

# AI Model Testing for the consumer industry

## Outsmarting risk with intelligent tactics

AI offers incredible opportunities to the consumer industry—and leaves you open to unique vulnerabilities. As your organization explores its power and reach, AI governance has never been more crucial.

What are the biggest risks? What can you do to help defeat them and validate your AI strategies? We recently surveyed leaders in consumer industry, asking about their testing approaches and uncovered applications that can guide you in mitigating potential challenges in your AI initiatives.



### What are key risks in AI models?

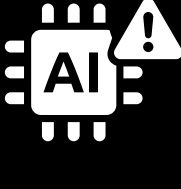
Its promise of efficiency, time-saving, and productivity add unquestionable value. But due to its great power, AI carries with it vulnerabilities.



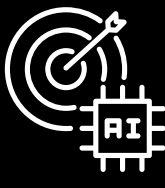
**Hallucination:** These incorrect or made-up outputs can lead to misinformation in customer interactions or product recommendations, impacting brand trust and customer satisfaction.



**Bias and Toxicity:** Harmful stereotypes and inappropriate content can infect AI models. They must be carefully managed to avoid reinforcing biases or producing content that could damage brand reputation or alienate your customers.



**Prompt Injection:** Malicious instructions in prompts in applications can lead to security breaches that compromise customer data and lead to unauthorized actions.



**Accuracy and Precision:** How reliable are your AI model outputs? Ensuring the veracity of outputs is vital for consumer-facing applications, where errors can directly affect customer experience and business outcomes.

## Pressure test your testing

What are some smart approaches organizations are using to validate their GenAI models?



### SME review

Spot-checking by subject matter experts (SME) is crucial. Despite scalability challenges, SMEs provide critical oversight to ensure AI outputs align with brand values and customer expectations.



### Statistical measurements

Quantitative assessment using accuracy and precision metrics are crucial, offering a data-driven approach to evaluating AI models—essential for maintaining high standards in consumer applications.



### LLM as a judge

Employing a secondary LLM to evaluate the output of your AI model enhances the scalability and robustness of validation processes in consumer sectors.

## Straight from the Industry: what leaders shared in our latest survey

Here's how leading organizations are using and validating GenAI applications.

### How are you using common AI applications?

**By text summarization and chatbots:** Most popular in consumer sectors, these tools are widely used to enhance customer service and streamline communication, crucial for maintaining competitive advantage.

**Code generation:** A close second behind chatbots, code generation aids in developing innovative consumer applications, speeding up the deployment of new features and services.

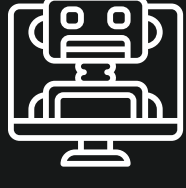
**Forecasting:** Finally, respondents shared that they are using machine learning to provide insights around inventory, financials, and marketing.

### How are you testing and monitoring?

Organizations are using anywhere from a handful (1-5) to well over twenty applications in consumer industries—how many are you using? The diverse application of GenAI technologies in consumer sectors underscores the need for comprehensive testing and validation to ensure reliability and effectiveness.



## Showing up in the real world—for real: more GenAI use cases



**Coding assistant:** This tool supports developers in creating consumer-facing applications by automating code generation and providing detailed explanations, enhancing efficiency and reducing errors.



**Marketing assistant:** This GenAI use produces personalized marketing content for consumer brands. An assistant helps consumer brands engage more effectively with their audiences, leveraging AI to enhance personalization and impact.

## Outsmart AI risks with intelligent tactics

AI and GenAI's promise of efficiency, time-saving, and productivity can be protected with key guardrails.



### Ongoing monitoring to help consumer models remain stable and perform as intended.

Continuous monitoring is crucial to maintain model reliability and adapt to changing consumer needs and preferences.



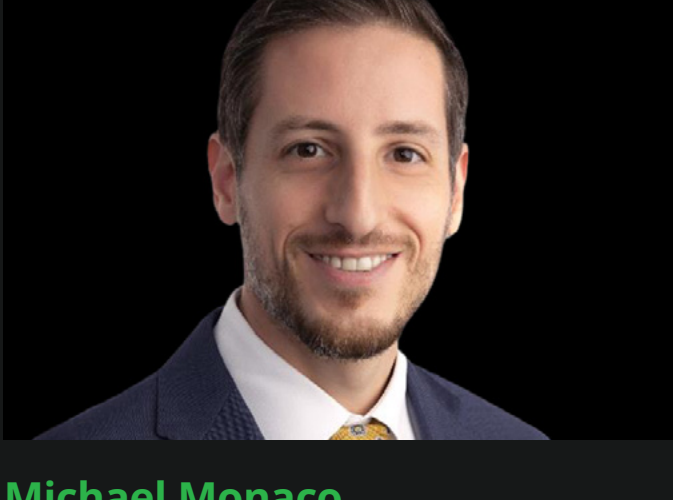
### Collaboration between model risk management and cybersecurity teams to get out in front of threats.

Effective management of GenAI models requires risk management and cybersecurity teams to work together to safeguard against potential threats and ensure comprehensive oversight.

## Make sure you trust the real thing—actual humans are standing by to take your questions today



**Clifford Goss**  
Partner  
Deloitte & Touche LLP  
cgoss@deloitte.com



**Michael Monaco**  
Senior Manager  
Deloitte & Touche LLP  
mimonaco@deloitte.com



**Kirat Dhillon**  
Senior Manager  
Deloitte & Touche LLP  
kidhillon@deloitte.com