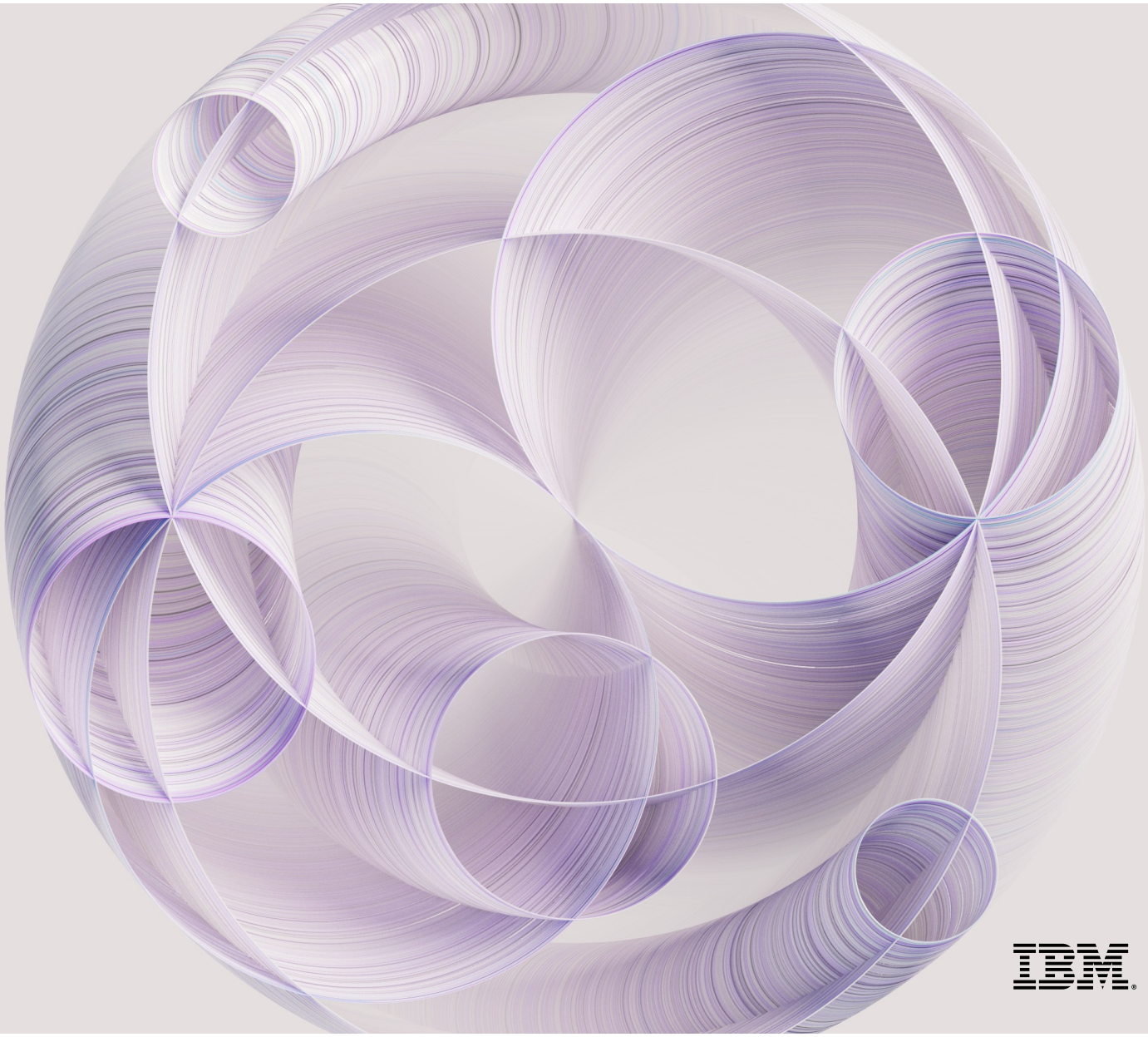
Accelerate responsible, transparent and explainable AI workflows



watsonx.governance

IBM

Foundation Models are bringing an inflection point in AI...

...but how enterprises adopt and execute will define whether they unlock value at scale

The speed, scope, and scale of generative AI impact is unprecedented

Massive early adoption

80%

of enterprises are working with or planning to leverage foundation models and adopt generative AI

Broad-reaching and deep impact

Generative AI could raise global GDP by

7% within 10 years

Critical focus of AI activity and investment

Generative AI expected to represent

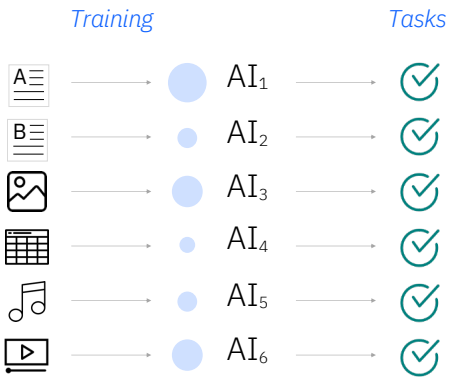
30%

of overall market by 2025

Sources: Statista; Reuters; Goldman Sachs; IBM Institute for Business Value; Gartner. Scale Zeitgeist: AI Readiness Report, a survey of more than 1,600 executives and ML practitioners

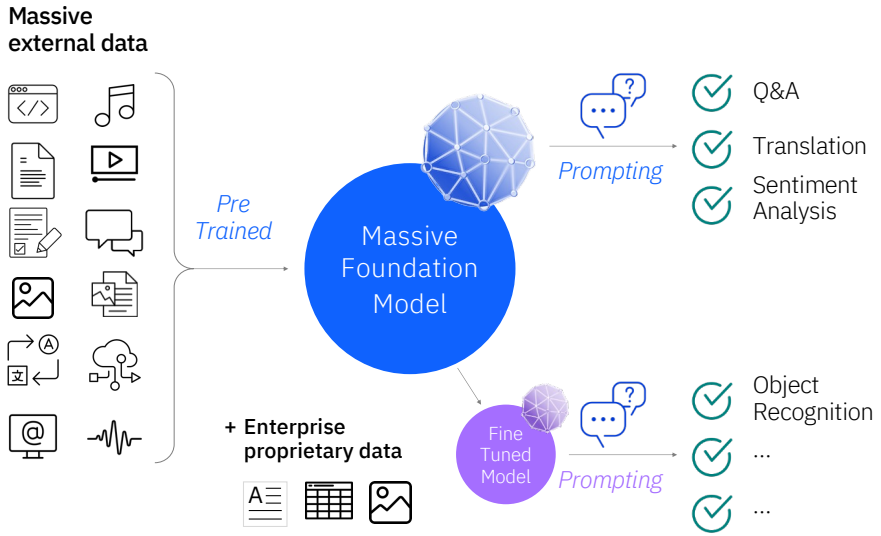
Foundation models establish a new paradigm for AI capabilities

Traditional AI models



- Individual siloed models
- Require task specific training
- Lots of human supervised training

Foundation Models



- Massive multi-tasking model
- Adaptable with little or no training
- Pre-trained unsupervised learning

Enhanced capabilities

- Summarization
- Conversational Knowledge
- Content Creation
- Code Co-Creation

Key advantages

- Lower upfront costs through less labeling
- Faster deployment through fine tuning and inferencing
- Equal or better accuracy for multiple use cases
- Incremental revenue. through better performance

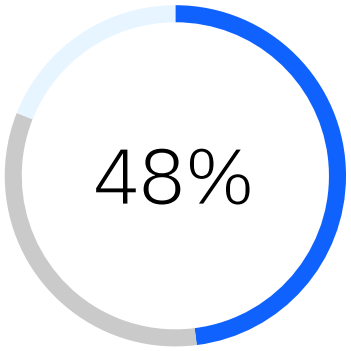
up to **70% reduction** in certain NLP tasks

Enterprise considerations

Business leaders face challenges in scaling AI across the enterprise with trust

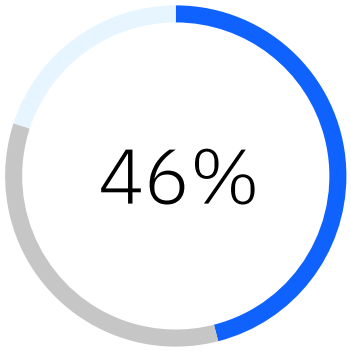
80% of business leaders see at least one of these ethical issues as a major concern

Explainability



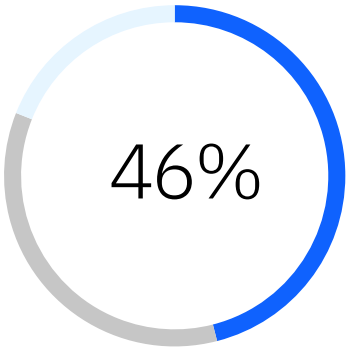
Believe decisions made by Generative AI are not sufficiently **explainable**.

Ethics



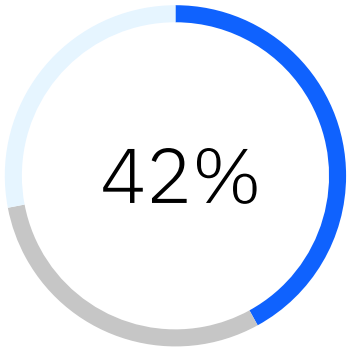
Concerned about the safety and **ethical** aspects of Generative AI.

Bias



Believe that Generative AI will propagate established **biases**.

Trust

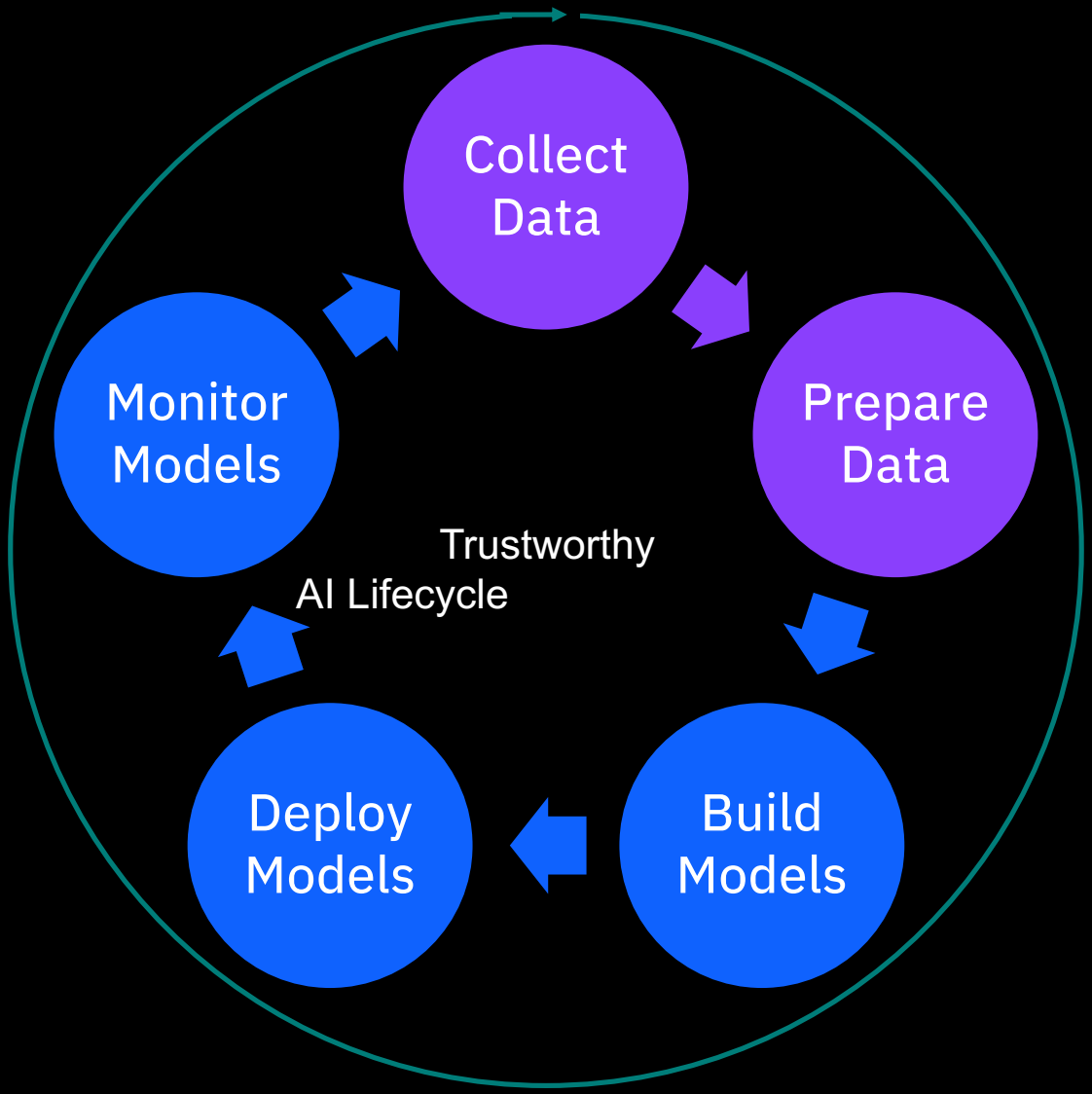


Believe Generative AI cannot be **trusted**.

Source: IBM IBV “Generative AI: The state of the market”, June 2023

Agree Neutral Disagree


Uncontrolled AI is a threat to your enterprise
Damage to your brand
Regulatory fines and failed audits
Lost revenue, higher expenses



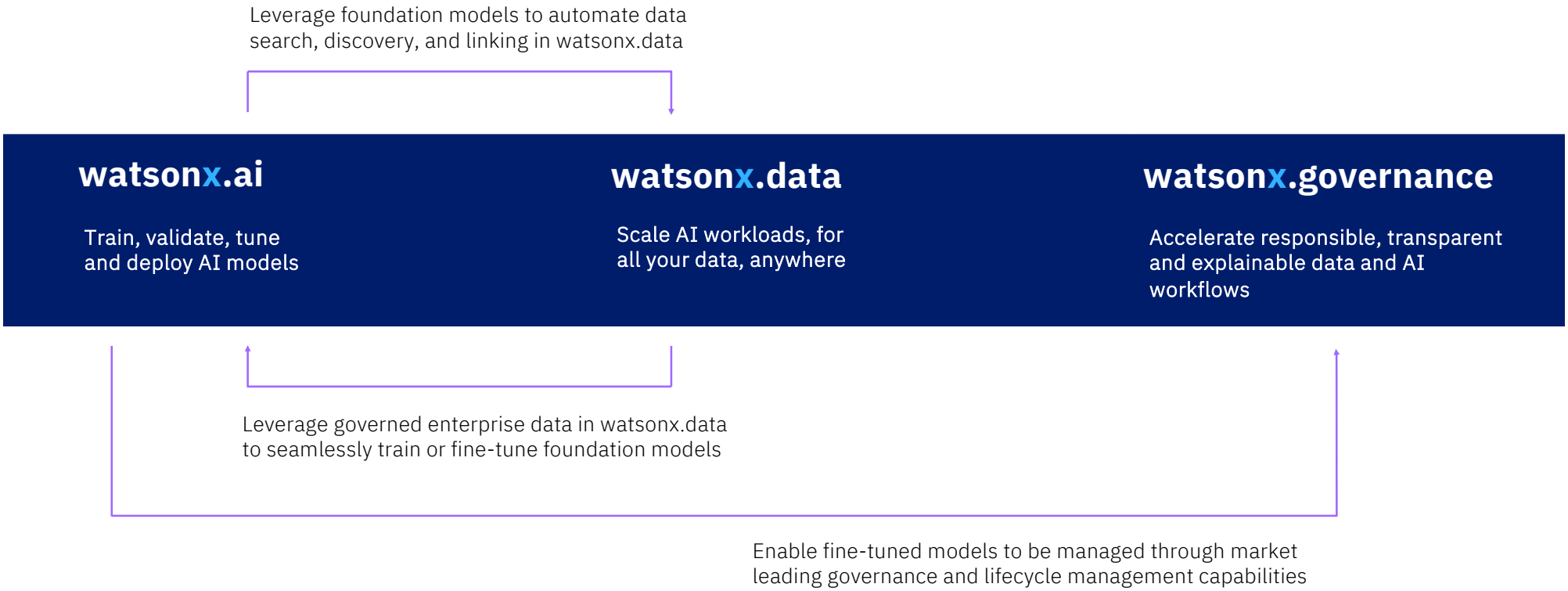
MLOps and Trustworthy AI

- **Data** : A complete view of quality data that is private, self-served and ready for analysis by multiple personas
- **Model** : MLOps infused with fairness, explainability, and robustness
- **Process** : Automation to drive consistency, efficiency and transparency for AI at scale

AI needs governance –
the process of directing,
monitoring and managing the
AI activities of an organization

Put AI to work with **watsonx**

Scale and accelerate the impact of AI with trusted data



IBM watsonx.governance

end-to-end automated AI lifecycle governance toolkit built to mitigate risk and improve compliance

watsonx.governance

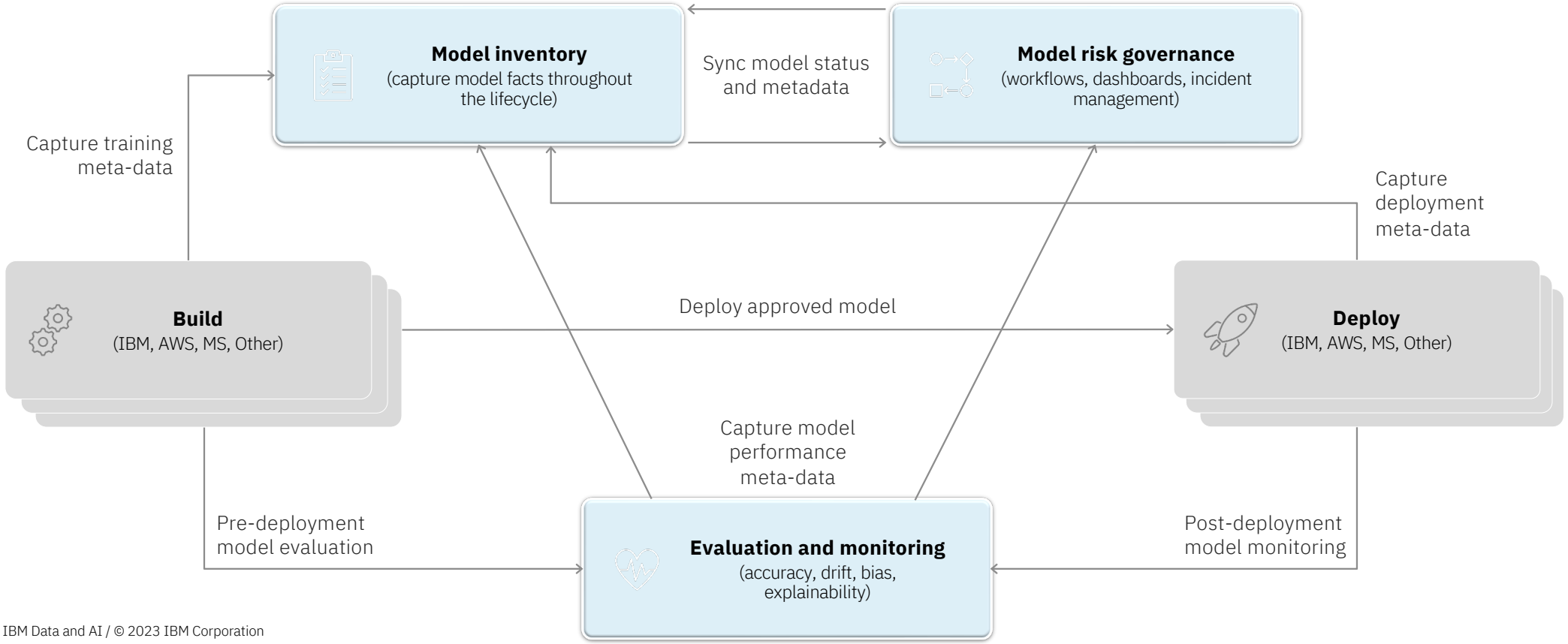
Accelerate responsible, transparent and explainable AI workflows

- Govern across the AI lifecycle.
- Automate and consolidate tools, applications and platforms.
- Capture metadata at each stage
- Support models built and deployed in 3rd party tools.

- Manage risk and protect reputation by automating workflows to ensure quality and better detect bias and drift.

- Adhere to regulatory compliance by translating growing regulations into enforceable policies.
- Helps you to operationalize AI with confidence

Govern across the AI lifecycle



In practice, that means...

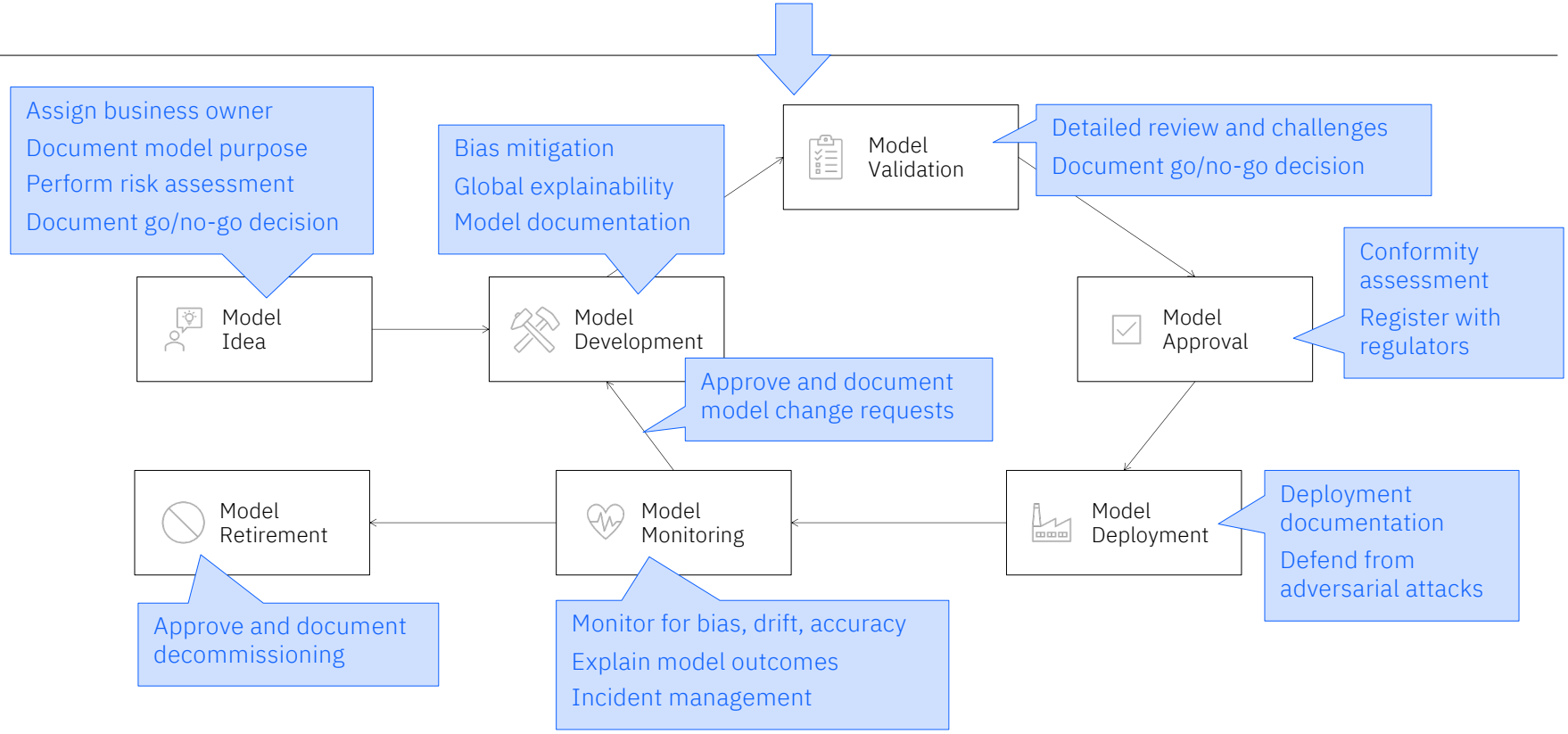
Macro level

Define legal obligations
Define relevant company policies
Extend enterprise risk framework

Articulate AI principles
AI Governance board & escalation process
Define roles & responsibilities

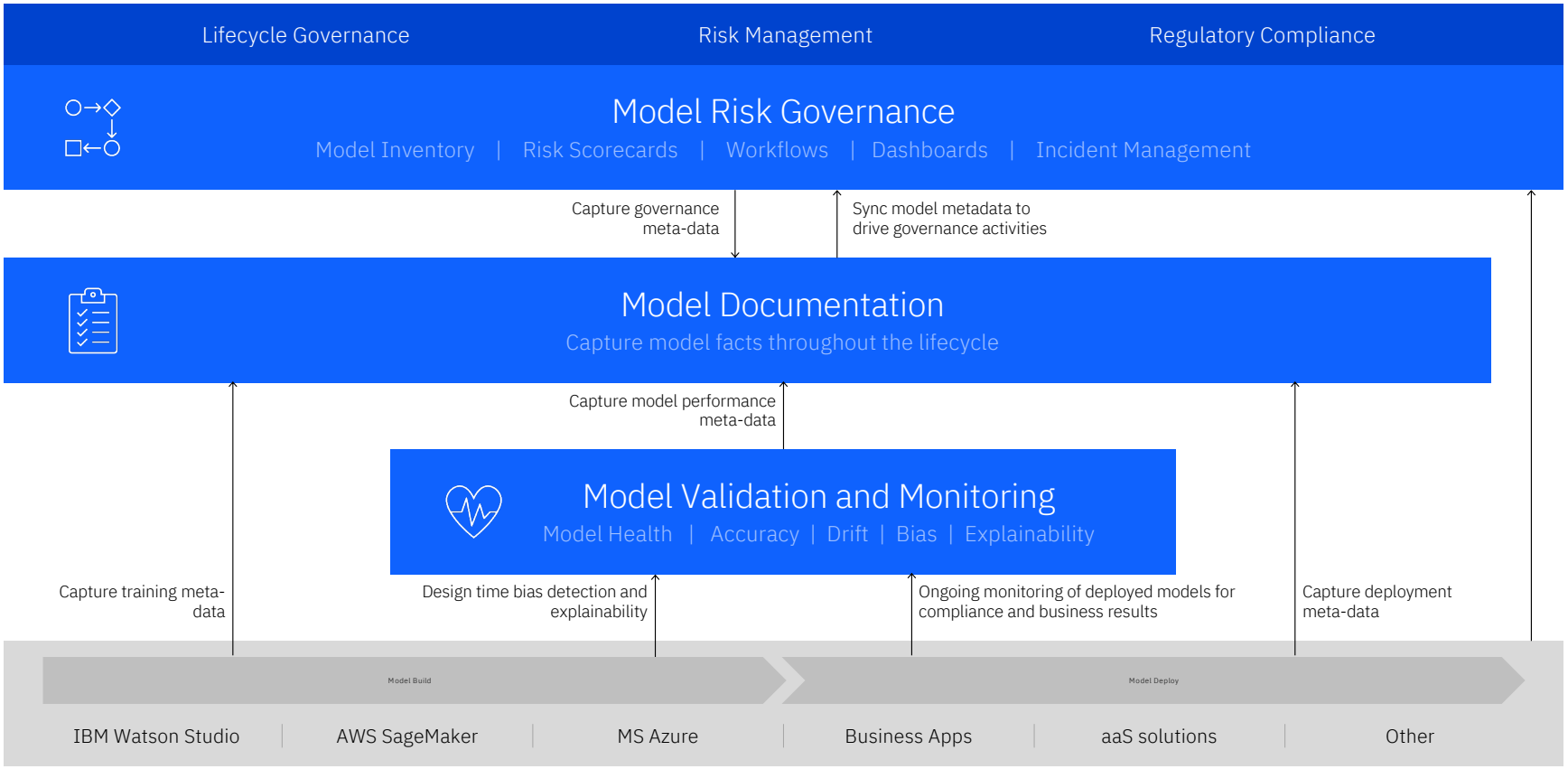
Awareness and skills
Audits
...

Micro level



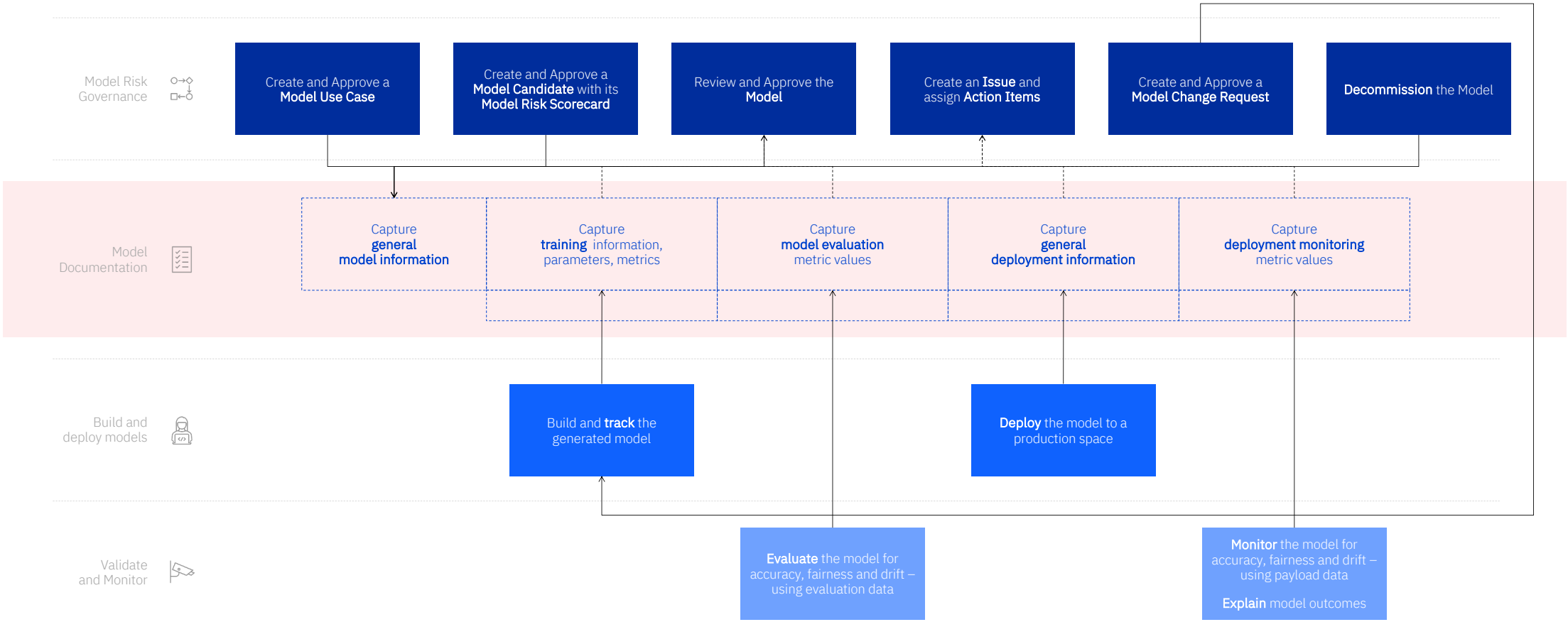
IBM AI Governance

- Model Owners
- Model Validators
- Audit Teams
- Compliance Teams
- Risk Management Teams
- Data Privacy Teams
- Principal Data Scientists



- Data Engineers
- (Citizen) Data Scientists
- MLOps
- ML Engineer

Lifecycle Governance – Model documentation



(boxes with solid fill colors are user actions, boxes with dashed outlines are what the integrated tooling does automatically)

IBM AI Governance

Role of the solution components in the three use cases

		Solution components				
		OpenPages			FactSheets	OpenScale
Use cases	Lifecycle Governance	Objects: •Model •Model Scorecard •Change Request •Review •Challenge	Workflows: •Model Request •Model Risk Assessment •Model Candidate •Model Development •Model Review •Model Change Request •Model Decommission	Dashboards widgets: •Models by Department •Models by Provider •Models by Lifecycle Stage •Model Reviews •Model Change Requests	•Capture training facts •Capture deployment facts •Capture evaluation facts	•OpenScale Python API to evaluate model during development •OpenScale MRM capability to evaluate model during validation •OpenScale production monitoring
	Risk Management	Objects: •Metric •Metric Value •Issue •Action Item	Workflows: •Metric Value Review •Issue Review •Action Item Approval	Dashboards widgets: •Policies •Models by Risk Tier •Model Metric Status •Development Accuracy	•Capture monitoring facts •Investigate metric alerts •Capture training facts for challenger models	•Post-deployment model monitoring •Investigate metric alerts •Pre-deployment evaluation of challenger models •Compare champion and challenger models
	Regulatory Compliance	Objects: •(Sub)Mandate •Policy •Requirement	Workflows: •Model Review on occurrence of major alert •Model checks before deployment to production, etc.	Dashboards widgets: •Mandates •Requirements •Obligations	•Model report export	•Model report export

What is automated between the AI Governance components?

User action	Automation
Create a new Model Use Case in OpenPages	• Create a factsheet for the use case • Sync general information from OpenPages to Factsheets • Create a hyperlink in OpenPages
Train a model in Python	• Factsheet Python client: Auto-logging model training facts from common open-source frameworks • OpenScale Python client: design time bias detection and explainability
Train a model in AutoAI	• AutoAI automatically saves model training facts in the model object (incl. Fairness metrics if you use that)
Track a model in Factsheets	• Create a new Model object in OpenPages (or map to an existing one) • Sync training metadata to OpenPages
Deploy a model from Watson Studio to WML	• Capture deployment metadata in Factsheets • Create a new Deployment object in OpenPages • Sync deployment metadata to OpenPages
Evaluate a model in OpenScale	• Capture monitoring metadata in Factsheets • Sync monitoring metadata to OpenPages
Evaluate a model using custom tech/metrics	• Capture external metrics in OpenScale • Capture monitoring metadata in Factsheets • Sync monitoring metadata to OpenPages

Risk is every one's business.

In today's turbulent environment, the need to take on risk with confidence is greater than ever before.

90%

of compliance leaders expect evolving business, regulatory, and customer demands to increase compliance-related operating costs by up to 30%.[1]

79%

of organizations report that keeping up with the speed of digital and other transformations is a significant risk management challenge.[2]

77%

of organizations recognize the need to upgrade their Third-Party Risk Management operating model.[3]

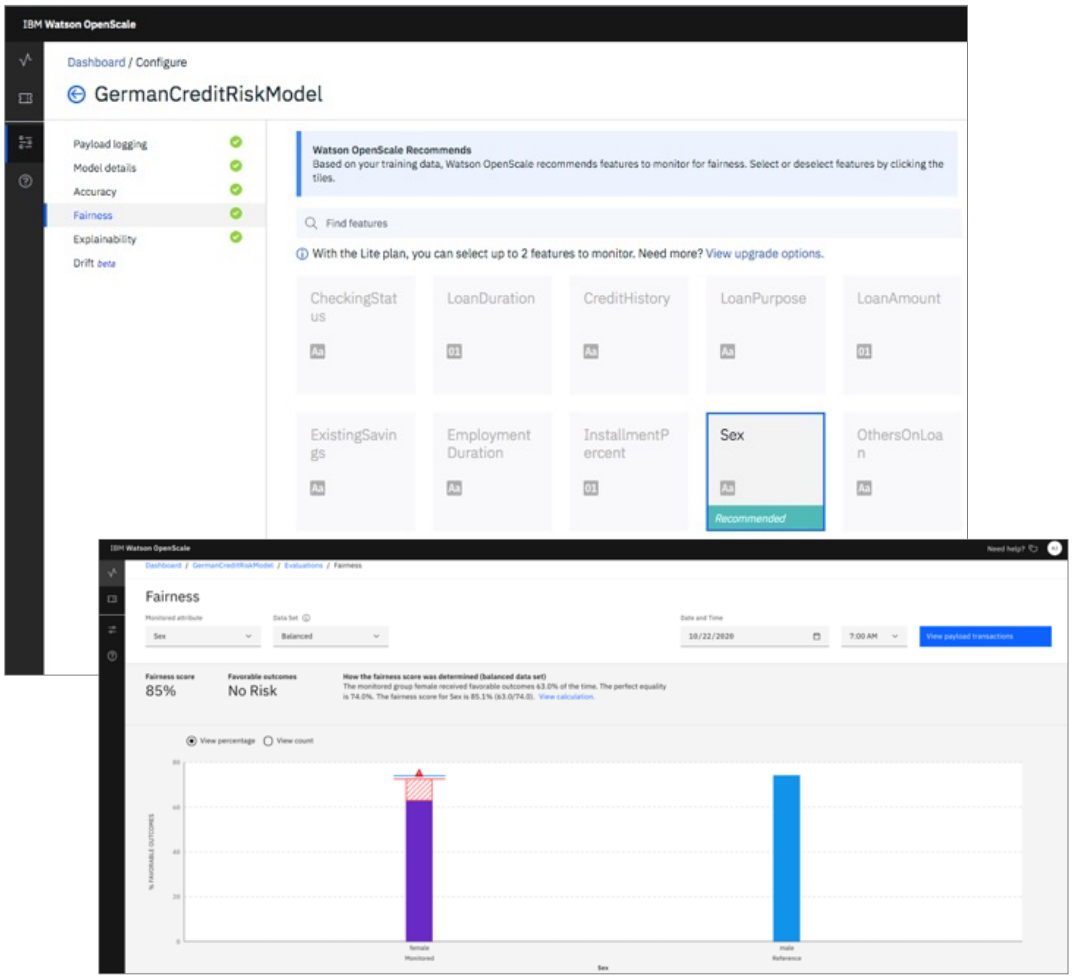
Source: 1. 2022 Compliance Risk Study Report, Accenture
Source: 2. 2022 Global Risk Survey, PWC
Source: 3. 2022 Third-Party Risk Management Outlook, KPMG

Bias Detection

Watson OpenScale enables enterprises to enforce fairness in their model's outcome by analyzing transactions in production and finding biased behavior by the model
It pinpoints the source of bias and actively mitigates the biases found in production environment

Value:

- Automatically recommend common protected attributes to monitor during production
- Detect biases in runtime in order to catch impacts on business applications and compliance requirements without time consuming, manual data analysis
- Metrics and data to help data scientists further troubleshoot issues in data sets or models
- Mitigate biases in runtime in order to enforce regulatory or enterprise fairness guardrails in real time

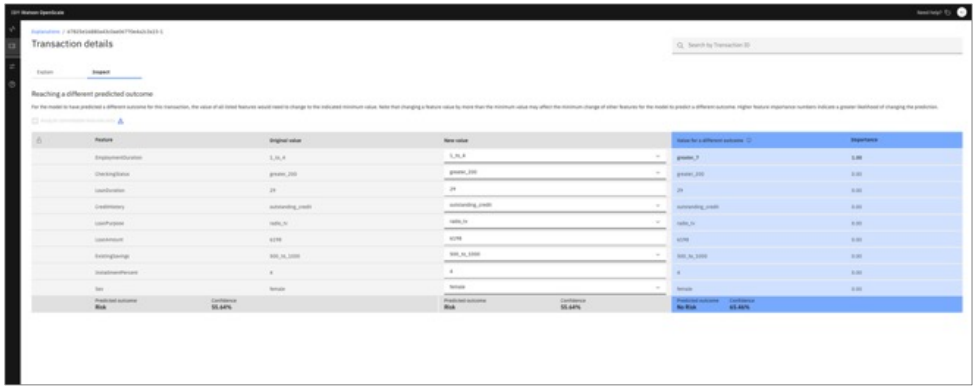


Explainability

Watson OpenScale records every individual transaction and drills down into its working to explain how the model makes decisions.
It provides a simple explanation that is user friendly and interactive

Value:

- Explain individual transaction level decisions made by the model in run time, including details about most important attributes and their values in order to assist in compliance and customer care situations
- Analyze individual transactions in a what-if manner in order to understand how model behavior will change in different business situations
- Provide global explanation for decisions made by the model using SHAP

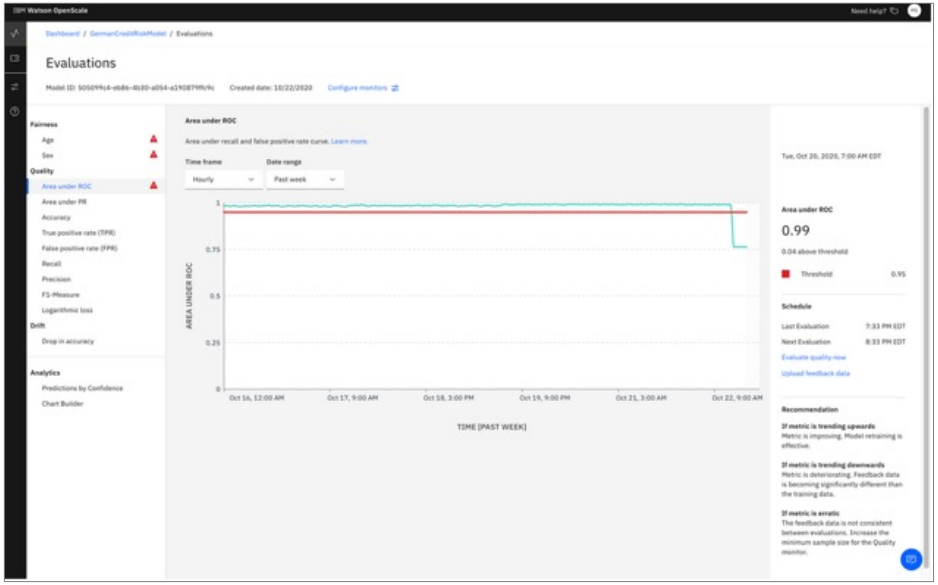


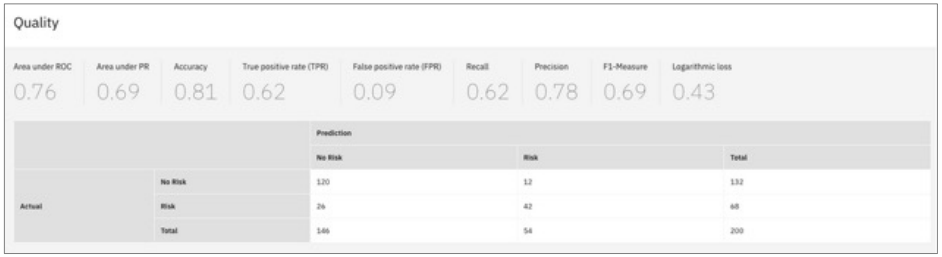
Quality

Monitor model quality metrics like accuracy, precision, recall, AUC, etc. and get alerts when the value goes beyond the set threshold
Can also define custom metrics to track quality for

Value:

- Observe and track quality metrics to determine how well your model predicts outcomes
- Use custom metrics to create and track user-defined metrics to assess the quality of models



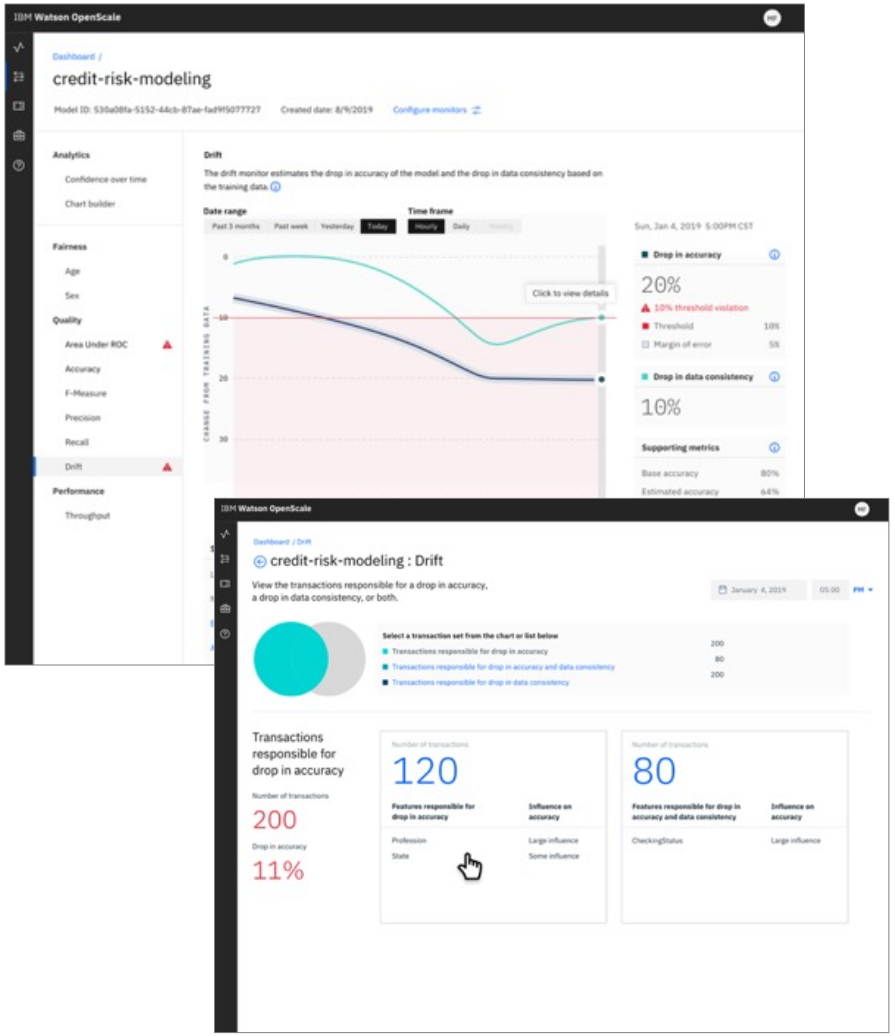


Drift

Automatically detect drifted transactions and pinpoint datapoints that contribute to drift
Identify drift in Accuracy as well as drift due to data consistency

Value:

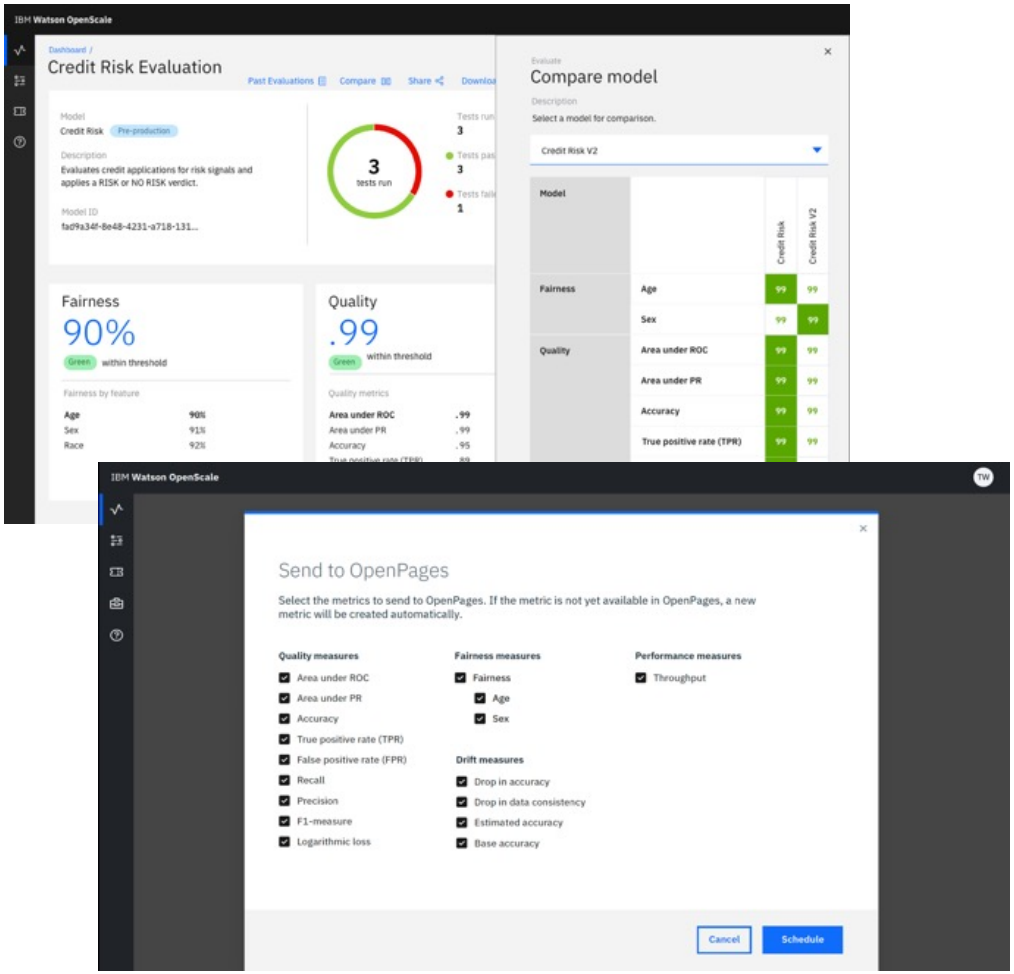
- Business environments are dynamic leading to “drift” in data and cause inaccuracies in model prediction. Monitor drift to make sure model is resilient and updated for such changing conditions
- Drill down into transactions causing drift to identify root cause and take appropriate actions



Model Validation/Model Risk Management

OpenScale enables enterprises to validate pre-production models before putting them into production to ensure they can be trusted to perform as intended.

- Validate pre-production models and generate reports of outcomes
- Enable customizable tests relevant to AI models
- Compare performance of models
- Automatically configure monitoring of production models to match pre-production settings.
- Synchronize results with Governance, Risk and Compliance (GRC) solutions (Initially with OpenPages Model Risk Governance)
- Integrate with AI Factsheets to capture metadata and track metrics



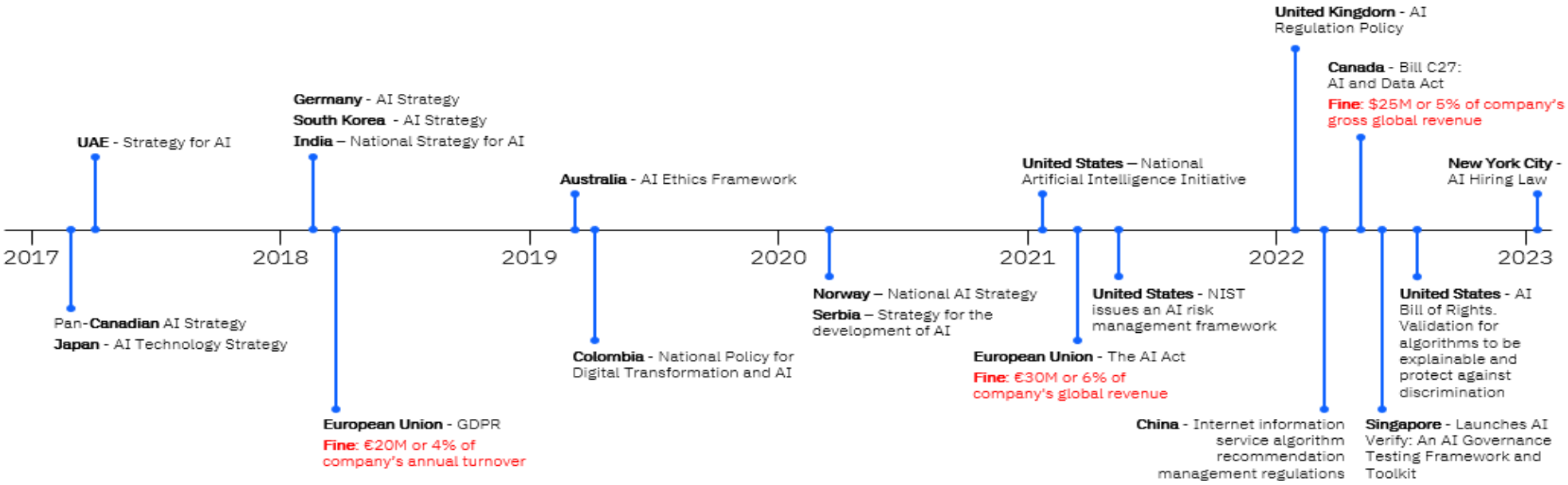
Policy Packs

- Pre-built assessments to give clients a “starting point” to aid in regulatory compliance
- Ability to determine **applicability** of model use cases against various regulations
- Receive **guidance** on how to comply
- Content is provided and curated through partnership with **regulatory experts**
- We will support selected regulations (e.g. NYC Local Law 144, SR-11-7, EU AI Act, etc.)

The proposed EU AI Act includes potential monetary penalties of €30M or 6% of company's global revenue[2]

Source:
2. EU set to ratchet up AI fines to 6% of turnover, Reuters

Adhere to growing and changing AI regulations



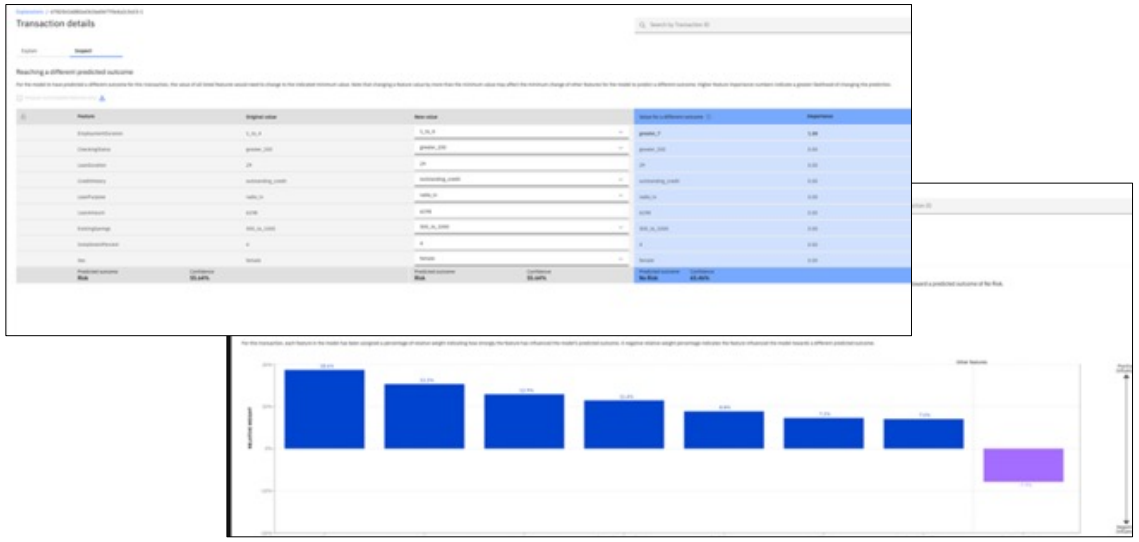
Adhere to growing and changing AI regulations

Improve AI regulatory compliance:

- Align processes and actions with your organizations' business goals and external legal and regulatory requirements.
- Automate the translation of AI regulations into enforceable policies and standards, quickly respond to change
- Provide persona-based dashboards and reporting to key stakeholders on the status of compliance

Document model facts using factsheets

- Support compliance audits, and requests for model details required by auditors, management, stakeholders and customers
- Explain transaction level decisions of models in runtime
- Understand how the model will behave in different business situations
- Facilitate subsequent enterprise validation based on model facts



Regions Bank – transparent AI with AI governance

Business challenge:

Regions Bank sought to address revenue loss to attrition among private wealth clients. Using trained models, the bank predicts customer risk to mitigate attrition.

Solution:

Enhance ML pipelines for Private Wealth Retention models to detect drift in model accuracy and data consistency in production data lake. The solution used IBM Expert Labs, IBM AI Governance, and IBM Watson Studio.

Results:

- Standardize monitoring of models in the end-to-end ML pipelines
- Continuous read on accuracy of predictions in production
- Automate retraining of models based on monitoring alert
- Identify wealth clients with low confidence predictions

Industry: Banking and Financial Markets
Geography: North America

Competitive overview

Traditional data science players

How is IBM better?

1. Most players lack GRC workflow capabilities such as OpenPages.
2. Automatic fact capture coupled with customizable reporting capability through Factsheets.
3. Monitoring and governance on these platforms tend to work with model developed on their platform only.

Niche players

How is IBM better?

1. Most niche players don't have a large or technical enough support and services team to guide customers.
2. Very few niche players focus on the E2E model lifecycle and lack a singular view for all types of model metadata.
3. IBM fuels its AI Governance stack with state-of-the-art algorithms and frameworks from IBM Research.

Open source

How is IBM better?

1. Most opensource tools tend to be very technical and tend to ignore the business user such as risk analysts.
2. Although state-of-the-art, opensource tools in the AI Governance space are difficult to maintain, requiring certain skills and resources.

Consulting firms

How is IBM better?

1. Consulting firms may develop one-off solutions that are catered to the customers current data science stack leading to heavily customized code, which is difficult to maintain. Leveraging consulting firms for end-to-end technological implementations is costly.