



Best Practice Tools

Examples supporting responsible AI maturity

September 2024

Guide to Strengthening Responsible AI Practices: Tools and Best Practices

This guide offers a carefully curated selection of tools and best practices designed to help companies enhance their Responsible AI (RAI) capabilities. Developed by leading mobile operators, these resources focus on the most critical and impactful areas within the GSMA RAI Maturity Roadmap.

While covering a subset of topics, these do span across all five key dimensions of the Responsible AI Roadmap: Vision, Operating Model, Technical Controls, Third-party Ecosystem, and Change Management and Communications. Each example has been chosen for its practical value in elevating your organisation's AI practices to the highest level of maturity.

This document is intended to complement the Responsible AI Roadmap, providing concrete examples and actionable insights for companies along their AI maturity journey. It forms part of a selection of documents that will enable organisations to better understand and implement responsible AI practices. Please also see [The GSMA Responsible AI Maturity Roadmap](#) and [Step-by-Step Guide](#).



Supporting Tools for Advanced Level of Dimension 1: **Vision**

Sub-Dimension 1.1: **RAI principles**

Best-practice RAI principles include:



FAIRNESS

Ensure fairness by preventing discrimination against individuals or groups, thereby avoiding adverse decisions or inferences



TRANSPARENCY AND EXPLAINABILITY

Be transparent about when an AI system is being used, what kind of data it uses, and its purpose, ensuring explainability



HUMAN AGENCY AND OVERSIGHT

Determine an appropriate level of human oversight and control of an AI system to prevent over-reliance



ACCOUNTABILITY

Establish a governance structure that clarifies responsibility for reporting and decision-making throughout the AI lifecycle, ensuring accountability



PRIVACY AND SECURITY

Respect and uphold an individual's right to privacy by ensuring personal data is protected and secure



ENVIRONMENTAL IMPACT

Design, develop, and deploy AI systems mindfully of their environmental impact throughout their lifecycle and value chain



SAFETY AND ROBUSTNESS

Ensure AI systems are safe, robust, and reliably operated in accordance with their intended purpose throughout their lifecycle

A mobile operator that has reached advanced-level responsible AI Maturity under sub-dimension 1.1 has:



Published RAI principles that are **embedded within their culture**



Consistently applied these principles across all operations through a **governance model**



Reinforced the principles through **organisation-wide training**



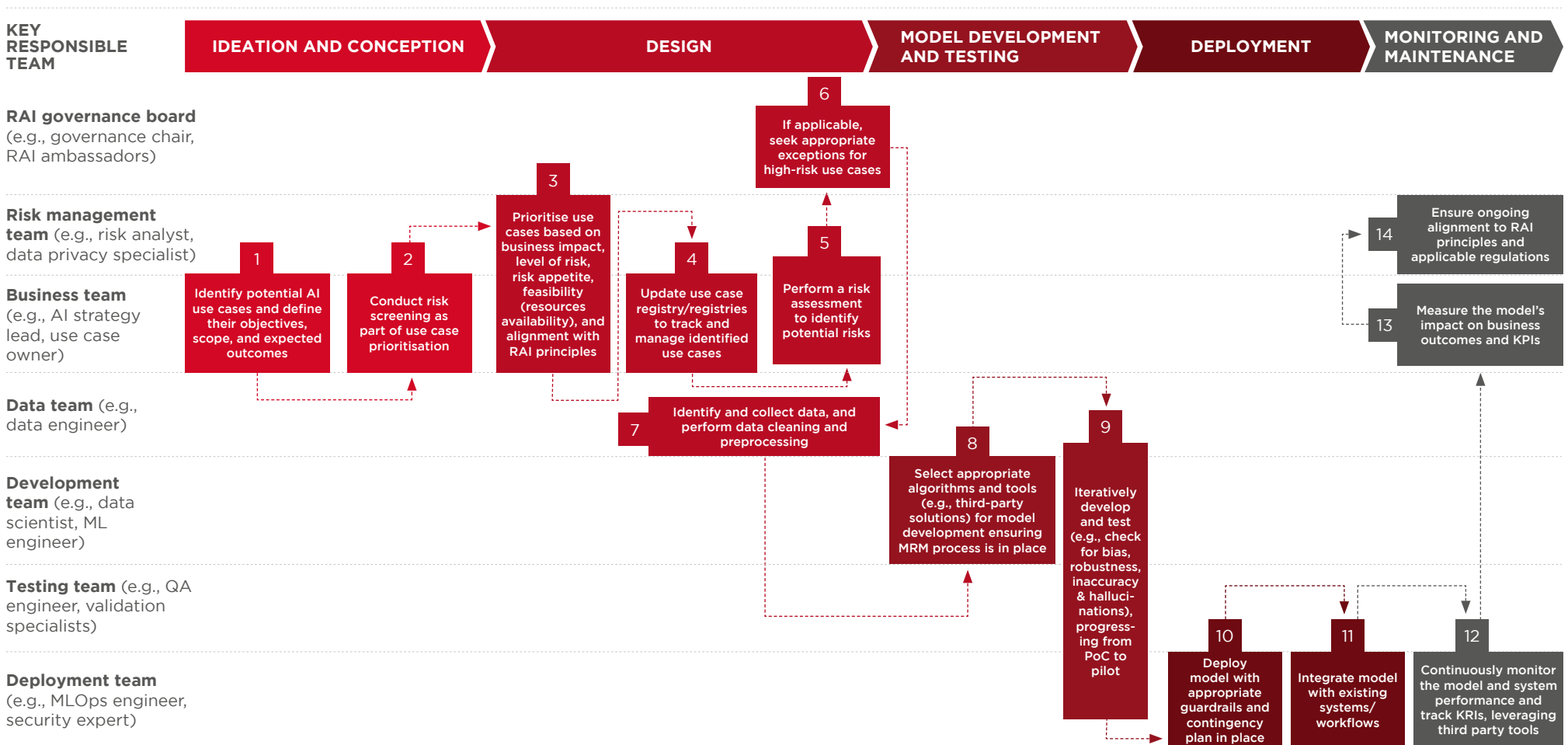
Conducted regular assessments, including questionnaires to evaluate use cases and updates to RAI principles, to **ensure compliance by third parties**



Supporting Tools for Advanced Level of Dimension 2: **Operating Model**

Sub-Dimension 2.2: Processes for identifying, assessing, and mitigating AI risks

Illustrative example of a risk management process along a use case lifecycle



Sub-Dimension 2.3: **Roles and responsibilities**

Every employee's role¹ in responsible usage of AI is to:



Adhere to policies and guidelines for the responsible use of AI within the organisation



Report any observed concerns in AI systems through designated channels



Refrain from using confidential data (personal or organisational) when interacting with AI systems



Participate in training programmes that promote RAI practices and enhance understanding of RAI usage





1. The role is not equivalent to an FTE; instead, one person can hold multiple roles

The below tables provide a comprehensive example of the specific RAI roles and responsibilities that can be allocated and created across an organisation.

TEAM	ROLE ¹	RAI-SPECIFIC RESPONSIBILITIES (NON EXHAUSTIVE)
EXECUTIVE LEADERSHIP	Across CxO (incl. CEO)	<ul style="list-style-type: none"> • Set the organisation's AI ambition and vision for RAI practices • Champion drafting of RAI principles and ensure alignment with organisational values
	Chief technology officer (CTO)	<ul style="list-style-type: none"> • Establish and oversee the technical parameters for RAI (e.g., safety benchmarks, interpretability standards), ensuring their effective operationalisation • Manage the procurement process and evaluate third parties to ensure that all AI technologies and partners meet the organisation's standards for RAI
	Chief risk officer (CRO)	<ul style="list-style-type: none"> • Ensure that AI initiatives adhere to established legal and RAI standards and guidelines
	Chief data officer (CDO)	<ul style="list-style-type: none"> • Establish data governance policies that prioritise RAI principles (e.g., privacy, security) • Ensure that data used in AI systems is responsibly sourced and managed
	Chief AI officer (CAIO)	<ul style="list-style-type: none"> • Develop and enforce AI governance frameworks that promote RAI practices • Provide strategic guidance on integrating RAI into AI development and deployment processes
	Chief information security officer (CISO)	<ul style="list-style-type: none"> • Enforce security protocols to protect AI systems from threats • Continuously monitor AI systems for potential security vulnerabilities, implementing measures to mitigate risks associated with AI deployment and usage
	Chief finance officer (CFO)	<ul style="list-style-type: none"> • Ensure appropriate funding for the development and implementation of RAI practices, including investments in necessary technologies, training, and compliance measures • Monitor and evaluate the financial impact of AI projects, ensuring they deliver sustainable value while managing risks associated with AI investments
	Chief human resource officer (CHRO)	<ul style="list-style-type: none"> • Organise and oversee training programmes to educate employees about the safe use of AI technologies, fostering a culture of responsibility and accountability in the adoption of AI • Ensure use of RAI practices in HR processes (e.g., recruitment, performance evaluation)

1. The role is not equivalent to an FTE; instead, one person can hold multiple roles

 **Net new role for RAI**

TEAM	ROLE ¹	RAI-SPECIFIC RESPONSIBILITIES (NON EXHAUSTIVE)
RAI GOVERNANCE BOARD²	 Governance chair (e.g., chief RAI officer)	<ul style="list-style-type: none"> • Lead the board in setting RAI standards and policies for AI initiatives, ensuring they are embedded throughout the organisation • Ensure adherence of any exceptions granted for high-risk use cases to the organisation's RAI principles
	 RAI ambassadors (i.e., representatives from various functions/domains of the organisation)	<ul style="list-style-type: none"> • Review AI use cases and perform in-depth evaluation of high-risk use cases • Offer domain-specific knowledge (e.g., product, tech, data) to assess implications and risks in AI projects ensuring appropriate use case prioritisation and alignment with CSR objectives • Provide input on how AI technologies can be used responsibly and serve as POC for RAI usage within their domain
	 RAI legal advisor/general counsel/chief legal officer	<ul style="list-style-type: none"> • Ensure that AI initiatives comply with both relevant local and global laws and regulations • Advise on legal risks and liabilities associated with AI deployment
	 AI ethics officer	<ul style="list-style-type: none"> • Ensure the organisation adheres to the RAI principles specifically focusing on promoting the well-being and interests of individuals affected by the AI technologies (e.g., users, employees and impacted communities) • Evaluate and balance the risks and benefits of AI use cases, and make recommendations on buy/build decisions for developing AI solutions
RISK MANAGEMENT TEAM	Risk analyst	<ul style="list-style-type: none"> • Conduct comprehensive risk assessments to assess and mitigate all types of risks associated with AI projects, prioritising use cases based on identified risks
	Data privacy specialist	<ul style="list-style-type: none"> • Implement data privacy and protection measures to protect sensitive information in AI systems • Ensure that data handling practices align with privacy regulations and best practices


1. The role is not equivalent to an FTE; instead, one person can hold multiple roles

2. The RAI Governance Board is an in-house team responsible for overseeing comprehensive RAI governance and compliance across the organisation

TEAM	ROLE ¹	RAI-SPECIFIC RESPONSIBILITIES (NON EXHAUSTIVE)
BUSINESS TEAM	AI strategy lead	<ul style="list-style-type: none"> • Prioritise AI use cases that align with RAI principles • Define strategies to measure the impact of AI initiatives on business outcomes
	Use case owner	<ul style="list-style-type: none"> • Ensure that AI use cases align with RAI principles and organisational values • Input use cases into registry/registries and complete required information, including risk screening • Assess implications and risks associated with specific use cases and assist in prioritising use cases based on impact and alignment with RAI principles • Ensure documentation is maintained for all aspects related to the use cases (e.g., model and data information)
DATA TEAM	Data engineer	<ul style="list-style-type: none"> • Implement data management practices that prioritise RAI principles (e.g., transparency, privacy, fairness) • Ensure that data pipelines are designed to mitigate bias and protect against unintended consequences
DEVELOPMENT TEAM	AI architect	<ul style="list-style-type: none"> • Design AI systems with built-in mechanisms for RAI principles (e.g., human agency and oversight, explainability, security)
	Data scientist	<ul style="list-style-type: none"> • Perform feature engineering that enhances model performance while avoiding risk concerns (e.g., fairness, transparency) • Conduct evaluations to understand how models perform across different demographic groups to ensure fairness
	Machine learning engineer	<ul style="list-style-type: none"> • Develop AI algorithms and models that prioritise RAI principles • Design systems to monitor models for drift and biases post deployment

1. The role is not equivalent to an FTE; instead, one person can hold multiple roles

 Net new role for RAI

TEAM	ROLE ¹	RAI-SPECIFIC RESPONSIBILITIES (NON EXHAUSTIVE)
TESTING TEAM	Quality assurance (QA) engineer	<ul style="list-style-type: none">• Test AI systems to ensure that they meet regulatory requirements• Identify and report potential biases, errors, and unintended consequences in AI models
	Validation specialist	<ul style="list-style-type: none">• Validate AI models (e.g., LLM models) against performance criteria to ensure their reliability and trustworthiness, and evaluate the fairness and transparency of AI systems through rigorous testing
DEPLOYMENT TEAM	MLOps engineer	<ul style="list-style-type: none">• Implement automated deployment pipelines that incorporate RAI practices• Ensure that AI models are deployed securely and with appropriate safeguards and feedback loops in place
	 Security expert	<ul style="list-style-type: none">• Implement security measures to protect AI systems from malicious attacks and unauthorised access, ensuring AI deployments adhere to security standards and best practices

1. The role is not equivalent to an FTE; instead, one person can hold multiple roles



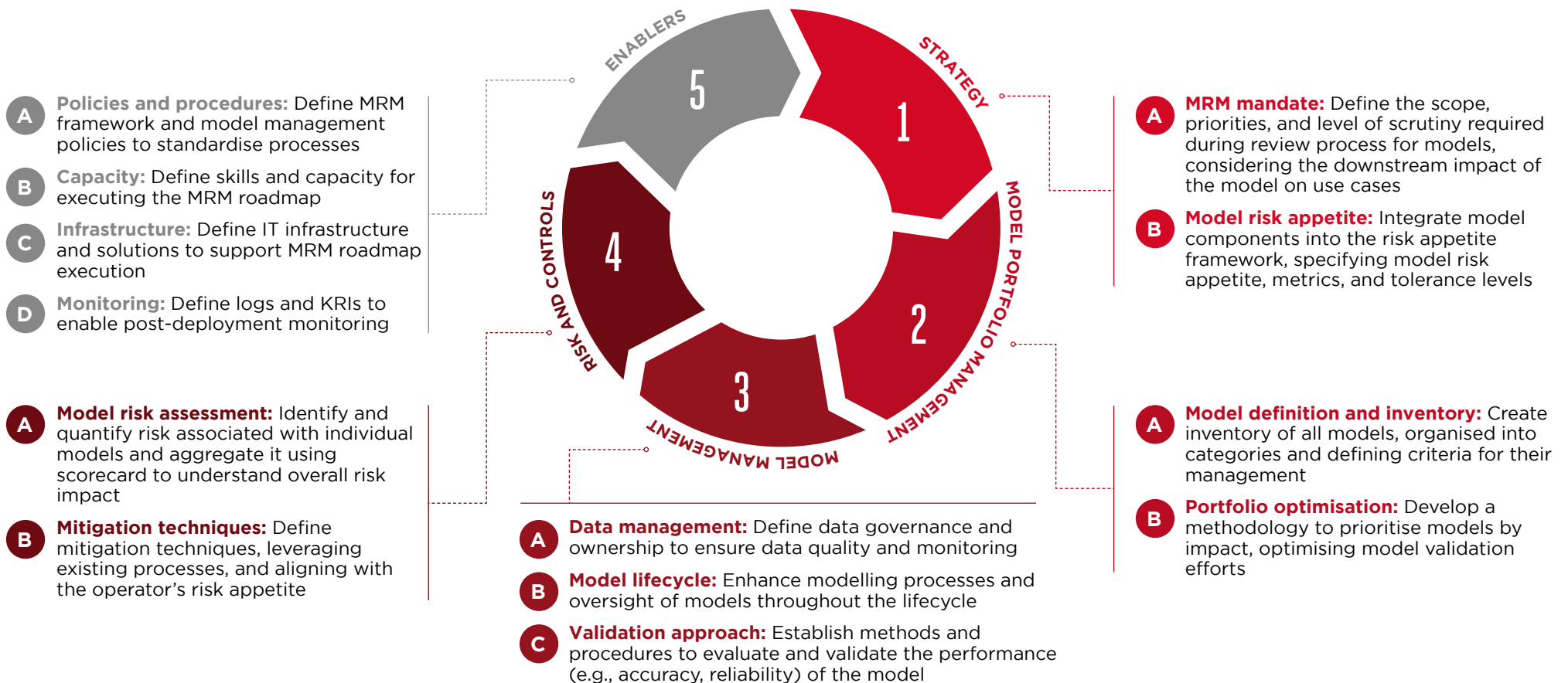
Supporting Tools for Advanced Level of Dimension 3: **Technical Controls**

Sub-Dimension 3.2: **Model Risk management (MRM)**

MRM practices should have the following components to effectively manage the AI model risks

NON-EXHAUSTIVE

MRM IS THE PROCESS OF IDENTIFYING, ASSESSING, MITIGATING, AND MONITORING RISKS ASSOCIATED WITH THE USE OF AI MODELS, MINIMISING POTENTIAL ERRORS OR RISKS







Sub-Dimension 3.3: **Control environment** (incl. technical guardrails)

Examples of technical controls and guardrails that can be deployed for each AI risk type

 **Net new role for GenAI**




NON-EXHAUSTIVE

AI RISKS	CONTROL DESCRIPTION	SPECIFIC EXAMPLES OF CONTROLS (NON EXHAUSTIVE)
DATA PRIVACY	Implement measures to protect sensitive data	<ul style="list-style-type: none"> • User access controls to appropriately restrict model and data access • Run use cases on private instances (e.g., locally hosted LLMs)
	Perform fairness and bias testing	 <ul style="list-style-type: none"> • Fairness testing tools to identify bias (e.g., Fairlearn, AI Fairness 360, VerifyML)
BIAS AND DISCRIMINATION	Ensure data used for training models is diverse and representative	<ul style="list-style-type: none"> • Data anonymisation • Collect data from wide range of sources
	Ensure diverse representation within teams that develop, deploy and audit AI systems	<ul style="list-style-type: none"> • Diversity and inclusion in development teams
	Leverage reporting mechanisms for users to escalate biased or discriminatory AI output	<ul style="list-style-type: none"> • Whistleblowing channels
IP INFRINGEMENT	Create regulatory compliance checklist	<ul style="list-style-type: none"> • Regulatory (GDPR, HIPAA) compliance checklist and assessments
INACCURATE OUTPUTS, LACK OF EXPLAINABILITY AND ACCOUNTABILITY	Create checkpoints to ensure content is accurate	<ul style="list-style-type: none"> • Human-in-the-loop review
	Generate use-case specific outputs	 <ul style="list-style-type: none"> • Context engineering through feeding use-case specific info  <ul style="list-style-type: none"> • Fine-tune model through Reinforcement learning from human feedback (RLHF) • Hyperparameter tuning by changing temperature and top values • Benchmark performance through KPIs, industry tools, and in-house models
	Enrich prompts to improve accuracy of responses	 <ul style="list-style-type: none"> • Automatic prompt engineering (i.e., proposes list of applicable prompts) • Prompt optimisation tools, prompt response

Examples of technical controls and guardrails that can be deployed for each AI risk type

 **Net new role for GenAI**

NON-EXHAUSTIVE

AI RISKS	CONTROL DESCRIPTION	SPECIFIC EXAMPLES OF CONTROLS (NON EXHAUSTIVE)	
INCORRECT/ MALICIOUS USE	Validate data and prompts in real time to safeguard from malicious inputs		<ul style="list-style-type: none"> • Prompt evaluation tools, prompt input allow/deny lists • Data quality checks (e.g., source citation)
	Perform adversarial testing to filter questionable responses		<ul style="list-style-type: none"> • Prompt injection/toxicity detectors • Model hardening using adversarial examples
SECURITY THREATS	Deploy guardrails to protect enhanced data security features		<ul style="list-style-type: none"> • Throttle incoming prompts to minimise denial of service/other security risks • Adapt existing security risk assessment framework for new attack vectors
	Prevent exploitation of vulnerable populations		<ul style="list-style-type: none"> • Assessment of working conditions for labour force tasked with reinforcement learning
WORKFORCE AND ENVIRONMENTAL HARMS	Track AI energy usage and carbon emissions		<ul style="list-style-type: none"> • Energy optimisation tools and carbon footprint tracking for foundation models
	Establish clear DEI objectives for AI workforce		<ul style="list-style-type: none"> • Institute fair hiring practices to mitigate bias and ensure a diverse AI workforce
	Upskill talent to ensure proper usage GenAI		<ul style="list-style-type: none"> • Training & continuous learning to ensure employees can properly and effectively use GenAI tools
THIRD-PARTY RISK	Evaluate third-party model risk and ethics		<ul style="list-style-type: none"> • Full-stack assessment of model vulnerabilities and potential ethics violations prior to deployment at scale
	Prevent sharing of proprietary data		<ul style="list-style-type: none"> • Information access controls to prevent sharing of proprietary data with public foundation models
	Implement measures to prevent vendor lock-in		<ul style="list-style-type: none"> • Lock-in risk assessment for vendor solutions



Supporting Tools for
Advanced Level of
Dimension 4:
Third-party Ecosystem

Sub-Dimension 4.1: **Third-party selection criteria and processes**

To select third parties, operators must:

1



Create a clear strategy for managing third-party risk (including, third-party concentration risk) with documentation of RAI selection criteria

2



Incorporate RAI selection criteria as part of the overall selection process, ensuring alignment with the internal RAI guidelines of the operator

3



Develop and regularly update selection criteria to align with evolving requirements

4



Conduct thorough due diligence (e.g., questionnaire) to verify third party's RAI practices

5



Define explicit RAI criteria in contracts (e.g., supplier's code of conduct), with SLAs for accountability and audit provisions

6



Maintain use case registry/registries with latest third-party data (as applicable)

Potential RAI-specific third-party selection criteria that could be leveraged to ensure a holistic evaluation process

THIS SLIDE DOES NOT CONSTITUTE LEGAL ADVICE

Does the third party¹... (note: this is potential criteria and not exhaustive)

● **Soft criteria²**

● **Hard criteria³**

Expertise

- Have personnel with expertise and experience in AI and RAI practices?
- Have the capability to respond appropriately during a risk crisis or audit?
- Provide training on RAI practices to their staff and offer educational resources to clients and users?

Responsible practices

- Publicly commit to RAI principles, ensuring rigorous adherence to AI practices?
- Implement measures to identify and mitigate AI risks with appropriate controls in their AI systems?
- Maintain version control and regularly conduct audits of their AI systems?
- Have testing and acceptance procedures to ensure compliance with RAI principles?
- Ensure their third-party providers follow RAI practices and offer reliable AI/data services?
- Have governance structures in place with clear lines of accountability (internal and external)?

Compliance

- Comply with relevant AI and data protection regulations?
- Have legal frameworks in place to handle AI-related disputes and responsibilities?
- Have any certifications to verify their responsible AI maturity?

Data management

- Have strong data privacy policies compliant with relevant regulations (e.g., GDPR, CCPA)?
- Implement security measures to protect data from breaches?
- Ensure the integrity of the data and conducts appropriate activities to govern data quality?
- Provide data that is safe and traceable?
- Ensure that data received follows a specific format and remains consistent over time?

1. Third-party ecosystem includes strategic partners: network platform and management, third party data providers/aggregators, LLM providers, Cloud providers, SaaS providers

2. Criteria allows for flexibility and adaptation 3. Essential criteria with which a third party must comply

Source: Digital Operational Resilience Act (DORA) - Regulation (EU), Risk management framework for procuring AI systems (by Center for Inclusive Change)



Supporting Tools for
Advanced Level of
Dimension 5:

Change Management and Communications





Sub-Dimension 5.2: Culture and change management

Advanced practices

Successful change management will build conviction and skills, and reinforce with role modeling and incentives

KEY PILLARS

EXAMPLE ACTION ITEMS

	Build understanding and conviction	<ul style="list-style-type: none">• Share change story and communication that clearly and consistently explain the why, what, when, and how of change (i.e., RAI practices)• Publish written documentation (e.g., RAI playbooks for role archetypes) that are easily accessible, and contain clear explanations to reinforce the change story
	Role model change	<ul style="list-style-type: none">• Ensure senior leaders actively adopt and enforce the new RAI practices (e.g., sponsor RAI initiatives through allocation of budget and resources)• Cascade down behaviour and role modelling to the front line (e.g., leverage active community of champions to share expertise across domains, organise forums for employees to interact with leadership)
	Develop skills required for change	<ul style="list-style-type: none">• Conduct simple and interactive skill building workshops and trainings (e.g., hackathons, trainings that bring RAI practices to life, showcase practical impact of non-compliance)• Promote informal coaching on RAI (e.g., provided by managers or central team)• Provide resources to build horizontal skills (e.g., regulatory awareness) and domain-specific skills (e.g., data governance and privacy practices)
	Reinforce with incentives	<ul style="list-style-type: none">• Create robust incentive structure linked to RAI-related OKRs/KRIs (e.g., incentives for red teaming/jailbreaking)• Offer a variety of incentives that are tangible (e.g., bonuses, additional PTO) and intangible (e.g., employee spotlights, awards, certifications from training)

For more information on the GSMA Responsible AI Roadmap visit our [website](#), watch the [video](#) or view the [The GSMA Responsible AI Maturity Roadmap](#) and [Step-by-Step Guide](#) documents.

You can also access the online tool to determine your organisations Responsible AI Maturity level [here](#).