Wipro's Policy

# Responsible Use and Development of Generative AI

**Purpose**

Generative AI (GenAI) technology is rapidly advancing, and tools are becoming increasingly available. While GenAI can provide significant benefits to an organization in terms of productivity enhancement and revenue generation, it also poses risks to privacy, cybersecurity, intellectual property, third-party/client engagements, legal obligations, and regulatory compliance.

This policy aims to forge a responsible usage, deployment and development of Generative AI across Wipro, harness the advantages that the GenAI technology bring whilst mitigating the risks and challenges, and provide guidelines for responsible usage, deployment and development of GenAI tools for internal use and client engagement.

**Scope**

This policy covers usage, deployment and development of Generative AI tools and technology.

This policy does not override any policy, process and guidance related to privacy, data protection, code of business conduct, intellectual property, and confidentiality. For example, any use case related to the deployment of Generative AI for internal utilization (e.g., to optimize HR processes, sales, and marketing campaigns, etc.), must undertake all the existing procedures, including security assessment, intellectual property and confidentiality due diligence, and privacy assessment when personal data is involved.

This policy is a living document as it reflects the fast-evolving nature of technology, which we embrace in a responsible, human-centric, sustainable and privacy preserving manner, in full adherence to the spirit of Wipro.

**Policy Details**

Generative AI refers to artificial intelligence systems that have the capability of generating new content (such as images, text, audio, codes, presentation, etc.) based on the data they were trained on. While Generative AI can revolutionize the way we operate, there are inherent risks associated with this technology.

These risks include but are not limited to:

- **Privacy and data protection.** Generative AI tools require large datasets thus pose issues related to data minimization and legal basis for processing data. These are core privacy law tenets now recognised all across the globe, and ones we must abide to. In addition, these tools present risks around fairness in both processing and outputs; opacity around the workings of generative AI may clash with transparency and informational right requirements. Finally, there are challenges around accuracy as generative AI tools may produce false information.

- **Security and confidentiality.** Through generative AI, attackers may generate new and complex types of malware, phishing schemes and other cyber dangers that can avoid conventional protection measures. Such assaults may have significant repercussions like data breaches, financial losses, and reputational risks. In

addition to personal data leakage risks, the use and development of Generative AI is also susceptible to data inference attacks, data poisoning and other forms of adversarial attacks that may compromise the security and confidentiality of data.

- **Intellectual property and enterprise proprietary information.** All aspects of Generative AI from the model, training data, prompts, to output pose IP risks namely:

    - *Infringement of IP Rights:* The training data used to train the Generative AI models could include copyrighted material and if the output of these models is the similar (or a derivative work) or in rare scenarios, the exact same as the input training data, then this could potentially infringe on copyright laws. Generative AI could inadvertently use or refer to trademarked products, brands or logos that could also be seen as infringing on those trademarks. The machine authored output could also be very realistic fake images, videos, or texts, which could be used to infringe on someone's copyright. The prompts used to interact with Generative AI systems may contain copyright-protected information, which could lead to infringement.

    - *Unclear Ownership*: The issue of IPR (Intellectual Property Rights) protection for machine authored content is an unclear, complex and evolving area of law. When integrated into Wipro or Client owned IP, it can result in downstream IP licensing and/or enforcement risks and challenges.

    - *Attribution:* AI generated content can be almost indistinguishable from human generated content and as such can lead to a risk of humans not getting due credit and attribution as the rights holder of their work.

- **Misinformation.** Generative AI can produce biased or inaccurate outputs and potentially lead to poor decision-making and legal or ethical consequences. Generative AI can also create content and represent facts even if they don't exist. In addition, the outdated data on which it is trained can lead to inaccurate predictions, poor recommendations.

- **Consumer Protection.** Businesses that fail to disclose usage of large language models to consumers run the risk of losing customer trust and be charged with unfair practices under various laws (e.g. the California chatbot law mandates that in certain consumer interactions, organizations must disclose clearly and conspicuously that a consumer is communicating with a bot.)

**Guidelines**

To mitigate the above-mentioned risks and to comply with internal policies and AI related legislation, the following guidelines must be adhered to in the (1) usage, (2) deployment and development of Generative AI tools and technology.

1. **Guidelines for Users**

    Users of Generative AI ("users") refers to personnels accessing and utilizing Generative AI tools for enhancement of their daily job or for client delivery. Users must adhere to the guidelines prescribed below.

    - Personal data should not be entered on Generative AI tools including but not limited to: ChatGPT, Bard, Co-Pilot, GPT-4, Dall-E. Personal data include names, addresses, phone numbers, or any other information that could be used to identify an individual. Note that personal data may differ across jurisdictions.

    - Users must avoid using language or content that may have proprietary customer, partners and Wipro's confidential information or mention of our customers, leadership team etc.

- Users must clearly indicate content that is generated by GenAI tool to avoid confusion with human-generated content. Users must acknowledge the source of any ideas or insights generated by the tool.

- GenAI tools can only be used for client projects if approved by clients or if the use is allowed as per client contract. Similarly, client enterprise data including personal details should not be used in Generative AI without client approval.

- Account teams should reach out to *Responsible AI taskforce* to obtain clearance for the usage Generative AI tools in all new and existing client engagement.

- GenAI tools that are not approved by Wipro for enterprise-wide use should not be used for any activity.

- Any Generative AI use case must be reviewed and approved by AI councils or Responsible AI taskforce (for high risk AI) before implementation or deployment. Users must submit a GenAI use case review request to *Responsible AI taskforce*.

- Any Generative AI use case involving personal data and/or has an impact to the data privacy rights of individuals must undergo a Data Protection Impact Assessment.

- APIs should not be used to transmit sensitive and/or confidential data to Generative AI systems.

**2. Guidelines for Deployers and Developers**

Deployers and Developers ("developers") of Generative AI refers to personnels engaged in the deployment or development of Generative AI which may include but not limited to:

a. Developing Generative AI Application using APIs (3rd Party),
b. Transfer learning from existing open source GPT models to inhouse GPT Models as the technology become accessible.

Developers of Generative AI must adhere to the following guidelines and to the guidelines prescribed to users.

I. General Guidelines

a. Developers must adhere to privacy and data protection laws and must ensure alignment to globally accepted frameworks such as OECD and NIST (National Institute of Standards and Technology) frameworks.

b. Developers must adhere to the defined rules around key requirements, such transparency, fairness (refer to Addendum 2. Privacy by Design Checklist for AI/ML completeness). A Data Protection Impact Assessment in collaboration with the Global Data Privacy team must be completed during the design phase or initial/ideation stages of the development of Generative AI.

c. Privacy by design approach must be adopted from conception stage to deal with issues such as data collection, legal basis, human oversight, informational rights, automated decision making.

d. Developers must comply with defined mechanisms for regular checks and must complete an audit of the system before deployment, and at regular intervals.

e. To mitigate security and confidentiality risk, deployers and developers must:
   - only use public data to train the AI system and maintain the anonymity of data
   - adopt the use of secure coding, protect access to AI System only to authorized users,
   - align data protection standards on training data and sensitive user input to encrypt and store securely,
   - perform regular security assessment to identify and address vulnerabilities in Generative AI system

f. Deployment and development GenAI tools for client delivery should be done only if approved by the client or if it is allowed as per client contract. Similarly, client enterprise data including personal details should not be used in Generative AI without client approval.

g. Account teams should reach out to *Responsible AI taskforce* for advisory on deployment or development of Generative AI tools in all new and existing client engagement.

h. Developers must comply with Intellectual Property policies and standards and ensure that there is no copyright infringement.

i. Should the system fall under the coverage of the EU AI Act (directly or indirectly marketed in Europe), developers must refer and adhere to the dedicated sections in the Act.

j. Generative AI-generated source code must undergo thorough testing and validation before integration into internal applications.

II. Generative AI-Generated Source Code for Internal Applications:

a. Developers responsible for incorporating Generative AI-generated code must maintain a comprehensive documentation trail to aid in debugging and troubleshooting.

b. Regular code reviews must be conducted to ensure adherence to coding standards, as well as to maintain code quality and security.

III. Generative AI generated code for shipment to clients/customers:

a. Clear differentiation must be made between Generative AI-generated code and human-generated code in customer-facing applications or products.

b. Customers must be informed about the utilization of Generative AI-generated code and any limitations associated with it.

IV. Generative AI-Generated Code by Third-Party Generative AI Services:

a. Thorough due diligence must be conducted before employing third-party Generative AI services, assessing their credibility, security measures, and compliance with applicable regulations.
b. Contracts or agreements with third-party Generative AI service providers must include clauses addressing data privacy, intellectual property rights, and liability.
c. Regular audits or assessments should be performed to ensure that third-party Generative AI services meet the enterprise's established standards.

V. Generative AI-Generated Code by Open-Source GPT Models:
   a. Usage of open-source GPT models must comply with the relevant licensing terms and conditions.
   b. Before using open-source GPT models in production environments, a comprehensive evaluation must be conducted to identify and address security vulnerabilities, bias, and ethical concerns.
   c. Proper attribution and acknowledgment must be provided for the open-source GPT models employed.

## **Other Responsibilities**

- Generative AI tools should not be used to create adverse effect to Wipro, customer, entities data and infrastructure.
- Issues or concerns, such as unauthorised access or data breaches should be reported in accordance with Wipro security incident reporting process.
- Users of Generative AI systems must always adopt a critical mindset and be able to validate the outcomes as such tools may compute inaccurate or false information.
- In case of any doubt reach out to the *Responsible AI taskforce* for assistance.

**Function Responsibilities**

| | |
|---|---|
| Generative AI Task Force | • Define and govern policy for the Responsible Use, deployment, development of Generative AI<br>• Drive awareness on the responsible use, deployment, development of Generative AI<br>• Leverage Wipro's investment to enable various functions and the use cases. |
| CTO / Lab45 | • Provide thought leadership and advisory on Generative AI technology and Wipro's GenAI offerings |
| BiTS | • Deploy tools and solutions and controls to actively detect and monitor any risk to Wipro or client data and IP as a result of usage of Gen AI.<br>• Report breaches to the Gen AI task force along with root cause analysis and mitigation actions taken<br>• Deploy cybersecurity solutions as defined by CISO |
| CISO | • Define cybersecurity controls to protect Wipro infrastructure from probable attack vectors and threats powered by generative AI.<br>• Define security controls to protect GenAI models and associated data against misuse and unauthorized use.<br>• Provide audit and governance over controls by BiTs on Gen AI<br>• Cyber security assessments during vendor onboarding to account for Gen AI risks |

| | |
|---|---|
| | • Provide training and awareness on cybersecurity risks around GenAI |
| Legal and Intellectual Property | • Define controls and processes to identify and mitigate risks to intellectual property as a result of usage, deployment and development of Gen AI.<br>• Provide advisory if any clauses on IPR and liabilities arising from usage or development of Gen AI should be addressed in client, vendor, contractor and partner contracts. |
| ERM | • Define an audit plan to check for usage, deployment and development of Gen AI during account audits. |
| Data Privacy | • Define a framework for assessing GenAI use cases that involves the use of personal information.<br>• DP assessments during vendor onboarding to account for Gen AI risks<br>• Provide training and awareness privacy risks around GenAI |
| Functions and Delivery Teams | • Function and delivery compliance teams are to perform checks and audit where GenAI is leveraged in their units. |

**Definitions:**

1. **Generative AI -** Generative AI refers to an artificial intelligence system that have the capability of generating new content based on the data they were trained on. Generative AI become extremely popular after the release of ChatGPT. However, it is worth highlighting that GenAI has a wide variety of applications beyond text. For example:
   - Text: Generative AI model can generate new text content, e.g., write essays, generate new code, or translate from one language to another. Such models are trained on massive amount of text data from various sources.
   - Image: Generative AI models can generate new images. For example, style transfer applications where you can upload your photo and it generated a Monet style portrait of your photo.
   - Audio: Generative AI models can generate new audio content. For example, generative AI models were used to complete Beethoven's $10^{th}$ Symphony.

2. **Natural Language Processing (NLP) -** is an arch of computer science techniques that enables computer to understand text and provide inference in the same way that human would.

3. **Large Language Model (LLM)** – LLMs represent a core component of NLP, a tool to enable AI to mimic human performance in understanding language. LLMs are large models (millions of parameters) that are trained on massive amount of data.

4. **Generative Pretrained Transformer (GPT)** – GPT refers to a subset of LLM models that uses an underlying Neural Network architecture called Transformers and is trained on a large body of data to perform wide variety of tasks such as text summarization, question answering, etc. GPT is the underlying model for ChatGPT.

5. **Transfer learning -** transfer learning refers to the process of transferring the knowledge of general models trained on large, general data sets to a more specialized models that aim to solve a specific problem of interest.

6. **Fine Tuning** - Fine tuning is the process/technique used for transfer learning and it refers to fine tuning the general "Pretrained" models by resuming their training on specialized data sets.

**Approvals/Escalation Matrix**

Any deviation or non-compliance to this policy must be immediately reported to *Responsible AI taskforce.*

Furthermore, should an employee be made aware of any suspicious activity involving the use of Generative AI, such as but not limited to unauthorised access, misuse, and unauthorized disclosure of data, he/she must report it immediately report it as a security incident.

**Revision History**

| Version | Revision Date | Reason for Change | Drafted/ Reviewed By | Approved By | Date Approved |
|---|---|---|---|---|---|
| 1 | N/A | Policy creation | Policy drafted by: GenAI Taskforce Committee members | General Counsel Chief Risk Officer | June 12, 2023 |
| 1.2 | July 22, 2024 | Changes in the approval/escalation matrix and contact information for queries and review requests. | Responsible AI taskforce | Chief Privacy and AI Governance Officer | July 22, 2024 |