Trusted AI and AI governance

Discussion paper

EY
Building a better working world

AGENDA

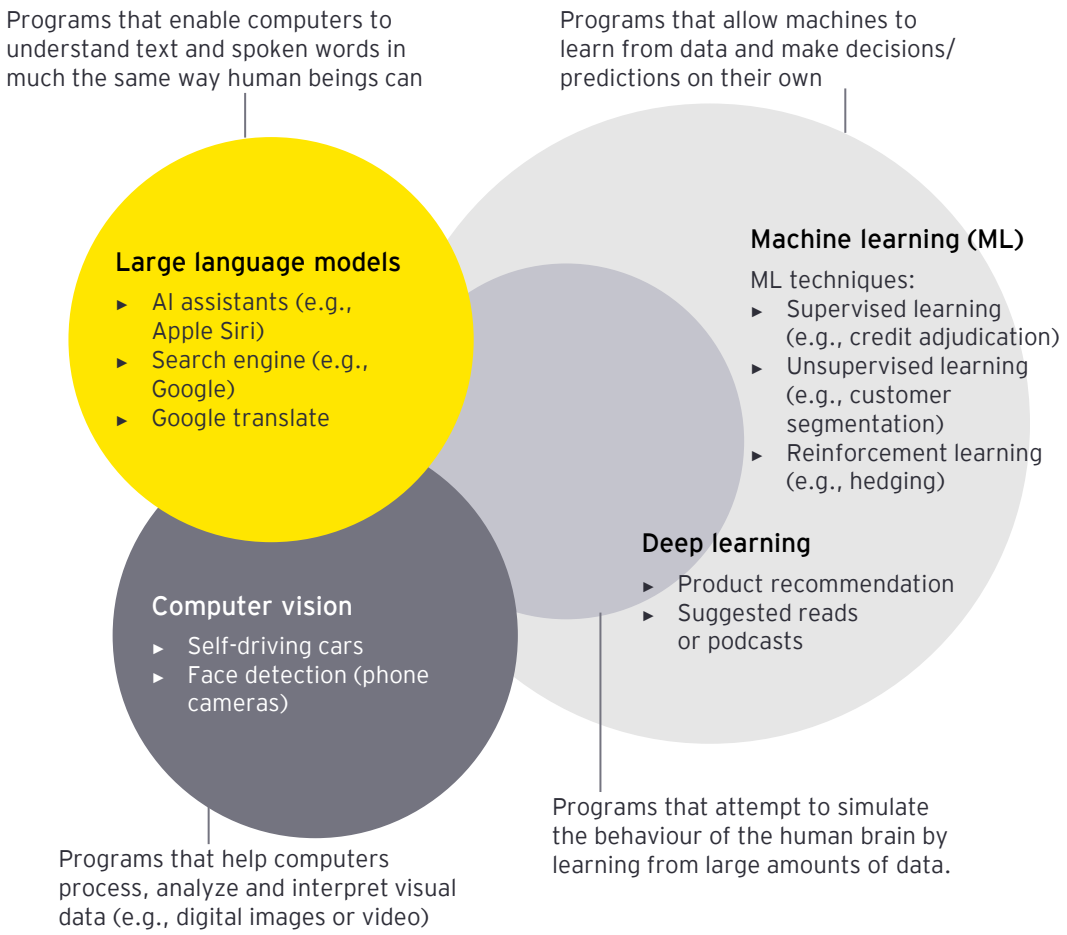
1. AI defined
2. Promise and risks of AI
3. AI regulatory landscape
4. AI ecosystem, risks and response
5. How do we manage these risks?
6. AI governance
7. AI model lifecycle
8. RACI Matrix
9. Risk Tiering
10. Model Inventory
11. Appendix: AI policy and procedures contents

EY
Building a better working world

AI defined

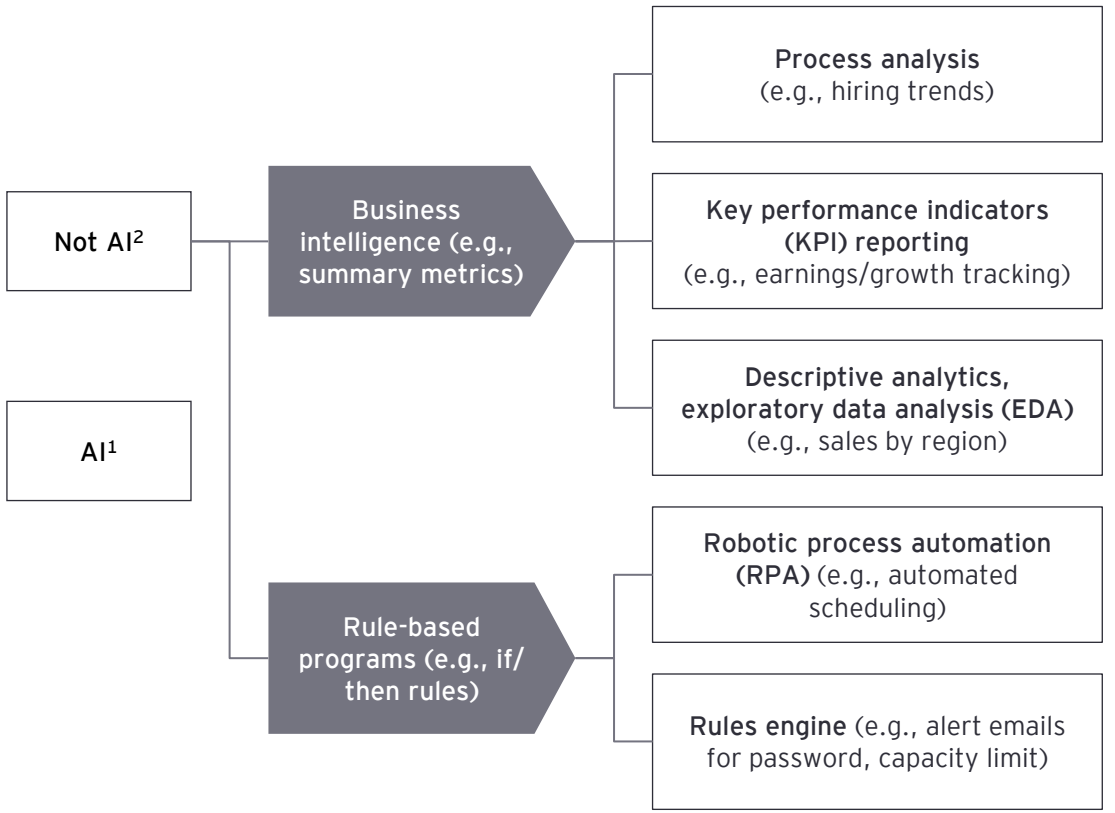
What is AI?

The application of computational tools to build models from examples, data and experience, rather than following preprogrammed rules.



What is NOT AI?

Computer software for which every action and outcome is defined or explicitly programed by humans

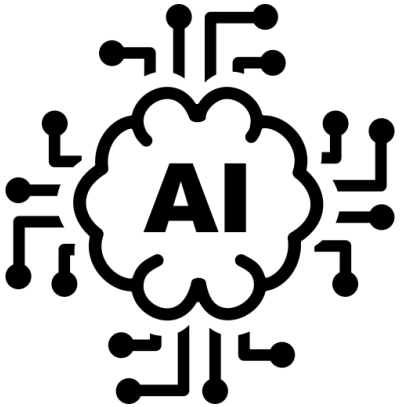


1. AI is a continuously evolving field and the list is meant to be illustrative.
2. Examples are meant to be illustrative and not exhaustive.

Promise and Risks of AI

Enablers

- More data
- Low-cost computing
- Better algorithms



Opportunities*

- Develop new product (e.g., enhancing weight management products)
- Transform business (e.g., targeted marketing, streamline supply chain)
- Improve customer experience & operational efficiency (e.g., intelligent recommendations, AI chatbot)

In the absence of proper controls, adoption of AI may expose the organization to potential failures and mishaps

Early experience with AI emphasizes the need for sound governance and risk management to safeguard against regulatory, reputational and business risks

November 2018: **BlackRock shelves neural net-based AI liquidity models** due to their inability to explain the results to senior management.

March 2016: **Tay, Microsoft's AI chatbot, gets a crash course in racism from Twitter.** Hours after launch, Tay makes racial and controversial tweets.

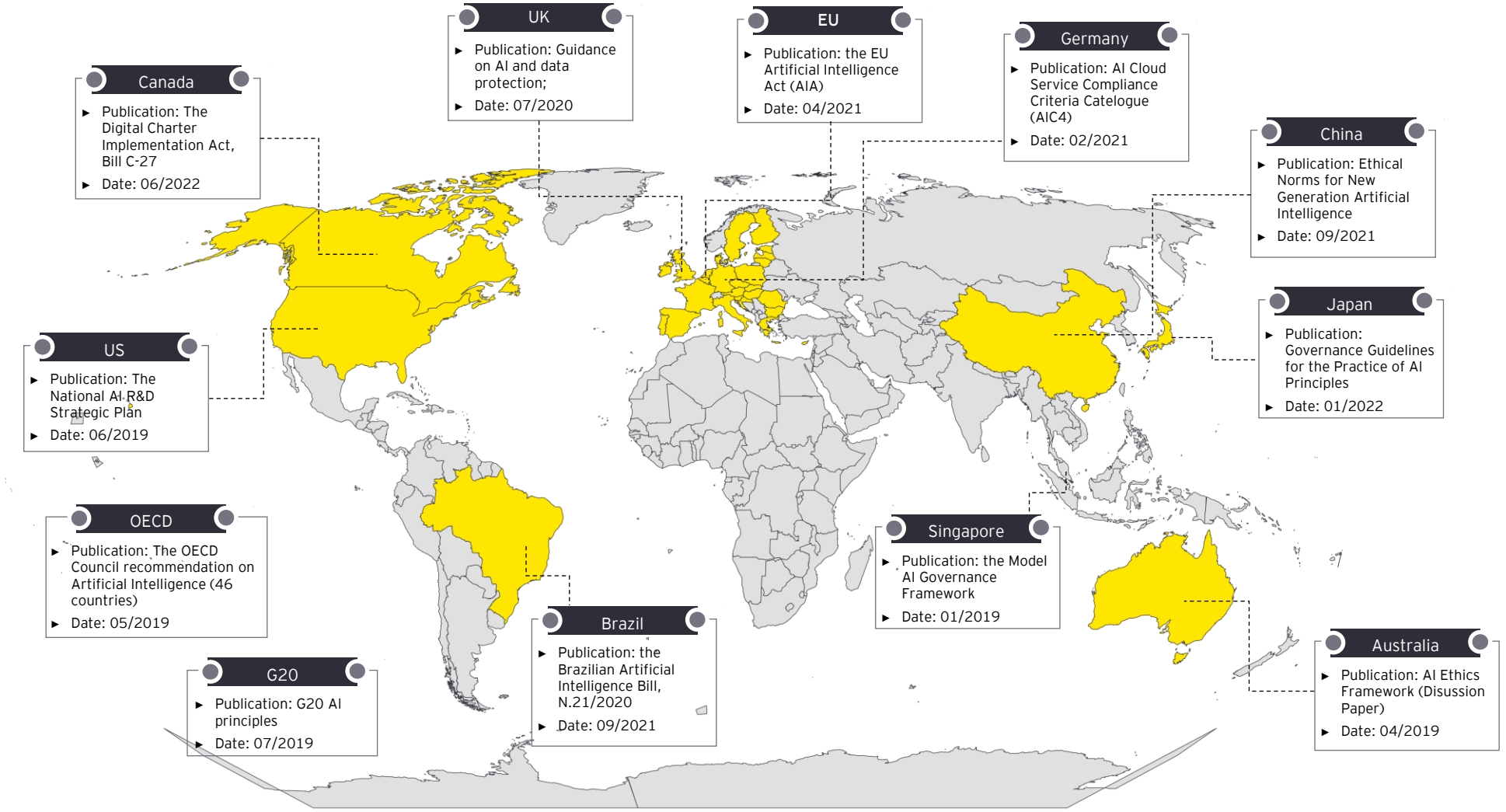
October 9, 2018: **Amazon** scraps **AI recruiting tool that showed bias against women.**

Late 2016: **Uber car runs six red lights.** Uber's self-driving AI technology relies on a highly complex system of AI.

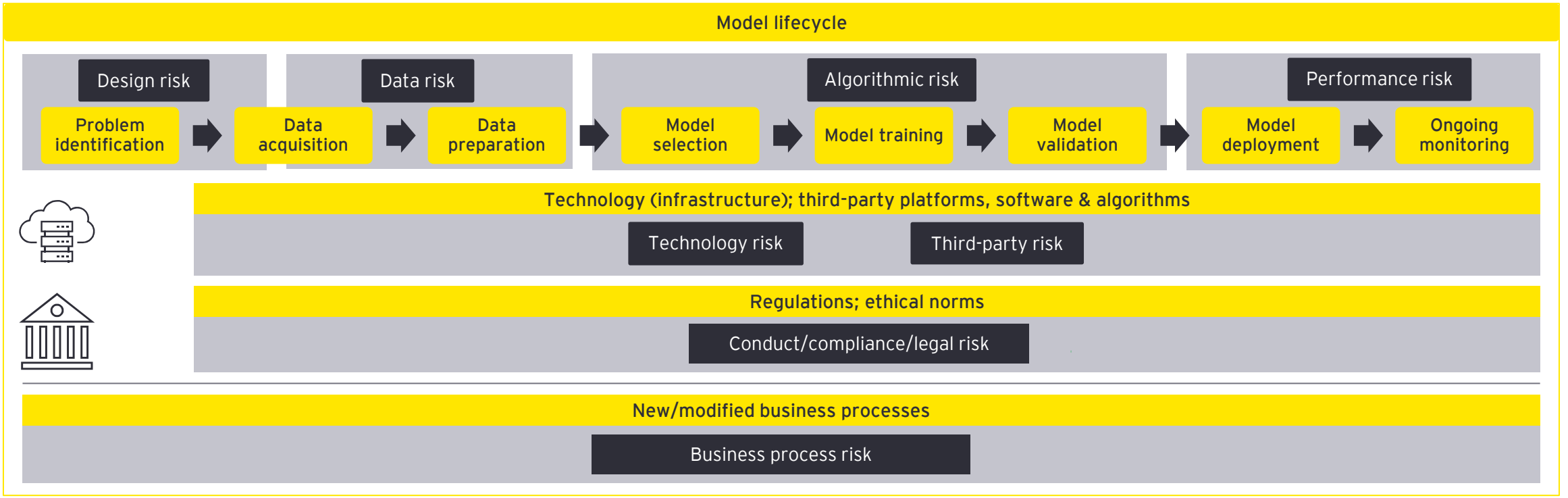
May 24, 2018: **Amazon Echo sent the conversation of a family** to a random person in their contacts.

AI Regulatory Landscape

Rising global guidelines/regulations on trustworthy AI signal urgency



AI ecosystem, risks and response



Managing these risks in the AI lifecycle is the key for enabling trustworthy AI as promoted by regulators

The Principles of Trustworthy AI

Unbiased
Is the AI system free from prejudiced assumptions and intended to drive positive social impact?

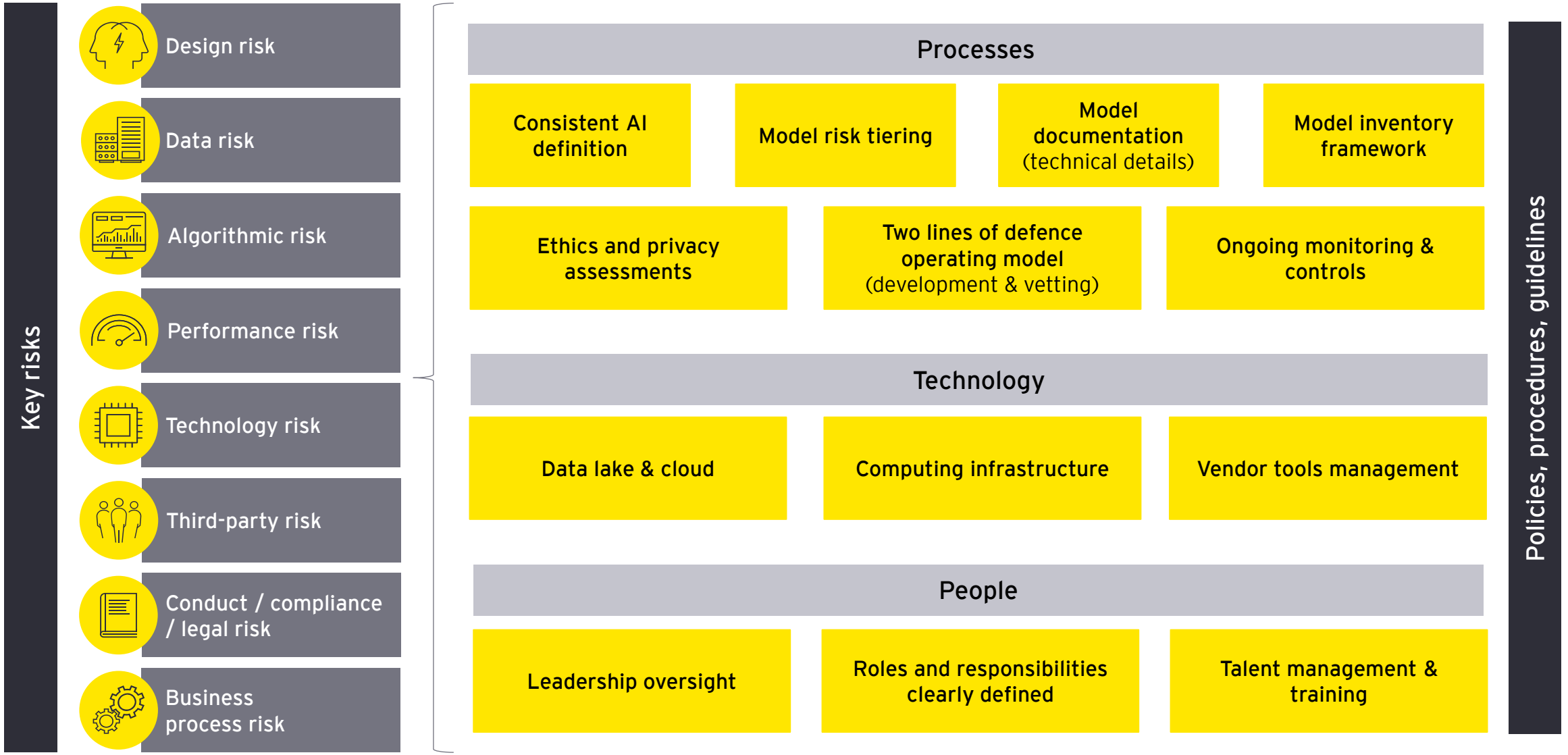
Resilient
Is the data feeding into the AI system secure from unauthorized access that may lead to incorrect outcome?

Explainable
Can a human understand, challenge, and validate the inner workings and results produced by the AI system?

Transparent
Do end users have knowledge and control on what data is being captured and how it is used?

Performance
Are the results from the AI system meeting stakeholder expectations consistently?

How do we manage these risks?



AI governance

The objective is to create **awareness** and **alignment** on key components of the AI governance framework across analytics, technology, risk, ethics and compliance, and privacy teams. This will shape how an organization builds out the 3 pillars of **processes**, **people** and **technology** to enable its AI (data and analytics) transformation journey with a focus on **trusted AI**. AI governance encompasses:

Policy
Why?

AI model risk management (MRM) policy

- Purpose and scope
- AI principles
- AI model risk management
- AI model risk reporting and execution

Procedure
How to do it?

AI model risk management (MRM) procedure

- Model initiation
- Model development
- Pre-implementation review
- Model approval
- Model implementation
- Post-implementation review
- Ongoing monitoring
- Ongoing review
- Model change
- Model decommissioning

Guidelines
Technical guidance

AI model development and validation guideline

- Model development guideline (for prioritized model types)
- Model validation guideline (for prioritized model types)

AI model lifecycle

Initiation

Initiation procedure

1. Identify the business problem (initiative)
2. Identify the need for AI model

Model development

Model development procedure

1. Model design
2. Model selection and build-out
3. Model testing and assessment
4. Ethics, compliance and privacy assessment

Pre-implementation review

Pre-implementation review procedure

1. Determine model risk tiering
2. Determine ongoing monitoring/review requirements
3. Model evaluation and verification
4. Review ethics, compliance and privacy assessments

Model approval

Model approval procedure

1. Approve new model/changes to existing model
2. Approve model risk tiering
3. Approve timelines for remediation action for model issues and exceptions

Model implementation

Model implementation procedure

1. Implement the model as validated
2. Set up adequate post-implementation change management and control

Model change & decommissioning

Model change procedure

1. Change assessment and actions
2. Change approval
3. Change implementation & monitoring

Model decommissioning procedure

1. Model retirement assessment
2. Retirement approval

Ongoing monitoring & review

Ongoing monitoring procedure

1. Check if the performance metrics are in the acceptance range

Ongoing review procedure

1. Review results from ongoing monitoring
2. Perform benchmark analysis, sensitivity analysis, etc. to assess appropriateness of the model
3. Determine if model change or decommissioning needed

Post-implementation review

1. Perform post-implementation testing to ensure the model is implemented as intended
2. Ensure that appropriate controls are implemented as planned

Illustrative roles and responsibilities matrix (RACI) for AI governance

R = Responsible (executing the task) **A** = Accountable (owner of the task) **C** = Consulted (key stakeholder who should be included in decision or work activity) **I** = Informed (needs to know of decision or action)

Tasks associated with AI risk management			Roles associated with AI Risk Management										
			1st line of defence (risk owners)				2nd line of defence (oversight and governance)						
AI lifecycle phase	Activity group	Activities	AI model develop-pers	IT ops (GTS)	AI model managers	AI model business owners	AI model validators	AI model validation owners	Ethics and compliance	Legal (privacy) counsel	Cyber-security	AI model risk committee	Enterprise risk management (ERM)
Initiation	Purpose identification	Identify the business problem				AR							
		Identify the need for AI model	C	C	C	AR							
Model Development	Model design	Design the AI model (scope and objective)	R		A	C							
	Inventory management	Register AI model	R		A	I							
	Model selection and build-out	Determine an appropriate methodology for use	R		A	I							
		Determine appropriate data to be used for model development	R	I	A	I			I	I			
		Select an appropriate model development platform	R	C	A	I					C		
		Program the necessary code for model implementation	R		A								
	Model testing and assessment	Perform appropriate testing to see if the underlying assumptions are plausible and that the model is programmed correctly	R		A	I							
		Assess the appropriateness of a model's use for the intended purpose	R		A	I							
	Ethics, compliance, and privacy assessment	Complete Ethics, compliance and privacy assessment	R	I	A	I			C	C			
		Review the questionnaire submission	C	I	C	I			C	AR		I	I
		Remediate ethics, compliance and privacy concerns	R	I	A	I			C	C			
Pre-implementation Review	Model risk tiering	Determine model risk tier based on risk factors	C	I	C	C	R	A	C			I	I
		Determine requirements for initial validation and ongoing model monitoring/review based on risk tier	C		C	C	R	A				I	
	Model evaluation and verification	Execute model validation, i.e., assess if all three model components – inputs, computation processes and outputs – are working as intended	C	C	I	I	R	A			C		
		Provide recommendations regarding model approval	C		C	C	R	A				I	I
	Review ethics and privacy assessments	Assess if the ethics and privacy assessments were completed accurately by the model developers	C	I	C	C	R	A	C	C		I	
		Verify if the remediation actions were appropriate to address the ethics and privacy concerns	C	I	C	C	R	A	C	C		I	

Illustrative roles and responsibilities matrix (RACI) for AI governance (cont.)

R = Responsible (executing the task) **A** = Accountable (owner of the task) **C** = Consulted (key stakeholder who should be included in decision or work activity) **I** = Informed (needs to know of decision or action)

Tasks associated with AI risk management			Roles associated with AI risk management										
			1st line of defence (risk owners)				2nd defence (oversight and governance)						
AI lifecycle phase	Activity group	Activities	AI model develop-pers	IT Ops (GTS)	AI model managers	AI model business owners	AI model validators	AI model validation owners	Ethics and compliance	Legal (privacy) counsel	Cybe-rsecurity	AI model risk committee	Enterprise risk management (ERM)
Model approval	Model approval	Approve AI model and the assigned model risk tiering	C		C	C	C	C	C	C	C	AR	I
Model implementation	Model implementation	Select an appropriate model deployment platform	R	C	A	C					C		
		Execute implementation and performance test	R	C	A	C					C	I	
		Document model	R		A	C					C		
Post-implementation review	Post-implementation testing and review	Perform post-implementation testing to ensure the model was implemented as intended	C	C	C	C	R	A				I	
		Check if appropriate controls are implemented	C	C	C	C	R	A			C	I	
Ongoing monitoring & review	Ongoing performance monitoring	Execute performance tests to check if the model performance is in the acceptable range	R	C	A								
		Report performance test results	R		A	I	I	I				I	
		Remediate performance test issues	R	I	A	C	I	I			I	I	
	Ongoing model review	Execute model review to assess if the model is still fit for the intended purpose	R	C	A	C			C	C	C		
		Report model review results	R		A	C	I	I				I	
		Review the model review results and recommend if any model change/decommission needed	C	I	C	C	R	A			I	I	I
Model change & decommissioning	Model change & decommission-ning	Approve any recommended model change/decommissions	C		C	C	C	C				AR	I
		Execute model change/decommissioning	R	I	A	C	I	I			I	I	I
Oversight & governance	Maintenance & reporting	Escalations and exceptions management and arbitration	C	C	C	C	C	C	C	C	C	AR	I
		Model risk management (MRM) adoption oversight and enhancements	C	C	C	C	C	C	C	C	C	AR	I

Model Risk Committee

- Chief Risk Officer (Chair)
- Leader, Enterprise Risk Management (ERM)
- Leaders, Ethics and Compliance
- Leaders, Privacy
- Leaders, Data & Analytics, AI
- Leaders, Cybersecurity

AI model risk tiering – model classification and assessment

AI model risk classification and assessment requirements			
Model risk tiering	**Initial validation**	**Ongoing monitoring**	**Ongoing review frequency**
High risk	Yes (high priority)	Monthly	Annual
Medium risk	Yes	Quarterly	Every two years
Low risk	Reduced scope validation	Annually	Every three years

Notes

- **Initial validation** refers to pre- and post-implementation review. Its objective is to review the conceptual soundness of the model methodology, model testing and verification, available documentation, etc.
- **Ongoing monitoring** is to confirm that the model continues to perform as expected over time after implementation, and to help identify the need for changes and enhancement should performance deterioration emerge over time (e.g., changes in market conditions or business activities, etc.)
- **Ongoing review** refers to periodic reviews of the model. Its objectives include reassessment of the continued appropriateness, performance of the model and any benchmarking analysis, back-testing metrics, sensitivity analysis, etc.; updating the model documentation accordingly.
- **Reduced scope validation** refers to a validation with reduced scope, with focus on documentation, conceptual soundness and performance, excluding thorough sensitivity analyses and benchmarking
- While the technical teams in the 1st and 2nd lines of defence are primarily responsible for the activities in the validation/monitoring/review stage, other teams, including ethics and compliance, privacy, IT/cybersecurity, etc., also play crucial roles in supporting these activities. More details on roles and responsibilities throughout the model lifecycle will be discussed in upcoming sessions.

Risk tiering framework – illustrative template

- Illustrative 3-tier risk matrix (low/medium/ high) to assess the performance risk tier of AI use cases based on **complexity** and **business impact**.
- The depth of initial validation and frequency of ongoing monitoring and ongoing review are adapted based on the model risk tier.

Performance risk tier		Business impact		
		Low	Medium	High
Complexity	Low	Tier 3	Tier 3	Tier 2
	Medium	Tier 3	Tier 2	Tier 1
	High	Tier 2	Tier 1	Tier 1

Model risk tiering	Initial validation	Ongoing monitoring	Ongoing review frequency
Tier 1	Yes (high priority)	Monthly	Annual
Tier 2	Yes	Quarterly	Every two years
Tier 3	Reduced scope validation	Annually	Every three years

Complexity assessment – illustrative template

Domain	Questions		Answers	Score	Max
Data source	1. How many data sources does the AI use case pulls data from?		A. One	0	2
			B. Two to three	1	
			C. Four or more	2	
	2. Does the AI use case use external data?		A. No	0	1
			B. Yes	1	
Input Data	3. Does the AI use case use unstructured data?	A. Yes		2	2
		B. No How many features does it use?	A. 1 to 20	0	
			B. More than 20	1	
	4. Was the AI use case trained with imbalanced dataset?		A. No	0	1
			B. Yes	1	
Methodology	5. What type of AI algorithms is used by the AI use case?		A. Highly interpretable algorithms (e.g., rule-based systems, linear/logistic regressions, decision trees etc.)	0	4
			B. Moderately interpretable algorithms (e.g., support vector machines, random forests, etc.)	2	
			C. Less interpretable algorithms (e.g., deep neural networks, support vector regression, gradient boosting etc.)	4	
	6. Are parameters dynamically adjusted (e.g., active learning, reinforcement learning)?		A. No	0	2
			B. Yes	2	
	7. Does the AI use case employ transfer learning?		A. No	0	1
			B. Yes	1	
Output	8. To what extent is additional processing required on the output of the AI use case?		A. None	0	2
			B. Somewhat	1	
			C. Significant	2	
Total score					**15**
Complexity					

Total score	Complexity
0 to 4	Low
5 to 9	Medium
10 to 15	High

Business impact assessment – illustrative template

Domain	Questions	Answers	Score	Max
Revenue	1. Does the AI use case aid in revenue-generating activities of the bank?	A. No	0	2
		B. Indirectly	1	
		C. Directly	2	
Operational efficiency	2. To what extent will the bank's operational efficiency be impacted if the AI use case does not perform as expected?	A. None	0	2
		B. Somewhat	1	
		C. To a large extent	2	
Customer experience	3. Is the AI use case impacting external clients or customers, as opposed to internal users?	A. No	0	2
		B. Somewhat	1	
		C. Major Impact	2	
Compliance	4. Is the AI use case used for control processes, such as internal/regulatory compliance?	A. No	0	2
		B. Yes, internal compliance	1	
		C. Yes, regulatory compliance	4	
Reputational risk	5. What would be the impact to the bank's reputation if the AI use case does not perform as expected?	A. None	0	2
		B. Moderate	1	
		C. Major	2	
Total score				12
Business impact				

Total score	Business impact
0 to 3	Low
4 to 6	Medium
7 to 12	High

AI model inventory

AI model inventory is intended to capture and track organization-wide AI assets on an ongoing basis.

1. AI model attributes

2. AI model version attributes

3. Issue management attributes

1

2

3

AI Model Inventory – AI Model Attributes — **Alignment on attributes** for the AI Model Inventory

Attribute Name	Definition	Attribute Type
AI Model Unique ID	Unique AI Model identifier	Generated
AI Model Name	Name of the AI Model	Text Box
AI Model Description including Purpose	Description and purpose of the AI Model	Text Box
AI Model Analytics Owner	Analytics Sponsor/Owner of the AI model	Drop-down
AI Model Business Owner	Business owner of the AI model	Multi-Select
AI Model Developer	Name of AI Model developer	Multi-Select
AI Model Provider	Name of provider / supplier of AI Model (can be internal or vendor supplied)	Drop-down: internal or external vendor (please specify)
AI Model Vendor Name (if applicable)	Name of the vendor if AI model is externally sourced, otherwise NA	
AI Model Validator	Name and team name of the AI Model Validator	Multi-Select
AI Model Programming Language	Internally developed AI Models - coding language used to develop AI Model; Vendor AI Models - vendor software that the AI Model is operated on	Drop-down
Data Source(s) Used	The repository from which data is sourced (e.g. data lake). *May be able to link to sources if existing organizational tools are used to implement AI Model Inventory*	Drop-down
Data Domain(s) Used	Data Domain(s) used for the model	Drop-down
External Data Source(s)	External Data Source(s) used	Text Box
AI Model Implementation System	Description about how the model is being served	Text box
AI Model Users	Name of the intended users/teams of the AI model	Drop-down
AI Model Approval Date	Date when approval was granted (version approval date)	Calendar, N/A if not applicable
AI Model Implementation Date	Date of initial implementation (first time AI Model was implemented)	Calendar, N/A if not applicable
AI Model Change Approval Date	Date when model change was approved	Calendar, N/A if not applicable
AI Model Change Date	Date when model was last updated / changed	Calendar, N/A if not applicable
Nature of Last Update/Change	Initial version / Major AI Model enhancement / Minor AI Model enhancement / Immaterial, if applicable	Drop-down, N/A if not applicable
AI Model Planned Decommission Date	Date at which the AI Model is expected to be retired if applicable	Calendar, N/A if not applicable
AI Model Decommission Approval Date	Date when model decommissioning was approved	Calendar, N/A if not applicable
AI Model Decommission Date	Date when model was decommissioned	Calendar, N/A if not applicable

AI Model Inventory – AI Model Version Attributes — **Alignment on attributes** for the AI Model Version

This AI Model Version table includes all attributes specific to each version of an AI Model in the AI Risk Management Inventory.

Attribute Name	Definition	Attribute Type
AI Model Unique ID	AI Model Version Unique identifier	Generated
Name	AI Model Version Name	Text Box
AI Model Version Unique ID	Version Number (integer)	Text Box
Last Ethics, Compliance and Privacy Review Date	Date of the most recent Ethics, Compliance and Privacy assessment	Calendar
Last Model Review Date	Date of the most recent model review	Calendar
AI Model Version Approval Date	Committee approval date of the version (should coincide with the approval date of the first review event associated with this version)	Calendar
AI Model Version Implementation Date	Date of AI Model version implementation date	Calendar
AI Model Version Change Approval Date	Date when model change was approved	Calendar
AI Model Version Change Date	Date when model was last updated / changed	Calendar
Comments	General comments	Text Box

AI Model Inventory – Issue Management Attributes — **Alignment on attributes** for the AI Issue Management

This Issue Management table includes attributes specific to Issue Management for escalated issues. These escalations may occur after the ethics, compliance, privacy assessment, implementation and deployment testing or the performance testing.

Attribute Name	Definition	Attribute Type
Issue Unique ID	System-generated	Generated
AI Model Unique ID	Identifier of the AI Model against which the issue is raised	Numeric
Title	Enter a short and descriptive title for the issue	Text Box
Issue Identifier – Coordinator	Person who escalated the issue	Search Box
Group that identified the issue	Analytics / Ethics / Privacy / Second-line Review	Drop-down
Issue Start Date	Date the issue was first communicated	Calendar
Source	'Performance Testing', 'Ethics Review', 'Privacy Impact Assessment' or 'Implementation Testing'	Drop-down
Issue Type	Logical grouping / theme of issue identified (*to be tracked for future enhancement of AI Governance policies, procedures, and guidelines*)	Drop-down
Description	Enter a description of the Issue	Text Box
Current Issue Severity	Select the severity risk rating (Low /Medium/High-Time sensitive)	Drop-down
Issue Owner	Enter name or team name of Issue Owner	Drop-down
Issue Status	Determine whether the issue is considered Open/Active or Closed	Drop-down
Action Plan Title	Enter a short and descriptive title for the Action Plan	Text Box
Action Owner	Team (e.g., model development) responsible for action plan input	Search Box
Action Owner Approver	Supervisor of the issue identifier who issued the finding	Search Box
Current Action Owner ERD (Year and Quarter)	Use the drop-down list to select the target year and quarter when the Action Plan is to be completed	Calendar
Specific Actions/Evidence required for Closure	Enter the specific actions/evidence/artifacts that are required as a pre-condition for Issue closure	Text Box
Action Plan Status	Action Plan Status captures the remediation status of the issue (not started, in-progress, blocked)	Drop-down
Closure Date	Date the issue was addressed and considered closed	Calendar
Comments	General Comments	Text Box

Table of contents for the AI policy

Scope

Entity and geographical scope
Mention the entity and geographies that are within the scope of this policy.

Definition of AI model
This will be based on the AI model definition discussed in the workgroups
Examples will be provided for guidance

Model risk management

Model risk principles and requirements
1. Model risk involves adverse financial and reputational consequences
2. Control on model risk requires an oversight authority for model approval

Model lifecycle
Risk is managed through model lifecycle – model initiation, development, independent review, approval, implementation, ongoing monitoring, model change and decommission

Two lines of defence
1. First line of defence owns and manages the risk, including model development, implementation and monitoring
2. Second line of defence identifies emerging risks, including model review and assessment

Model risk management framework
1. Model initiation and model development
2. Model review and implementation
3. Monitoring and governance

Summary of roles and responsibilities
Management of model risk is a joint responsibility of model owner, developer, user and independent risk management function.

Policy violation
Policy violations can occur through either model exceptions or due to non-compliance of existing models.

Materiality
Materiality in a model is the potential impact of model, including financial, strategic and operational

Independent Effective Challenge
Effective challenge refers to critical reviews by informed parties including the validators.

Model Risk Reporting and Execution

Model risk appetite
Risk appetite is amount of risk the organization is willing to assume to meet its desired objectives

Escalation
AI Risk Committee will escalate matters to Management Risk Committee as aligned with the risk escalation guideline

Committee table
List the committees involved in the process and their roles outlined

Use of external resource
Model development and model validators can use external help to complete their tasks as long as they abide by the organization's policies, and they supervise and approve the work

Training and policy attestation
Annual training with respect to model risk management must be completed by individuals carrying out model-related roles and responsibilities

Table of contents for the AI procedure

Model initiation and development

Model Initiation
1. Identify the business problem (initiative)
2. Identify the need for AI models

Model development
1. Model design
2. Model selection and build out
3. Ethics and privacy assessment
4. Model testing and assessment

Model documentation
Model development document captures model design, processing, evidence of industry practice, academic theories, rationale for data choice, model limitation and aligned with risk

Model review and implementation

Pre implementation review
1. Determine model risk tiering
2. Determine ongoing monitoring/review requirements
3. Model evaluation and verification
4. Review ethics and privacy assessments

Model risk tiering
Model owners are required to quantify model risk as per risk tiering guidelines, and model reviewers evaluate it

Model approval
1. Approve new model/changes to existing model
2. Approve model risk tiering
3. Approve timelines for remediation action for model issues and exceptions

Model implementation
1. Implement the model as validated
2. Set up adequate post-implementation change management and control

Post implementation review
1. Perform post-implementation testing to ensure the model is implemented as intended
2. Ensure appropriate controls are implemented as planned

Monitoring and governance

Ongoing monitoring
1. Check if the performance metrics are in the acceptance range

Ongoing review
1. Review results from ongoing monitoring
2. Perform benchmark analysis, sensitivity analysis etc. To assess appropriateness of the model
3. Determine if model change or decommissioning needed

Model change
1. Change assessment and actions
2. Change approval
3. Change implementation & monitoring

Model decommissioning
1. Model retirement assessment
2. Retirement approval

Vendor model management
A vendor model is managed similarly as an internal model. The owner is responsible for documentation, etc., and a validator for review

Model inventory management
Procedures for maintaining the centralized inventory for all models. The second line of defence will do periodic inventory reconciliation

Appendix

RACI

Model inventory

EY | Building a better working world

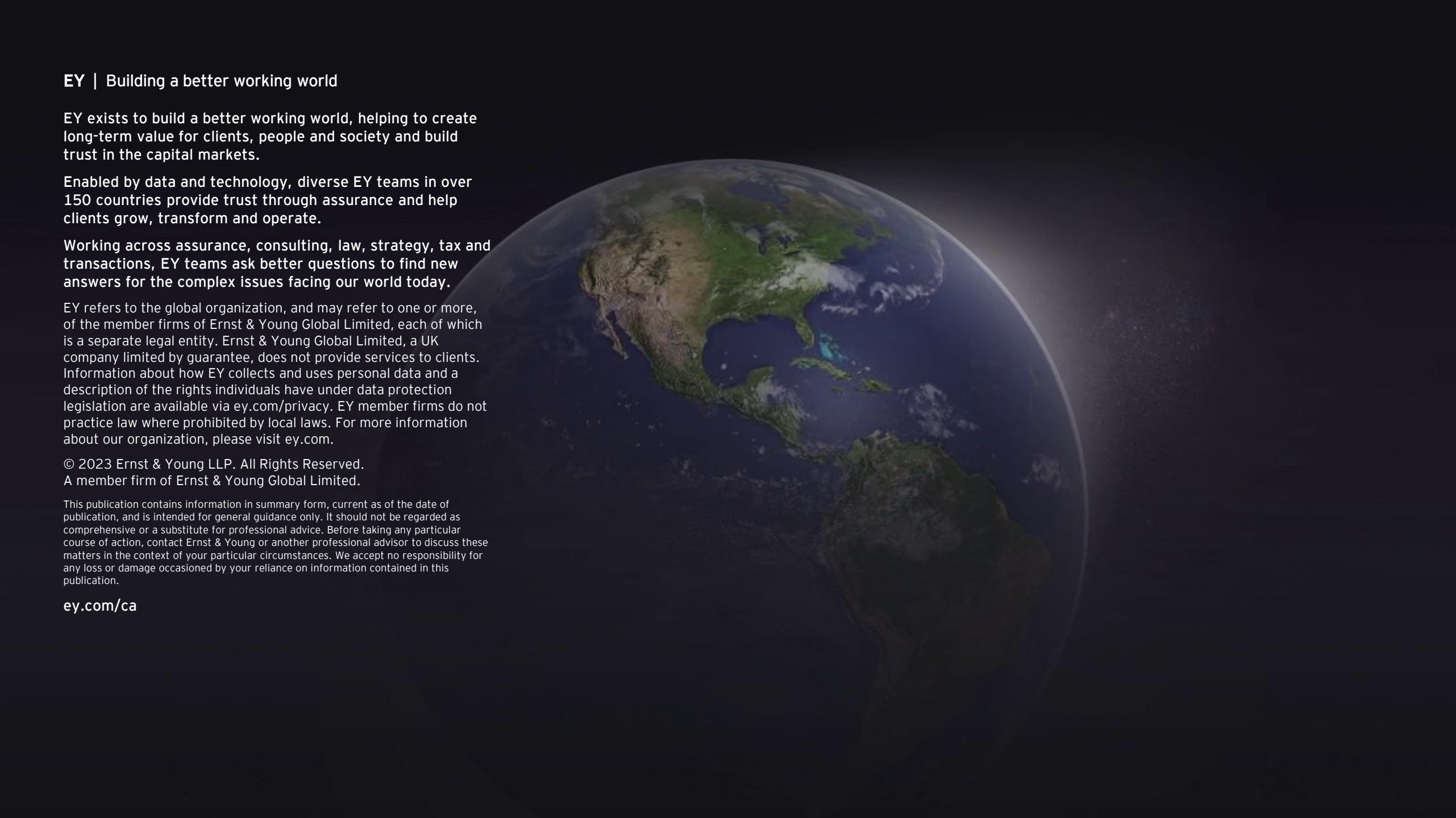
EY exists to build a better working world, helping to create long-term value for clients, people and society and build trust in the capital markets.

Enabled by data and technology, diverse EY teams in over 150 countries provide trust through assurance and help clients grow, transform and operate.

Working across assurance, consulting, law, strategy, tax and transactions, EY teams ask better questions to find new answers for the complex issues facing our world today.

EY refers to the global organization, and may refer to one or more, of the member firms of Ernst & Young Global Limited, each of which is a separate legal entity. Ernst & Young Global Limited, a UK company limited by guarantee, does not provide services to clients. Information about how EY collects and uses personal data and a description of the rights individuals have under data protection legislation are available via ey.com/privacy. EY member firms do not practice law where prohibited by local laws. For more information about our organization, please visit ey.com.

ey.com/ca



Authors

Mario Schlener

Partner, Lead Financial Services Risk Management Practice and Enterprise Risk Strategy, EY Canada

EY Global FS Risk Technology, Alliance, Innovation Lead

mario.schlener@ca.ey.com

Yara Elias, Ph.D.

Senior Manager, AI Risk Lead, Financial Services Risk Management, EY Canada

yara.elias@ca.ey.com

Anil Sood

Senior Manager, AI Governance Lead, Financial Services Risk Management, EY Canada

Anil.Sood@ca.ey.com

Kiranjot Dhillon

Senior Manager, AI Risk, Financial Services Risk Management, EY Canada

kiranjot.dhillon1@ca.ey.com

Liang Hu, Ph.D.

Manager, Responsible AI and AI Risk , Financial Service Risk Management, EY Canada

liang.Hu@ca.ey.com

Rasoul Shahsavarifar, Ph.D.

Manager, AI Risk, Risk Consulting, EY Canada

Rasoul.Shahsavarifar@ca.ey.com