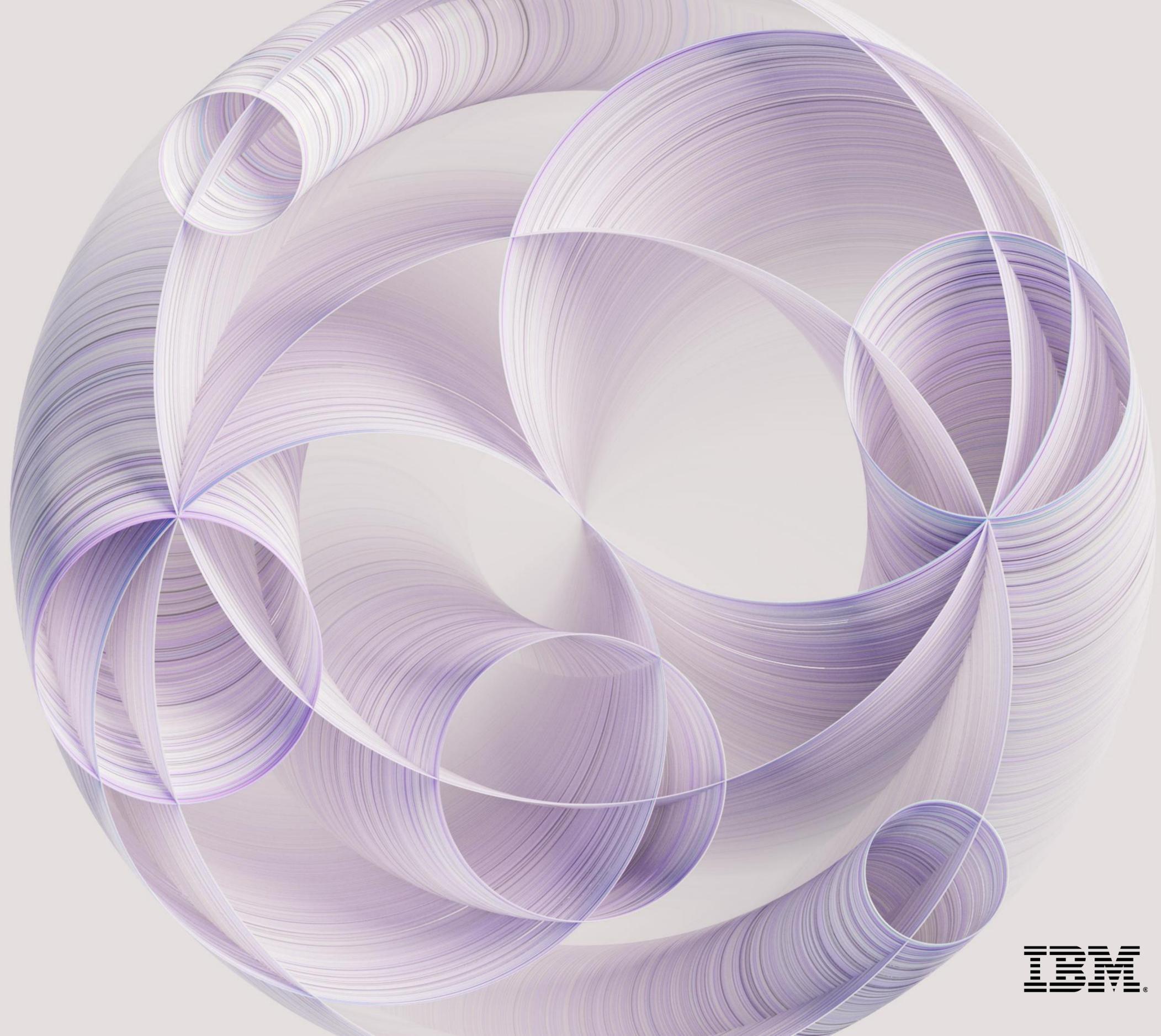


Accelerate
responsible,
transparent and
explainable AI



watsonx.governance enhances these use cases by (1/2):

AI Lifecycle Mgt.

→ **Monitor** hate, abuse, profanity, PII for both inputs and generated output

→ **Inventory** automate documentation and develop an inventory of AI assets.

→ **Evaluate** evaluation of AI with evaluation metrics via API

RAG-based bots

→ **Assess quality** of generated answers through quantitative measures of grounded but also attribution to the source text

→ **LLM-as-a-judge** can be used to generate RAG metrics with Python SDK to measure model performance of RAG tasks.

→ **Root cause evaluation** can now analyze prompt template evaluation results for RAG tasks with root cause analysis.

HR Automation

→ **Detect and mitigate bias** for ML models used in hiring decisions and maintain

→ **Manage risk** by easily identifying potential risks through assessments – including regulatory applicability assessments

→ **Explainability** enables scenario analysis to better assess and understand model outputs

watsonx.governance enhances these use cases by (2/2):

Audit

- **Streamline audit process** to meet regulatory requirements of model documentation
- **Automate capture of meta data** over time to facilitate management and compliance
- **Inventory and track** ML models, Generative AI apps, agents and models from concept through the lifecycle

Customer Care

- **Confirm agents deliver quality results** both immediately and long term after placed in production without overhead
- **Monitor chats for red flags** of toxicity, personal info, or off-topic conversations
- **Flag when responses are outside** of social-norms and summaries are factually incorrect further reducing human intervention

BPO

- **Monitor customer-specific process documents** for Personal Identifiable information with customizable alerts
- **Monitor NLU text models** for drift, relevancy, and more for hands-off monitoring
- **Support BPO activities** across finance, HR, Procurement and more groups with dashboards and customizable workflows

