

Summary of Six Principles of DGPSI_AI

DGPSI-AI Principle-No	Principle
1	Unknown Risk is a Significant Risk
2	Behind every AI algorithm there shall be one human for accountability
3	Every Privacy Notice covering an AI Process involved in processing of personal data shall be accompanied by an Explainability disclosure.
4	Use of every AI Process shall be validated by a document justifying the technical, operational and economical need both at the level of the Data Fiduciary and the Data Processor with unconditional indemnity to the data principal.
5	Every AI process shall document the specific guardrails to secure the processing against Dark Patterns, Neurological manipulation and physical harm to any data principal.
6	The responsibility of the AI deployer as a "Fiduciary" shall ensure all measures to safeguard the society from any adverse effect arising out of the use of the AI.



Nine Model Implementation Specifications of DGPSI-AI for Deployers

Following is the list of suggested Nine Model implementation specifications under DGPSI AI for deployers who are data fiduciaries under DPDPA.

MIS-AI No	Specification
1	The deployer of an AI software in the capacity of a Data Fiduciary shall document a Risk Assessment of the Software. During the process the Data Fiduciary shall also obtain a confirmation from the vendor that the software can be classified as 'AI' based on whether the software leverages autonomous learning algorithms or probabilistic models to adapt its behaviour and generate outputs not fully predetermined by explicit code. This shall be treated as DPIA for the AI process.
2	The DPIA shall be augmented with periodical external Data Auditor's evaluation at least once a year.
3	Where the data fiduciary in its prudent evaluation considers that the sensitivity of the "Unknown Risk" in the given process is not likely to cause significant harm to the data principals, it shall create a "AI-Deviation Justification Document" and opt not to implement the "Significant Data Fiduciary" obligations solely as a reason of using AI in the process.
4	Designate a specific human handler on the part of Deployer-Data Fiduciary to be accountable for the consequences of the use of AI in personal data processing. By default the DPO/Compliance officer will be accountable. However, the "Process Owner" envisaged under the DGPSI framework and Process based compliance could be an alternate designate.
5	Document the human handler for the AI on behalf of the licensor through the licensing contract and if the developer has hardcoded the accountable person for the AI in the Code, the same may be recorded in the licensing contract.
6	The deployer shall collect an authenticated "Explainability" document from the developer as part of the licensing contract indicating the manner in which the AI functions in the processing of personal data and the likely harm it may cause to the data principals.
7	The deployer shall develop a "AI Justification Document" before adopting an AI led process for processing personal data coming under the jurisdiction of DPDPA justifying the use of AI and exposing the data principals to the unknown risks from technical and economical perspectives.
8	Document an assurance from the licensor that

	1. the AI software is adequately tested at their end for vulnerabilities, preferably from the third party auditor. The document should state that the "When deployed for data processing, the AI Software is reasonably secured against vulnerabilities that may adversely affect the confidentiality, integrity and availability of data and the Privacy principles where the data processed is "Personally identifiable data". 2. The document shall also mention that sufficient guard rails exist to protect the Data Principals whose data may be processed by the deployer. 3. The document shall also mention that the AI has been tested and is free from any malware that may affect other systems or data owners.
9	The Deployer of an AI shall take all such measures that are essential to ensure that the AI does not harm the society at large. In particular the following documentation of assurances from the licensor is recommended. 1.The AI comes with a tamper-proof Kill switch. 2.In the case of Humanoid Robots and industrial robots, the Kill Switch shall be controlled separately from the intelligence imparted to the device so that the device intelligence cannot take over the operation of the Kill Switch. 3.Where the kill switch is attempted to be accessed by the device without human intervention, a self destruction instruction shall be built in. 4.Cyborgs and Sentient algorithms are a risk to the society and shall be classified as Critical risks and regulated more strictly than other AI, through an express approval at the highest management level in the data fiduciary. 5.Data used for learning and modification of future decisions of the AI shall be imparted a time sensitive weightage with a "Fading memory" parameter assigned to the age of the observation. 6. *Ensure that there are sufficient disclosures to the data principals about the AI risk*

13 Implementation Specifications for DGPSI-AI to be mandated by Deployers on the AI Vendors

MIS-No (DGPSI_AI-Developer)	Description
1	The AI developer shall generate an "Explainability" document for the AI that explains the Algorithmic function of the Model embedded in the software using the key principles of transparency as per enclosed format. (refer footnote)
2	The AI developer shall provide the Business Contact details of the "Human Handler" of the AI model responsible for its functioning as part of the licensing contract
3	The AI developer shall document the Training and Testing process adopted for the development of the model.
4	The AI developer shall document a Risk Assessment of the model indicating its susceptibility to third party security compromise and the potential harm to the user or data principals whose personal data may be processed as well as the society at large.
5	The AI developer shall document the Guardrails incorporated in the AI model to mitigate the security risks
6	The AI developer shall document the default configuration of the model.
7	The AI developer shall incorporate a set of comprehensive instructions to the users on any re-configuration or re-training that may be suggested, required or is expected in its normal use
8	The AI developer shall incorporate a set of emergency handling instructions in case the AI has a risk of hallucination or going rogue.
9	The AI developer shall incorporate a "Kill Switch" which is reasonably tamper proof.
10	The Kill Switch shall be configured so as not to be accessible by the Model and shall be controlled in a separate chip or circuit that can be accessed independently with a provision that if the Model tries to access the Kill Switch, the Algorithm should self-destruct.
11	The AI model shall be audited by an independent third-party auditor using an acceptable audit standard.
12	If the AI is classified a "Critical Risk" taking into consideration its autonomy, the type of sensitive data it is expected to process and the automated decision-making capability that could affect the society, a

	post implementation behaviour monitoring shall be made available at the choice of the deployer.
13	The AI developer shall document the use of AI agents as part of its work force and its likely impact on the AI algorithm/Model developed

For any clarifications, contact Naavi founder of www.naavi.org and Chairman of Foundation of Data Protection professionals in India (www.fdppi.in)

This is released in public interest and is requested to be used responsibly. Suggestions for modifications are welcome. FDPPI will be open to improvements and a special committee for the purpose would be available at FDPPI.

Comments can be sent to dgpsi@naavi.org .