NICE Actimize

Insights Article

Defining Thoughtful and Responsible Generative AI

Generative AI solutions have enormous potential to transform financial crime and compliance programs. However, this opportunity may pose complex implications with catastrophic effects, ranging from reputational damage to regulatory fines to financial loss.

Innovation and advancements in generative AI are occurring at breakneck speed, while regulations move at more of snail's pace. In the absence of industry-wide regulations, it will be up to solution providers to build a responsible, effective and ethical generative AI framework.

Our analysis and investigation of generative AI has helped us to define our methodology, as we take a thoughtful and responsible approach in leveraging this expansive and emerging technology. We have established three foundational pillars to ensure we deliver the best-in-class generative AI solutions:

1. **Privacy and Data Security**
2. **Explainability**
3. **Large Language Model (LLM) Governance**

Privacy and Data Security

The information entrusted to financial institutions (FIs) makes handling and safeguarding data from unauthorized access a critical function. When considering a generative AI solution, firms must ensure a comprehensive privacy and security strategy to effectively protect sensitive information and confidentiality.

Large language models, or LLMs, require extensive data to identify, label and train. By exposing sensitive and confidential data (including PII) to LLM models can have significant repercussions, including, but not limited to, unintentionally recreating or sharing personal client details. And any use of third-party extensions will only exacerbate privacy risks.

Data security when using generative AI is integral to safeguarding confidential information and maintaining trust in the deployment of the solutions. Employing technical and procedural measurements will be needed to prevent unauthorized access, data breaches or other malicious activities.

When building solutions using emerging technology it's imperative to partner with a solution provider that has extensive experience and expertise in data privacy.

Explainability

Explainability, or the ability to understand and interpret the data used to train a generative AI solution, is crucial for building trust. Model architectures are generally opaque, with responses based on the technology's best predictions given it's training, the prompt received, and the massive amounts of data available to the LLM. But what was the basis behind the output, and how did the model reach its conclusion? A lack of explainability is a significant limitation. It can be a major barrier to generative AI adoption.

When looking at model governance, explainability is needed for risk management to ensure accountability and compliance with regulations. For end users, including investigators, regulators and both internal and external auditors, explainability offers the ability to understand the path taken by the solution to generate its response.

To adhere to regulatory guidelines and ensure users understand the model outputs, explanations are required. Achieve Generative AI explainability by:

- Incorporating transparency into the model's training and parameters
- Including test outcomes for every change made
- Prioritizing explainability and interoperability in the modeling process

Large Language Model Governance

In the evolving landscape of financial crime and compliance, LLMs present a unique governance challenge. Ensuring the precision and quality of generative AI outputs, responding adaptively to training data as emerging regulatory guidelines surface, and meticulously monitoring for inherent biases and hallucinations in model behavior are critical to upholding the integrity and efficacy of the solution. Measure, Adapt and Monitor is the mantra for ensuring effective LLM governance.

- **Measuring the quality of outputs**
  Only through a thorough analysis of the generative AI outputs can we effectively understand how well the solution is performing. Constant analysis of AI-generated outputs is vital in the regulated financial crime and compliance sector. Firms must measure and continuously reassess the quality and precision of these outputs. Employing pre-and post-implementation metrics is essential to gauge the LLM model's effectiveness and make necessary adjustments.
- **Adapting to emerging regulatory guidance**
  While FIs are already facing stringent regulatory guidelines around model validation and risk assessment, policymakers have not yet issued regulations on the use of generative AI, leading to ambiguity. In the absence of specific generative AI regulations, FIs face a dilemma: implement potentially non-compliant solutions or risk inaction. To address this risk firms should partner with a trusted solution provider who has deep insights into current regulations, expertise within highly regulated global environments, and knowledge of how to be flexible. This allows for the probability of enhanced regulatory reporting, which is crucial.
- **Monitoring negative outputs**
  Addressing inherent biases in training data and mitigating the risk of model hallucinations are intertwined challenges that require a comprehensive approach to data management and model oversight.

Bias in AI, stemming from unrepresentative, homogenous or skewed training data, can lead to distorted outputs, harming an institution's reputation and financial standing. To mitigate this, firms and solution providers should adopt responsible data management practices. This includes ensuring proper data sampling sizes, confirming data relevance, and incorporating human oversight to minimize bias risks.

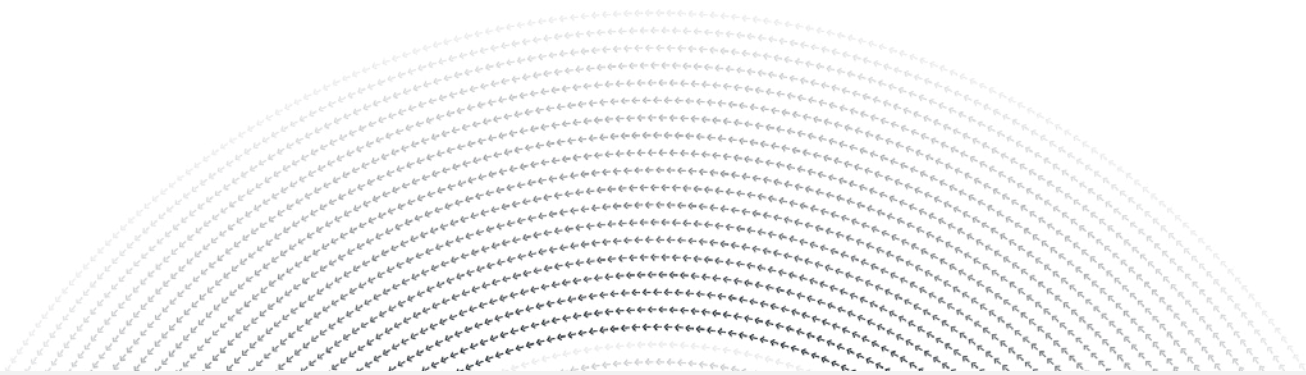
Hallucinations are illogical and unintentional outputs from poorly trained models that could elicit ethical concerns, diminish trust and create legal implications. By employing a combination of diverse training data, feedback loops, human oversight, and robust explainability, we can reduce the incidents of hallucinations.

Defining Trusted and Thoughtful Generative AI

At NICE Actimize, we're taking a holistic approach to generative AI that enhances decision-making, employs ethical AI practices, and ensure full transparency. When it comes to financial crime investigations, we believe that the most effective and responsible solutions blend generative AI complemented by human oversight—the AI provides enhanced intelligence and insights, while the human expertise directs critical decision-making and outcome assessments.

Explore our latest Generative AI solutions

Schedule a demo



About NICE Actimize

NICE Actimize is the largest and broadest provider of financial crime, risk and compliance solutions for regional and global financial institutions, as well as government regulators. Consistently ranked as number one in the space, NICE Actimize experts apply innovative technology to protect institutions and safeguard consumers and investors assets by identifying financial crime, preventing fraud and providing regulatory compliance.

The company provides real-time, cross-channel fraud prevention, anti-money laundering detection, and trading surveillance solutions that address such concerns as payment fraud, cybercrime, sanctions monitoring, market abuse, customer due diligence and insider trading.

www.niceactimize.com