section 35 NDPA.
[35] Section 37(1) NDPA.
[36] Section 37(3) NDPA.

prudent to designate a timeframe for that purpose. It is also prudent to document every right request and resolution of the request, and where such request is rejected, the reason should be documented properly. Where the exercise of data subjects' rights would be difficult to facilitate, this should be documented in detail in the DPIA.

xi. Exemption of Application:

The NDPA exempts the application of the Act in several circumstances listed under Section 3(2). For example, a DPIA is not mandated by the NDPA, where processing is carried out by a competent authority for the prevention, detection, or investigation of a crime, or the prevention and control of a national public health emergency, national security, etc. The NDPC is also empowered to prescribe types of personal data and processing that may be exempted from the application of the NDPA.[37]

e. Conclusion:

It is without peradventure that the processing of personal data in an AI system falls within the ambit of the NDPA in Nigeria except in limited circumstances. Consequently, AI system developers, data controllers, and data processors utilizing AI systems are unequivocally bound to comply with the NDPA's provisions. However, ensuring compliance presents a significant challenge due to the inherent complexities associated with AI utilization. Nonetheless, data controllers and processors are strictly obligated to adhere to Nigerian data privacy laws throughout the entire lifecycle of any AI system, from the initial conceptualization and design phases to its deployment and operationalization. This obligation commences even before the first line of code is written by an AI developer, or an AI system is procured for organizational use. Consequently, AI developers must actively embrace data privacy by design and default approaches to mitigate potential privacy risks inherent in AI use, and thus facilitate robust compliance with the NDPA.

[37] Section 3(3) NDPA.