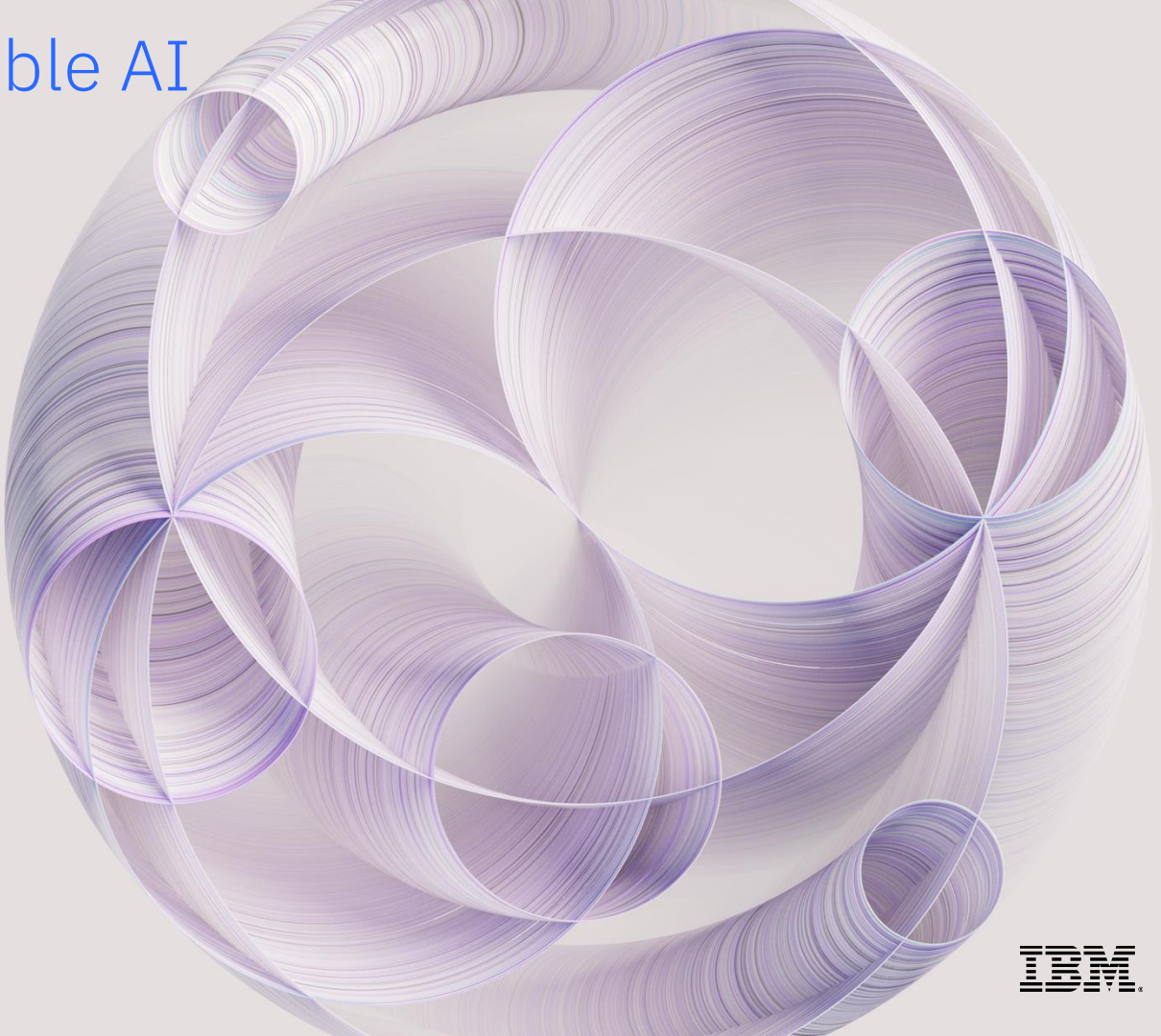
Accelerate Responsible AI with End to End AI Governance



Melanie Brunache
IBM Worldwide Data and AI Team
November 2024

Why AI Governance now?

AI is becoming widely adopted but....

BlackRock shelves unexplainable AI liquidity models

Risk USA: Neural nets beat other models in tests, but results could not be explained

YouTube sued for using AI to racially profile content creators

They claim YouTube's algorithms discriminate against black users

Cigna Sued Over Algorithm Allegedly Used To Deny Coverage To Hundreds Of Thousands Of Patients

/ Users have been reporting all sorts of 'unhinged' behavior from Microsoft's AI chatbot. In one

the study from earlier this year claimed the higher audit rate for Black taxpayers is due to a flawed artificial intelligence algorithm relied on by the IRS to decide who gets audited.

The Washington Post
Democracy Dies in Darkness
Get 1 year for $29

Apple Card algorithm sparks gender bias allegations against Goldman Sachs

RETAIL OCTOBER 10, 2018 / 4:04 PM / UPDATED 2 YEARS AGO

Amazon scraps secret AI recruiting tool that showed bias against women

Over-Segmenting In Financial Services Is So Over - Bye, Bye

Google AI chatbot threatens student asking for homework help, saying:

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

Operationalizing AI governance requires a strategic approach

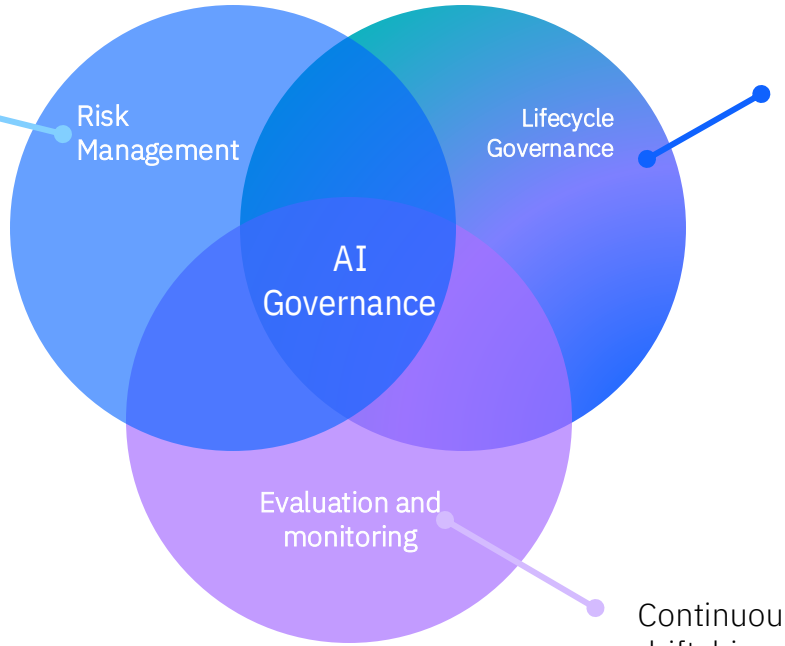
Common challenges	Getting to an ideal state
Manual work can lead to costly errors, drawn out model lifecycles and lead to employee burnout	**Automate** AI governance activities to streamline processes
Overhead and missed opportunities resulting from **multiple, disparate, tools and interfaces not optimized for AI**	**Enhance suboptimal tooling,** automate and consolidate
Governing AI is not a **one-size-fits-all approach**, need to govern across hybrid AI	**Use a single set of governance policies and workflows** across platforms and applications
AI is becoming a team sport with many stakeholders, **lack of tools for collaboration and communication** can draw out model deployment and	**Drive visibility across the organization** with automated collaborative tools, customizable dashboards and reports

IBM watsonx.governance

A toolkit to help accelerate responsible, transparent and explainable data and AI workflows

3 Core Pillars

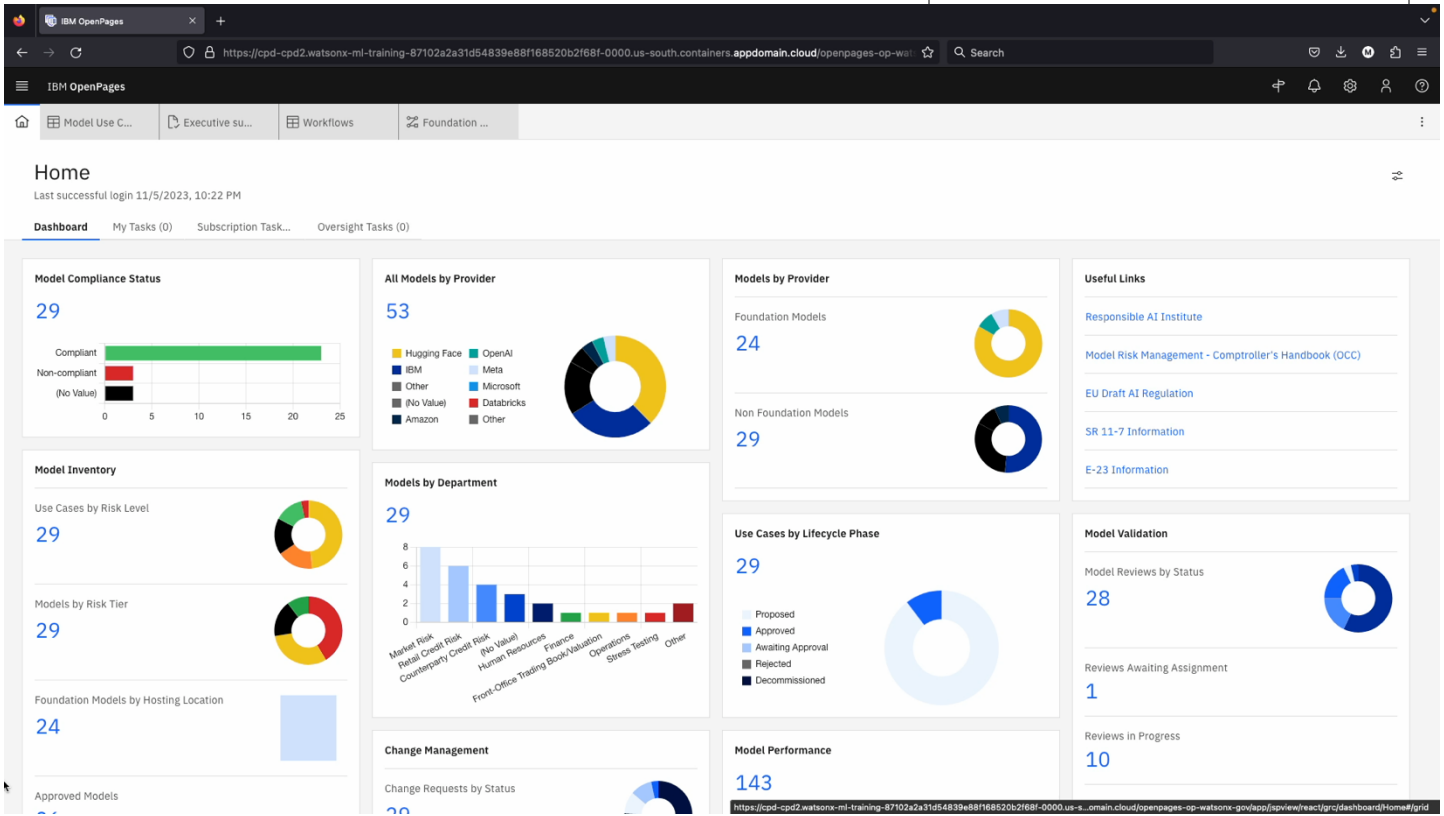
Workflows, dashboards for risk & issue management



Capture model facts throughout the entire lifecycle

Continuous monitoring for accuracy, drift, bias, explainability, and performance

Risk Management



What is it?

Single glass pane view of the enterprise's model risk landscape.

Manage and automate the activities around attestation, review, validation, change management and issue resolution of AI models

Why is it important?

Meet the compliance requirements of model-focused regulations across regions and jurisdictions.

Reduce governance costs.

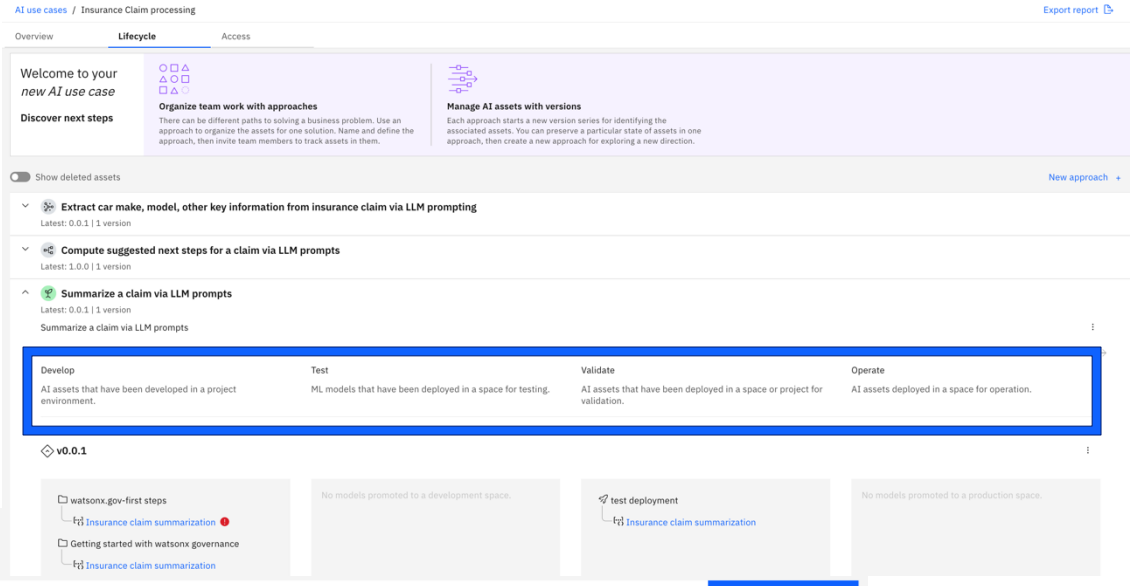
Maintain clear roles and responsibilities.

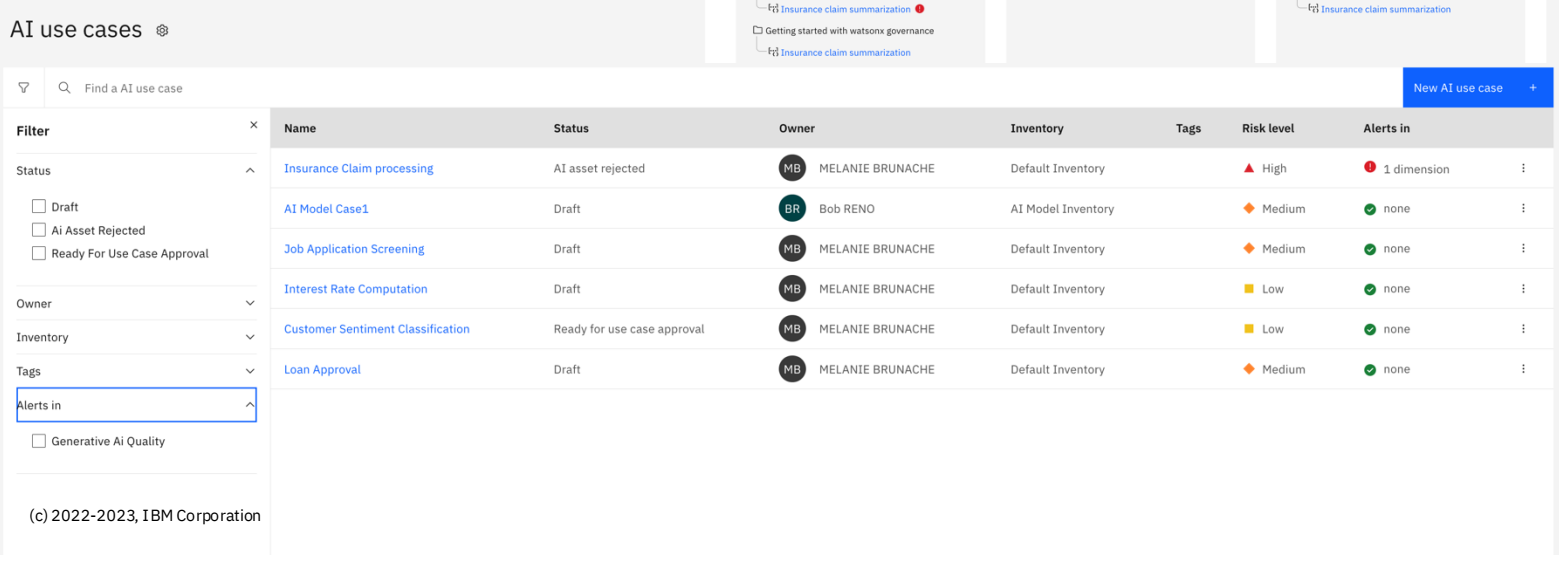
Lifecycle Governance

What is it?

Tracking the lifecycle of models from cradle to grave.

View factsheets for model assets that track lineage events and facilitate efficient ModelOps governance.





Why is it important?

Reduce manual efforts to document models.

Increase transparency of models.

Evaluation and Monitoring

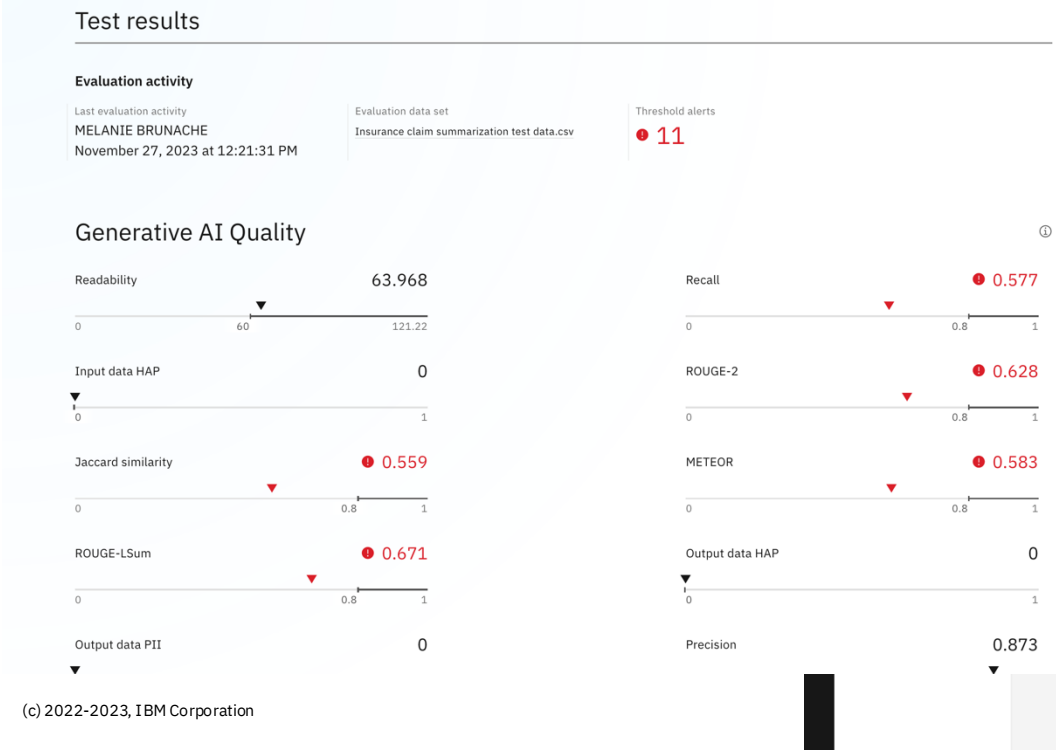
What is it?

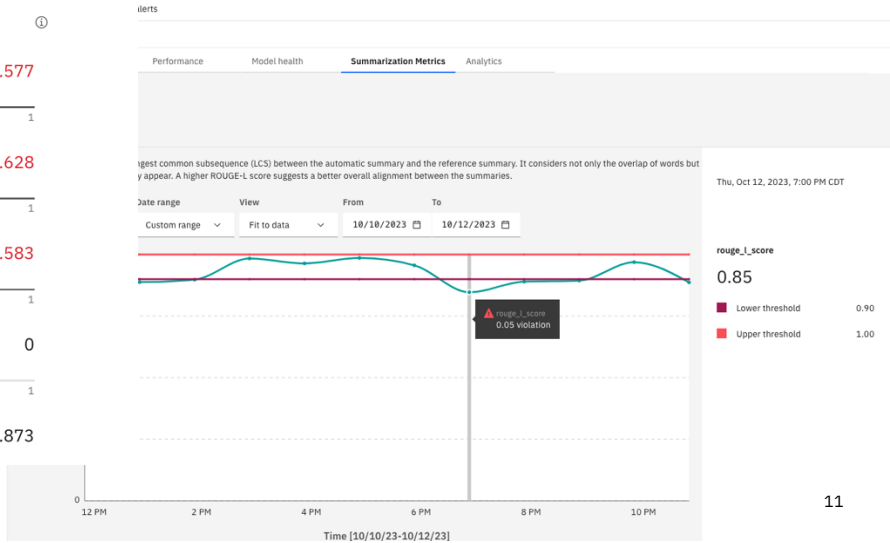
As data scientists are building their models, they can apply fairness and explainability techniques directly from within their Python notebook.

Why is it important?

Identify and mitigate data and model bias as early in the process as possible.

Explain complex models at a global level as well as a transactional level





Who are the users?

AI Governance Roles

Chief Risk Officer, Chief Data Officer and other C-level executives

Lead/Sponsor AIG Initiatives

Risk management team

Implement MRM framework, reporting

Data science team

Build, deploy, remediate

MLOps and enterprise engineering team

Validate, monitor

Learn more!

watsonx.goverance
click through demos

Turn phone
sideways if
accessing from iOS!

Try watsonx.governance
for free

IBM AI Risk Atlas

Watch the demo video

Securing your prompts from
Adversarial Attacks using
IBM watsonx.governance

Demo