View on Web

Walking The Generative AI Tightrope - Innovation Amid Data Invasion

1st Oct,2022

This might shock you beyond measure. But there is every possibility of it happening. ChatGPT can divulge prompts and the history of your conversations to other users. Such an occurrence is unanticipated. Unintentional too. A Generative AI model like ChatGPT can inadvertently ingest sensitive and confidential data or Personally Identifiable Information (PII) and regenerate it later. According to a Statista survey from 2022, 42% of those surveyed were concerned or very concerned about their online data. Further, a Statista marketing report from 2023 states that 73% of respondents use Generative AI tools as part of their work.

The rapid adoption of Generative AI applications, coupled with consumer privacy concerns, warrants businesses to be aware of how it affects data privacy and what they can do about it. If not developed responsibly, generative models may replicate copyrighted data or media, reveal confidential conversations, or propagate biases. But before we talk about the guardrails, it's worth exploring the workings of Generative AI, the role of user data, and understanding the privacy vulnerabilities.

Role of User Data in Generative AI

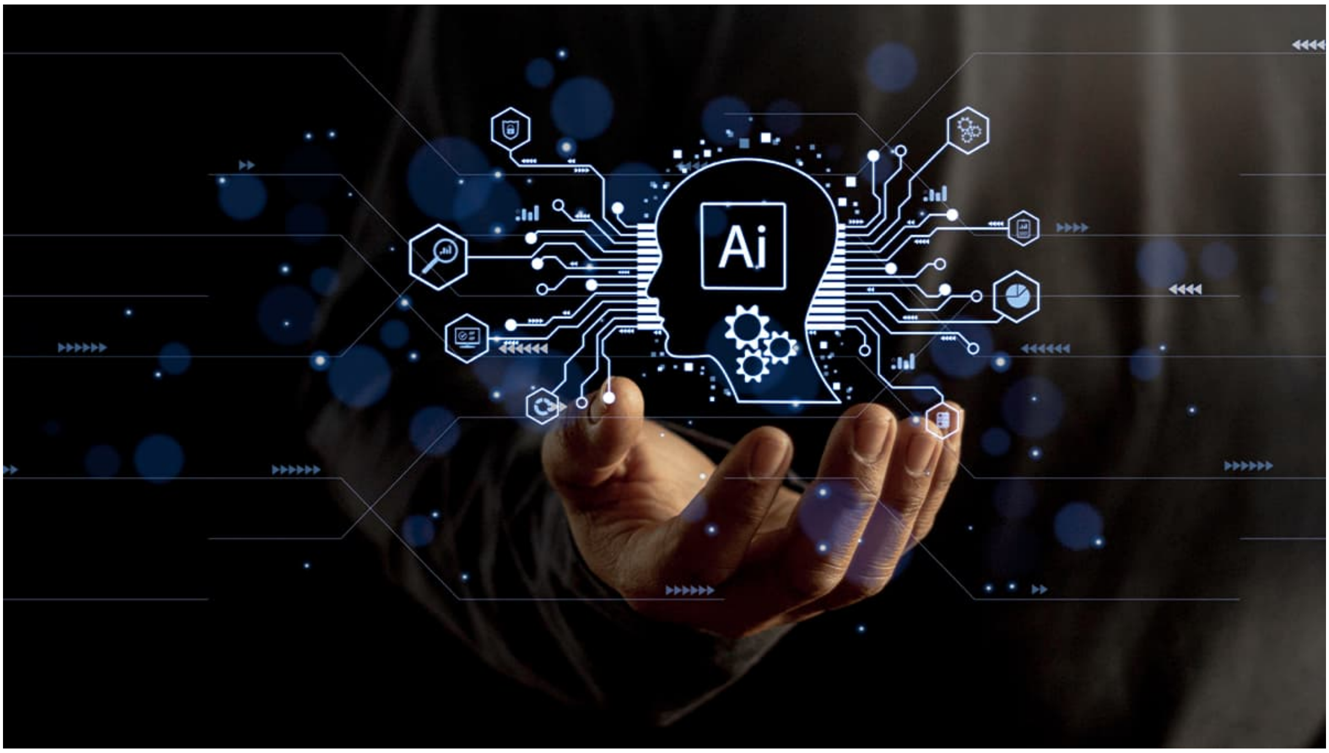
Generative AI is the Picasso of **Artificial Intelligence (AI)**, but instead of painting mind-bending abstract art, it creates new and original things using patterns it has learned. It is a master of imitation infused with innovation. It learns from massive data to generate new and exciting patterns, styles, and structures. An algorithm for Generative AI can benefit significantly from user data collected from various sources such as social media, online browsing habits, and personal preferences. Besides understanding and replicating human behavior, preferences, and creative output, it can also help these algorithms perform better. However, this symbiotic relationship between Generative AI and user data raises important concerns regarding data security and privacy.

Getting The Hang of Privacy Concerns

With Generative AI, privacy concerns are as trendy as avocado toast. The collection, storage, and misuse of personal data worry users. Who wants their secrets floating around in cyberspace? The use of Generative AI raises questions about consent, transparency, and how our data is used without our consent. Generative AI is becoming more sophisticated, and so are the methods of data breaches and privacy vulnerabilities. Malicious actors and hackers are always looking for ways to compromise data security. The more user data is collected and stored, the greater the risk of unauthorized access.

Balancing Innovation and Privacy Protection

Finding the sweet spot between innovation and privacy protection is like walking a tightrope with a feather in hand. Striking the right balance requires collaboration between AI developers, regulators, and users. By adopting privacy-by-design principles, implementing user-centric consent frameworks, and providing transparent information about data usage, we can navigate the Generative AI landscape without sacrificing our privacy.



How can Companies Put Strong Guardrails to Check Data Theft?

One major guardrail that any company can implement is to refuse to share the data collected by their AI systems with any third party. They can also put in place the safeguards below to combat data breaches:

- Perform robust testing to identify harmful biases, privacy risks, and policy violations. Monitor models continuously after deployment - this is an iterative process.
- Control unsafe content generation using block lists, allow lists, and suppression lists.
- Review high-risk outputs before releasing them using "human-in-the-loop" approaches.
- Provide accessible user reporting channels for flagging policy-violating or abusive AI behaviors.
- Retrain models frequently based on the latest norms and values.
- Disengage problematic model versions and roll back to a previous safe checkpoint. When ChatGPT users reported problems, OpenAI removed web browsing functionality.

How CSM Tech is Ensuring Data Privacy in Its Generative AI Services

Data privacy and security are key concerns in Generative AI. **To solve specific business problems, we analyze data thoroughly, extracting only essential information. During data extraction and storage, we employ encryption techniques to ensure that data cannot be read by unauthorized parties. Strict access control based on roles is enforced for repository data. For both training and deploying models, we use on-premise infrastructure exclusively, reducing exposure to external data.**

To protect individual privacy, personal data is anonymized before model training. After the model is trained, a well-defined data retention and deletion policy is followed to minimize data exposure. Our regular audits and assessments ensure continuous data security protection by proactively identifying and mitigating potential threats.

The Future - Juggling Disruption with a Laser Focus on User Privacy

We can make Generative AI advancements ethically sound and respect user privacy by proactively addressing privacy concerns. As Generative AI advances and reshapes various industries, it's essential to address privacy concerns. Technology companies, policymakers, and society as a whole must work together to strike a balance between Generative AI and user data protection. Implementing robust data protection measures, adopting ethical frameworks, and updating regulations to deal with Generative AI's unique challenges will protect user privacy.

AUTHOR:

Jayajit Dash

Senior Manager- Corporate Communications (Marketing)