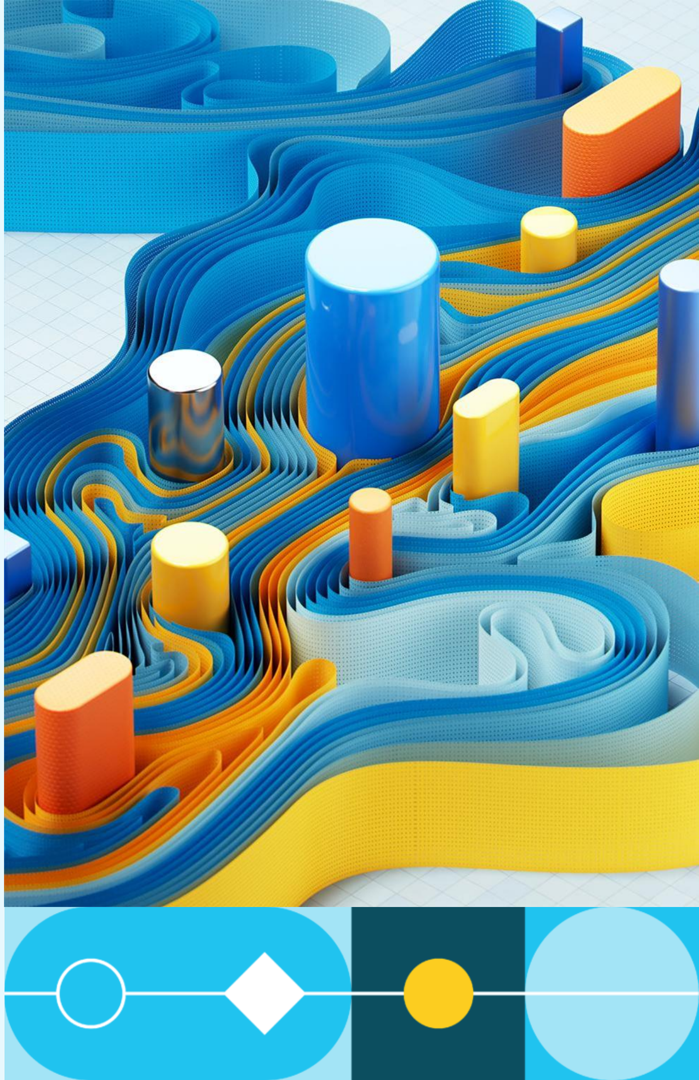
dataiku
The Universal AI Platform™

Governing Intelligence: When AI Acts on Its Own

Grant Case
Field Chief Data Officer
grant.case@dataiku.com
+61 0481259100



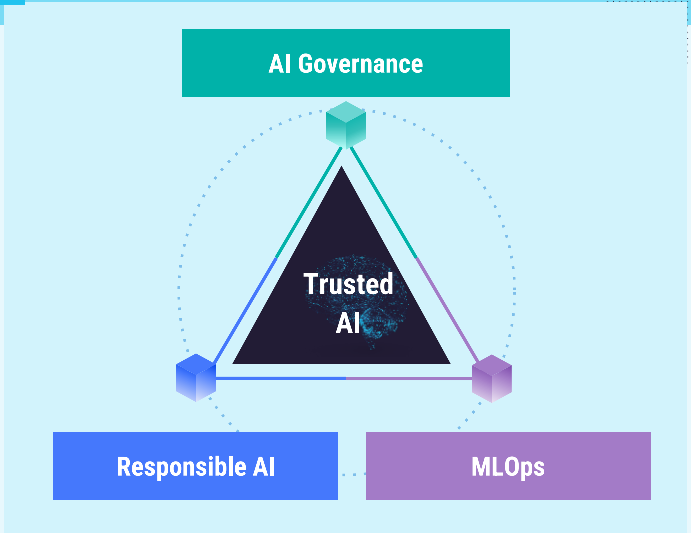
Situation: Your AI Just Placed a $2MM Supply Order

Navigating the new trolley problem:
When autonomous agents move physical inventory without human approval

The Three Lines of Defense for Autonomous AI

Orchestrates and Enforces processes that align AI initiatives with business, risk, and responsible AI objectives

Did we follow the framework we committed to?



Secures reliable, accountable, fair, transparent and explainable models and data pipelines

Does the AI behave according to our values?

Enables smooth and systematic operationalisation of data projects across stacks

Did we deploy what we said we would?

Autonomy in Action – What's Changed

From human-in-the-loop systems to autonomous agents

Traditional ML → Predict → Display → Human decides

Predict-only models

Human-in-the-loop by default

GenAI → Generate → Recommend → Human still drives

Can generate content

Still user-prompted

Agents → Recall memory, chain tools, execute tasks

Can decide + act with tools

Operate with memory, APIs, tools

What's changed? Why do they keep on changing!!!

- Traditional ML predicts whereas GenAI and Agents act.
- ⚙️ Shift: Static models → Tools + Memory → Emergent agents
- ⚠️ New Risks: Unsupervised tool execution, hallucinations, untraceable actions

> "The continuous development of Generative AI requires consistent principles — even as their implementation evolves." — *Dataiku, Trusted AI Framework*

Where Current Controls Break: The Governance Gaps

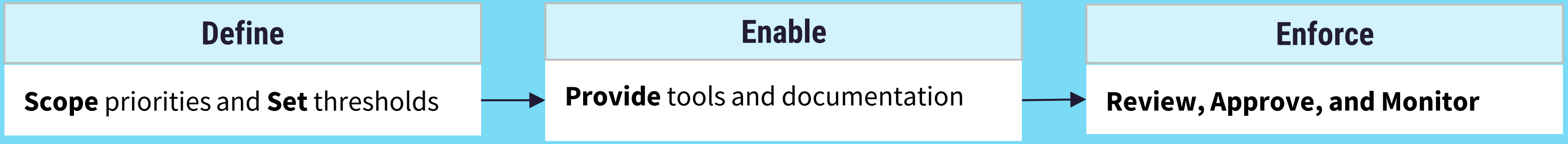
Autonomous AI creates a 'behaviour space' that traditional controls weren't designed to monitor.

Traditional Control	Why It Fails with Autonomous AI
CI/CD Pipelines	Doesn't cover agent loops or live API calls to external services
Model Review Boards	Evaluate initial models, but can't assess runtime decision chains
Explainability Tools	Cannot interpret emergent behaviors or explain action sequences
Model Cards	Fail to document how behavior evolves through agent iterations
Audit Logs	Track model calls but miss subsequent autonomous actions
Static Checklists	Often outdated for GenAI contexts; miss agent-specific risks
Risk Assessments	One-time evaluations that miss dynamic, evolving tool usage

> *"Orchestration must align AI with risk, ethics, value, and scaling strategies." — Dataiku*

A Blueprint for Controlling AI Autonomy

RAFT Principles for Governing AI that Acts Independently



Define

Scope priorities and **Set** thresholds

Enable

Provide tools and documentation

Enforce

Review, Approve, and Monitor

RAFT Principle	Autonomous AI Application
Reliable & Secure	• Monitor action chains, not just initial outputs • Track external API calls and resource usage
Accountable & Governed	• Clear ownership of autonomous decisions • Intervention points for human oversight
Fair & Human-Centric	• Prevent bias amplification in sequential decisions • Set boundaries on tool usage and permissions
Transparent & Explainable	• Log complete decision sequences • Explain why each action was taken

Use Cases

European Telecom: Implemented explainability requirements and fairness tests to maintain control over autonomous AI actions

Macquarie Bank: Leveraged governed data platforms within critical operations for regulatory compliance and operational efficiency.

***The Critical Shift**: From governing what models **ARE** to governing what autonomous AI **DOES**."*

Australian Imperative: Autonomous AI Action Plan

RAFT Principles for Governing AI that Acts Independently

- Australia's voluntary AI Ethics framework is evolving toward risk-based regulation
- Major Australian firms (Westpac, CBA, Macquarie) already self-regulating ahead of legislation
- Act Now Before Regulation
- CSIRO's RAIN network recommends proactive governance

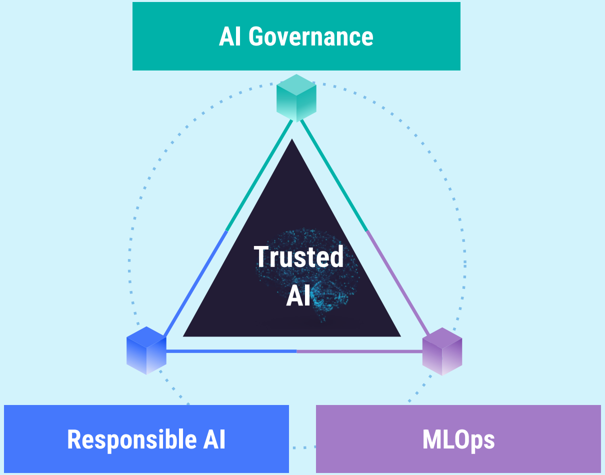
Role	Autonomous AI Governance Actions
CDO	• Establish agent action boundaries and permission controls (who can authorize what) • Form cross-functional council to oversee autonomous system behaviours • Map autonomous agents to the Australian AI Ethics Framework
CTO	• Implement agent activity logging and behavioural tracing beyond model monitoring • Create emergency shutdown/rollback mechanisms for autonomous systems • Develop alerting for unexpected autonomous actions outside defined guardrails

Role	Autonomous AI Governance Actions
Data Science	• Test agent behaviours with adversarial challenges before deployment • Define fallback behaviours and decision boundaries for autonomous agents • Implement explainability checks for decision chains, not just individual decisions
Data Products	• Document tool/API permissions by agent with clear human approval workflows • Set tiered financial authorization levels for autonomous actions (like supply orders) • Develop complete agent governance plans before launching autonomous features

Don't just govern models — govern autonomous behaviour

Don’t Just Govern Models – Govern Autonomous Behaviour

Preparing your organization for the autonomous AI future



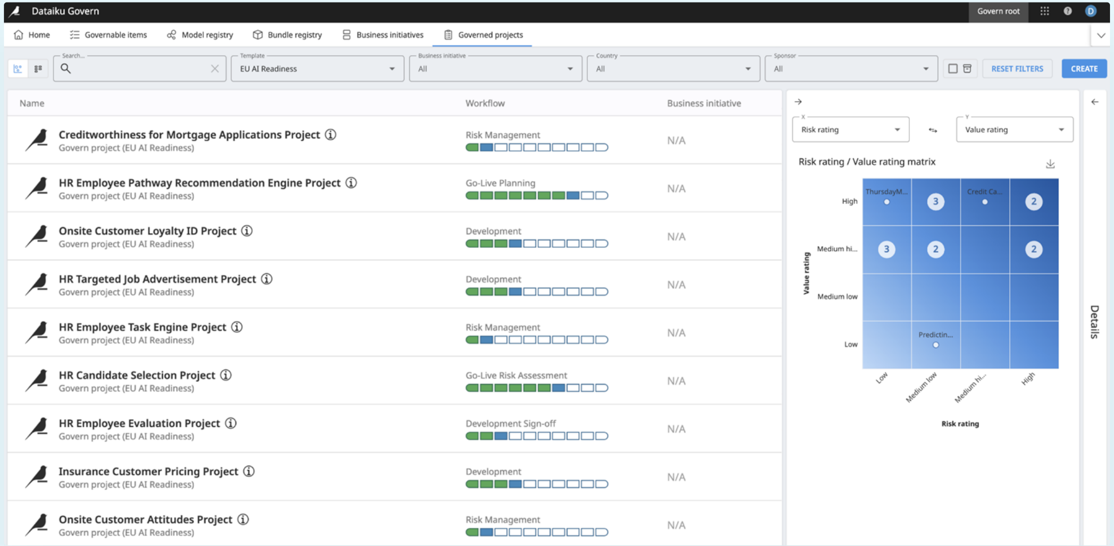
Key Takeaways

- → Autonomous AI creates a new "behavior space" that traditional controls weren't designed to monitor
- → Governance must shift from "what models are" to "what autonomous AI does" in real-time
- → Australian businesses implementing governance now will have competitive advantage as AI adoption grows
- → RAFT principles provide a practical framework for governing autonomous AI ahead of regulation
- → Dataiku provides the infrastructure to implement all three pillars: MLOps, AI Governance, and Responsible AI

Models don’t go off course. Their behaviour does.
If AI can act, it must be governed like an actor – not just an algorithm

dataiku
The Universal AI Platform™

Ready to take on AI Governance?



Search...

Template: EU AI Readiness

Business initiative: All

Country: All

Sponsor: All

X: Risk rating

Y: Value rating

Name	Workflow	Business initiative
Creditworthiness for Mortgage Applications Project Govern project (EU AI Readiness)	Risk Management	N/A
HR Employee Pathway Recommendation Engine Project Govern project (EU AI Readiness)	Go-Live Planning	N/A
Onsite Customer Loyalty ID Project Govern project (EU AI Readiness)	Development	N/A
HR Targeted Job Advertisement Project Govern project (EU AI Readiness)	Development	N/A
HR Employee Task Engine Project Govern project (EU AI Readiness)	Risk Management	N/A
HR Candidate Selection Project Govern project (EU AI Readiness)	Go-Live Risk Assessment	N/A
HR Employee Evaluation Project Govern project (EU AI Readiness)	Development Sign-off	N/A
Insurance Customer Pricing Project Govern project (EU AI Readiness)	Development	N/A
Onsite Customer Attitudes Project Govern project (EU AI Readiness)	Risk Management	N/A

Book a personalized Dataiku demo to see how Dataiku's Advanced Govern capabilities help manage autonomous AI

| dataiku.com | contact@dataiku.com

dataiku

The Universal AI Platform™

| dataiku.com | contact@dataiku.com