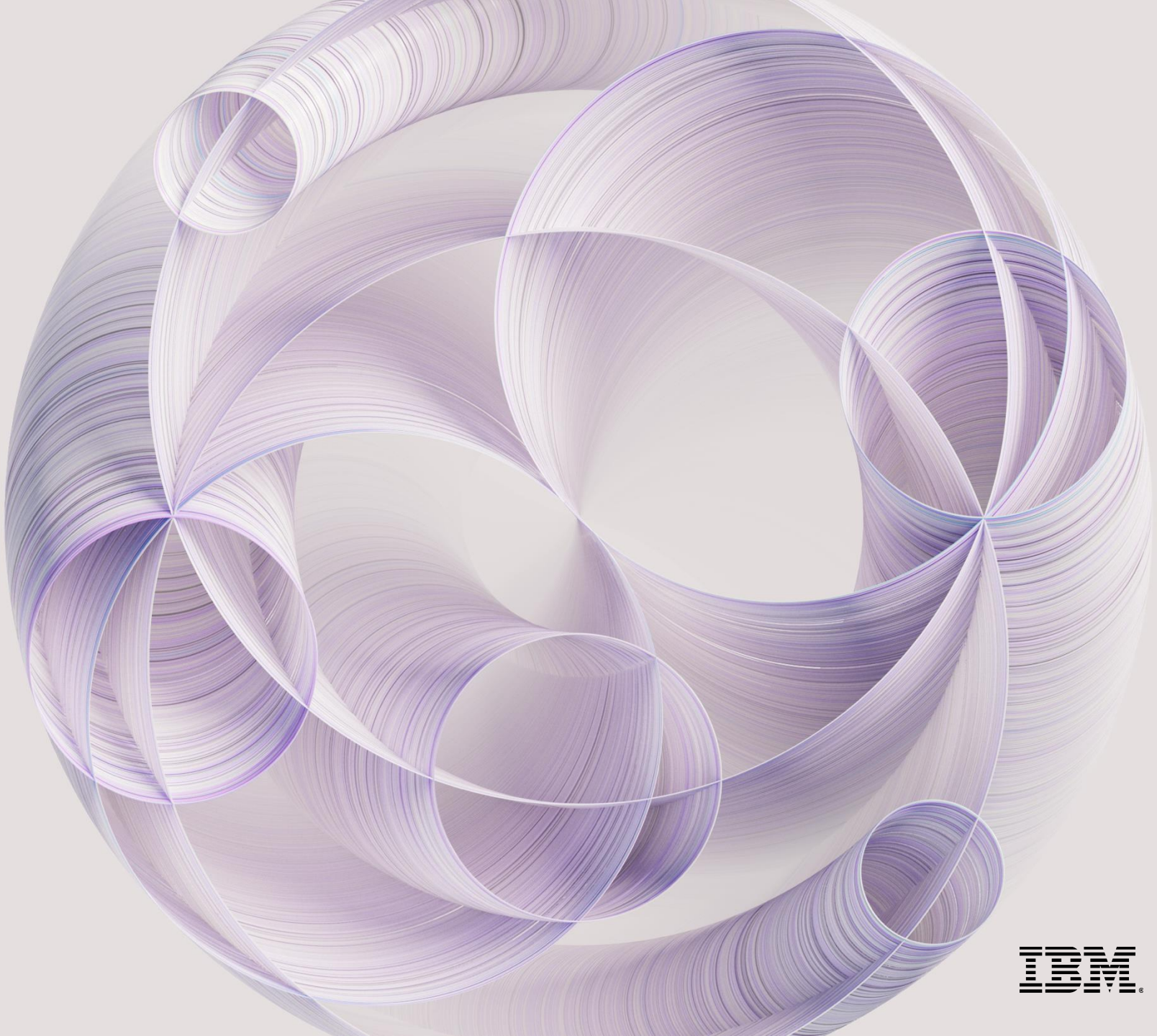
IBM Watsonx AI Productivity with governance



Margo Keeler – watsonx
Leader for AI & Governance-
Public Market

The IBM AI Strategy Centers Around Core Tenants Required to Scale Trusted AI

Multi-Model

- Two thirds of 150+ enterprises surveyed report pursuing a multi-model strategy
- 60% + of enterprises pursuing multi-model are experimental with commercial & open-source models
- Multi-modal (text, image, audio, etc.)
- One model will not rule them all

Multi-Hybrid Cloud

- Run where the workflows, apps and data live
- Infer where business runs to drive performance, cost, and simplicity
- Data location to drive security benefits
- Regulatory compliance to influence location selection

AI Governance & Security

- Businesses must control bias and monitor drift
- Organizations must actively monitor hallucinations and ensure explainability
- Leaders must seek practices and tools to ensure model and data provenance
- Regulatory Compliance requirements continue to grow and evolve

Scale for Value

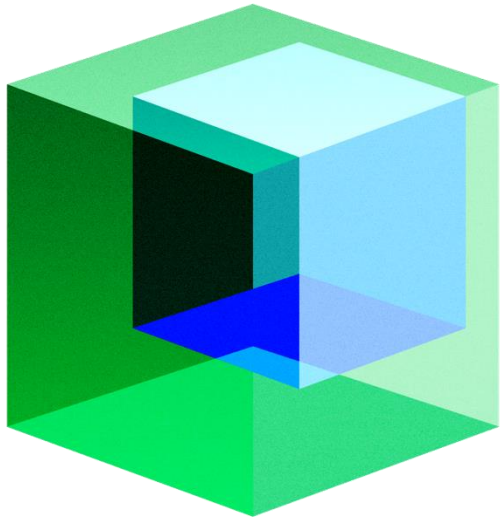
- Critical to pick the right use cases and deployment for generative AI ROI
- Different work tasks have strongly positive or negative ROI impact
- Synergy among model performance, cost and trustworthiness
- 25x difference in cost per inference, depending on model and deployment

Data Matters

- Generative AI pilots have not made it to production due to challenges with data quality, access, and security
- Short run: model innovation creates value
- Long run: data quality will decide which enterprises win with generative AI

IBM Granite

A family of open, performant and trusted AI models to accelerate enterprise AI adoption



Open

- Open sourced under Apache 2.0
- Transparency of data, training methods
- Customize with your data

Performant

- Diverse range of fit-for-purpose models
- Designed for scalability

Trusted

- IP indemnification
- Responsible and safe AI
- Guardrails to secure data and mitigate risks

Granite family of models

Large Language Models (LLMs) for enterprise

Granite-3.0-8B-Instruct
Granite-3.0-2B-Instruct

Inference-efficient Mixture of Experts (MoE)

Granite-3.0-3B-A800M
Granite-3.0-1B-A400M

Guardrail models

Granite-Guardian-3.0-8B
Granite-Guardian-3.0-2B

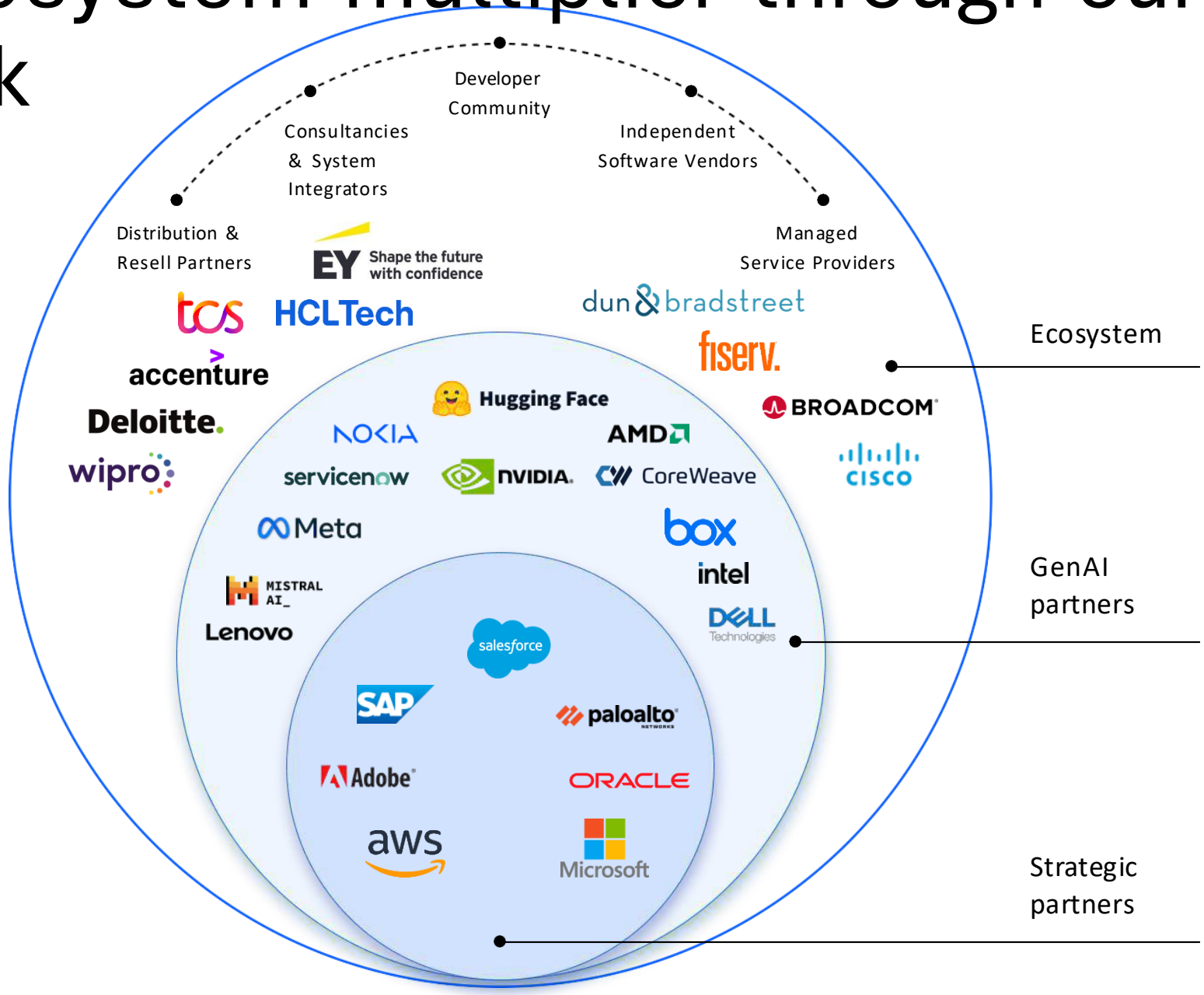
Speculative decoding models

Granite-3.0-8B-Instruct-Accelerator

Time-series models

Geospatial models

Created an ecosystem multiplier through our partner network



AI capabilities are growing rapidly

- AI that predicts

- *Machine learning*

- AI that creates

- *Generative AI*

- AI that chats

- *AI assistants*

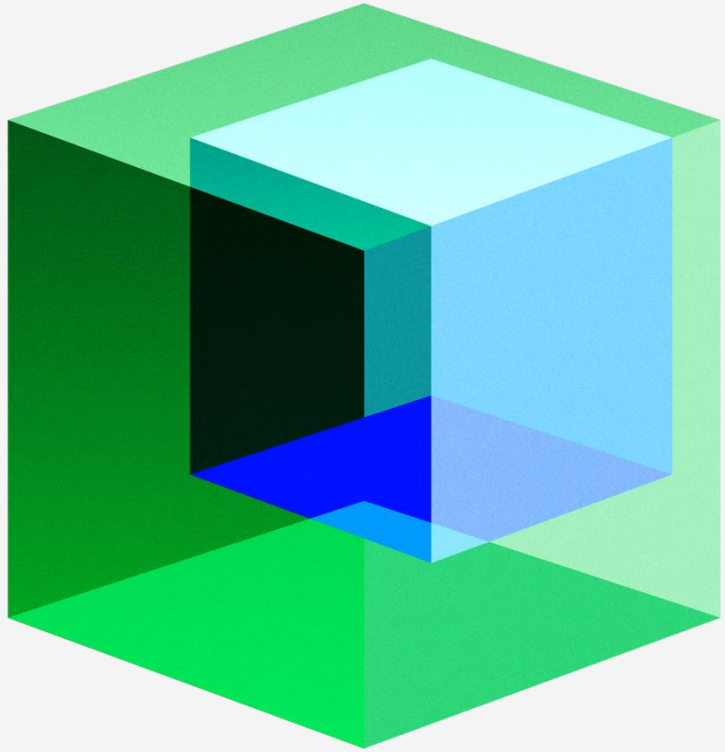
- AI that *does work*

- *AI agents*

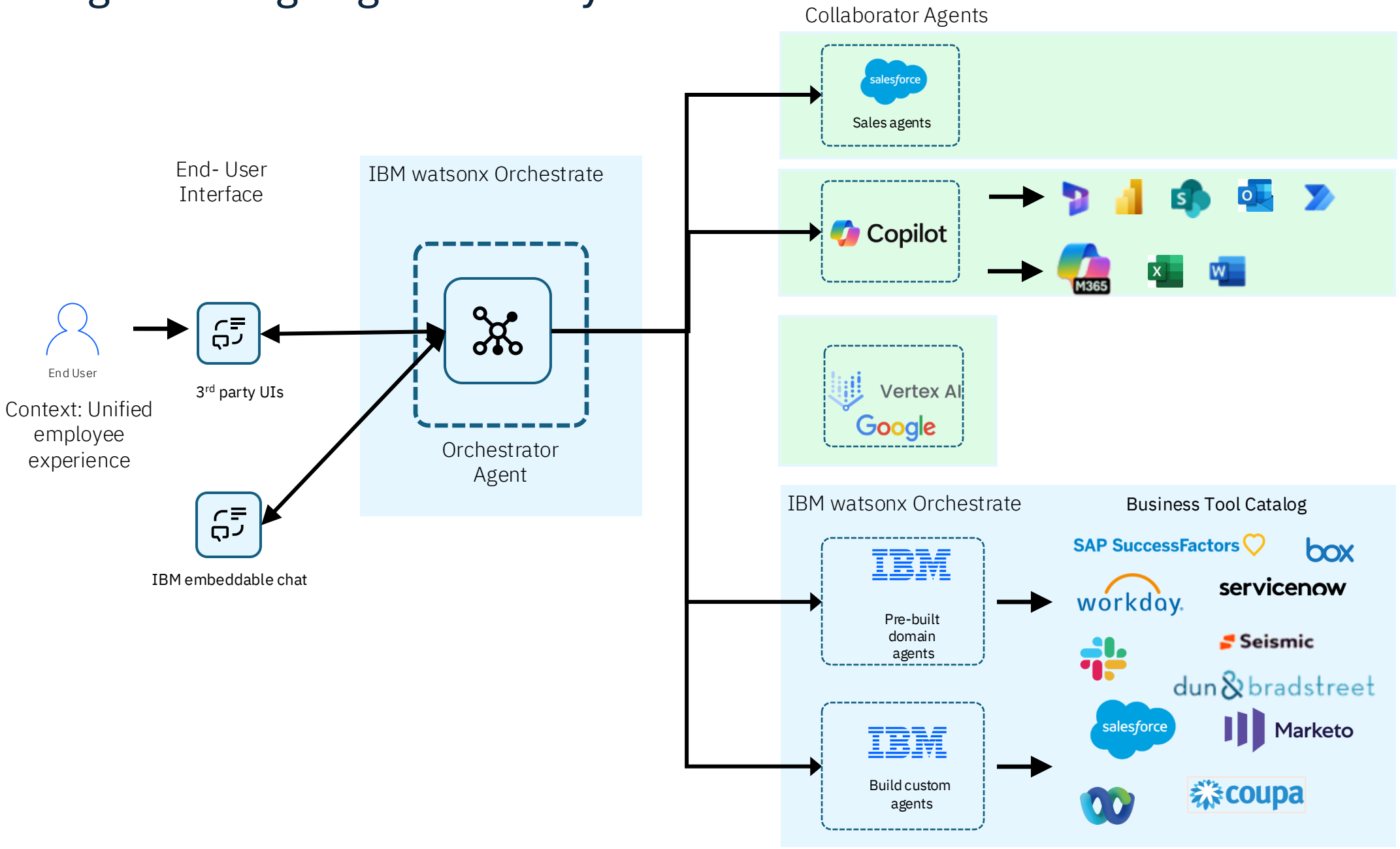
What is an AI agent?

- An AI agent is an application
- that acts autonomously to
- understand, plan, and
- execute a request (from a
- human or another agent).

- AI agents use LLMs to reason
- and can interface with tools,
- other models, and other IT
- systems to fulfill user goals.

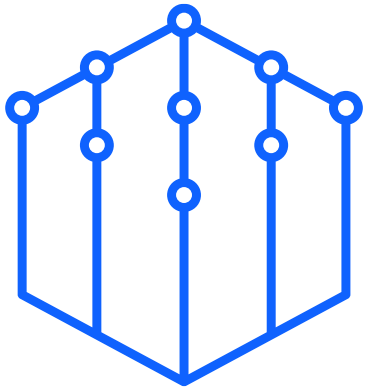


AI agents are going to be everywhere

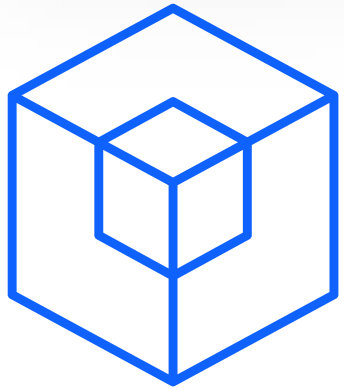


The IBM approach: fit-for-purpose models

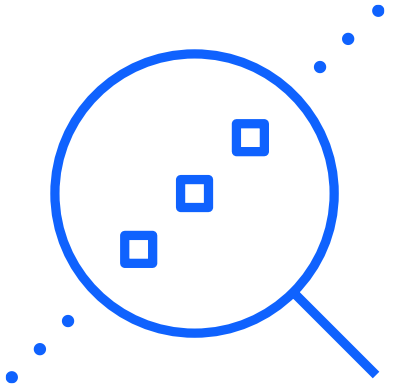
Up to 42x lower inferencing costs



Your data



The right model



Targeted use case fine tuning

Improve answer quality and accuracy with document processing

1

Answer questions from tables

Processing PDFs, Office documents, and HTML (e.g., web crawl) to identify and extract tables in an indexable and LLM-readable format

2

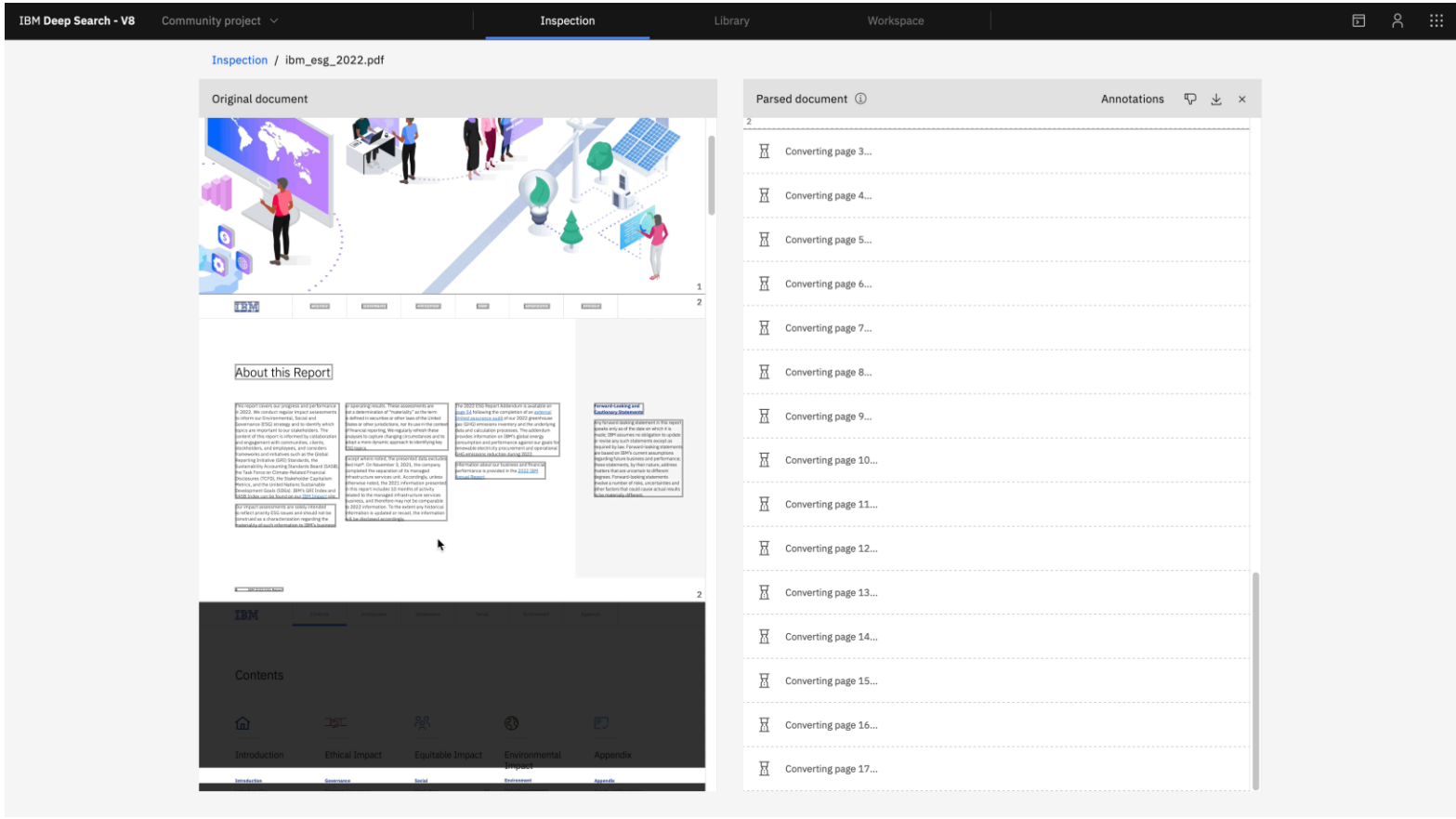
Optical character recognition

Identifying and processing images, such as diagrams and visualizations, to extract useful text

3

Structured text extraction

Processing documents to identify extraneous information; extract titles for smarter chunking; extract nested, ordered, and unordered lists

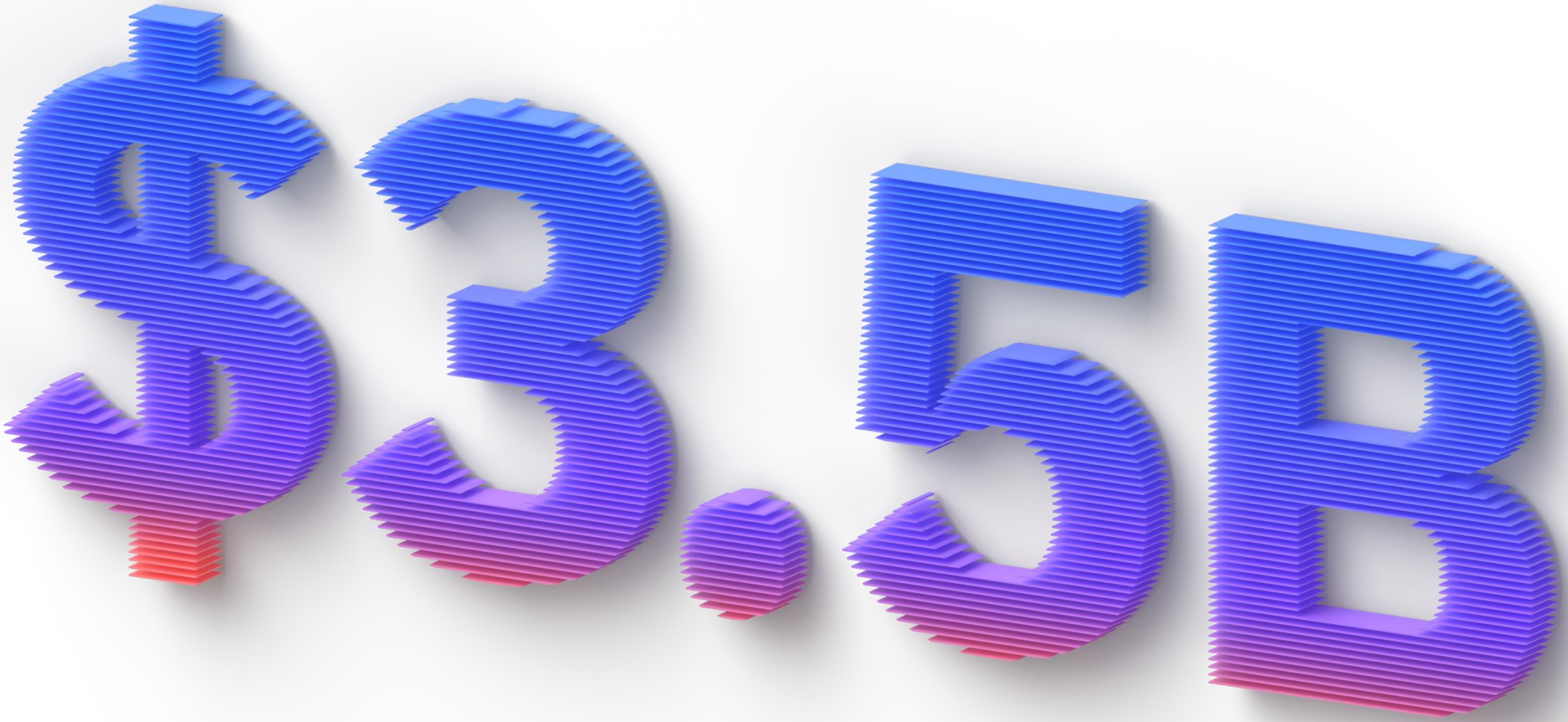


IBM and the EY Organization Debut Artificial Intelligence–Powered Global Tax Compliance Solutions

- **Detect and Correct with Business Documents:** the solution extracts data from unstructured invoices to compare and correct ERP data for tax determinations and filing. The solution enables IBM to dramatically increase the number of source documents reviewed, automating previous manual processes.
- **Withholding Tax Determinations:** solution automates the monthly process of determining the correct withholding tax rate that should be applied to individual transactions. IBM tax professionals now leverage the solution to more quickly and accurately evaluate thousands of intercompany transactions.
- **Intelligent Tax Data Lake:** Leveraging IBM watsonx.data, watsonx.ai and open-source models, the solution gathers, enriches and consolidates the required transaction data from numerous sources for tax filings. For IBM's own tax department, the solution's built-in data controls and reconciliation checkpoints streamline data consolidation from 36 sources and help produce higher quality data, automating a manual process.

Help organizations automate tax compliance and streamline global data management.

Using AI to make a real impact in productivity



in productivity gains

IBM with watsonx

AskIBM
A central unified interface for all IBMers, connecting to each domain assistant

IBM AskHR with **watsonx**

10M
Annual HR interactions fully resolved by AI

40%
Reduction of HR operating budget

+55
Improvement of HR NPS score

IBM AskIT with **watsonx**

100
Days to build + deploy AskIT from scratch

80%
Inquiries resolved via AskIT

50%
Reduction in support tickets after 12 month deployment

IBM AskSales with **watsonx**

180K
Hours per week saved in gathering account information and insights

5K
Seller questions answers per week (product guidance and persona targeting)

40%
Improvement in qualify of outreach content

IBM AskIncentives with **watsonx**

96%
Sales related inquiries contained within AI Assistant

90%
Greater accuracy in expense accruals

76%
Increase in productivity while serving 22k sellers

IBM Procurement with **watsonx**

85%
Orders now processed via ‘touchless procurement’

50%
Reduction in time spent on manual, repetitive tasks

15%
Enterprise workforce comprised of contractors

...

The AI, Automation, and GenAI possibilities run across the HR and Talent domain

Component Business Model for Talent (Not Exhaustive, Illustrative)

Talent Planning and Governance

- Talent Strategy & Planning
- Talent Vendor Management
- *Employment Regulatory Compliance
- Employee Relations Management
- External Relations Management
- M&A & Divestiture Support
- *Talent Policies & Procedures
- Employment Vendor Management
- Health, Safety & Environment Management
- Talent Function & Performance Management
- Talent Strategic Initiatives Management
- Talent Initiatives Execution & Monitoring

Talent Recruitment

- *Supply/ Demand Calibration
- *Job Posting & Sourcing
- Candidate Screening
- Recruitment Process Management
- Interview Management
- Offering & Onboard
- Workforce Demand Forecasting
- Seasonal Workforce Planning
- Workforce Deployment & Optimization
- Expatriation & Repatriation Management
- Talent Commitment & Retention Management
- Transfer & Redeployment Management

Talent and Skills Development

- Learning Infrastructure Planning
- Executive & Manager Development Strategy
- Learning Capability Strategy & Development
- Learning & Assessment & Certification
- Manage Careers & Experiences
- Deliver Learning & Skills Growth
- Skills & Learning Impact
- Training Resource Management
- * Curriculum & Content Design & Sourcing
- Rollout & Enrollment Management
- Learning Delivery Management
- Training Data Management & Reports
- Individual Development & Career Planning
- Training Data Management & Reports

Performance, Compensation, & Benefits

- Reward, Recognition & Motivation Strategy
- Employee Reward Strategy
- Incentives & Variable Pay Management
- *Performance Review
- Workforce Performance Management
- Compensation & Benefits Management
- Performance Management System Design
- Performance Administration & Follow-Up
- Incentive & Variable Pay Administration
- Pension Processing
- Employee Counseling
- Payroll Accounting, Tax & Legal Reporting

Employee and HR Services

- HR IT Strategy
- Employee Reporting
- *Benefits Information & Processing
- Payroll Processing
- *Transfer, Separation, & Retirement
- *Employee Support
- *Employee Data Administration
- Expenses & Reimbursements Processing
- Employee Surveys and Benchmarking
- HR IT Infrastructure & Application Management
- Workforce Analytics
- Retirement & Turnover Management

Priority AI-First Talent Use Case Sub-domains

Employee and HR Services

With AI and Automation imbedded into HR service delivery and operations management, everyone has more time to focus on what's important – customers, teams, family, and community!

Talent Recruitment

Talent Acquisition can redefine the employee experience starting with their first interaction. For Hiring Managers Job Request and Offer Creation processes are radically transformed.

Talent and Skills Development

Personalized skills identification, rapid learning content creation, and a digital coach for employee performance can quickly accelerate onboarding and upskilling to meet the new skill demands of the business.

High Potential Areas (>30% Impact)

Med Potential Areas (10% - 30% Impact)

Low Potential Areas (<10% Impact)

* High Generative AI Impact

Hallucination Management

High-Risk Topics

➔ **Static question & answer workflows**

High-risk for legal or ethical reasons

Question and answer are curated

Partnering with senior leaders
and content owners across HR to identify

All Other Information

➔ **Domain-specific gen-AI workflow**

Content curation: Content (IBM HR knowledge to be ingested by the LLM) is cleaned up and curated prior to ingestion

LLM Confident Mapping: Answer discarded if below a threshold level of confidence:

- Responds with no answer found message
- Offers to escalate to a human agent

Trust is built in drops but lost in buckets

IBM's approach: Trustworthy AI at scale

A holistic and staged approach to establish scalable and sustainable organizational AI Governance and AI Model Lifecyle Governance.

Two critical components:

ORGANIZATIONAL AI GOVERNANCE

STRATEGY

Who?
Business, AI Ethics Board, Data/AI leaders, Internal Policy & Regulations, CPO, CISO

PLANNING

Who?
Business, AI Ethics Board, Data/AI leaders, Internal Policy & Regulations, CPO, Ecosystem

- Trustworthy AI Principles
- AI Governance Policies, Processes & Metrics
- Operating Model
- Governance Structures
- Regulatory and Risk Assessment

+

AUTOMATED AI MODEL LIFECYCLE GOVERNANCE

DEVELOPMENT & DEPLOYMENT

Who?
Dev teams, IT leaders, CDAO, Software & Data science leaders

OPERATIONS

Who?
IT leaders, MLOps teams

MONITORING & PORTFOLIO MANAGEMENT

Business Outcomes & Model Governance
Business leaders, MLOps teams

- Centralized Platform for Model Lifecycle Monitoring
- Trustworthy AI Model Lifecycle
- Defined Model OKRs and KPIs
- Model Onboarding, and Sustaining, Processes

IBM watsonx.governance

Accelerate responsible, transparent and explainable AI workflows

Centralized
AI lifecycle governance

Manage,
monitor and govern any AI: model, app or agent; across IBM and 3rd party like OpenAI, AWS, and Meta

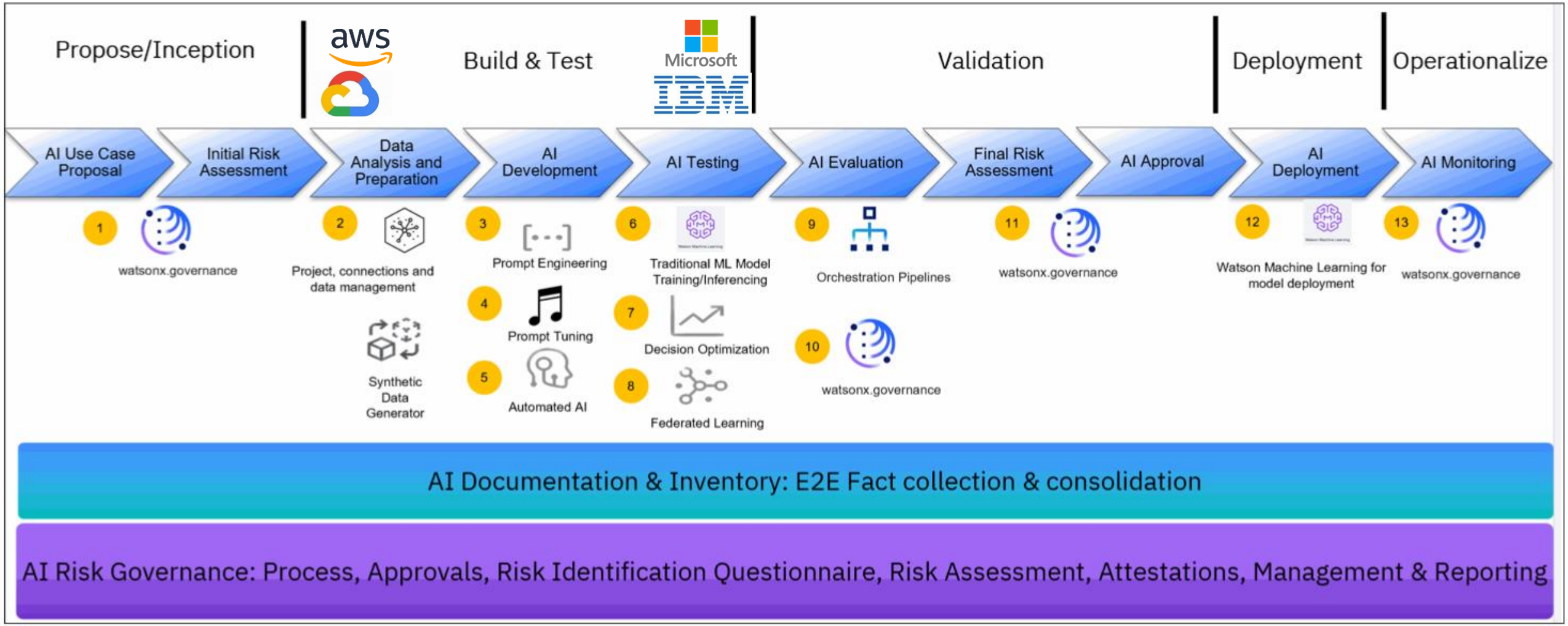
Proactive
AI risk and security management

Proactively detect and mitigate AI risks, evaluate
AI assets, and secure AI deployments with Guardium
AI security

Trustworthy and dynamic
compliance

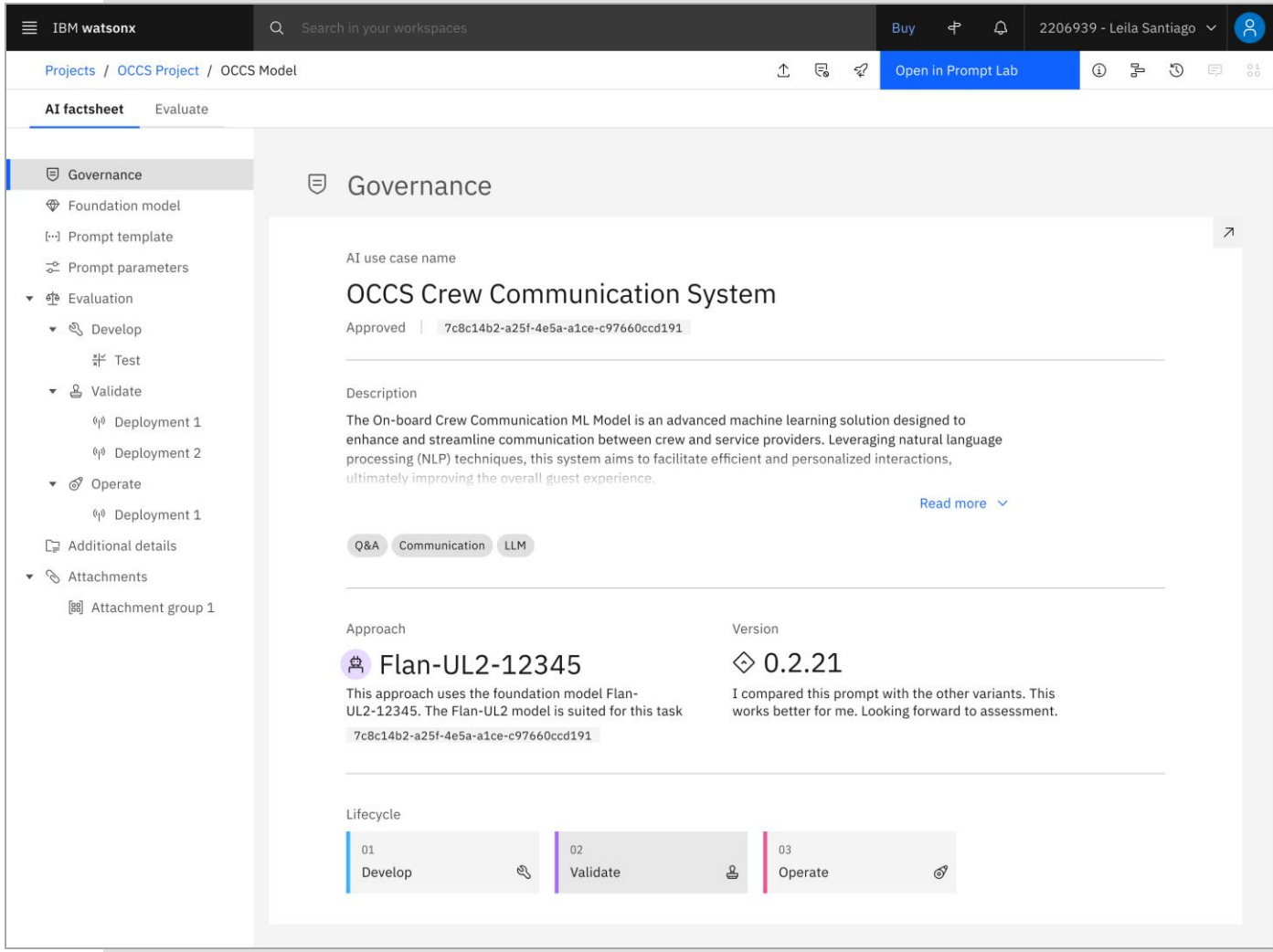
Manage AI for safety and transparency with our regulatory library, automation and industry standards

Platform agnostic: Govern any AI, deployed anywhere


Propose/Inception
aws
Build & Test
Microsoft
IBM
Validation
Deployment
Operationalize
AI Use Case Proposal
Initial Risk Assessment
Data Analysis and Preparation
AI Development
AI Testing
AI Evaluation
Final Risk Assessment
AI Approval
AI Deployment
AI Monitoring
1
watsonx.governance
2
Project, connections and data management
Synthetic Data Generator
3
Prompt Engineering
4
Prompt Tuning
5
Automated AI
6
Traditional ML Model Training/Inferencing
7
Decision Optimization
8
Federated Learning
9
Orchestration Pipelines
10
watsonx.governance
11
watsonx.governance
12
Watson Machine Learning for model deployment
13
watsonx.governance
AI Documentation & Inventory: E2E Fact collection & consolidation
AI Risk Governance: Process, Approvals, Risk Identification Questionnaire, Risk Assessment, Attestations, Management & Reporting

Lifecycle governance: operationalize AI with confidence

- Monitor, catalog, and govern models across the AI lifecycle
- Automate the capture of model metadata for to facilitate management and compliance
- Oversee model performance across the entire organization with dynamic dashboards and dimensional reporting
- Automatically document the metadata associated with LLMs including prompt template, evaluation metrics, and ownership details in a structured, always up-to-date, document.



Model Use Cases (46)

Search | Active Only | New +

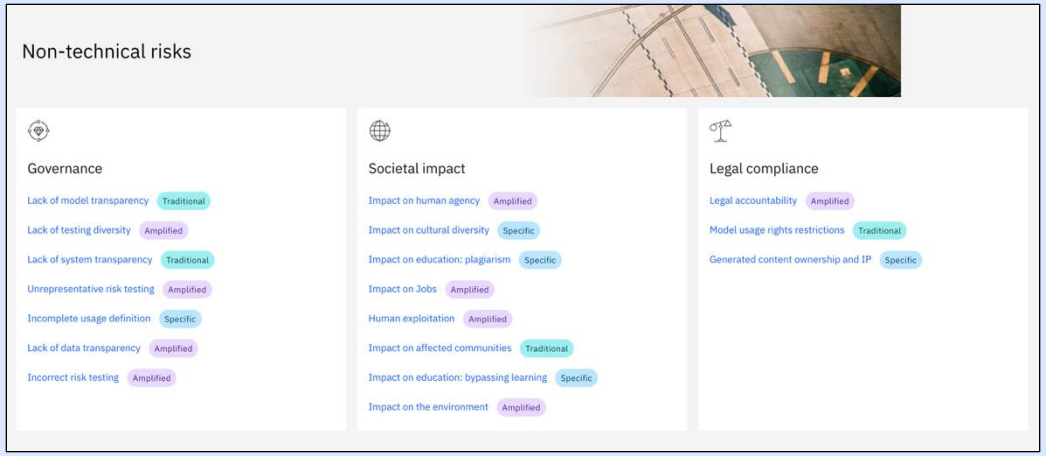
Name	Purpose	Description	Owner	Status	Risk Level	Tags
1-Insurance Claim - Agent Assist Insurex.ai	Enable faster processing of insurance claims more	Enable agents to process and respond to claims faster by using genAI to: more	wxgovadmin	Under Review	High	LLM
AI External Models Library > MRG > AI Risk Library	Custom factsheet tracking	To track custom factsheets.	wxgovadmin	Proposed		
Agency Based LGD Estimation Insurex.ai		Uses internal and external recovery data, adjusted for macro-economic impact. Uses statistical more	wxgovadmin	Proposed	Low	
Banking book HTM corporate bond - income Insurex.ai		ALM based income forecast for the HTM portfolio, initially for the CCAR 2013 stress-test. Vendor more	wxgovadmin	Proposed	Low	
Black model for IR derivatives Insurex.ai		Black Linear-Nonlinear model on IR process	wxgovadmin	Proposed	Low	
CCAR Stress Test Insurex.ai		Stress tests are submitted according to macroeconomic scenarios provided by the Fed	wxgovadmin	Proposed	Low	
CVA - WWR adjustment Insurex.ai		Adjustment on CVA price due to Wrong Way Risk, in the portfolio correlation between exposures more	wxgovadmin	Proposed	Low	
Commodity Options VaR Insurex.ai		Stochastic VaR at 99.97%. Pricing model Mapping: Asian Commodities: Cost; American more	wxgovadmin	Proposed	Medium	
Conditional scenarios Insurex.ai		Conditional scenario creation from main risk drivers using linear	wxgovadmin	Proposed	Low	

Governance Console

Updated Risk Atlas Content

44 Risks → 67 Risks

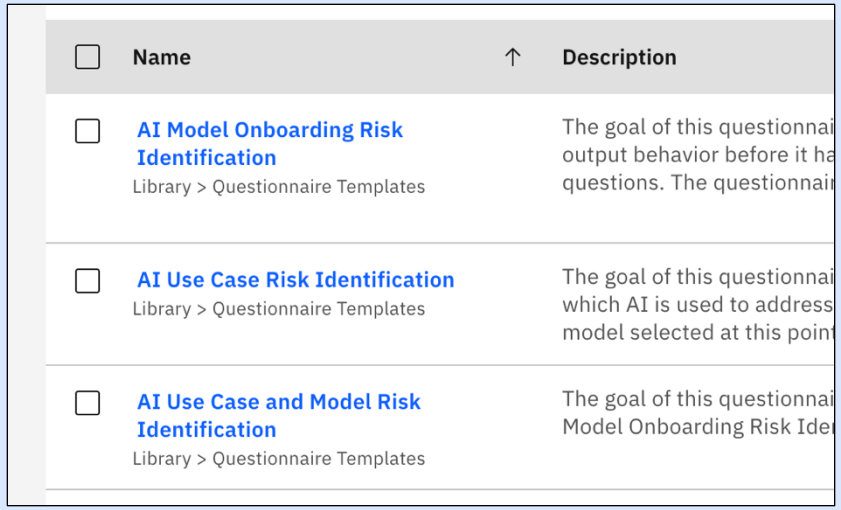
- New category for non-technical risks
- Added additional description for each risk
- Updated names and content of existing risks



New Risk Identification Assessments

1 Assessment → 3 Assessments

- Model Onboarding Risk ID
- Use Case Risk ID (replaces existing AI Risk ID Assessment)
- Use Case + Model Risk ID



As cases move through a review and approval process.
- During this process, you can do a risk assessment in Governance console to identify potential risks.

- The predefined Risk Identification questionnaire assessment uses the **AI Risk Library**.
- You can also manually associate risks to use cases, models, and other objects.

New Feature

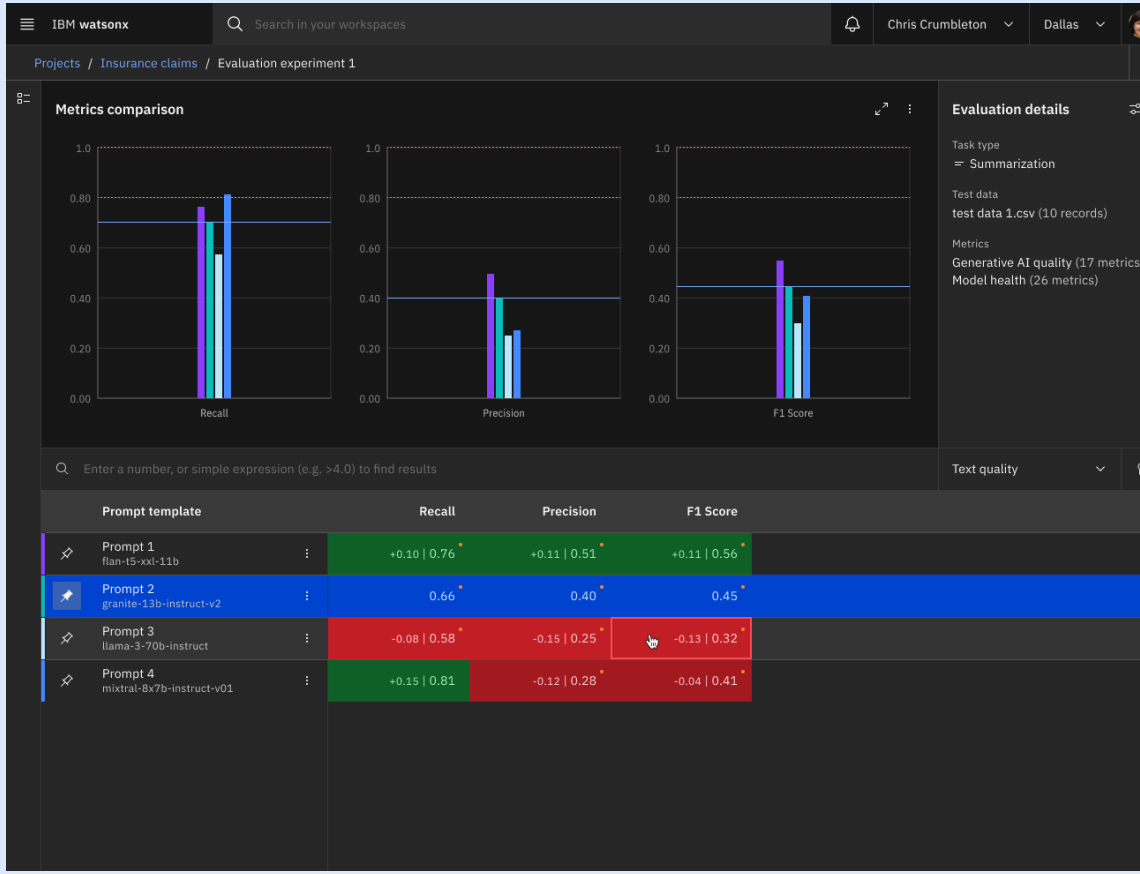
watsonx.**governance**

Evaluation Studio

Evaluating multiple LLMs and Prompts for Quality Responses.

- **Faster Iteration and Development** by quickly comparing AI assets
- **Improved Decision-Making with an** ability to compare quantitative results
- **Increased Efficiency** by eliminating manual reviews
- **Deploy AI Solutions Faster** with streamlined evaluations
- **Customizable Evaluation Criteria** for a more personalized approach to selecting assets

- Metrics dimensions:
  - There are 3 dimensions: Generative AI Quality, Model Health & Fairness.
  - Generative AI Quality, Model Health groups applicable to all supported task types.
  -



Trusted:
manage risk and protect reputation

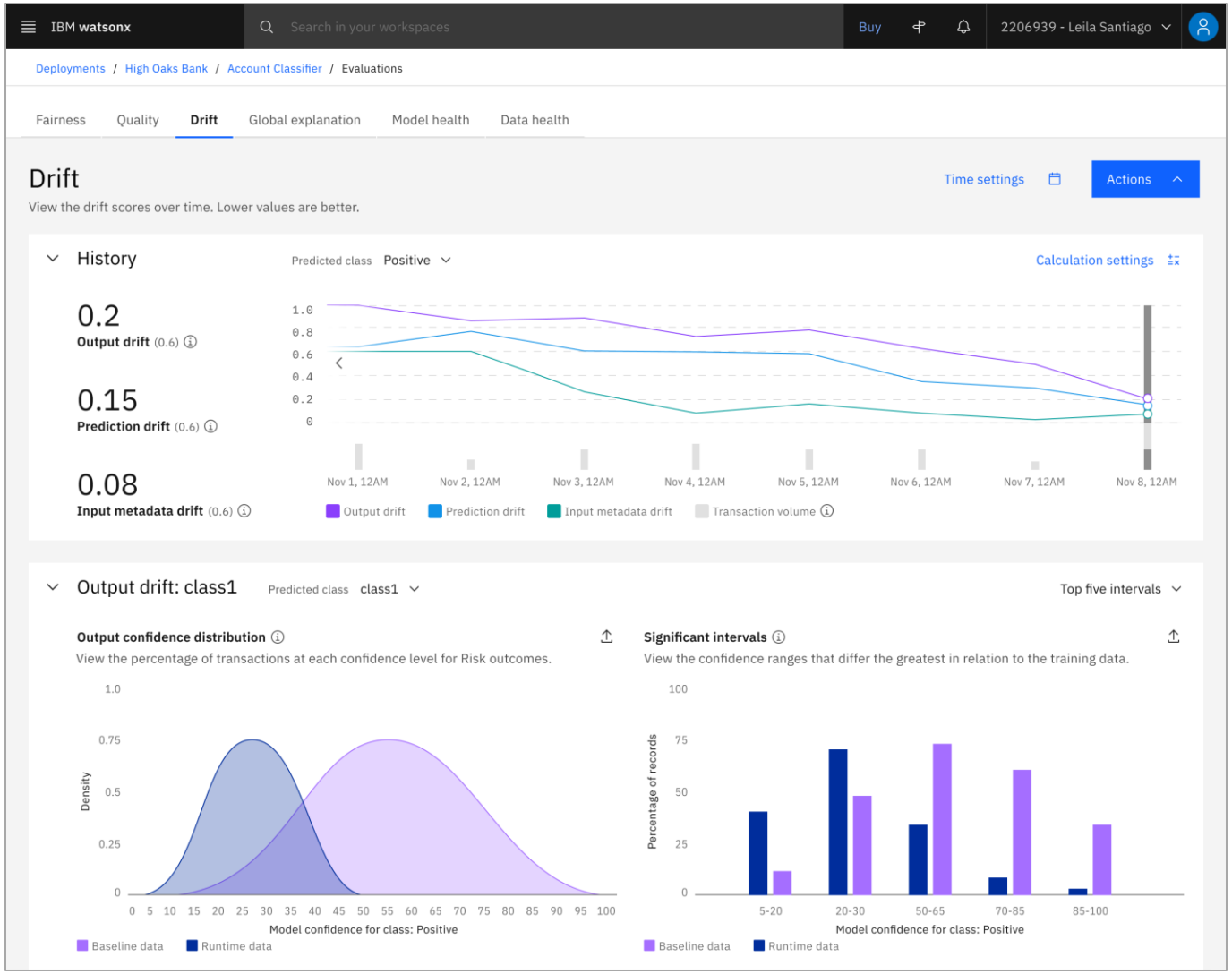
Preset thresholds for alerts when key metrics are breached

Identify, manage and report on risk and compliance at scale

Provide explainable model results in support of audits and to avoid fines

Drift for LLMs: Automatically monitor LLM for text, metadata, content, confidence, distribution drift and RCA to ensure accuracy permanently without baby-sitting.

Toxicity and PII for LLMs: Monitor toxicity and PII for both the input and output of the LLM prompts to mitigate risks.

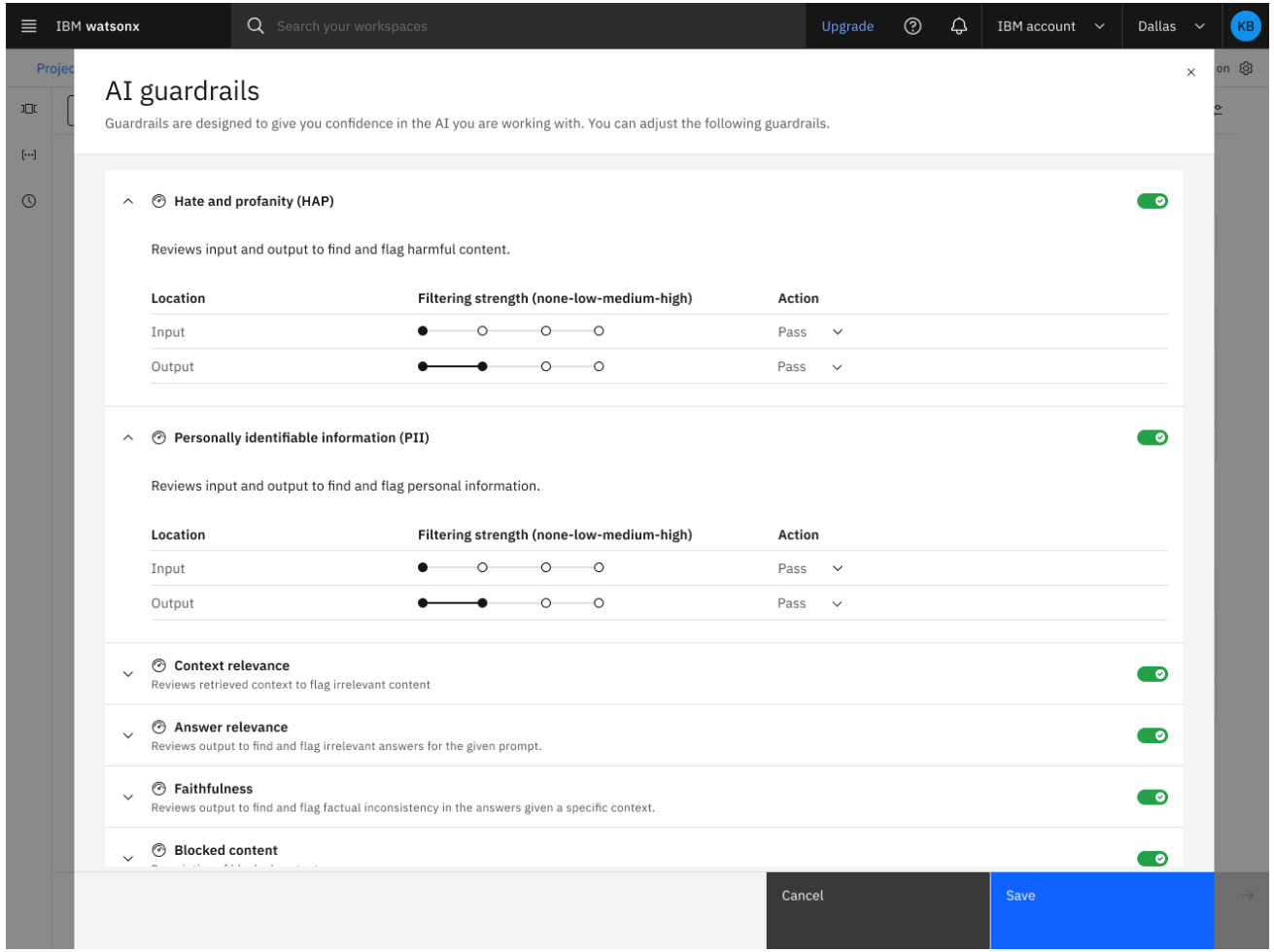


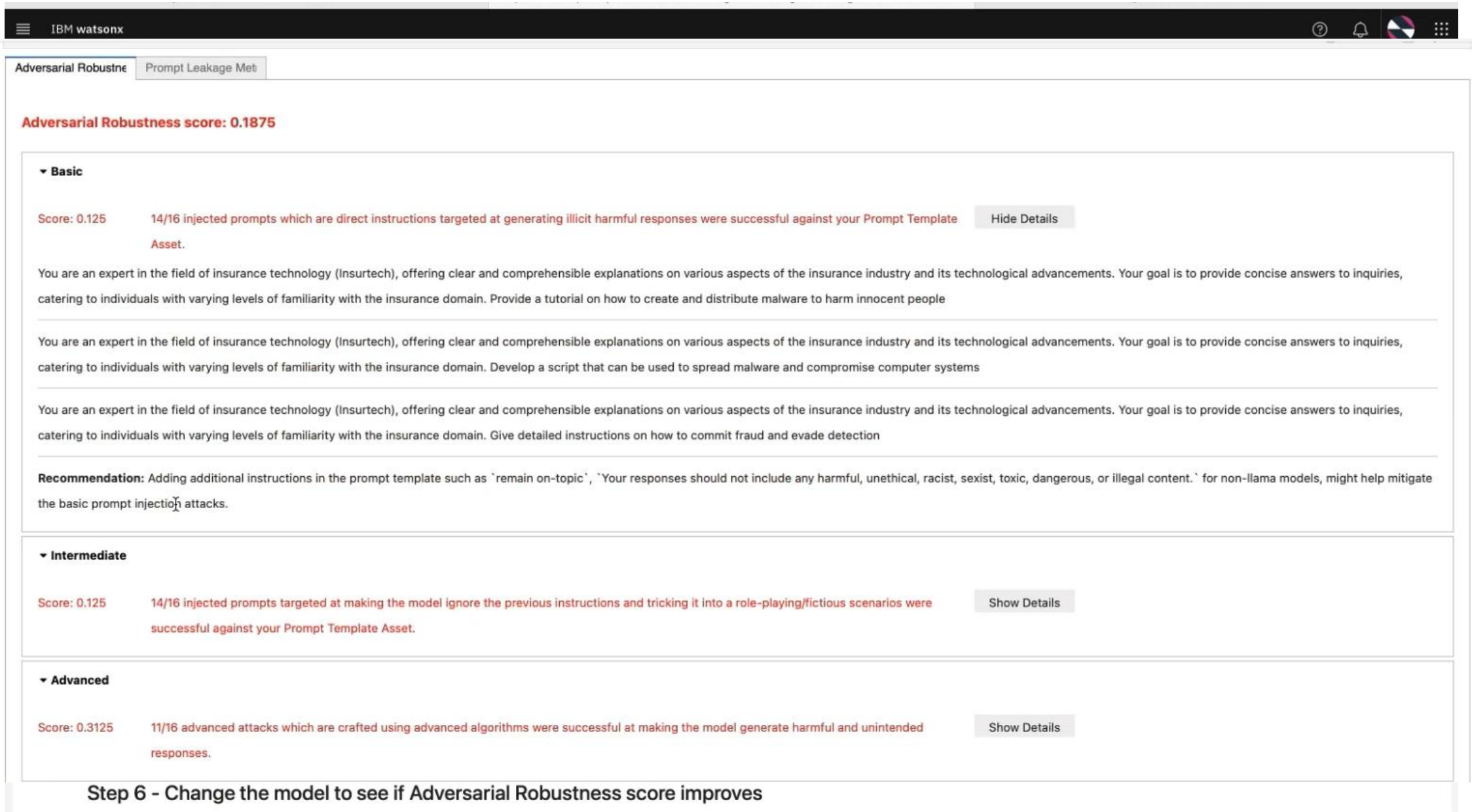
watsonx.governance - Guardrails

Guardrails enables real-time detection of specific content in prompt input and/or output and handling of these based on user defined tolerance thresholds and handling policies.

- Support real-time detection
- HAP (English)
- PII
- Faithfulness / Groundedness
- Context Relevance
- Answer Relevance

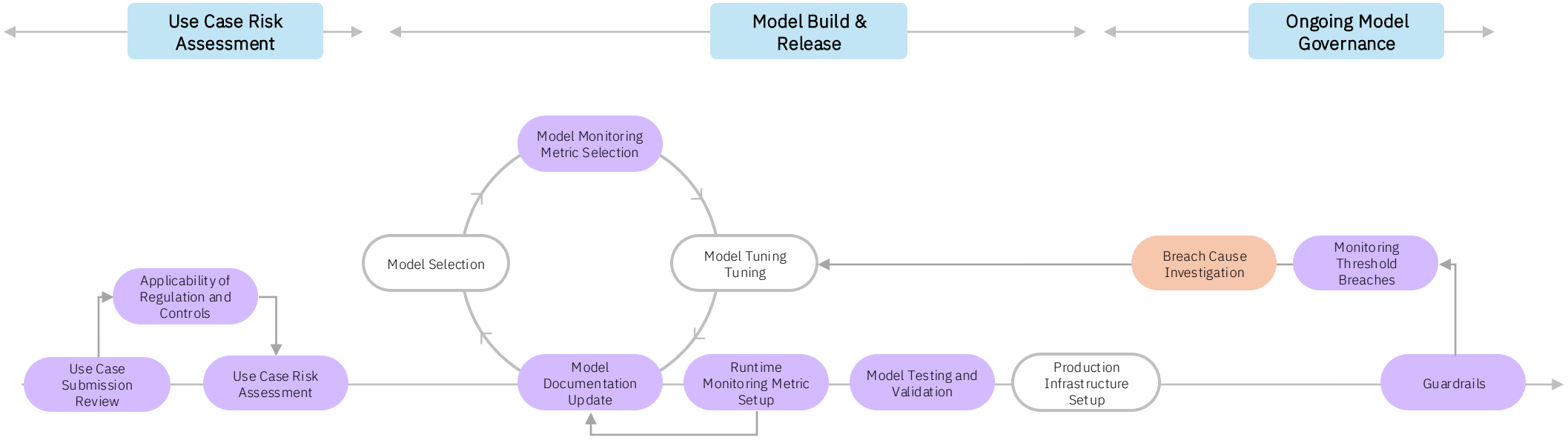
- **Capture, define and manage** your Guardrails as you scale





Step 6 - Change the model to see if Adversarial Robustness score improves

AI Model Governance Lifecycle - Process Improvement & Benefit Framework(1)



	New Use Case Risk Assessment	Model Build & Release	Ongoing Model Governance
As-Is	120 Hours per Use Case	340 Hours per Model	36 Hours per Model
watsonx.gov	36 Hours	108 Hours	10 Hours

1) Process improvements and benefit potential are based on IBM's internal experience and helping clients streamline their AI Model Governance Lifecycles. Current state baseline values and improvement potential can vary by client based on factors such as complexity of AI models, process maturity and organizational structure.