How can Trusted AI accelerate business transformation?

A roadmap to Trusted AI through re-imagined Governance & Risk Management

EY PoV - May 2024

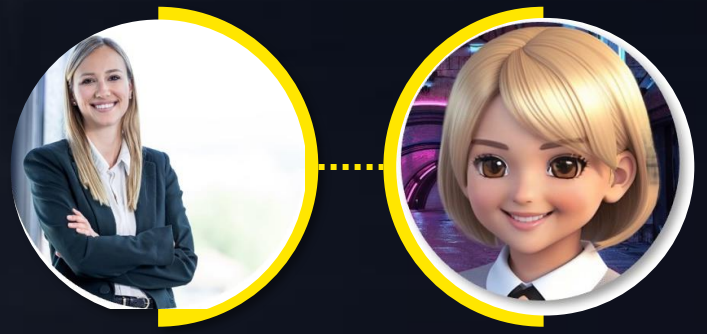
The better the question. The better the answer.
The better the world works.

Building a better working world

Meet our speakers

Iuliia Metitieri

Iuliia is a Senior Manager in EY's Risk Consulting practice in Switzerland and the Responsible AI solution lead in Switzerland.



Janine Luginbühl

Janine is a Consultant in EY's Risk Consulting practice in Switzerland, focusing on Responsible AI & AI Governance.

Agenda

1. What is Trusted AI and why is it important?
2. What are common challenges on the way to Trusted AI?
3. How to establish a Trusted AI Governance and Risk Management?
4. Roadmap to Trusted AI: how to get started?

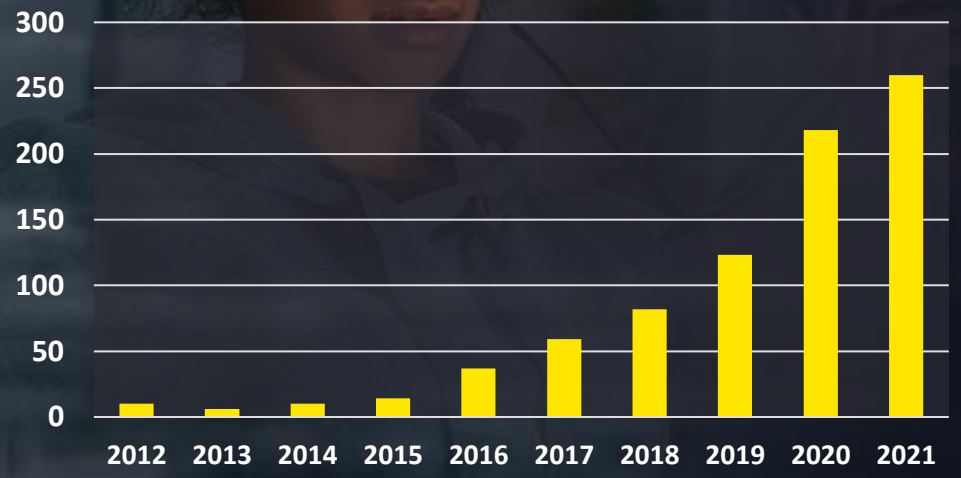
1

What is Trusted AI and why is it important?

Why do we talk about Trusted AI?

Trustworthiness is a prerequisite for people and societies to develop, deploy and use AI systems.

Annual reported AI incidents and controversies



Source: AIAAIC Repository (2022) via AI Index (2023).

Early experience with AI emphasizes the need for sound AI governance and risk management in order to safeguard against a variety of related 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 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 2018 Amazon Echo:

sent the conversation of a family to a random person in their contacts.

What makes AI trustworthy?

Eight interrelated **AI attributes** to create an AI ecosystem that is beneficial, trustworthy and effectively compliant.

Accountability

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

Bias & Fairness

Minimizing biases in AI algorithms to prevent discrimination against any individual or group.

Explainability

Decision criteria and output of AI systems can be reasonably understood, challenged, and validated by human operators.

Privacy Protection

Safeguarding personal and sensitive data processed by AI systems.

Reliability

Ensuring AI systems are stable and function as intended under various conditions.

Security

AI systems are secured from unauthorised access, and resilient against adversarial attacks.

Sustainability

Considerations of the impacts include physical, social, economic, and planetary well-being.

Transparency

Ensuring that AI systems and their decisions can be understood by end-users and stakeholders.

What is the value of Trusted AI for organizations?

1. Regulation

The regulatory landscape surrounding AI is rapidly evolving across regions — clients need to prepare for impending obligations or else risk record fines.

- EU AI Act
- Non-compliance can result in fines of up to €35m or 7% of global turnover.

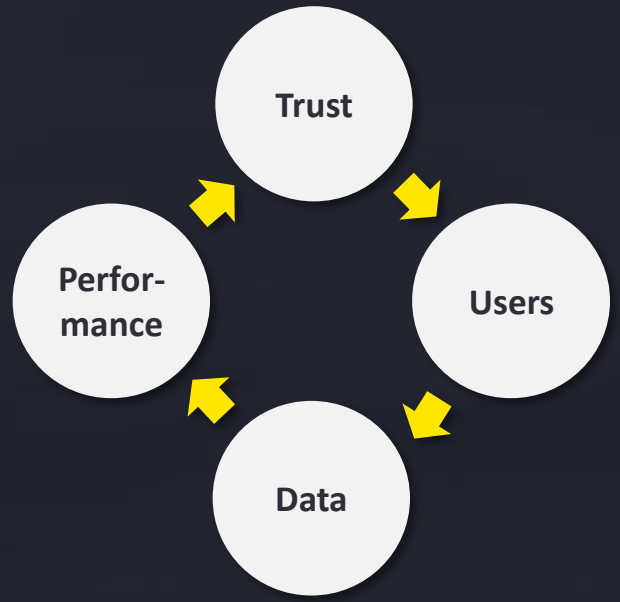
2. Reputation

Unethical AI practices leading to societal harm can carry serious reputational risks for organizations, translating into commercial losses and brand value degradation.

- Algorithmic discrimination
- Hallucinations
- Data breach
- Data misuse
- Ethics-washing

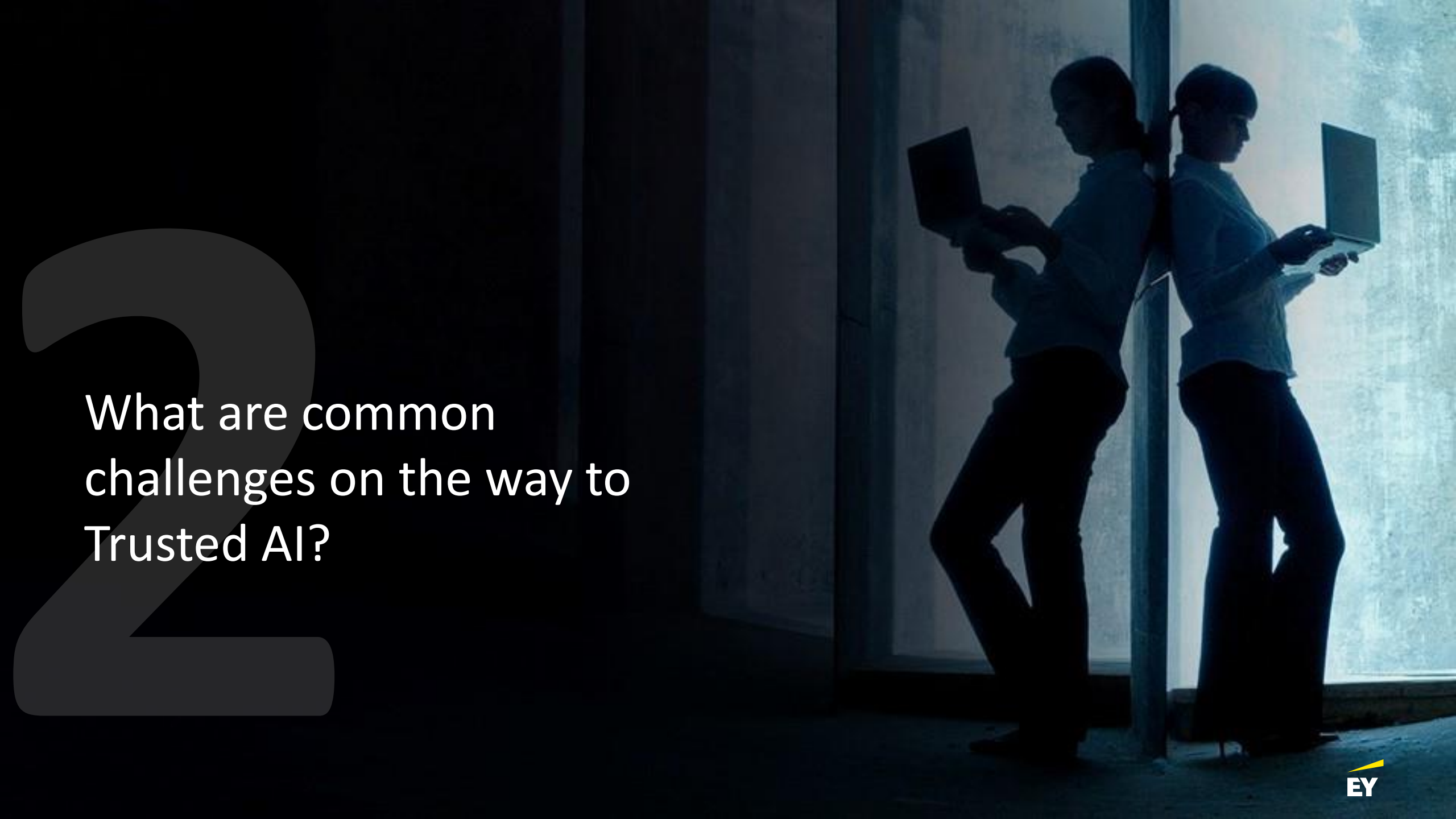
3. Realisation

Trusted AI is not just about risk — it's about realizing value. There is a close relationship between highly trusted and high-performing technologies.



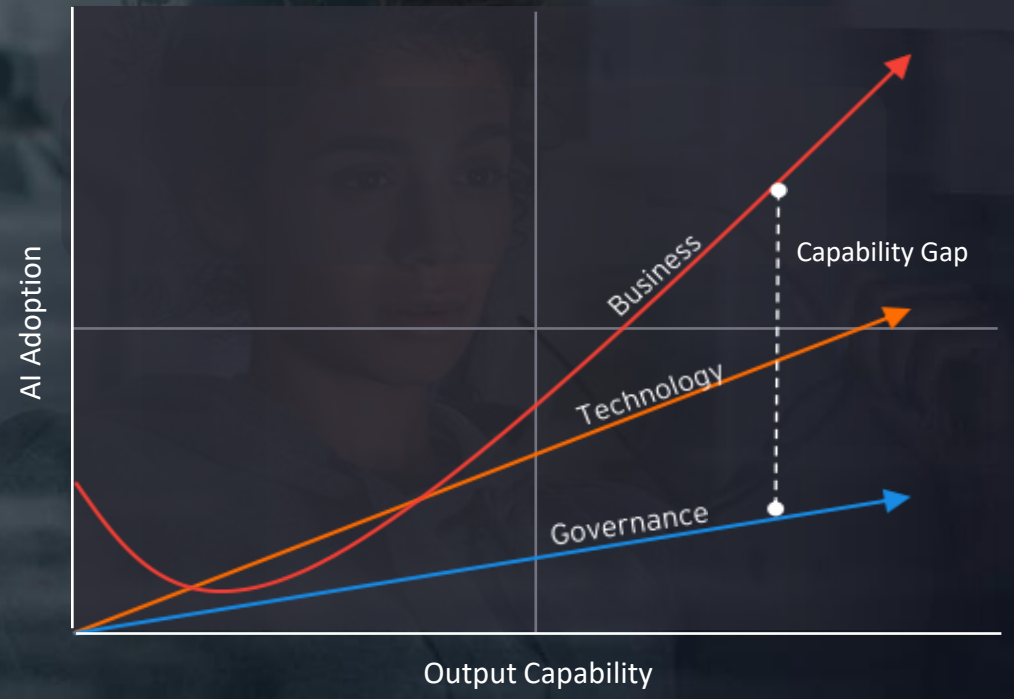
2

What are common challenges on the way to Trusted AI?



How to manage the challenge of Confident Adoption at Speed and Scale?

AI Capability & Adoption Relationship



Market growth will accelerate

Hence, AI can be the true game-changer

But the growth in AI adoption and technological advances is only one part of a bigger story

The speed of AI development is far outpacing the regulations. Regulatory bodies across the world are working on closing the gap.

What are some of the lessons learned from the early days of Governing AI?

1. Misalignment between an organization's AI strategy & initiatives
2. AI governance can be a multiplier for compliance fatigue
3. Translation of AI principles to engineering requirements is missing
4. AI lacks proper profiles and technology enablement
5. The pace and nature of AI require new risk and control mechanisms

What does the general AI maturity across industries look like?

Level 4 'Leaders'

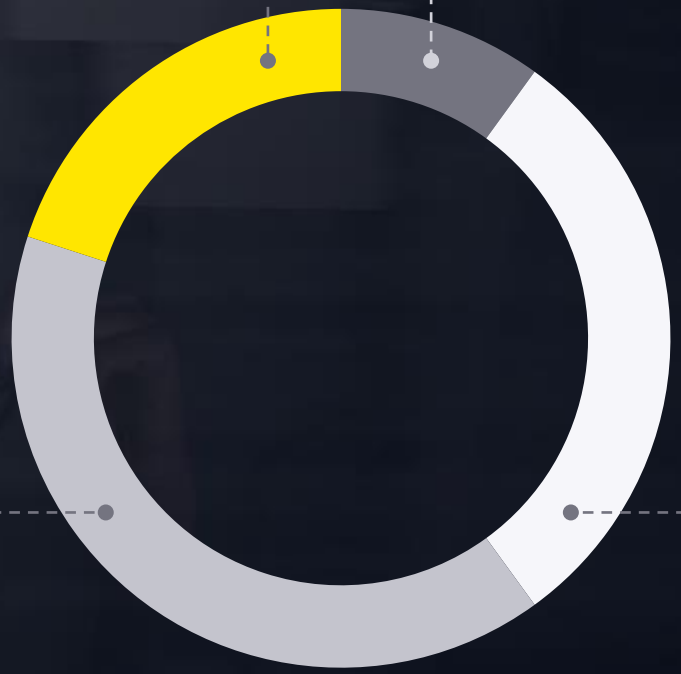
20%

- Trusted AI Principles
- AI Policies
- AI Capabilities
- AI Operating Model
- Integrated Tools & Technology

Level 1 'Wakeup'

10%

- Trusted AI Principles
- AI Policies
- AI Capabilities
- AI Operating Model
- Integrated Tools & Technology



Source: Privacy and AI Governance Report, IAPP, 2023.

Level 3 'Transformers'

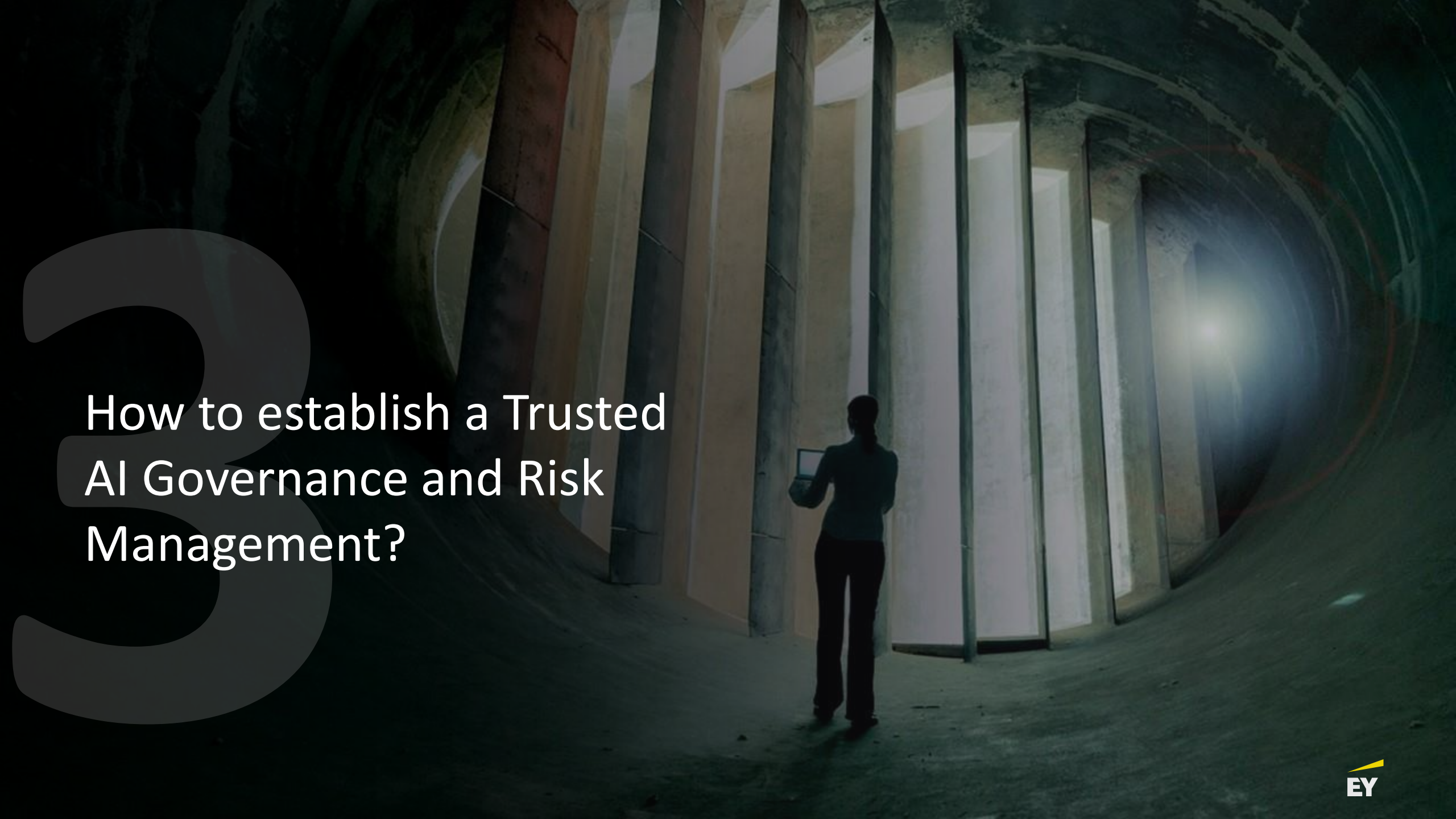
40%

- Trusted AI Principles
- AI Policies
- AI Capabilities
- AI Operating Model
- Integrated Tools & Technology

Level 2 'Starters'

30%

- Trusted AI Principles
- AI Policies
- AI Capabilities
- AI Operating Model
- Integrated Tools & Technology



3

How to establish a Trusted AI Governance and Risk Management?

What does a holistic Trusted AI framework look like (EY PoV)?

At the core of Trusted AI Framework are the Trusted AI Principles, implemented throughout the AI Lifecycle and four Pillars of Trusted AI on enterprise level as well as on individual AI solution level.

Trusted AI Principles

Trusted AI Pillars

Enterprise-level

1. AI governance

Establishes a foundation for agile, vigilant and purposeful trusted AI program that aligns with your organization's strategic vision and objective to drive competitive advantage

2. AI risk management enablement

Utilizes a holistic, systematic and flexible set of iterative processes that help drive early detection, evaluation, quantification and continuous management of risk throughout the AI lifecycle

Solution-level

3. AI Confidence Index

Provides a measurable methodology to quantify risks and enable confidence in the data and technology, process and people components that make up your AI ecosystem based on your trusted AI principles

4. AI ModelOps

Provides a holistic and technology-enabled approach to streamline end-to-end operations throughout the AI lifecycle. It promotes scalability through building and deploying AI solutions consistently

What does a holistic Trusted AI governance model look like (EY PoV)?

TRUSTED AI GUIDING PRINCIPLES

AI Strategy | AI Policy | AI Governing Body | AI Inventory | AI Data Governance

AI Organizational Structure | AI Roles and Responsibilities | AI Competencies and Capabilities | AI Interaction Model

AI Lifecycle Management

- AI Evaluation & Prioritization
- AI Lifecycle Processes & Procedures
- Documentation & Records
- Incident Reporting and Information

- AI Regulatory scanning
- AI Risk and Control Framework
- AI Compliance Monitoring
- AI 3rd-Party Risk Management

AI Risk Management

TRUSTED AI PROGRAM MAINTENANCE

Objectives and KPIs | Change Management | Program Independent Evaluation | Program Continuous Improvement

How can AI Risk Management be enabled?

AI risk management should provide a comprehensive suite of methodologies or tools to support AI risk management framework at the enterprise level.

AI Risk Management Enablement

- **AI Risk Oversight**
- **AI Risk Assessments**
- **AI Risk Tiering**
- **AI Risk Mitigation and Controls**
- **AI Solution Inventory**

AI governance risks | AI operational risks | AI model risks | AI strategic risks

What is the AI Solution Confidence Level and how can it be used?

The AI confidence index quantifies the understanding of risk related to an AI solution for continuous monitoring across the AI lifecycle and aggregation on enterprise level. To achieve that, Trusted AI dimensions are weighted by their business impact.

Solution level - Illustrative

Strategy and Value enablement

Business Impact

- Compliance
- Non-Financial Impact
- Financial Impact

Institution-level

- AI Strategy

Portfolio-level

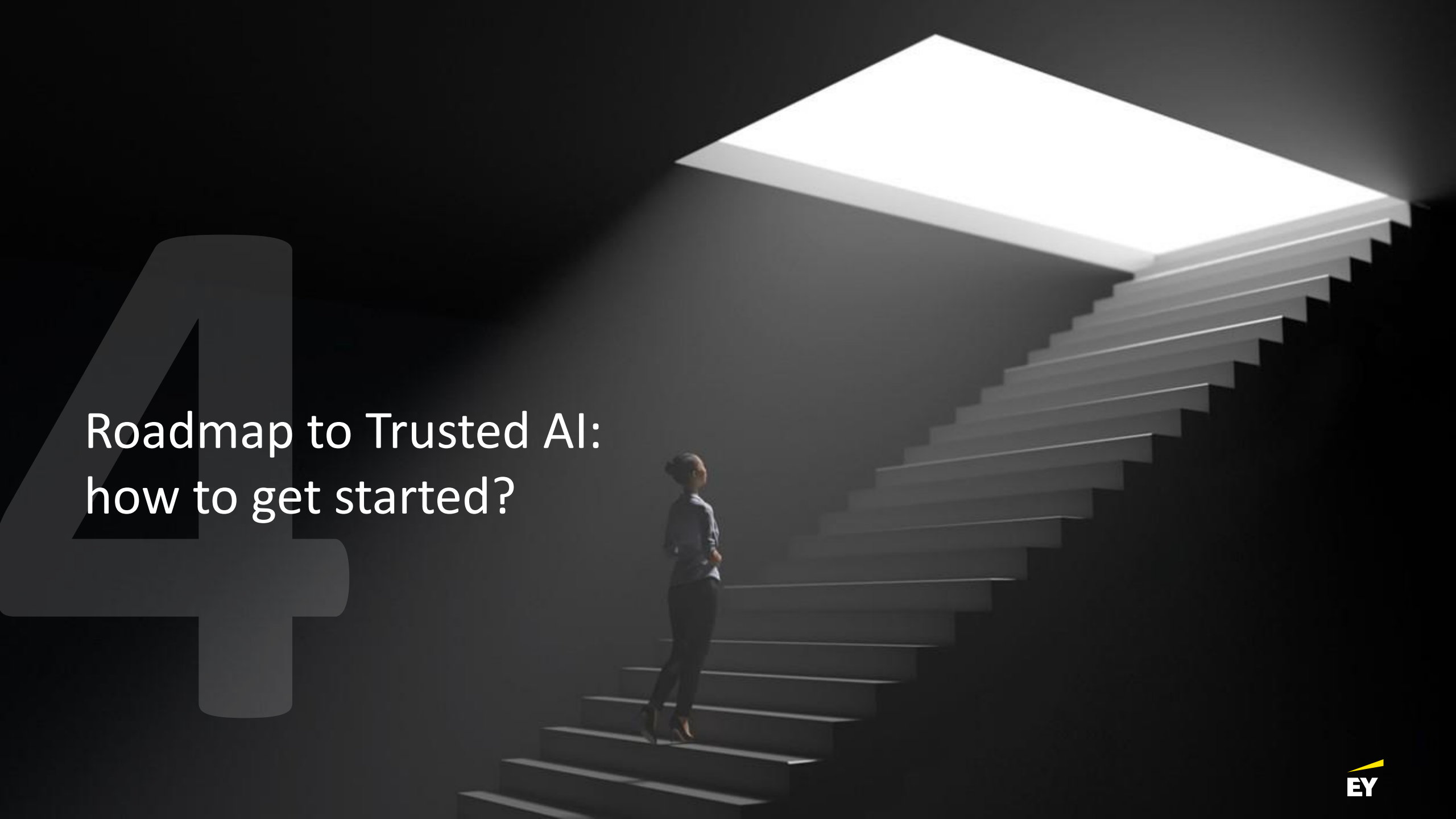
- Risk Distribution
- Enhancement Prioritization

Solution-level

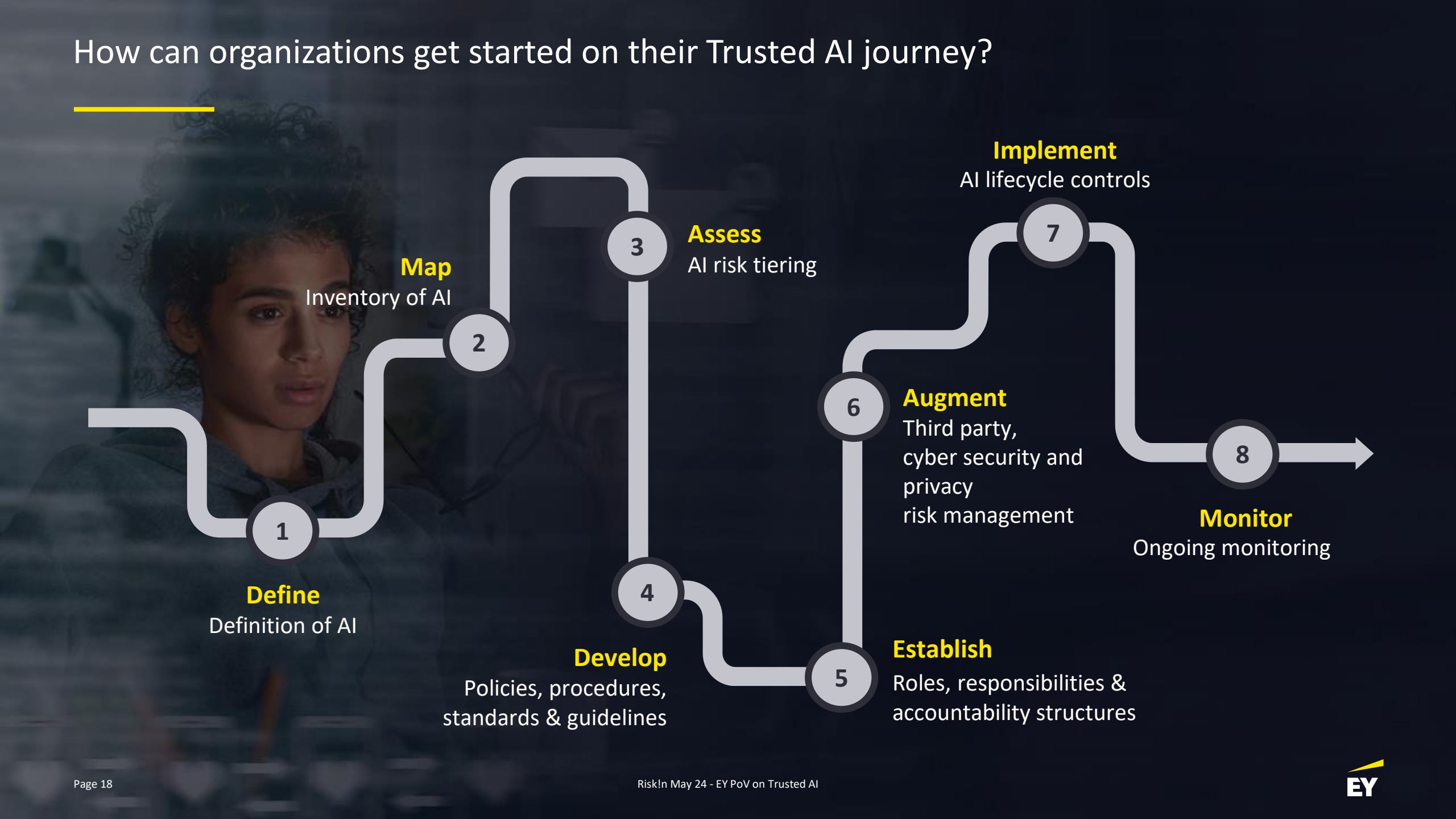
- Deployment Readiness
- Monitoring and Mitigation

4

Roadmap to Trusted AI: how to get started?



How can organizations get started on their Trusted AI journey?



EY | Building a better working world

About EY
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.

EYG no. 006995-22Gbl
ED NONE

In line with EY's commitment to minimise its impact on the environment, this document has been printed on paper with a high recycled content.

ey.com