BUILDING AI GOVERNANCE FRAMEWORKS

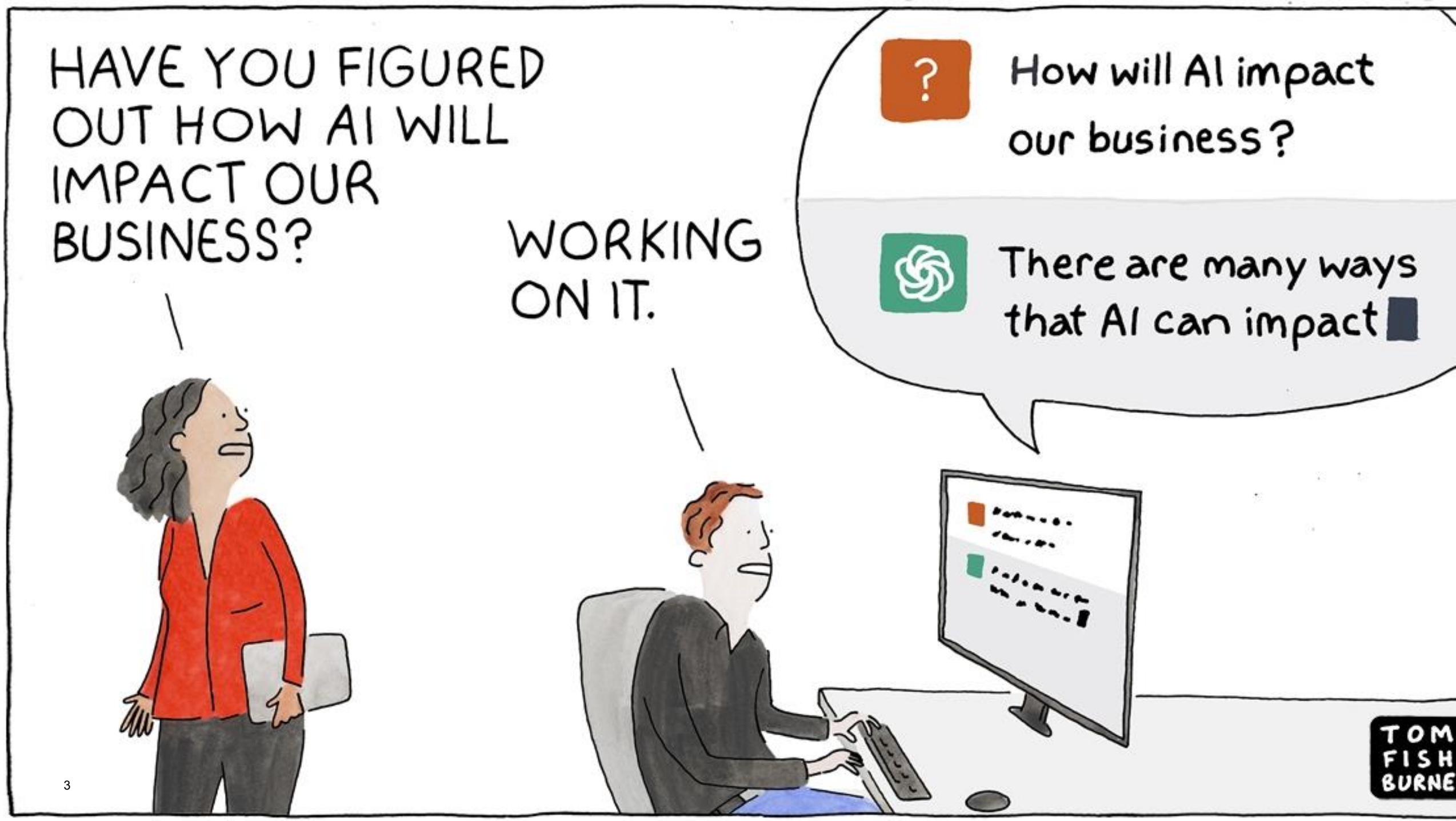
March 2026

protiviti®

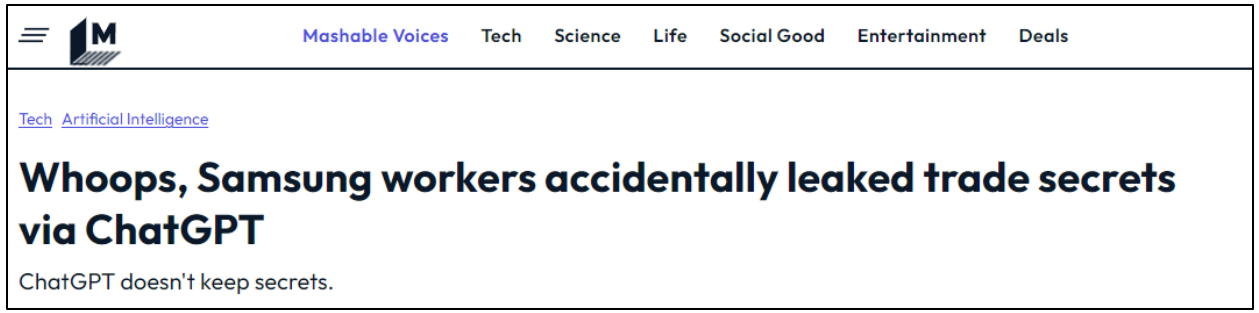
Global Business Consulting

Introduction

protiviti®


HAVE YOU FIGURED OUT HOW AI WILL IMPACT OUR BUSINESS?
WORKING ON IT.
How will AI impact our business?
There are many ways that AI can impact
TOM FISH BURNE

Examples of what can go wrong



protiviti

Building an AI Governance Framework

protiviti®

Leading Companies Approach AI Strategically, Grounded in Responsibility

Envision

- What are the best opportunities for AI and how do I start?
- What tools/technologies/ platforms should I use?
- How do we balance opportunity and risk?
- What AI is already in place/in use?

Establish

- Do I have appropriate data and is my environment ready?
- Is the use case technically feasible and risk profile commensurate?
- Do I have the right buy-in and are we ready for change?

Execute

- How do I best integrate into my existing technology stack?
- How do I prepare my data and technology environment for AI?
- How do I ensure Responsible AI principles are followed?

Evolve

- How can I continuously test, tune and adapt my AI?
- How will I monitor and protect my AI solutions and outcomes?
- How do I accelerate AI innovation without compromising risk management?

Enable

Responsible AI Principles

Target Operating Model and Governance Framework

What organizational structures and skillsets do I need for AI innovation and risk management?

What legal / regulatory / privacy / security / ethical concerns do I need to consider?

How will our AI governing bodies and functions interact with one another?

Flexible, Dynamic Resourcing Model

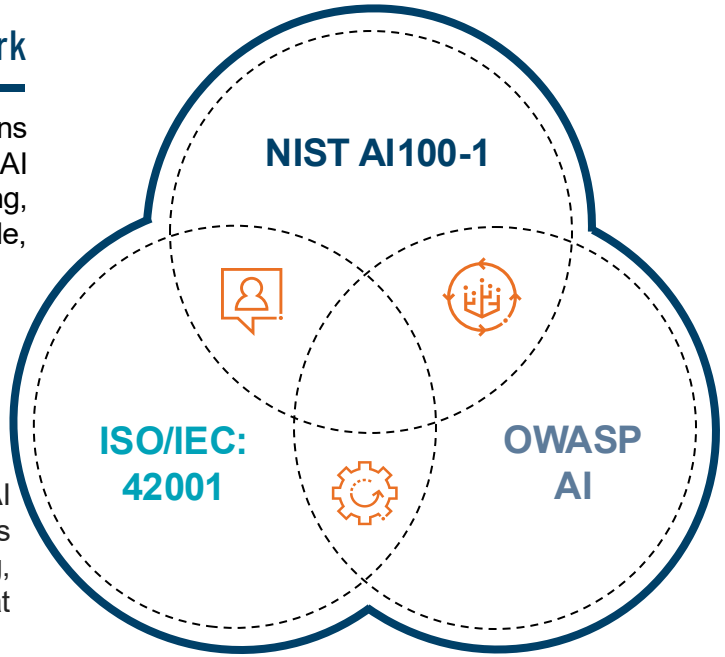
Recognized Emerging Standards for AI Governance

NIST AI Risk Management Framework

Voluntary guidance designed to help organizations manage the risks associated with AI systems. The NIST AI RMF provides a structured approach to identifying, assessing, and mitigating risks throughout the AI lifecycle, from design to deployment and even decommissioning.

ISO/IEC: 42001/2023

The first international standard for the governance of AI Management System (AIMS). This standard provides requirements for establishing, implementing, maintaining, and continually improving an AIMS within organizations that develop or use AI.



Open Web Application Security Project (OWASP)

OWASP's AI initiatives aim to empower organizations and individuals involved in AI development and deployment with actionable guidance and resources to ensure the secure and responsible use of AI. OWASP addresses key vulnerabilities and security concerns in AI applications, such as prompt injection, sensitive information disclosure, and supply chain vulnerabilities.

Predominant Themes

Risk Management	**Transparency & Explainability**	**Accountability & Governance**	**Metrics & Monitoring**
Trustworthiness & Ethical AI	**Data Privacy & Security**	**Fairness & Bias Mitigation**	**Resilience & Robustness**

protiviti

Key AI Governance Pillars

01

AI Governance Policies & Standards

Organizational AI policies that include guidance on the responsible and ethical use of AI, AI lifecycle management, and risk and compliance management are established.

02

AI Governance Steering Committee/Board

A cross-functional AI Steering Committee meets periodically to review existing and proposed AI deployments and provide input on responsible, ethical, and trustworthy use of AI within the organization.

03

Risk Management Framework

Processes, procedures, and practices are in place to determine the needed level of risk management activities based on the organization's risk tolerance.

04

AI Inventory

An inventory of AI systems used or owned by the organization (including key characteristics for each system) is established, updated as new solutions are developed/procured, and reviewed on a periodic basis.

05

Data Protection & Governance

Standards for the appropriate scope, use, source, protection, and format of data during the training, validation, testing, or use of an AI system are defined, enforced, and reviewed on a periodic basis.

06

Third-Party Risk Management

Processes have been established to assess and monitor Third Party AI vendors, including risks around data storage, data usage, data handling, and privacy provisions.

07

Regulatory Compliance

Legal and regulatory requirements are appropriately addressed in organizational policies and procedures. Any required modifications to policies and procedures are completed timely and communicated to affected stakeholders.

08

Metrics and Monitoring

Deployed AI models are continuously monitored to confirm expected functionality, identify potential issues, evaluate operational data or model outputs, and identify drift. Any unexpected outcomes or issues are reviewed, documented, and monitored through resolution.

09

Technical Guardrails & Controls

Technical guardrails and controls have been established and are applied throughout the AI lifecycle to ensure principles of responsible AI are embedded into relevant AI deployments.

10

Training & Awareness

The organization’s personnel and partners receive AI training on a periodic basis to verify understanding of related policies, procedures, and agreements.

Components of Enhanced AI Governance

These artifacts and tools form the foundation of your ongoing AI Governance function.

1 AI Governance Charter & Framework

Formalization of the AI Governance approach and providing high level guidance across the organization on how to interact with the AI Governance Committee.

2 AI Governance Operating Model

Defines how AI use cases will be governed by structuring its processes, activities, and decision-making workflows, integrating RACI information to clarify roles, ensure alignment and achieve efficiency.

3 Workflows – Approvals, Changes, Monitoring

Workflows designed to show flow through for approvals and other AI related tasks such as change mgmt. & monitoring.

4 AI Risk & Control Matrix

Standard set of AI Risks and Controls that should be leveraged when reviewing individual use cases. This informs on required controls for each use.

5 Governing Bodies RACI

Responsibility matrix showing key parties, required tasks, and overall organization on execution for AI Governance.

6 AI Glossary

Agreement on common terms to ensure alignment. Examples are as simple as what constitutes AI (e.g., ML, Generative, other).

7 AI Use Case Catalog

Capture of key characteristics of each use case, and ultimately the inventory of all AI use cases with details on ownership, approvals and other key facts.

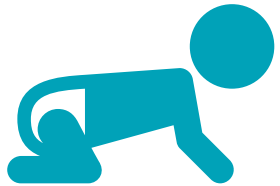
8 AI Use Case Questionnaires

Used to collect characteristics for different AI Use Cases and helps to collect individual for routing of use cases / groups involved in approval.

Common AI Observations

Observation	Impact	Recommendations
Lack of standardized AI definitions	• Misalignment of AI projects with business objectives • Over characterization of AI use cases	• Create an AI governance steering committee with cross-functional representation • Conduct workshops to define AI for the organization
Tension between innovation and risk	• Conflicting priorities among departments leads to missed opportunities for competitive advantage	• Develop an AI strategy document that align with business and technology strategies • Develop Key Performance Indicators (KPI) and Key Risk Indicators (KRI) to monitor the design and implementation lifecycle of AI applications
Non-alignment with traditional risk frameworks (i.e., Model Risk Management v NIST AI)	• Ineffective management of risks associated with AI applications • Bifurcation of model & AI policies	• Update risk management policies to include AI-specific risks • Conduct regular model validation and bias assessments • Create unified or mapped MRM & AI Risk Framework
Uncertainty on aligning AI practices globally or locally	• Inconsistent AI practices and difficulty complying across different regions • Unequal treatment of acquisitions	• Develop global AI policy and standards to align with ethical principles and legal requirements • Allow regional teams to customize policy and standards based on local regulations
Shadow AI currently being used	• Deployment of noncompliant AI applications that are biased or unfit for use	• Create intake and approval process for all AI initiatives • Develop training to help employees understand AI technologies, associated implications, and best practices
Lack of centralized tracking for AI use cases	• Challenges governing and managing applications • Unable to report to regulators or third parties	• Deploy a structured tool to inventory AI applications and use cases • Mandate intake and approval process for all AI initiatives (including initiatives with third parties)
Conflating perceived high risk with actual legal risk	• Increased administrative burden and reduced efficiencies	• Establish AI risk tolerance with risk considerations attributed to all phases of the AI lifecycle • Develop and leverage risk control matrix to differentiate between perceived and actual risks
Transparency & Explainability	• Opaque black box decision making challenge regulators, consumers and companies in meeting explainability and transparency obligations	• Develop Process and outcome-based rationales for input and output variables, to better understand roles of components and impact on decisioning to enable explanations for various parties

Crawl-Walk-Run Approach to AI Governance & Risk Management



CRAWL

Establish visibility and basic governance

- Form **AI Governance Committee** and define roles
- Create **AI Acceptable Use Policy**
- Build **AI Use Case Inventory**
- Implement **basic risk assessment template**
- Add **AI clauses in vendor contracts**
- Launch **AI literacy training**
- Set up **simple reporting and incident escalation**

WALK

Formalize processes and expand controls

- Develop **AI Risk Management Framework**
- Introduce **quantitative risk scoring** and TEVV (Testing, Evaluation, Verification, Validation) practices
- Establish **data governance standards** (lineage, consent, quality)
- Implement **vendor due diligence** and monitoring
- Create **AI risk dashboards** for leadership
- Begin **internal audits** of AI systems
- Expand training to include **ethics and compliance**

RUN

Optimize and embed AI risk management

- Establish **AI Risk Management Working Group**
- Integrate **AI risk controls** into enterprise risk management (ERM)
- Deploy **automated monitoring and alerting tools** for AI performance and compliance
- Implement **continuous TEVV** and bias detection
- Formalize **regulatory compliance** program
- Establish **advanced vendor assurance** programs
- Foster **culture of responsible AI** innovation

The Future of Governance Operating Models – Use AI, Govern AI, Enable People

As organizations fully integrate AI/ML technologies, what does operations look like when the highest-value predictive, generative, and agentic use cases are fully deployed?

Emphasis on developing a comprehensive operating model that enables governance functions within organizations to navigate these rapid advancements by preparing their workforce, establishing the necessary technical infrastructure, and systematically identifying and deploying high-impact AI/ML use cases.

1

Accelerating AI/ML-Enabled Transformation

The objective is to help enterprises rapidly transform their operating model to seamlessly embed AI/ML capabilities into daily operations, ensuring long-term adaptability and resilience.

Key functions such as **technology, legal, compliance, risk, audit, data governance, privacy, and security** play a crucial role in AI adoption. These functions can either act as **accelerators** or **bottlenecks** to AI deployment.

Provide the fastest path to transforming these governance functions into enablers of AI adoption, ensuring enterprises can scale AI initiatives with confidence.

2

Leveraging a Scalable and Responsible AI Governance Framework

Approach focuses on developing a **framework, methodology, and execution strategy** that enables organizations to:

- **Rapidly transform governance functions** to facilitate AI/ML adoption with precision and agility.
- **Establish a robust AI use case pipeline**, ensuring comprehensive oversight while tailoring governance to the **risk profile of each use case**.
- **Balance speed and rigor**, allowing organizations to deploy AI at scale without compromising **compliance, security, and ethical standards**.

3

Empowering the Workforce for AI Adoption

- Beyond governance, **people enablement** is essential for AI success. **This transformation strategy** ensures that organizations **equip their workforce** with the necessary skills, tools, and cultural mindset to embrace **AI-driven operations** fully.
- By integrating **AI/ML into governance and workforce enablement**, organizations can move from **experimentation to enterprise-wide adoption at unmatched speed and scale.**

Source(s): UpGrad

Deploying Responsible AI

protiviti®

AI Governance Lifecycle Framework

Below outlines an illustrative AI Governance Lifecycle / Framework with underlying risks, controls, and supporting Governance elements (People, Process, Technology). This is an illustrative framework and is not intended to be exhaustive, but rather to be right-sized and form fit for your organization.

1	2	3	4	5	6	7
Business Need & Context	**Data Understanding & Preparation**	**Solution Design & Development**	**Testing Evaluation Validation & Verification**	**Integration & Deployment**	**Monitoring & Optimization**	**Contingency Plans & End of Life Management**

Status Reporting & Escalation when behavior drifts from plan

Communicate & Educate the strengths & weaknesses of the AI in question, understanding technical details, incorporating controls into AI design, credible challenge

Holistic | **Multidimensional** | **Systematic** | **Actionable**

Identifying Value & Evaluating Readiness of AI

Value Identification

Where should we apply AI?

Key Focus:

- Do we have a strong understanding of the strengths of AI?
- What areas of the business offer the greatest potential?
- What job roles & How many? Cost Basis?
- What levers of value will we pull? (productivity improvement, etc)? How will we measure and how fast could we deliver results?

Data

Do we have the data and is it ready?

Key Focus:

- What data will be needed?
- Where is it located?
- Is it organized for usage and scalable?
- What governance/security procedures do we have in place?
- Is that data appropriate for use by AI (e.g., considering privacy, security, quality, bias)

Skills

Do we have the talent we need to unlock the value?

Key Focus:

- What AI skills do we have in our organization?
- Where are they and how many?
- Is there alignment between our skills and our identified value hypothesis?
- What is our training/dev plan?

Ecosystem

Where to build/buy/partner?

Key Focus:

- How closely does our value hypothesis align to existing or emerging solutions?
- Are we actively engaged with the AI ecosystem? Big & small?
- Do we have a framework on how to evaluate build, buy or partner decisions?

Experimentation

Do we have a capacity to test & learn quickly?

Key Focus:

- Do we have a technology environment available?
- Do we have a Rapid Prototype team/approach?
- Are we leveraging best practices in Design Thinking / Agile / Lean / Innovation?

Change Management

Do we have an organized plan to execute?

Key Focus:

- What is the Change Readiness of our Org?
- What is most important to achieve buy in?
- How do we communicate our plan?
- How will we collect, evaluate, and act on feedback?

Understanding AI: Approaches to Building

	Out-of-the-Box (OOTB) Services and Solutions	Hybrid Model	Custom Model
	Leverage pre-built models without influencing training or integrating data	Start with pre-built, extensible models and solutions, and perform enterprise-specific training cycles leveraging relevant data	Build an AI-ML model from the ground up, including model type selection and design, fully custom training cycles and data
Pros	Cost-effective, quick ramp-up, little to no specialized skills required	Optimizes training cycle efficiencies, incorporates business-specific terminology and/or knowledge, medium ramp-up time	Fully customized solution optimized for an individual use case, built enterprise-specific, full control of data lineage and storage
Cons	Generic responses and capabilities, cannot incorporate business- specific terminology or knowledge	May require specialized skillsets, training data may be sent to various 3rd party data centers	Longest timeline to deploy, requires specialized data science skillsets, requires most maintenance
Best Suited for	General domains such as common object detection, speech to text, FAQ bots, etc.	Most enterprise-ready applications, such as conversational AI, predictive analytics, nuanced object detection	Highly regulated and/or sensitive data, unique use cases, specialized enterprise requirements

Prompt Guardrails

Prompt guardrails are mechanisms used to control and guide the behavior of LLMs and Generative AI systems. They are essential to mitigate risks and ensure that these models generate outputs that align with ethical, safety, and content guidelines. Below are details on what prompt guardrails are, specific examples, and how they address risks:

1 Ethical & Responsible Content Guidelines

- These guardrails set limits on generating content that may be offensive, discriminatory, or harmful.
- **Examples:** Prohibiting hate speech, offensive language, and discriminatory content.
- **Risk Mitigation:** Ensures that the AI model does not generate content that promotes harm, discrimination, or unethical behavior.

2 Fact-Checking & Misinformation Prevention

- These guardrails focus on verifying information and preventing the spread of false or misleading content.
- **Examples:** Ensuring that the AI does not generate false medical information, conspiracy theories, or misinformation.
- **Risk Mitigation:** Reduces the likelihood of LLMs generating inaccurate or harmful information.

3 Privacy & Confidentiality

- These guardrails are to protect personal or sensitive information from being generated or leaked.
- **Examples:** Preventing the model from sharing personal data or confidential information.
- **Risk Mitigation:** Safeguards against privacy breaches and unauthorized disclosure of sensitive data.

4 Compliance with Laws & Regulations

- These guardrails ensure that AI-generated content adheres to legal and regulatory requirements.
- **Examples:** Avoiding content that promotes illegal activities or violates copyright laws.
- **Risk Mitigation:** Helps organizations and developers avoid legal liabilities associated with AI-generated content.

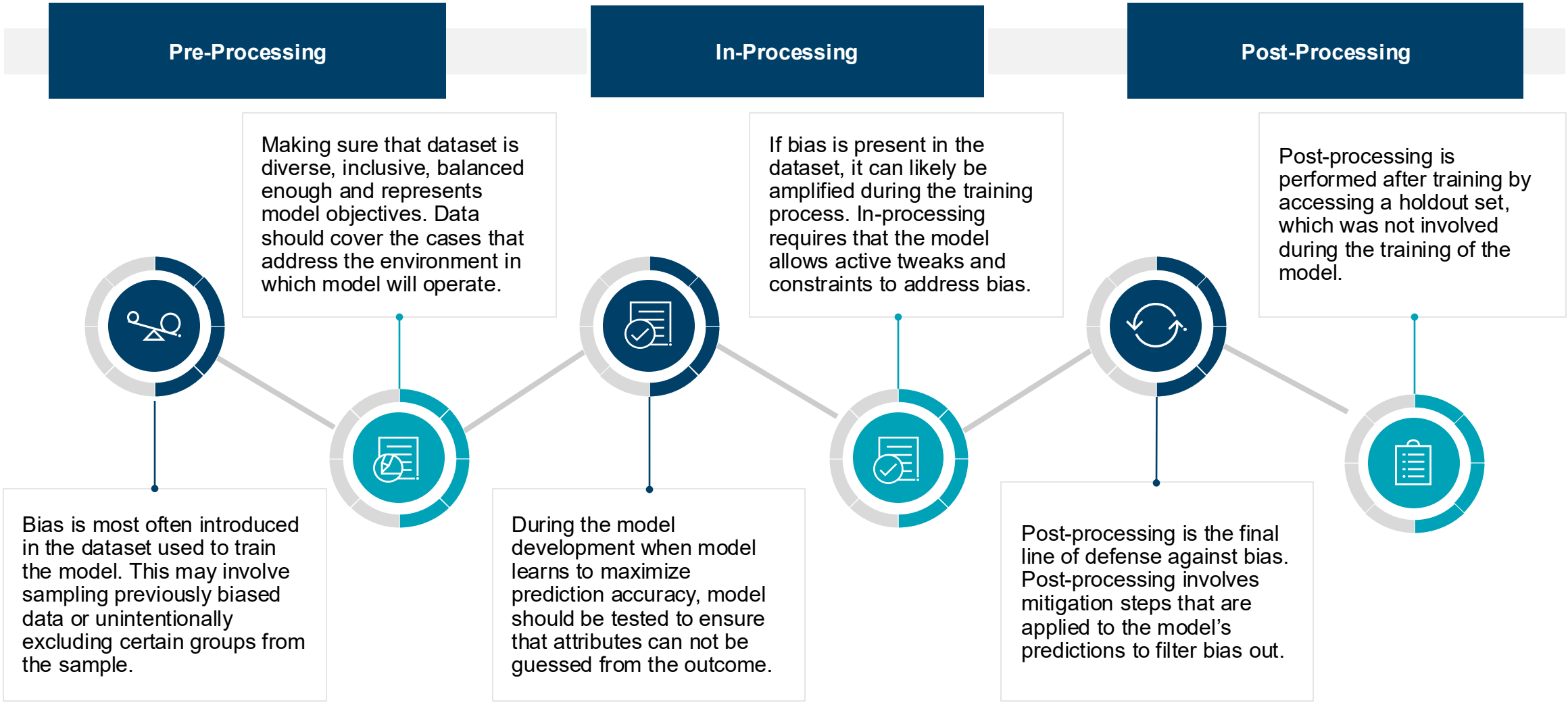
5 Bias & Fairness

- These guardrails aim to reduce bias in AI-generated content and promote fairness.
- **Examples:** Implementing techniques to debias responses, avoid stereotypes, and ensure equitable outputs.
- **Risk Mitigation:** Addresses concerns related to AI-generated content that may perpetuate biases or inequalities.

6 Safety & Harm Prevention

- These guardrails are designed to prevent the generation of content that encourages self-harm, violence, or illegal activities.
- **Examples:** Ensuring the model does not provide instructions for harmful actions or engage in dangerous behavior.
- **Risk Mitigation:** Protect users and the public from potentially harmful content generated by LLMs.

Mitigation of Bias



protiviti

Data Governance for Effective AI Governance

Data governance provides the necessary infrastructure and guidelines for effective data management and ensures that these data practices align with ethical standards, legal requirements, and societal expectations.

Data Lifecycle

- Planning to collection to deletion
- Data governance provides the foundation upon which AI governance can be built.

Responsible AI Systems

- Data Availability
- Data Quality and Integrity
- Data Standards and Interoperability
- Representative Data

Inherent Risks

- Regulatory Compliance
- Risk Management

Ethical AI Systems

- Stakeholder Engagement
- Consent
- Dispute Resolution

Holistic Approach

- Technology Agnostic
- Common Foundation

Implementation and Standardization

- Standards on use and interpretation
- Due Diligence
- Applications to AI

Questions

protiviti®