EY.ai Confidence Index

Bolster value through responsible AI adoption

20.799

[47.381 78.611]
H - 85 | 8594 - 9053]

90.467

61.800

EY
Building a better
working world

An AI Confidence Framework requires grounding in a key purpose, values and principles

Purpose – The Why?

The primary purpose of the AI Confidence Framework is to enable the organization to:

▸ **Compete** by enhancing speed and effectiveness of AI adoption

▸ **Comply** with evolving regulations

▸ **Protect** the organization and its stakeholders from downside risks
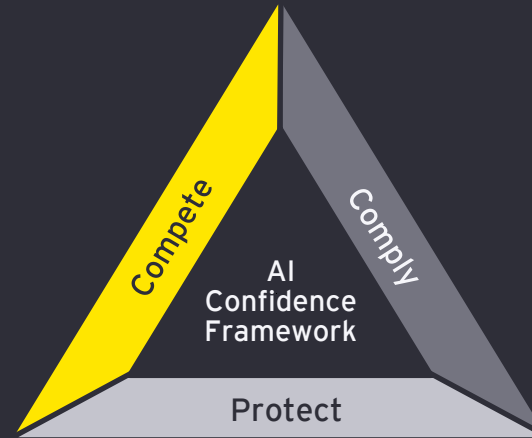
Values – The What?

The AI Confidence Framework is grounded in the following values:

▸ **Purposeful Design**

▸ **Vigilant Supervision**

▸ **Agile Governance**

This enables the enterprise to achieve the purpose of the program keeping ethics in focus and aligning the AI strategy with your organizational values.
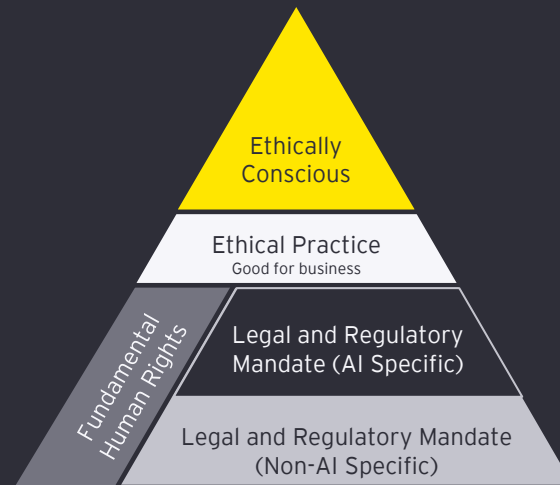
Principles – The How?

The AI Confidence Framework is delivered through the enablement of the following Principles:

▸ Accountability

▸ Bias and Fairness

▸ Explainability

▸ Privacy (Data Protection)

▸ Reliability

▸ Security

▸ Sustainability

▸ Transparency

▸ Compliance

EY

EY.ai Confidence Index helps organizations reap the benefits of responsible AI adoption

EY.ai
Confidence Index

Value proposition

▸ The EY.ai Confidence Index enables confidence in the data and technology, process and people pillars making up your AI ecosystem.

▸ It supports enhanced decision making and more efficient operations through reliability and explainability, and it promotes responsible AI by enhancing transparency and privacy, through measurable confidence levels.

▸ Each assessment is customizable and contextual, balancing quantifiable metrics like solution performance with qualitative aspects like user transparency.

Specifications

Based off of our Principles of Responsible AI:

▸ Accountability   ▸ Reliability

▸ Bias and Fairness   ▸ Security

▸ Explainability   ▸ Sustainability

▸ Privacy (Data   ▸ Transparency
  Protection)
                        ▸ Compliance

Users

▸ Data and AI, innovation, and emerging technology functions

▸ Risk, ethics, legal, and compliance functions

Assets

▸ AI solution and portfolio level development, testing and monitoring

▸ Supported by organization level AI governance, including AI policy, procedure

EY

The EY.ai confidence index builds and executes an AI confidence framework tailored to the organization, to promote AI adoption and innovation

Illustrative
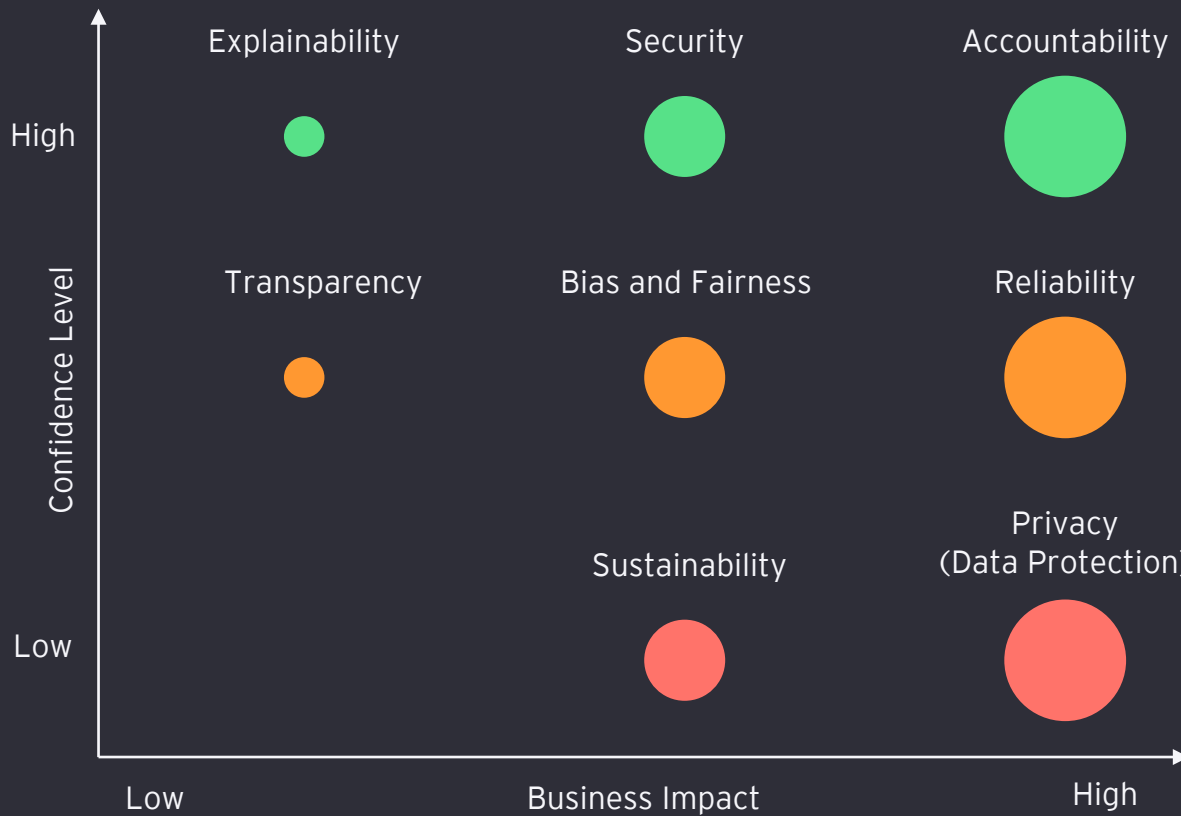
Quantifying our understanding of risk in the AI solutions with continuous monitoring across its lifecycle, based on our Responsible AI dimensions weighted by their business impact

Confidence Level

High	Medium	Low

Relevance based on business impact
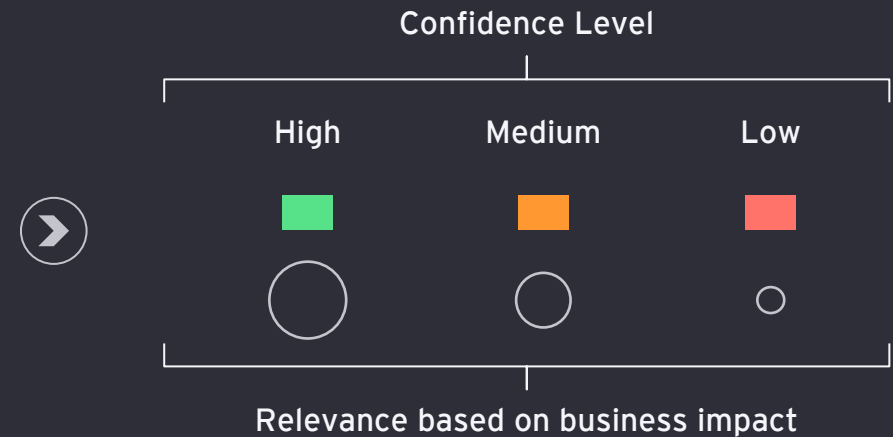
The ninth principle, compliance, is considered as part of Relevance, along with financial and non-financial impact (operational impact, citizen experience)

EY

Example of EY's fairness solution applied on loan adjudication fairness assessment

- Assess the fairness confidence index for a credit loan adjudication AI model during the model development phase
- Explore the principle of equalized odds to identify whether the outcomes are unfair relative to gender
- Remediate any unfair outcomes using the threshold optimization technique and compute the confidence index subsequently

Methodology

Group Fairness Assessment:

Equalized odds looks at the difference between false positive rates (FPR) and true positive rates (TPR) across the unprivileged group (Female or F) and the privileged group (Male or M).

Group Fairness Mitigation:

The threshold optimization technique reduces the Equalized Odds difference to 0.2%. That said, the overall area under the curve (AUC) of the model drops by a relative 1%

Unmitigated Model Performance On Test Data Set

Unmitigated accuracy, FPR, FNR by group

Mitigated Model Performance On Test Data Set

Mitigated accuracy, FPR, FNR by group

Results

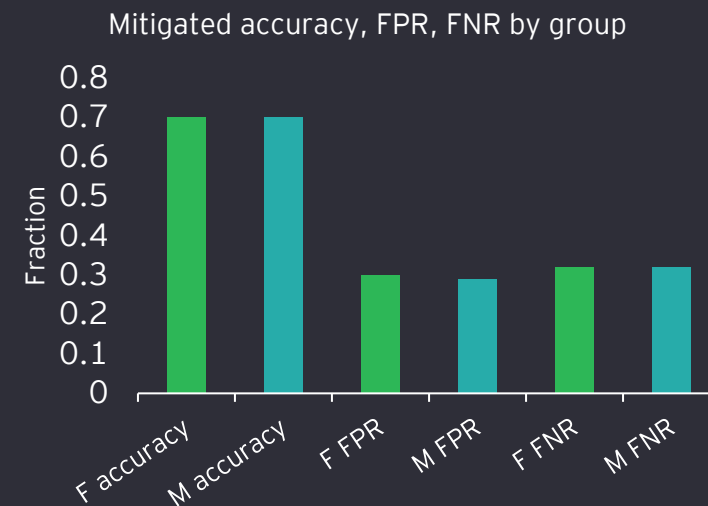- At the first iteration, the model results in unfair outcomes relative to gender (a 6.3% equalized odds difference between the female group and the male group)
- After mitigation, the threshold optimization technique reduces the equalized odds difference to 0.2%, helping remediate unfair outcomes. This results in a relative drop of 1% in performance.
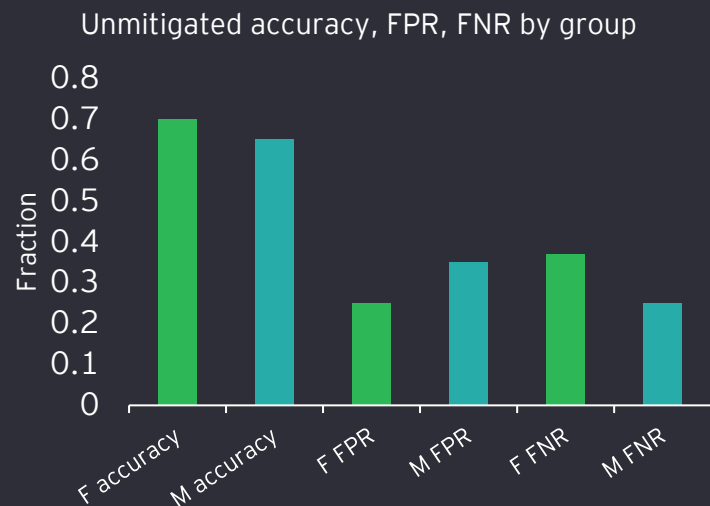
Dashboard representation of the EY.ai confidence index at the solution level

▸ Each slice (1/8) of the octagon represents an individual RAI principle

▸ The colour on each wedge indicates the confidence level of each principle

▸ The extent of coverage on each wedge depicts the relevance on business impact

▸ If the RAI principle is not applicable to the business impact, only the confidence level is indicated with a border around the wedge

Business Impact

High Relevance

Medium Relevance

Low Relevance

Not Applicable

Confidence Level

High Confidence

Medium Confidence
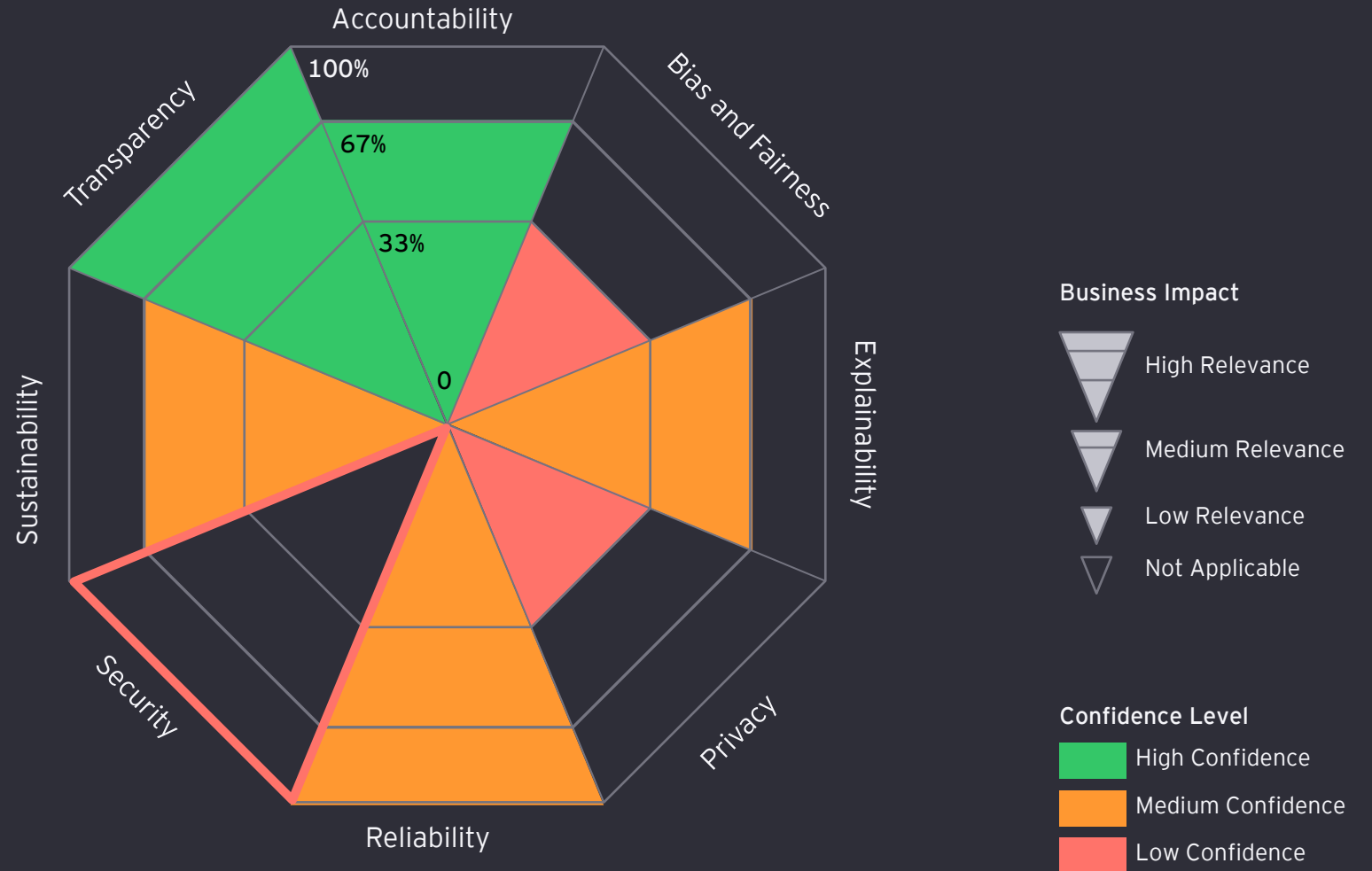
Low Confidence

Action plan based on the enterprise's risk appetite and the Ey.ai confidence index

The AI confidence index can be used in conjunction with the enterprise's risk appetite, that is the level of risk that it is willing to accept while pursuing its objectives. The enterprise's risk appetite can be translated into a matrix as illustrated next*, enabling the go/no-go decision at a solution level and monitoring triggers, by checking whether the AI confidence index results do not meet the confidence levels outlined based on business impact.

Illustrative

Business Relevance	⬭ High	⬭ Medium	◦ Low
Accountability	High	High	Medium
Bias and Fairness	High	Medium	Medium
Explainability	High	Medium	Low
Privacy	High	High	Medium
Reliability	High	High	Medium
Security	High	High	Medium
Sustainability	Medium	Medium	Low
Transparency	Medium	Medium	Low

Confidence Level

- 🟩 High
- 🟧 Medium
- 🟥 Low

Relevance based on business impact

- ⬭ High
- ⬭ Medium
- ◦ Low
- · Not Applicable

*The matrix may need to be tailored per business function.

EY.ai confidence index enables understanding of AI risk at solution, portfolio and institution level

The AI confidence index can be used to better understand the AI risk distribution in a bottom-up approach, assessing AI risk for each solution individually, and subsequently consolidating into a portfolio-level and institution-level view of AI risk. As a result, institutions are enabled to define, measure and mitigate AI risk to ensure compliance and alignment with the institution's risk appetite*.

1 Solution-level

At a solution level, the confidence index helps AI teams assess:

a. whether a solution is ready to be deployed

b. whether the monitoring results are satisfactory or an appropriate mitigation plan needs to be triggered

2 Portfolio-level

At a portfolio level, the confidence index dashboard helps business and risk teams understand and action:

a. the risk distribution across Responsible AI dimensions, with a view on the business impact

b. priority mitigations or enhancements based on the institution's top priorities

3 Institution-level

At institution level, the confidence index dashboard enables CXOs to:

a. have a bottom up view of the institution's AI risk at various levels of granularity

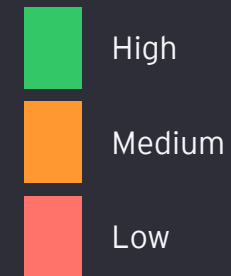b. align the AI strategy with the institution's risk appetite

EY

EY.ai confidence index dashboard representation at the portfolio level

Once the AI confidence index is assessed at solution level, it can be aggregated across the portfolio (per business line for example or per function) for a consolidated view of the portfolio's AI confidence weighted by business impact. This can also be done at institution level, enabling stakeholders to understand the level of AI confidence across the institution's solutions.

	Low	Medium	High	N/A
Accountability	% { ° ○ ◯ }	% { ° ○ ◯ }	% { ° ○ ◯ }	%
Bias and Fairness	% { ° ○ ◯ }	% { ° ○ ◯ }	% { ° ○ ◯ }	%
Explainability	% { ° ○ ◯ }	% { ° ○ ◯ }	% { ° ○ ◯ }	%
Privacy	% { ° ○ ◯ }	% { ° ○ ◯ }	% { ° ○ ◯ }	%
Reliability	% { ° ○ ◯ }	% { ° ○ ◯ }	% { ° ○ ◯ }	%
Security	% { ° ○ ◯ }	% { ° ○ ◯ }	% { ° ○ ◯ }	%
Sustainability	% { ° ○ ◯ }	% { ° ○ ◯ }	% { ° ○ ◯ }	%
Transparency	% { ° ○ ◯ }	% { ° ○ ◯ }	% { ° ○ ◯ }	%

Confidence Level

- High (green)
- Medium (orange)
- Low (red)

Relevance based on business impact

- ◯ High (%)
- ○ Medium (%)
- ° Low (%)
- • Not Applicable

EY.ai confidence index discussion paper

EY
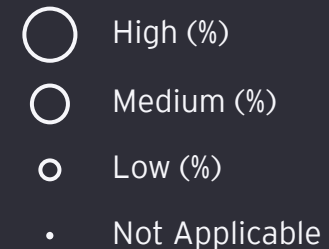
AI solution risk tiering informing the depth and frequency of solution testing

Enterprises have an opportunity to better optimize their resources by adapting the depth and frequency of solution testing to the solution risk level. An AI solution is assigned a low, medium, or high risk tier based on factors such as financial, non-financial, and compliance business impact, informing the depth and frequency of solution testing. Solution testing includes:

AI Solution Risk Tiering

Initial vetting refers to pre and post implementation review using the AI confidence index.

Ongoing Monitoring is to confirm that the solution continues to operate as expected over time after implementation, and to help identify the need for changes and enhancement should the AI confidence index deteriorate over time.

Ongoing Review refers to periodic reviews of the solution. Its objectives include re-assessment of the continued appropriateness of the solution and any benchmarking analysis, back-testing metrics, sensitivity analysis.
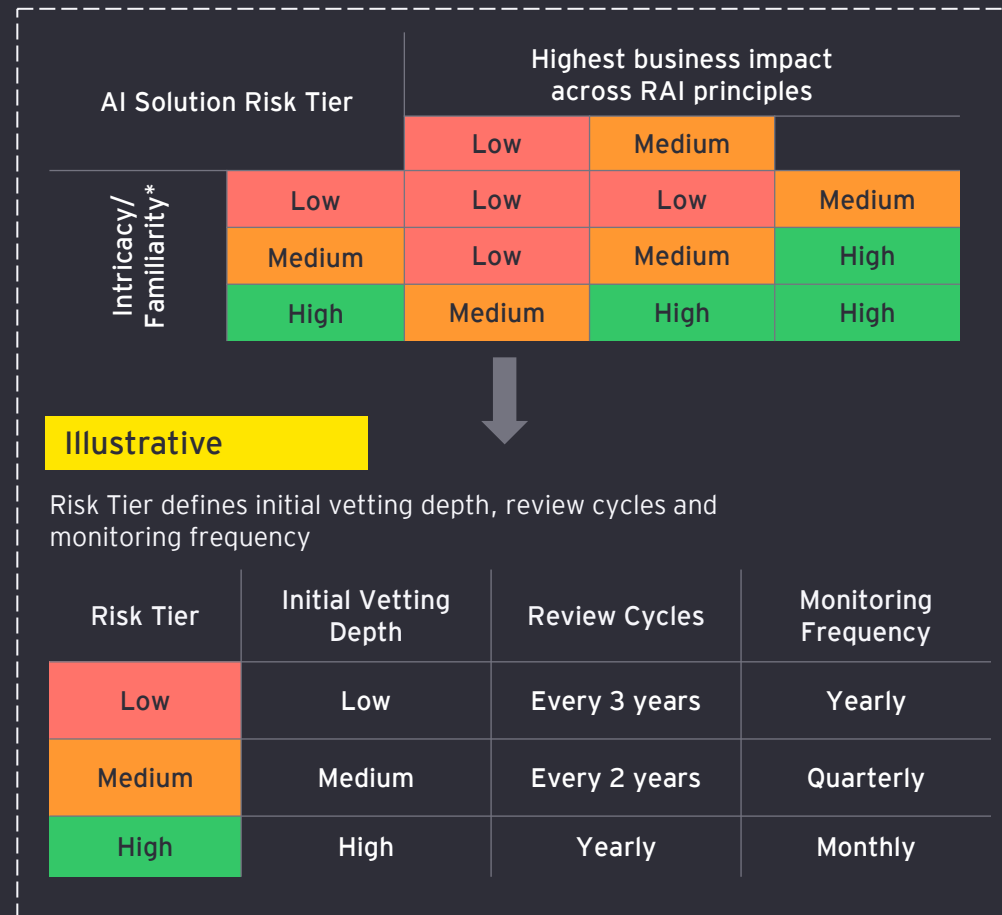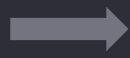
*As assessed by a supplementary questionnaire.** For a complex solution, the reliability RAI principles will reflect business relevance and confidence. They will be monitored at a depth/frequency commensurable with the business impact, and guardrails will be in place based on the institution's risk appetite (previous slide).

E-23 Regulatory Guideline, OSFI, Canada, Section 4.1 Model risk materiality
Senior Management should implement an appropriate model risk materiality classification scheme applicable to the relevant models. Its design and approval should be integrated with the governance structure for model approval. Size and complexity of model inventories may require, as appropriate, separate governance structures.

AI Solution Risk Tier — Highest business impact across RAI principles

AI Solution Risk Tier		Low	Medium		
Intricacy/ Familiarity*	Low	Low	Low	Low	Medium
	Medium	Low	Medium	High	
	High	Medium	High	High	

Illustrative

Risk Tier defines initial vetting depth, review cycles and monitoring frequency

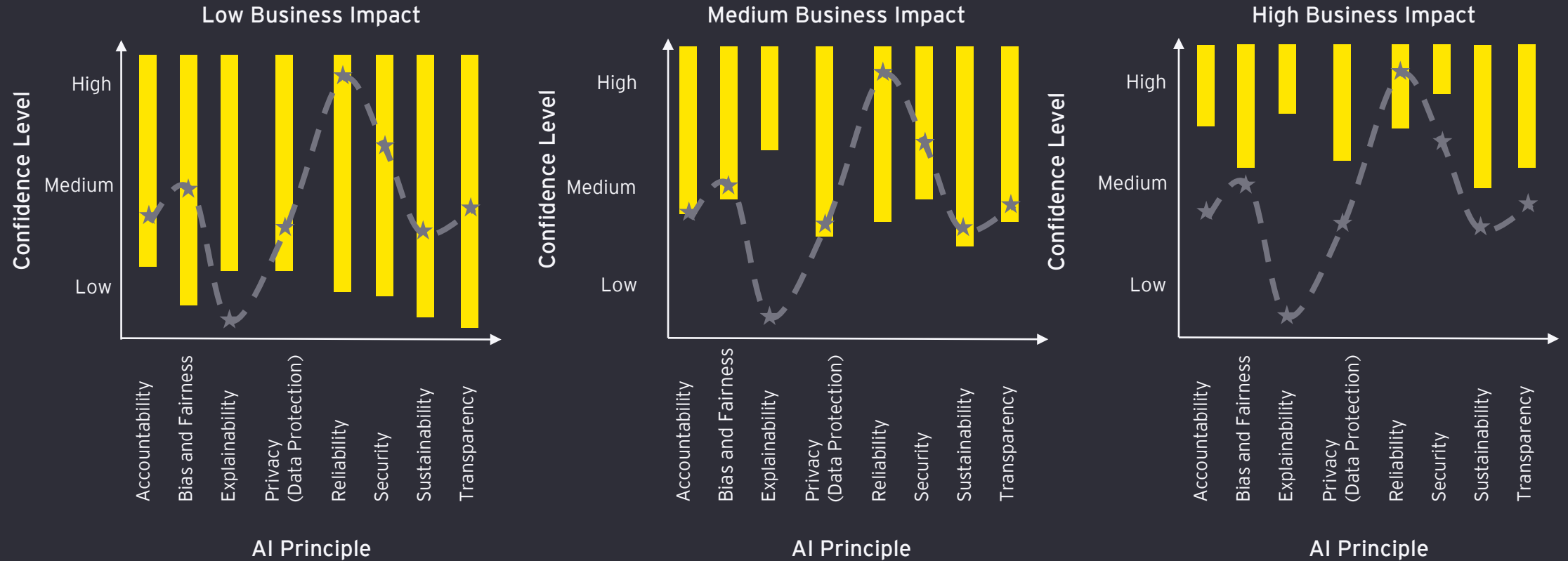| Risk Tier	Initial Vetting Depth	Review Cycles	Monitoring Frequency
Low	Low	Every 3 years	Yearly
Medium	Medium	Every 2 years	Quarterly
High	High	Yearly	Monthly

EY

A4: AI solution vetting additional triggers

After an AI solution is successfully vetted pre-deployment and moves into production, ongoing oversight is required. Changes to an AI solution or its environment in production may trigger additional vetting, to ensure the continued adequacy of the AI solution.

	AI Solution Vetting Triggers	Business Impact Changes*	Confidence Level Changes
System Changes	Data changes (sources or features)	Privacy, Transparency	All but Accountability
	Model changes	x (Risk Tiering - Intricacy)	All but Accountability
Environment Changes	Technology platform changes	x	All but Accountability
	Regulation changes	All affected	Depends on regulation
	Strategy changes	Potentially all	All
	Change/Expansion of Usage (e.g., portfolio acquisition)	Potentially all	All
Incidents	Model failure/harms (raised concern or occurred)	Potentially all	All affected dimensions

*Changes to the Business Impact may lead to a change in risk tiering, which in turn may trigger further assessments of the confidence dimensions.

EY.ai confidence index discussion paper

The EY.ai confidence index builds and executes an AI confidence framework tailored to the organization, to promote AI adoption and innovation

Illustrative

Low Business Impact

Confidence Level

- High
- Medium
- Low

AI Principle

Accountability · Bias and Fairness · Explainability · Privacy (Data Protection) · Reliability · Security · Sustainability · Transparency

Medium Business Impact

Confidence Level

- High
- Medium
- Low

AI Principle

Accountability · Bias and Fairness · Explainability · Privacy (Data Protection) · Reliability · Security · Sustainability · Transparency

High Business Impact

Confidence Level

- High
- Medium
- Low

AI Principle

Accountability · Bias and Fairness · Explainability · Privacy (Data Protection) · Reliability · Security · Sustainability · Transparency

Tolerance range for AI principle across organization   ★ Average value for AI principle across organization

EY

Benefits of the EY.ai confidence index

Rolling up the confidence index across AI solutions to portfolio level provides a view of confidence distribution across Responsible AI principles.

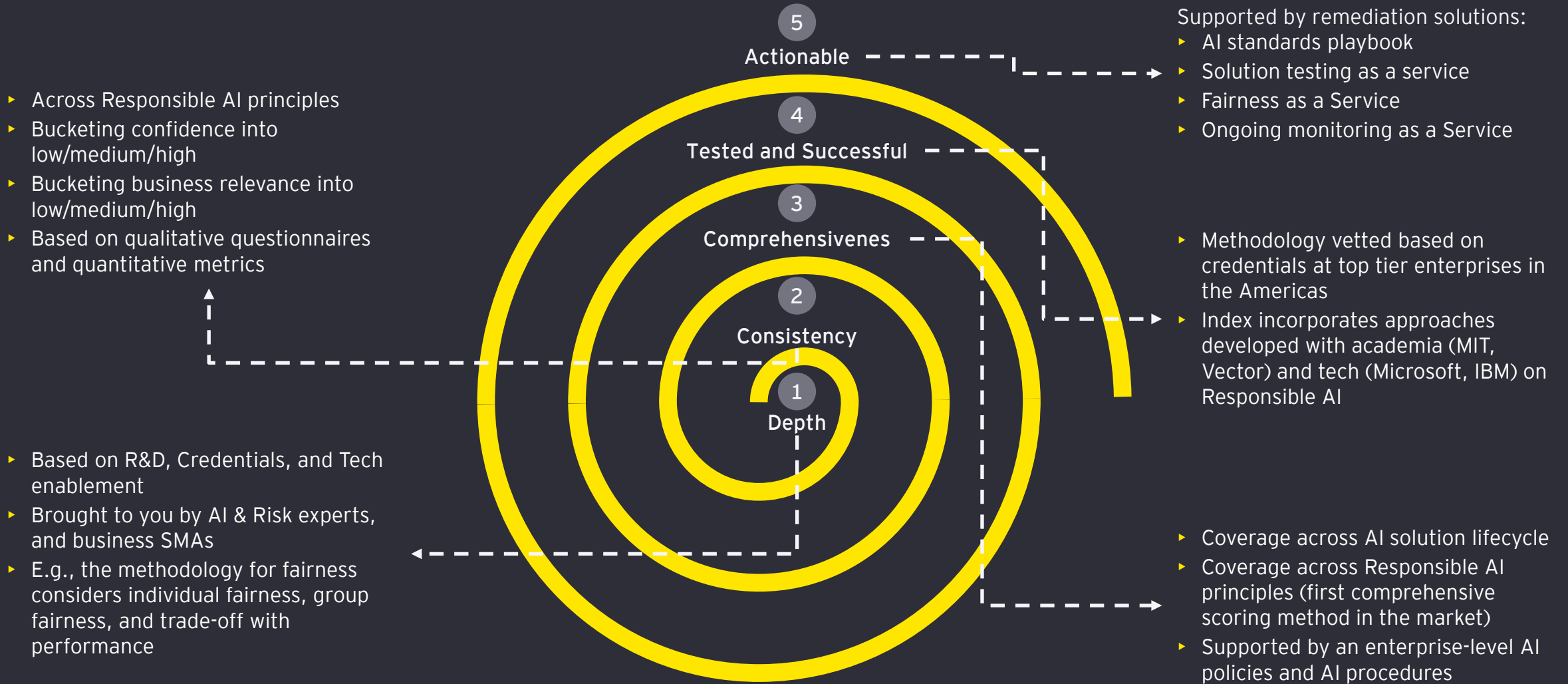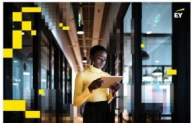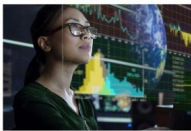The confidence landscape weighted by the business impact relevance enables the prioritization of enhancements and mitigations.

Continuous testing ensures AI output compliance with regulatory requirements.

The enterprise level confidence index dashboard can be contrasted with the enterprise's risk appetite to refine the AI strategy.

In-production confidence testing ensures AI models remain appropriate for use or else triggers an alert.

Confidence testing promotes better Citizen experience through transparency, reliability, privacy, and fairness.

03 Measurable targets

04 Cost-effective risk mitigation

02 Ongoing solution assurance

05 Regulatory compliance

10 Alignment with business objective

AI Governance

Value Creation

01 Accelerated solution deployment

06 Enhanced decision-making

Better citizen experience **09**

Improved operations **07**

User trust enhancement **08**

Confidence testing allows analytics teams to efficiently and effectively assess whether the solution is ready to be in production.

Confidence testing ensures AI models yield reliable, responsible outputs, enhancing decision-making and market competitiveness.

Confidence testing supports efficient and effective operations, based on reliable and explainable outputs.

Providing transparent AI reasoning through continuous confidence testing bolsters user trust.

EY

EY.ai Confidence Index market differentiators

5

Actionable

4

Tested and Successful

3

Comprehensivenes

2

Consistency

1

Depth

- Across Responsible AI principles
- Bucketing confidence into low/medium/high
- Bucketing business relevance into low/medium/high
- Based on qualitative questionnaires and quantitative metrics

- Based on R&D, Credentials, and Tech enablement
- Brought to you by AI & Risk experts, and business SMAs
- E.g., the methodology for fairness considers individual fairness, group fairness, and trade-off with performance

Supported by remediation solutions:
- AI standards playbook
- Solution testing as a service
- Fairness as a Service
- Ongoing monitoring as a Service

- Methodology vetted based on credentials at top tier enterprises in the Americas
- Index incorporates approaches developed with academia (MIT, Vector) and tech (Microsoft, IBM) on Responsible AI

- Coverage across AI solution lifecycle
- Coverage across Responsible AI principles (first comprehensive scoring method in the market)
- Supported by an enterprise-level AI policies and AI procedures

EY.ai confidence index discussion paper

EY

EY.ai Confidence Index service offering differentiators

Leveraging generative AI for model documentation automation

AI-enabled early warnings signals framework

AI models ongoing monitoring discussion paper

Vetted based on client experience and credentials

‣ AI governance build out in Financial Services, Healthcare, Nutrition, Manufacturing

‣ AI and Gen AI Risk solution provider across all 5 (non-audit) top tier banks in the Americas

‣ AI solutions spanning AI techniques and methodologies, across business functions

Confidence Aggregation methodology is EY's IP

‣ Playbook of 50+ cutting edge solutions

‣ Agile and ongoing innovation e.g., first Reinforcement Learning model in Banking in North America, RL hedging at top tier bank

‣ Helped Microsoft integrate individual fairness in their tech stack

‣ Compared to most recent academic publications (e.g., Paolo Giudici et al. "Artificial Intelligence risk measurement, Expert Systems with Applications"), our solution is much more advanced on coverage, depth, and rigor

The confidence index is the foundational capability to support strategy decisions

‣ Investors: improves ROI on AI and financial KPIs

‣ C-suite: improves operational efficiency by strategic AI deployment

‣ Customers: improves service levels and product quality through transparent AI use

‣ Employees: improves internal processes and builds workforce cohesion around AI use

The Confidence Index is the risk measurement basis for a comprehensive framework for AI governance

‣ Clear KPIs and metrics for measurement and monitoring of AI solutions

‣ Complements XOps tools and AI standards and regulations from strategic standpoint

‣ Hierarchical analytics and decision support framework

Solution ⇒ Portfolio ⇒ Organization

Calculation Methodology for each principle is based on Rigorous scientific foundation and based on applied research

‣ 6-year investment in R&D

‣ 25+ peer-reviewed publications with collaborators from MIT, UC Berkeley, UofT, Mila, Vector, Databricks, Microsoft, IBM

‣ Methodology based on ~110+ academic research papers

‣ 50+ PhD and SMEs on AI, technology, ethics and privacy, compliance, sustainability, business line experts

The methodology is developed to support the solution offering across regions and sectors

‣ Incorporates gold standard AI regulations across the world supported by collaboration with regulators (publication on fairness with Singapore authorities, participation in sessions with Canadian Regulator OSFI, support to Ontario Securities Commission on AI for Capital Markets)

‣ Applies across sectors with a possibility to tailor based on enterprise input and risk appetite
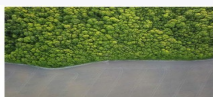

Responsible AI and AI governance discussion paper

Identifying, measuring, and mitigating individual unfairness | Assessing and mitigating unfairness in credit models

Data quality management discussion paper

Artificial intelligence ESG stakes discussion paper

EY

EY.AI Confidence Index: Coverage of existing standards and regulations

Corporate Governance Policies

‣ Comprising a suite of policies, frameworks and procedures, corporate governance serves as the guiding force for decision-making, risk management, and stakeholder relationships.

‣ Policies delineating the ethical use of AI, frameworks addressing risk assessments and mitigation strategies, and procedures ensuring transparency in AI related decision-making collectively contribute to a governance structure that fosters trust and accountability.

Voluntary Guidelines and Standards

‣ Complementing corporate governance, voluntary AI guidelines and standards provide additional layers of guidance and best practices; developed by international organizations, industry bodies and expert consortia, these guidelines offer a global perspective on ethical AI design and deployment.

Examples

‣ **OECD AI Principles***
‣ Asilomoar AI Principles
‣ IEEE's Ethically Aligned Design
‣ Montreal Declaration for a Responsible Development of Artificial Intelligence
‣ **NIST AI Risk Management Framework***
‣ OECD Framework for Classification of AI Systems
‣ White House's voluntary commitments from leading AI companies
‣ **Canada's generative AI code of conduct***
‣ Responsible Artificial Intelligence Institute

Laws and Regulations

‣ As AI technologies advance, governments and regulatory bodies worldwide are introducing legislated frameworks to address the ethical, societal and legal implications of AI.
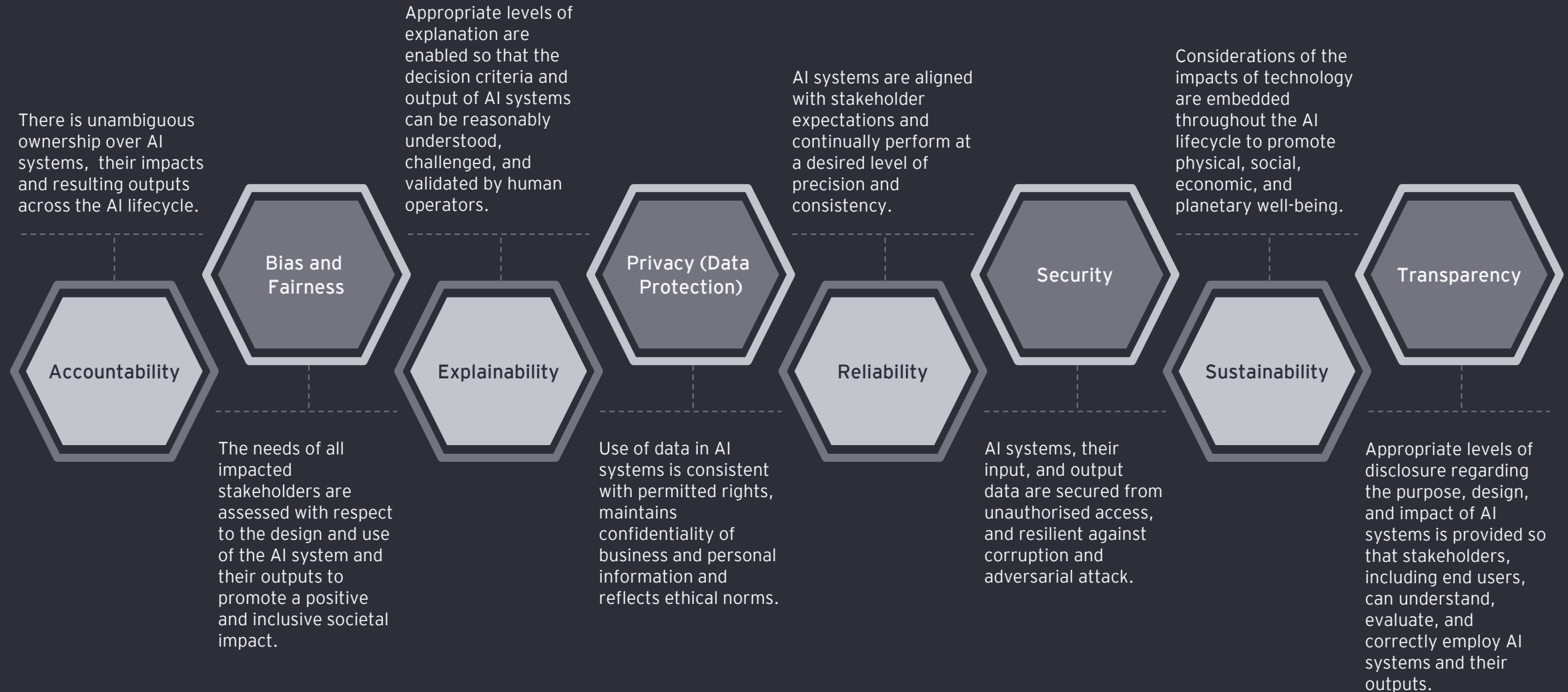
Examples

‣ **AI Act (EU)**
‣ **The Artificial Intelligence and Data Act (Canada)***
‣ NYC Local Law 144 of 2021 (regulating automated employment decision tools)
‣ American Data Privacy and Protection Act, Section 207 (US)
‣ Internet Information Service Algorithm Recommendation Management Regulations (China)
‣ ISO/IEC JTC 1/SC 42
‣ IEEE P7000 series of standards projects
‣ CEN/CENELEC standards development
‣ RAI Institute's Certification Program for AI Systems

Responsible AI principles applied for the EY.ai confidence index

There is unambiguous ownership over AI systems, their impacts and resulting outputs across the AI lifecycle.

Appropriate levels of explanation are enabled so that the decision criteria and output of AI systems can be reasonably understood, challenged, and validated by human operators.

AI systems are aligned with stakeholder expectations and continually perform at a desired level of precision and consistency.

Considerations of the impacts of technology are embedded throughout the AI lifecycle to promote physical, social, economic, and planetary well-being.

Bias and Fairness

Privacy (Data Protection)

Security

Transparency

Accountability

Explainability

Reliability

Sustainability

The needs of all impacted stakeholders are assessed with respect to the design and use of the AI system and their outputs to promote a positive and inclusive societal impact.

Use of data in AI systems is consistent with permitted rights, maintains confidentiality of business and personal information and reflects ethical norms.

AI systems, their input, and output data are secured from unauthorised access, and resilient against corruption and adversarial attack.

Appropriate levels of disclosure regarding the purpose, design, and impact of AI systems is provided so that stakeholders, including end users, can understand, evaluate, and correctly employ AI systems and their outputs.

EY.ai confidence index discussion paper

EY

Business impact domains calculated in the EY.ai confidence index

Domain	Scope
Financial Impact	Level of adverse financial impact (i.e., decrease in revenue and/or increase in cost) the client could face if the AI Model does not perform as expected.
Operational Efficiency	Impact on the operational efficiency of the client if the AI Model does not perform as expected.
Citizen Experience	Impact of AI model on Citizens and/or internal users.
Compliance	Ensures the design, implementation and use of AI systems and their outs comply with relevant laws, regulations, and professional standards.

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.com