LEAN COMPLIANCE

AI GOVERNANCE, GUARDRAILS, & LAMPPOSTS

ELEVATE COMPLIANCE WEBINARS

Raimund Laqua, PMP, P.Eng
ray.laqua@leancompliance.ca



Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

ABOUT LEAN COMPLIANCE

Our mission – Helping you stay between the lines and ahead of risk

Ensuring Mission Success Through Compliance

Helping you Stay Between the Lines and Ahead of Risk



Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

LEAN COMPLIANCE

ELEVATE COMPLIANCE HUDDLES

Mondays @ Noon on Zoom



ELEVATE COMPLIANCE HUDDLE

Elevating Compliance to Higher Standards

MONDAYS @ NOON ON ZOOM

Principles-to-Practice

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

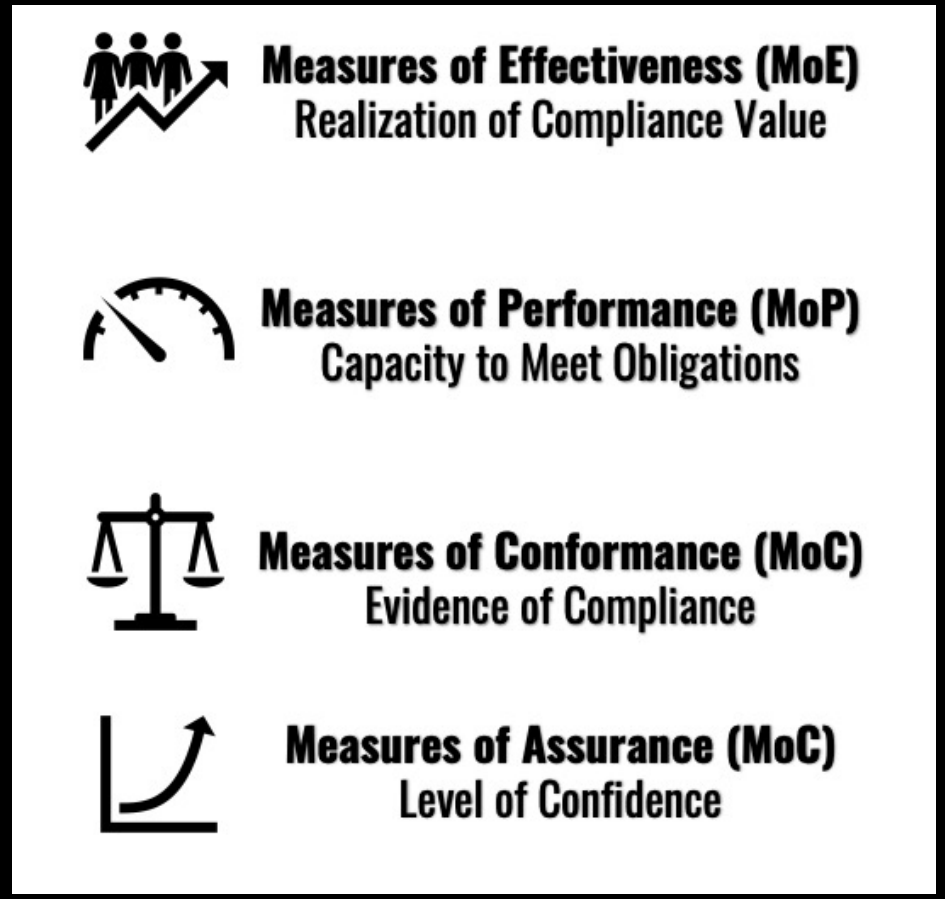
MONTHLY ELEVATE COMPLAINCE WEBINARS

Next month

ELEVATE COMPLIANCE USING METRICS THAT MATTER MOST

ELEVATE COMPLIANCE WEBINARS

Raimund Laqua, PMP, P.Eng
ray.laqua@leancompliance.ca



Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

BIG IDEA

When it comes to AI, we are frogs in a boiling pot of water



"The fundamental difference between AI and other technologies is that it has the potential to become smarter than its creators."

- Stuart Russell

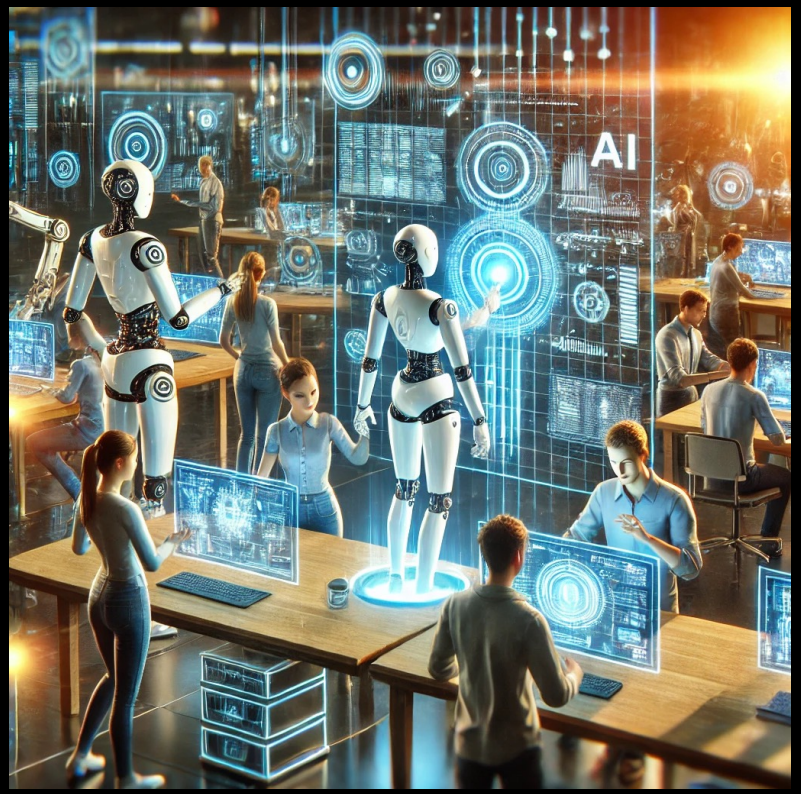
"We need to find meaning in a world where AI can do everything we can do, but better."

– Elon Musk

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

BIG IDEA

AI is the beginning of a new era of industrialization



How do you govern when the rules of the game are changing?

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

BIG IDEA

In a landscape that is uncertain and unpredictable, traditional governance may not be enough

Industry 5.0



Industry 4.0

You will need more than policies and principles; you will need guardrails and lampposts too

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

OUTLINE

1. Why AI is a Different Kind of Risk
2. The Problem with AI Regulation
3. What is Governance?
4. Methods of AI Governance
5. A Program to Govern AI

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

1. WHY AI IS A DIFFERENT KIND OF RISK

What kind of risk is AI, and why this matters



Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

1. WHY AI IS A DIFFERENT KIND OF RISK

The opportunities and threat of AI



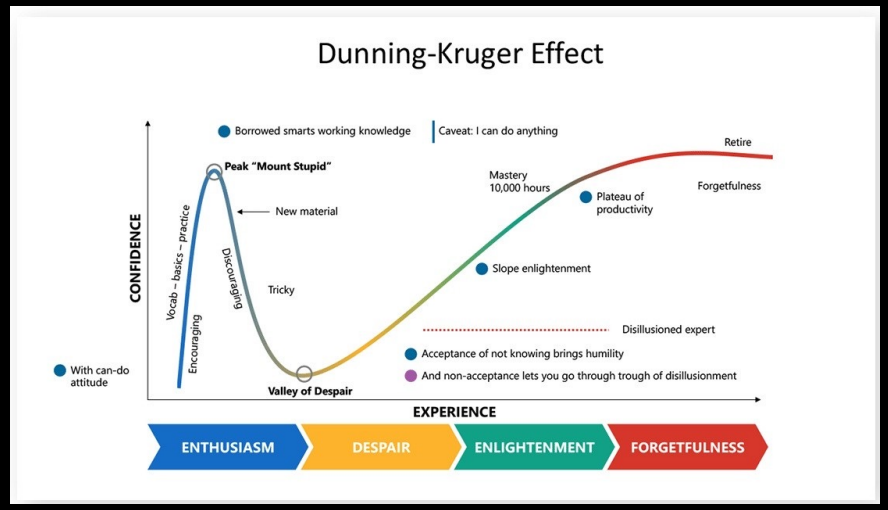
AI will …

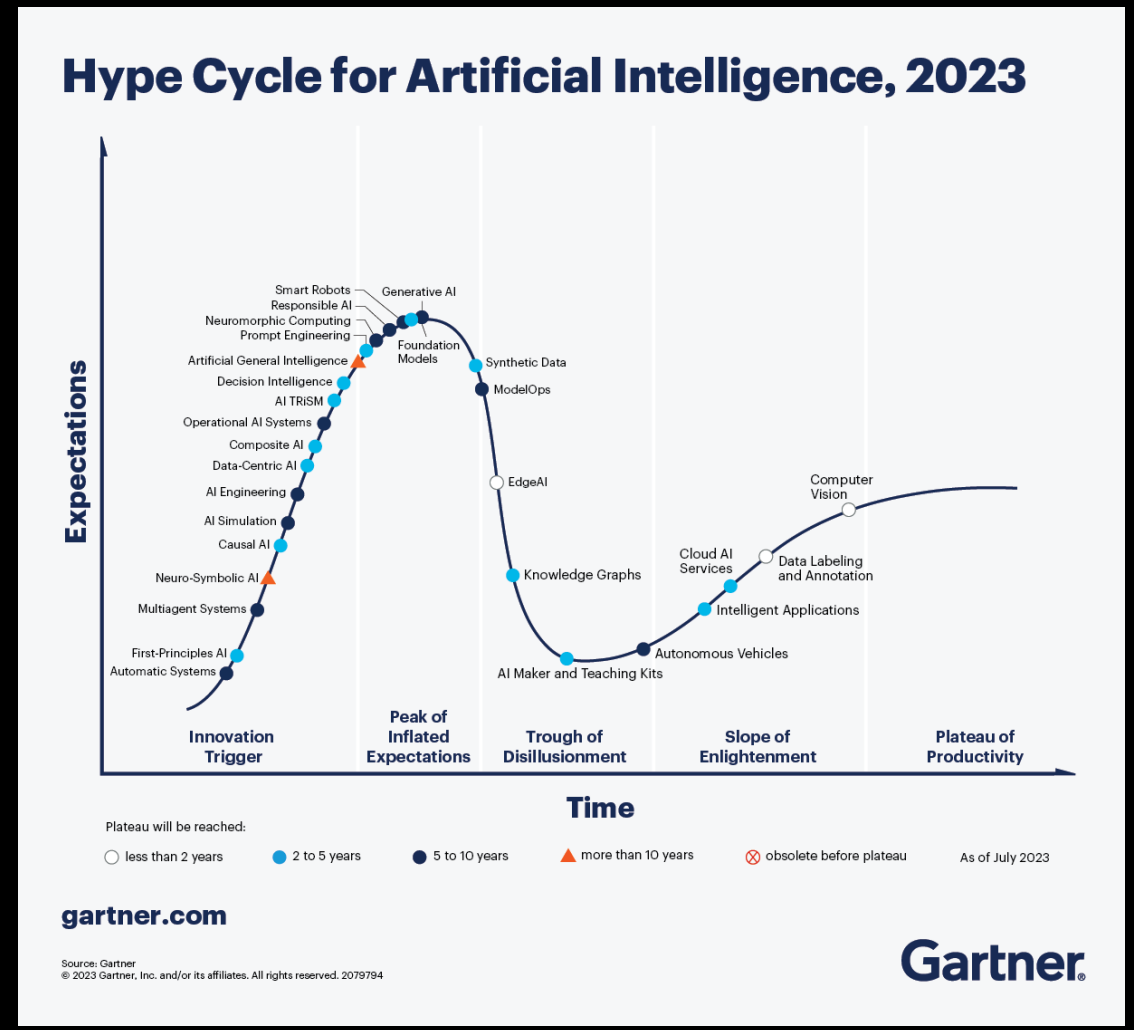
- Solve critical challenges facing the world: healthcare, climate change, economic growth, etc.
- Eliminate jobs, and businesses
- Transform work
- Create new jobs, technologies, and opportunities
- Attack businesses, organizations, critical infrastructure, and governments
- Increase surveillance and monitoring
- Accelerate the creation of gene editing tools
- Turn businesses into autonomous robots
- Double data center energy consumption by 2030 (1-2% to 3-4%)

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

1. WHY AI IS A DIFFERENT KIND OF RISK

Negotiating the hype

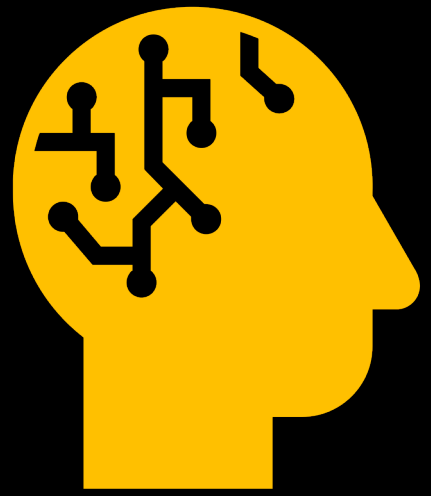




Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

1. WHY AI IS A DIFFERENT KIND OF RISK

What is artificial intelligence?



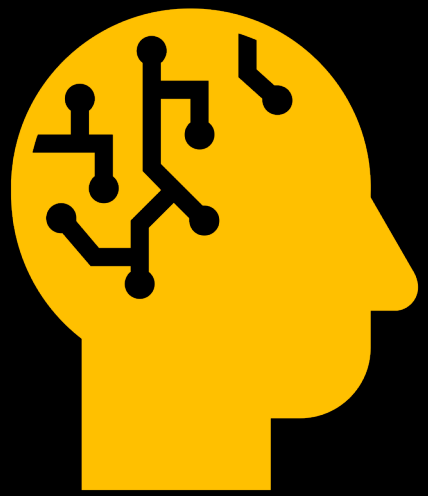
Artificial Intelligence (US DOJ Compliance Program Evaluation Update):

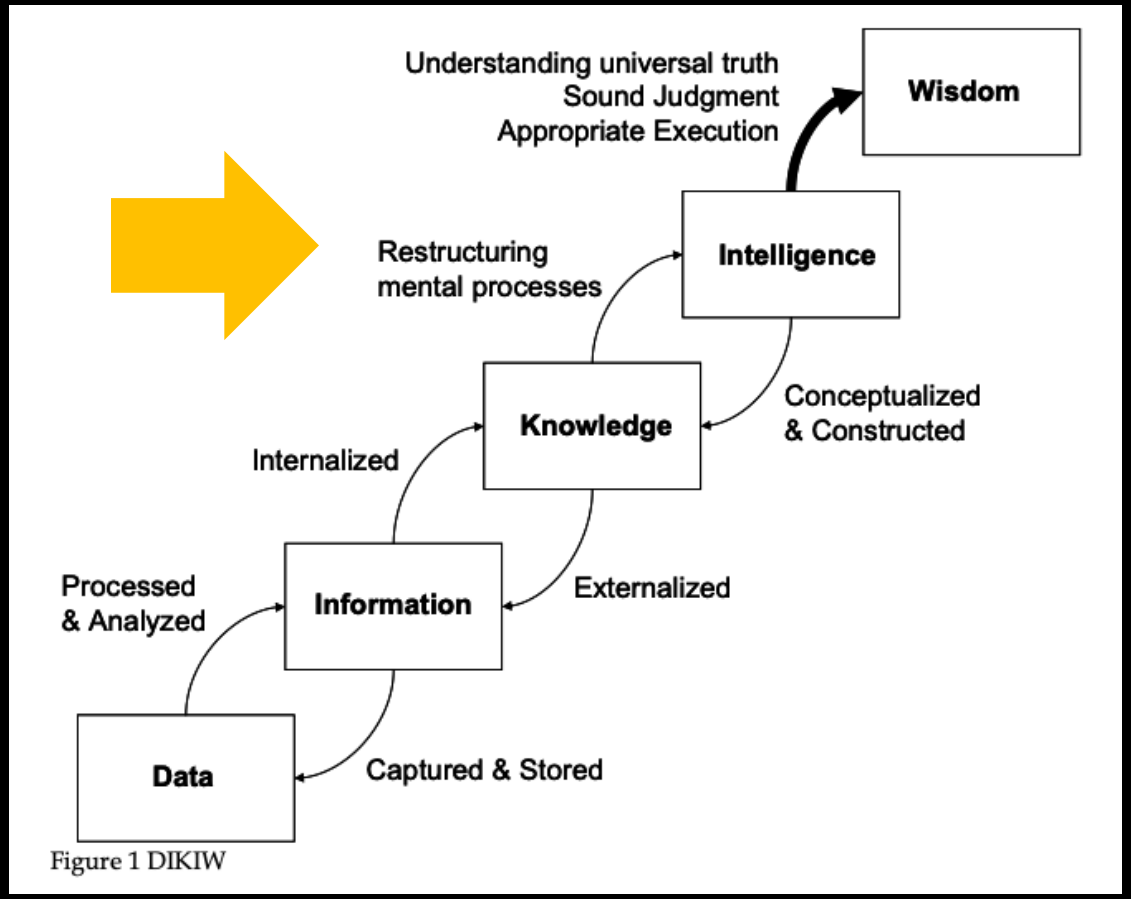
1. Any artificial system that performs tasks under varying and unpredictable circumstances without significant human oversight, or that can learn from experience and improve performance when exposed to data sets.
2. An artificial system developed in computer software, physical hardware, or other context that solves tasks requiring human-like perception, cognition, planning, learning, communication, or physical action.
3. An artificial system designed to think or act like a human, including cognitive architectures and neural networks.
4. A set of techniques, including machine learning, that is designed to approximate a cognitive task.
5. An artificial system designed to act rationally, including an intelligent software agent or embodied robot that achieves goals using perception, planning, reasoning, learning, communicating, decision making, and acting.

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

1. WHY AI IS A DIFFERENT KIND OF RISK

What is intelligence?





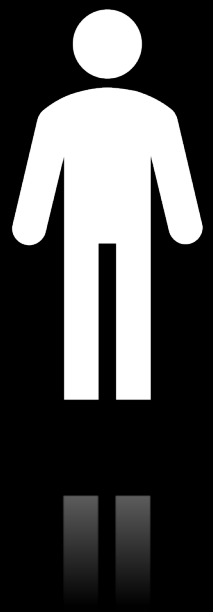
DIKIW: Data, Information, Knowledge, Intelligence, Wisdom and their Interrelationships, 2013, Anthony Liew

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

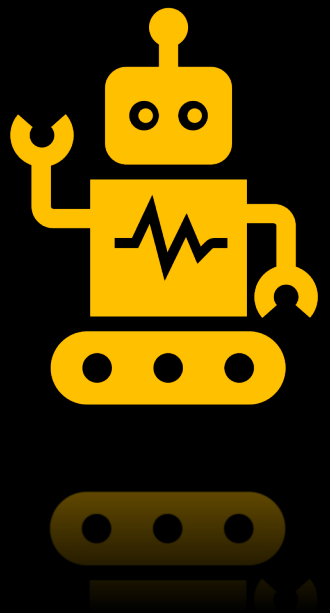
1. WHY AI IS A DIFFERENT KIND OF RISK

Risk – What Risk?

RISK



NO RISK

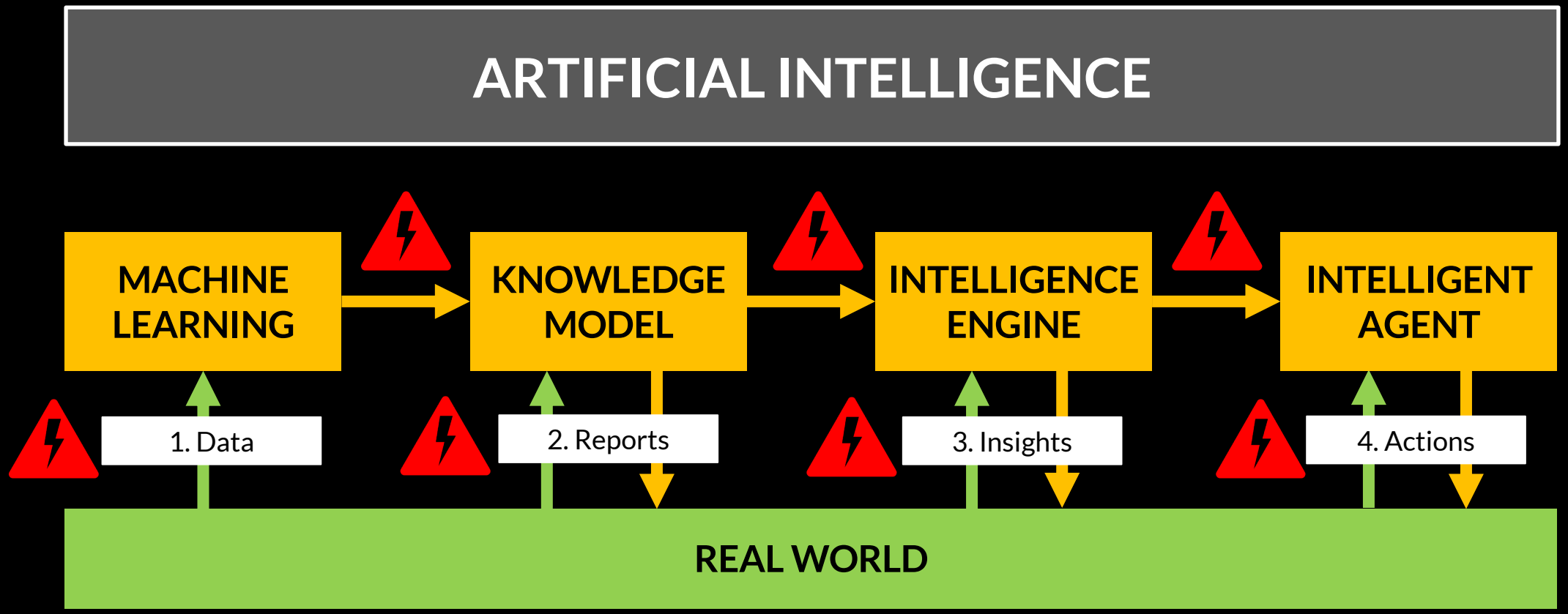


Some claim technology, especially AI, creates no threat.
It's only people that use AI who create the risk.

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

1. WHY AI IS A DIFFERENT KIND OF RISK

Inherent Risk

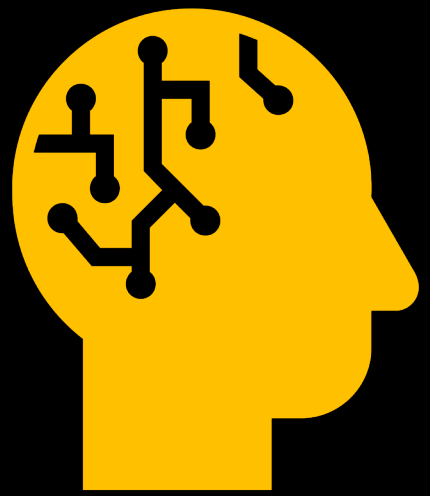


Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

1. WHY AI IS A DIFFERENT KIND OF RISK

The elephant in the room - Emerging Risk

What if you built intelligent machines that help you create better intelligent machines?

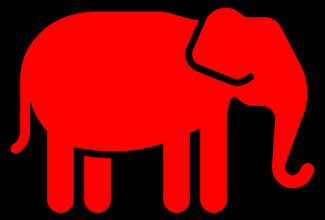


=

FIELD OF STUDY (SCIENCE)

+

INTELLIGENT MACHINES (ENGINEERING)



Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

2. THE PROBLEM WITH AI REGULATION

Will it be too late, and will it be enough?



Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

2. THE PROBLEM WITH AI REGULATION

AI regulation is slowly being introduced around the world

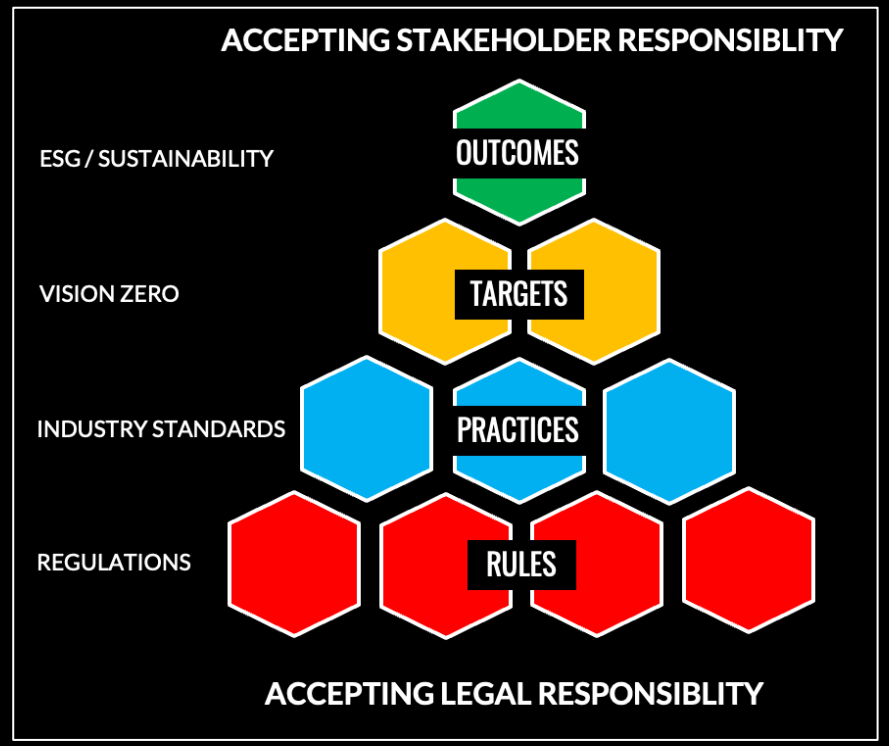
- **European Union's AI Act (proposed in 2021):** This was a comprehensive attempt to regulate AI systems based on their risk level. It was still under negotiation as of early 2024.
- **China's AI regulations:** China implemented several regulations, including the "Interim Measures for the Management of Generative AI Services" in 2023.
- **U.S. AI Bill of Rights (2022):** While not legally binding, this outlined five principles for the design, use, and deployment of AI systems.
- **EU's Digital Services Act and Digital Markets Act (2022**): These regulations, while not specifically about AI, have implications for AI-driven platforms and services.
- **NIST AI Risk Management Framework (2023):** The U.S. National Institute of Standards and Technology released this voluntary framework to help organizations address AI risks.
- **Various state-level regulations in the U.S.:** For example, laws regarding facial recognition technology and automated decision systems.
- **Canada's Artificial Intelligence and Data Act (proposed in 2022):** This was part of broader legislation aimed at regulating high-impact AI systems.
- **ISO/IEC 42001:2023 - Artificial intelligence — Management system**: Published in February 2023, this international standard provides a framework for organizations to develop and maintain an Artificial Intelligence Management System (AIMS).

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

2. THE PROBLEM WITH AI REGULATION

The AI regulatory terrain is complex

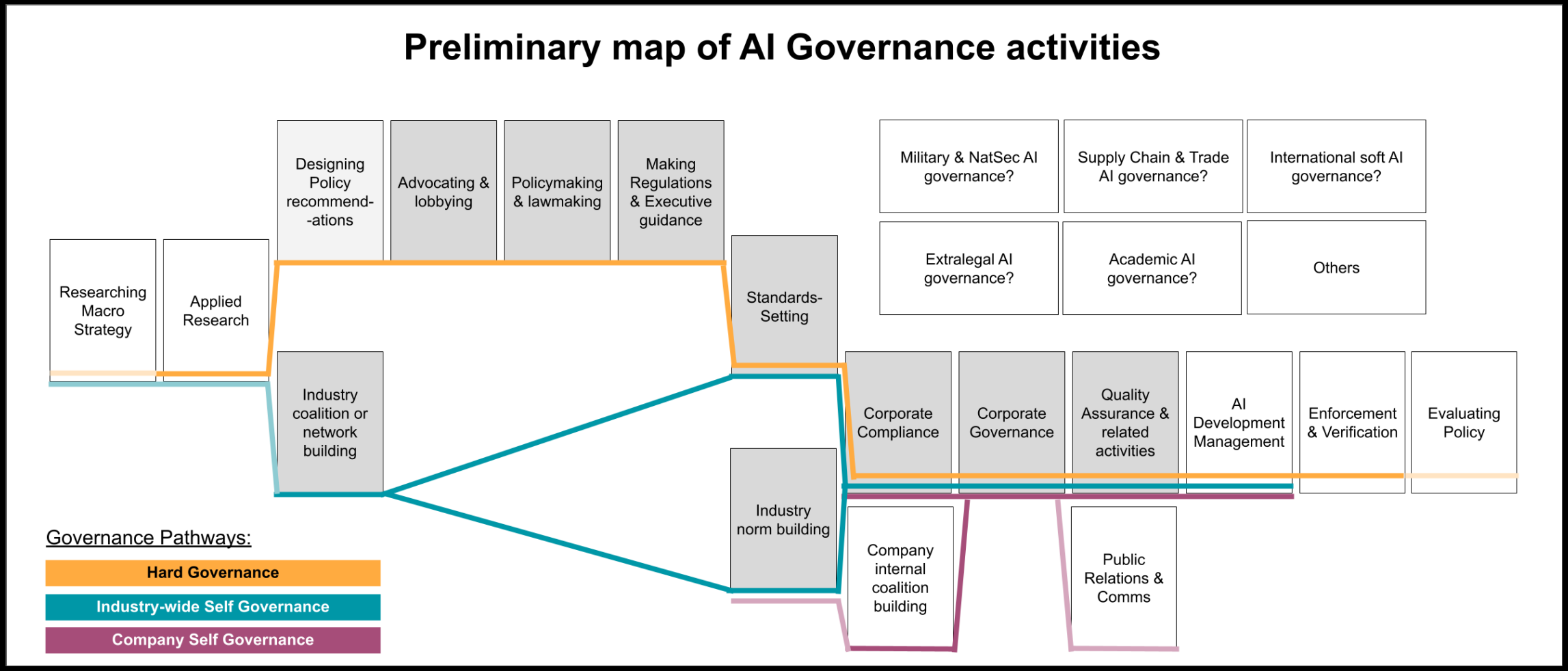




Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

2. THE PROBLEM WITH AI REGULATION

It will take time, which is something we don't have

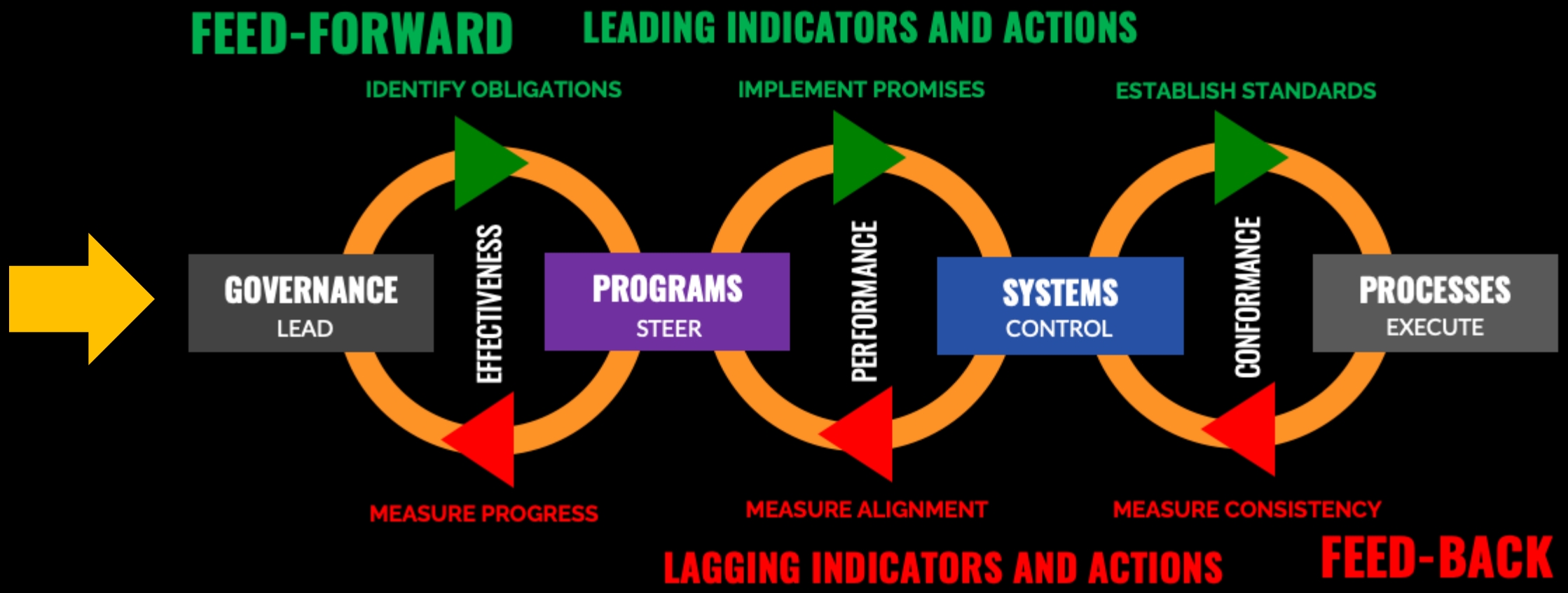


Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

https://forum.effectivealtruism.org/posts/tmxkRFx6HyhhvHdz4/a-map-to-navigate-ai-governance

2. THE PROBLEM WITH AI REGULATION

You will need to self-regulate, you need Operational Governance



Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

3. WHAT IS GOVERNANCE

How do we define it?



Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

3. WHAT IS GOVERNANCE

What do we mean by governance?

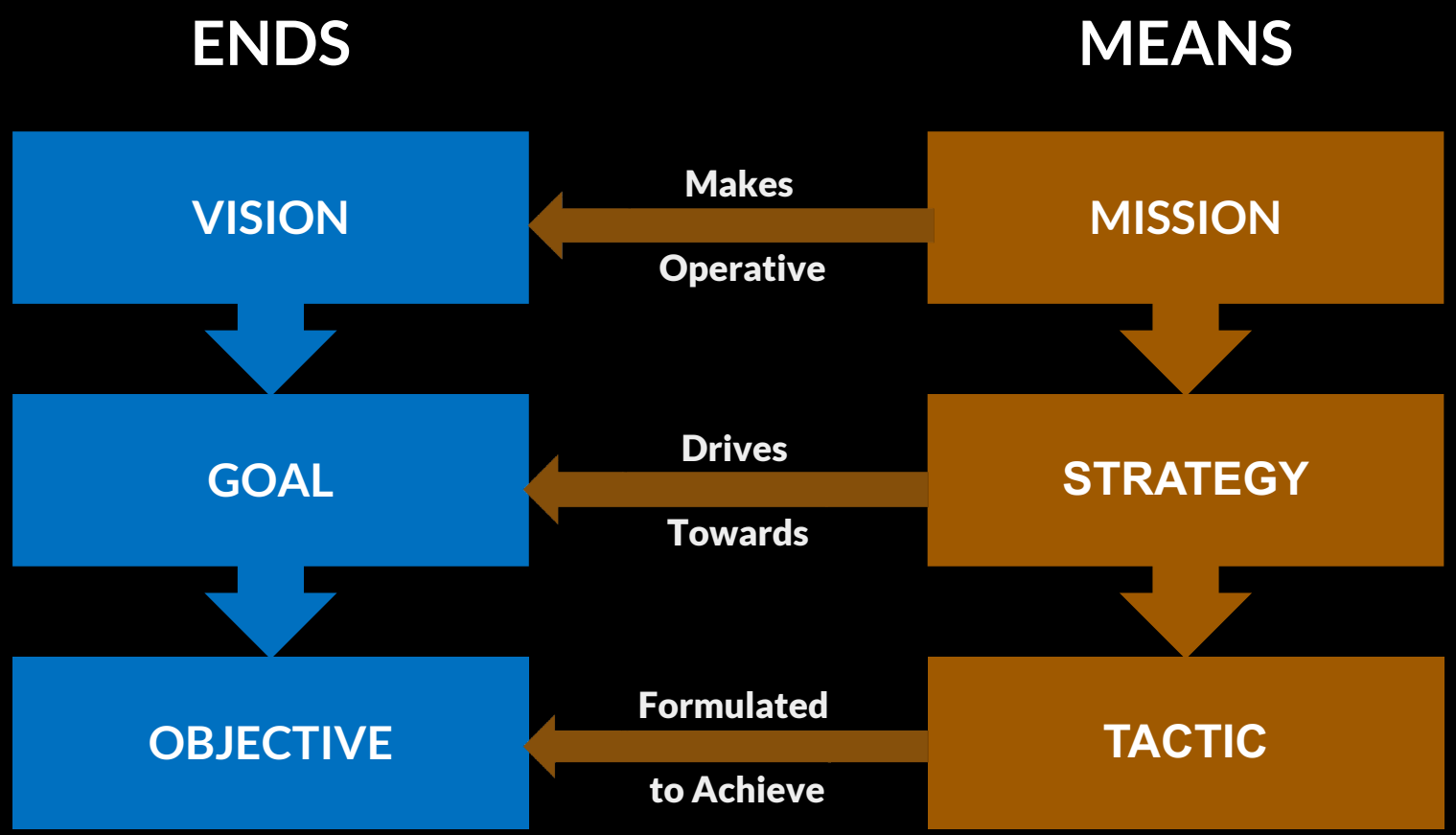
Governance: to steer, to pilot, to guide

Governance is more than oversight, it is the act of **steering**

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

3. WHAT IS GOVERNANCE

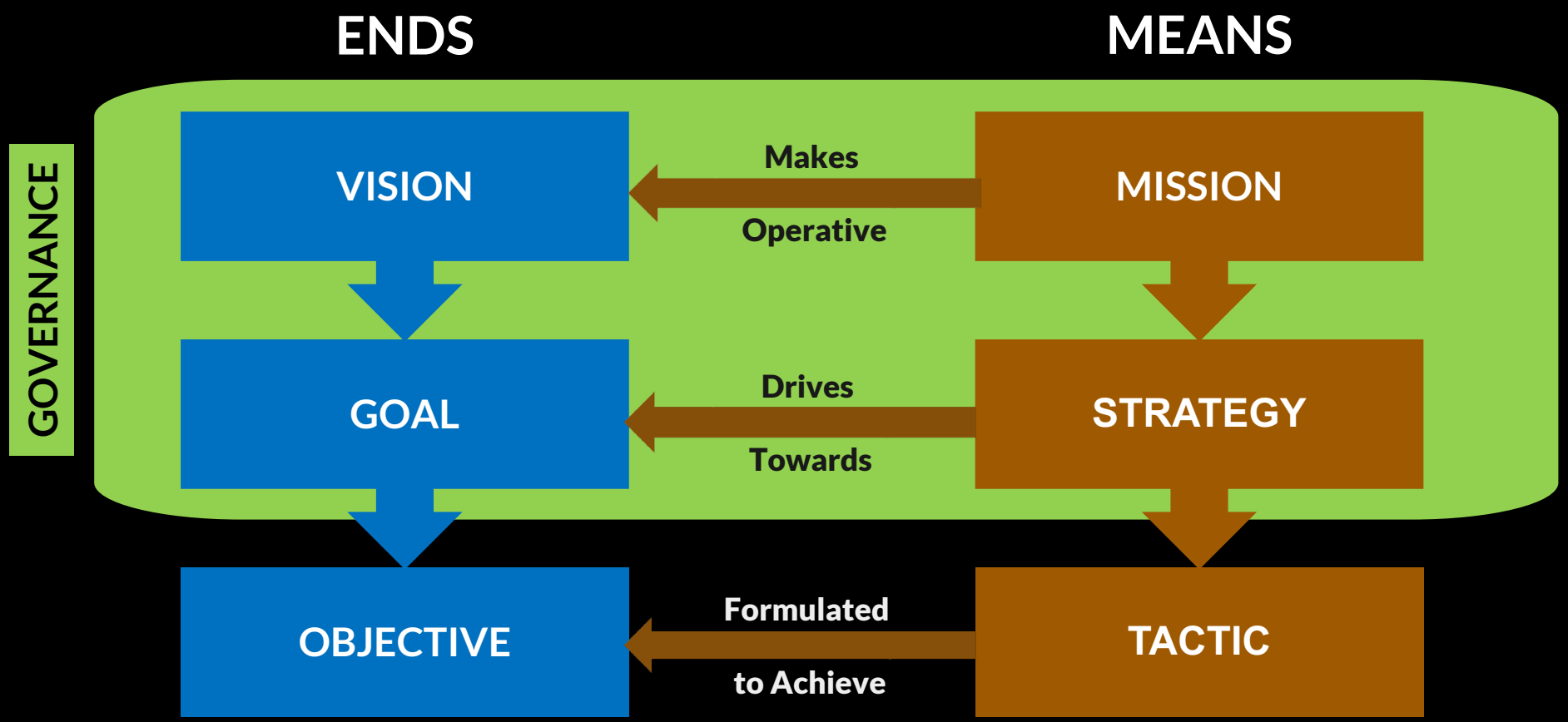
What are we steering?



Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

3. WHAT IS GOVERNANCE

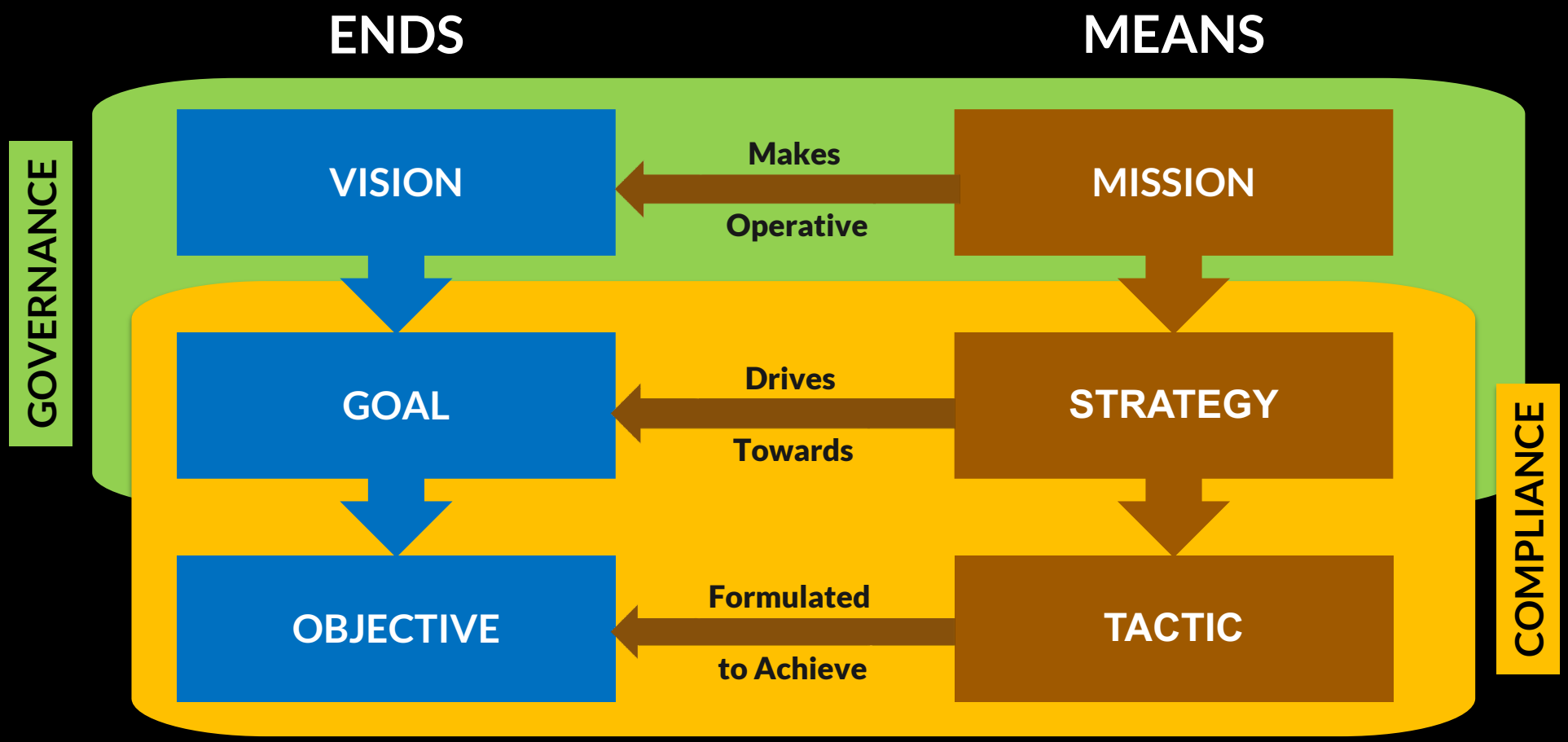
Governance steers by aligning the means with the ends



Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

3. WHAT IS GOVERNANCE

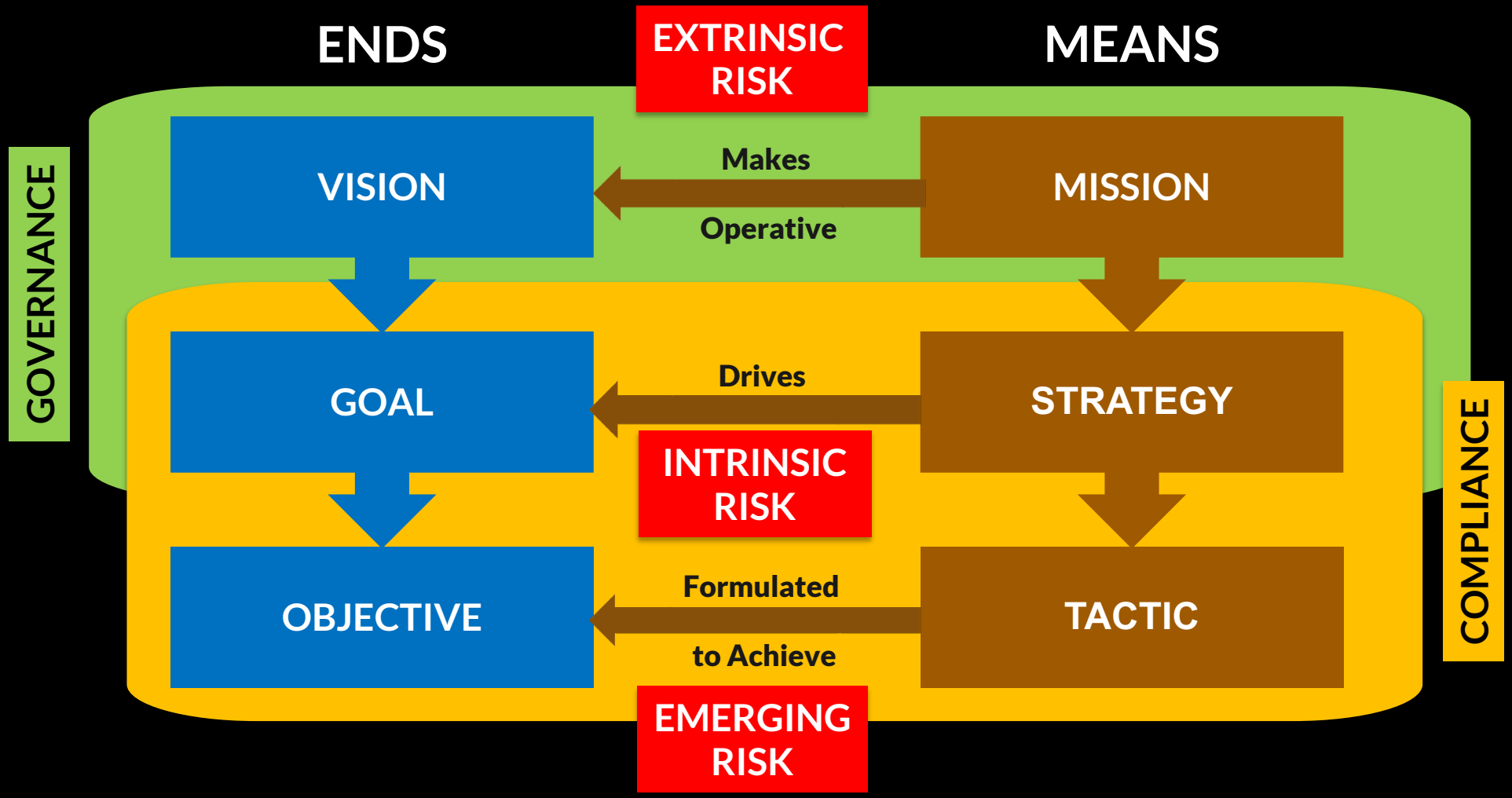
Compliance bridges operational objectives with organizational outcomes



Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

3. WHAT IS GOVERNANCE

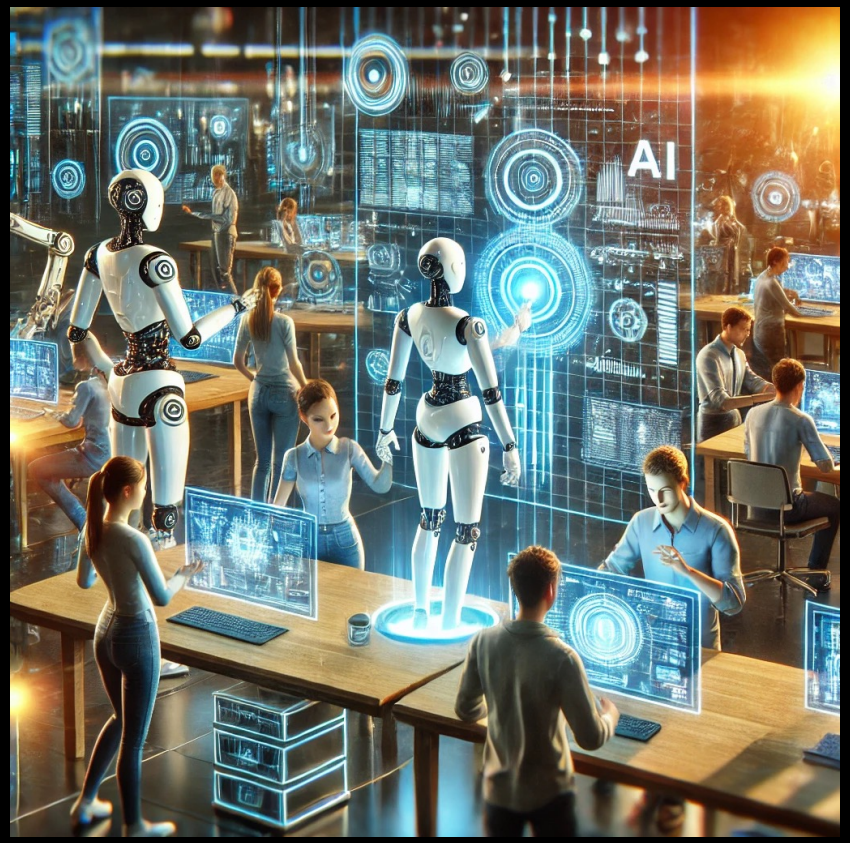
All this happens in the presence of uncertainty

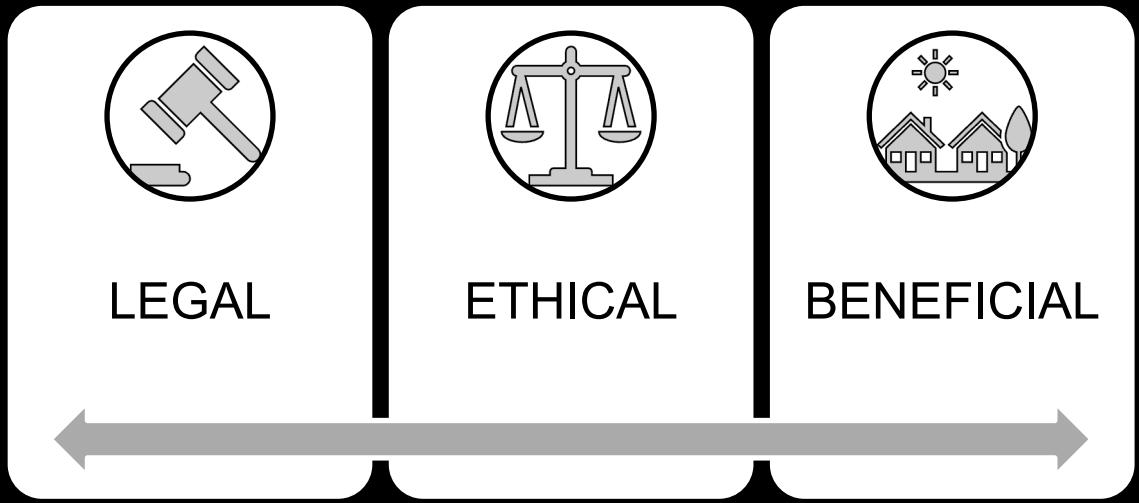


Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

3. WHAT IS GOVERNANCE

Three lines you cannot cross: illegal, unethical, and loss



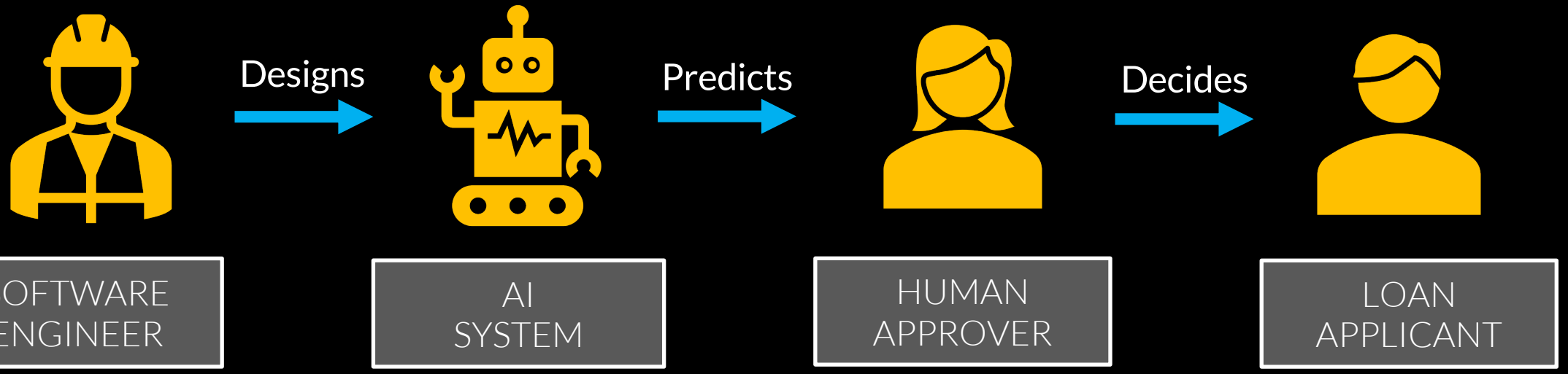


1. AI must be **LEGAL**
2. AI must be **ETHICAL**
3. AI must be **BENEFICIAL**

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

3. WHAT IS GOVERNANCE

An Example

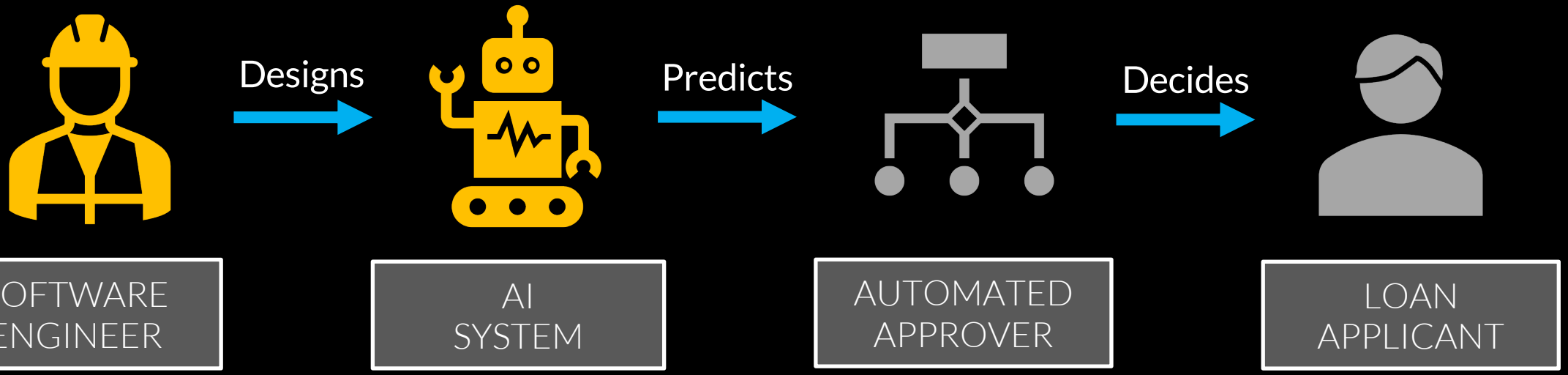


individuals should not be subject to a decision that is based solely on automated processing (such as algorithms) and that is legally binding, or which significantly affects them. – GDPR

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

3. WHAT IS GOVERNANCE

How will you govern this situation?



- Is the Engineer in violation of their Engineer's Professional Code of Ethics?
- Did the Engineer fulfill their duty to their employer
- Did the employer fulfill their ethical and legal obligations?
- Where the rights of the loan applicant violated?

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

4. METHODS OF AI GOVERNANCE

Policies, principles, guardrails and lampposts



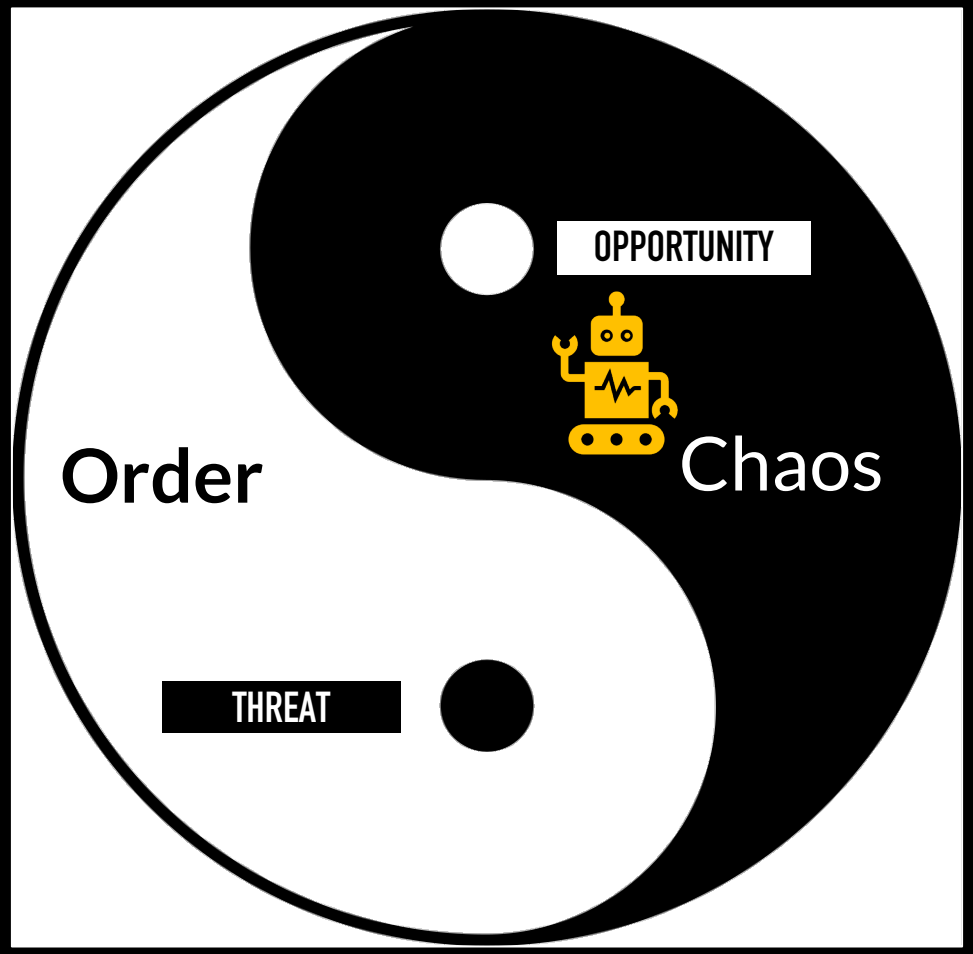
Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

4. METHODS OF AI GOVERNANCE

Which best describes the nature of AI? Order or Chaos

ORDER

- The place you are when what you do works
- Predictable
- Occasional chaos
- Deterministic
- Routine
- Underwhelming



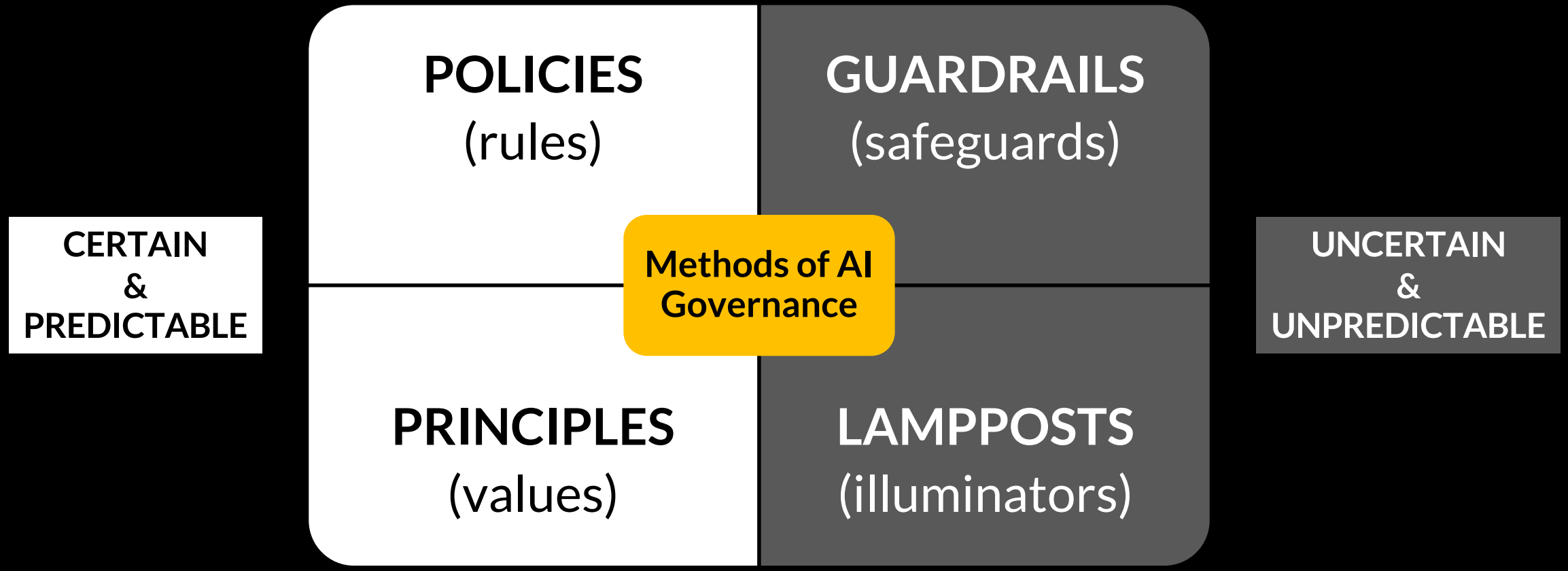
CHAOS

- The place you are when what you do doesn't work
- Unpredictable
- Occasional order
- Random
- Abnormal
- Overwhelming

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

4. METHODS OF AI GOVERNANCE

Policies & Principles + Guardrails & Lampposts



Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

4. METHODS OF AI GOVERNANCE

When the terrain is certain and predictable



POLICIES

- Establishes the rules of the game
- Identifies promises & commitments
- Identifies outcomes and consequences
- Defines mandatory code of conduct



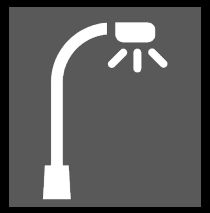
PRINCIPLES

- Sets core values and guiding beliefs
- Identifies strategies and approaches
- Identifies methods that are true for all time
- Defines advisory course of action

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

4. METHODS OF AI GOVERNANCE

When the terrain is uncertain and unpredictable



LAMPPOSTS

- Keeps you heading in the right direction
- Lights your path and exposes risks
- Makes what is hidden, visible
- Helps find opportunities



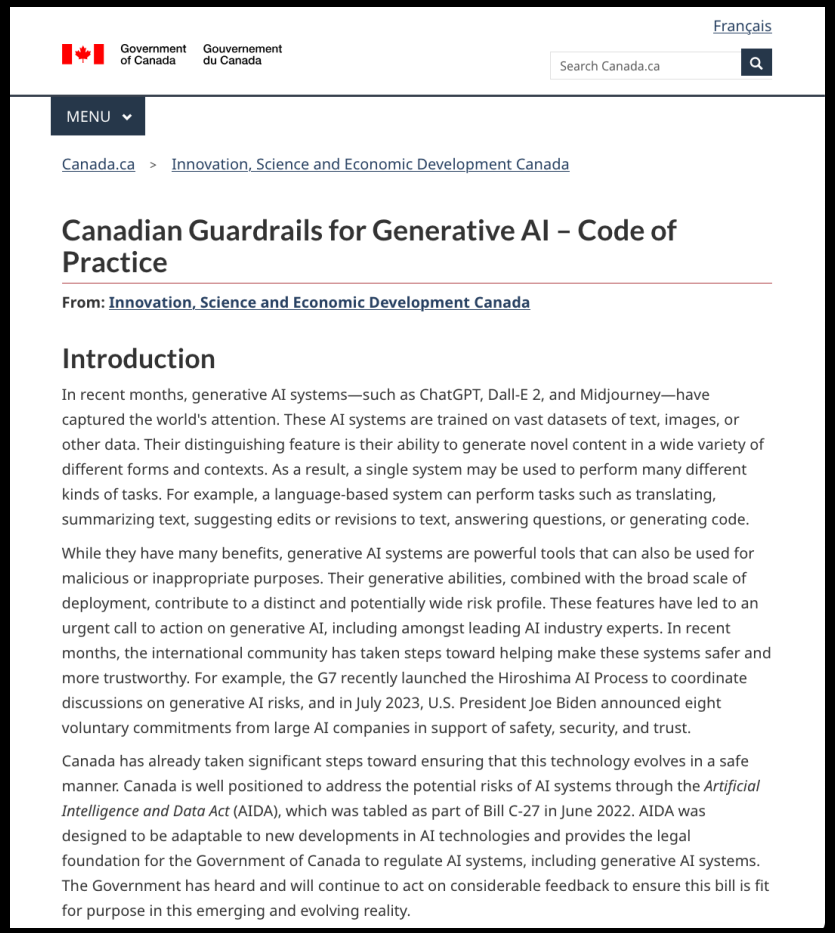
GUARDRAILS

- Keeps you away from danger
- Safeguards what is valuable
- Makes what is dangerous, safe
- Helps avoid loss

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

4. METHODS OF AI GOVERNANCE

Canadian "Guardrails" for Generative AI – Code of Practice



Français

Government of Canada | Gouvernement du Canada

Search Canada.ca

MENU

Canada.ca > Innovation, Science and Economic Development Canada

Canadian Guardrails for Generative AI – Code of Practice

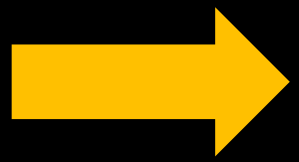
From: Innovation, Science and Economic Development Canada

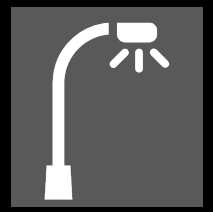
Introduction

In recent months, generative AI systems—such as ChatGPT, Dall-E 2, and Midjourney—have captured the world's attention. These AI systems are trained on vast datasets of text, images, or other data. Their distinguishing feature is their ability to generate novel content in a wide variety of different forms and contexts. As a result, a single system may be used to perform many different kinds of tasks. For example, a language-based system can perform tasks such as translating, summarizing text, suggesting edits or revisions to text, answering questions, or generating code.

While they have many benefits, generative AI systems are powerful tools that can also be used for malicious or inappropriate purposes. Their generative abilities, combined with the broad scale of deployment, contribute to a distinct and potentially wide risk profile. These features have led to an urgent call to action on generative AI, including amongst leading AI industry experts. In recent months, the international community has taken steps toward helping make these systems safer and more trustworthy. For example, the G7 recently launched the Hiroshima AI Process to coordinate discussions on generative AI risks, and in July 2023, U.S. President Joe Biden announced eight voluntary commitments from large AI companies in support of safety, security, and trust.

Canada has already taken significant steps toward ensuring that this technology evolves in a safe manner. Canada is well positioned to address the potential risks of AI systems through the *Artificial Intelligence and Data Act* (AIDA), which was tabled as part of Bill C-27 in June 2022. AIDA was designed to be adaptable to new developments in AI technologies and provides the legal foundation for the Government of Canada to regulate AI systems, including generative AI systems. The Government has heard and will continue to act on considerable feedback to ensure this bill is fit for purpose in this emerging and evolving reality.





OR



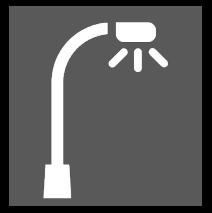
LAMPPOST

GUARDRAIL

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

4. METHODS OF AI GOVERNANCE

#1 Safety





Safety

Safety

Safety must be considered holistically throughout the system's lifecycle, with a broad view of potential impacts, particularly with regard to misuse. Given the wide range of uses of many generative AI systems, their safety risks must be assessed more broadly than those of systems with more specific uses.

Developers and deployers of generative AI systems would:

- **identify the ways that the system may attract malicious use** (e.g., use of the system to impersonate real individuals; use of the system to conduct "spearfishing" attacks) and take steps to prevent this use from occurring.

Developers, deployers, and operators of generative AI systems would:

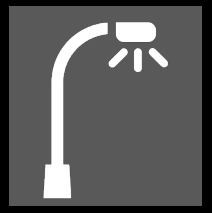
- **identify the ways that the system may attract harmful inappropriate use** (e.g., use of a large language model for medical or legal advice) and take steps to prevent this from occurring, such as by making the capabilities and limitations of the system clear to users.

Is **Safety** a **Lamppost**, **Guardrail** or **Both**?

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

4. METHODS OF AI GOVERNANCE

#2 Fairness and Equity





Fairness and Equity

Fairness and Equity

Due to the broad datasets on which they are trained and the scale at which they are deployed, generative AI systems can have important adverse impacts on societal fairness and equity through, for example, perpetuation of biases and harmful stereotypes. It will be essential to ensure that models are trained on appropriate and representative data, and provide relevant, accurate, and unbiased outputs.

Developers of generative AI systems would:

- **assess and curate datasets** to avoid low-quality data and non-representative datasets/biases.

Developers, deployers, and operators of generative AI systems would:

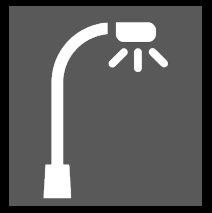
- **implement measures to assess and mitigate risk of biased output** (e.g., fine-tuning).

Are **Fairness and Equity Lampposts**, **Guardrails** or **Both**?

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

4. METHODS OF AI GOVERNANCE

#3 Transparency





Transparency

Transparency

Generative AI systems pose a particular challenge for transparency. Their output can be difficult to explain or rationalize, and their training data and source code may not be publicly available. As generative AI systems become more sophisticated, it is important to ensure that individuals realize when they are interacting with an AI system or with AI-generated content.

Developers and deployers of generative AI systems would:

- **provide a reliable and freely available method to detect content generated by the AI system** (e.g., watermarking); and
- **provide a meaningful explanation of the process used to develop the system**, including provenance of training data, as well as measures taken to identify and address risks.

Operators of generative AI systems would:

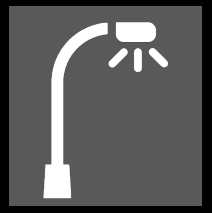
- **ensure that systems that could be mistaken for humans are clearly and prominently identified as AI systems.**

Is **Transparency** a **Lamppost**, **Guardrail** or **Both**?

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

4. METHODS OF AI GOVERNANCE

#4 Human Oversight and Monitoring





Human Oversight and Monitoring

Human Oversight and Monitoring

Human oversight and monitoring of AI systems are critical to ensure that these systems are developed, deployed, and used safely. Because of the scale of deployment and the wide range of potential uses and misuses of generative AI systems, developers, deployers, and operators need to take particular care to ensure sufficient human oversight and mechanisms to identify and report adverse impacts before making systems widely available.

Deployers and operators of generative AI systems would:

- **provide human oversight in the deployment and operations of their system,** considering the scale of deployment, the manner in which the system is being made available for use, and its user base.

Developers, deployers, and operators of generative AI systems would:

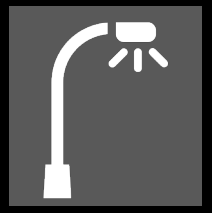
- **implement mechanisms to allow adverse impacts to be identified and reported** after the system is made available (e.g., maintaining an incident database), and commit to routine updates of models based on findings (e.g., fine-tuning).

Are **Human Oversight and Monitoring Lampposts**, **Guardrails** or **Both**?

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

4. METHODS OF AI GOVERNANCE

#5 Validity and Robustness





Validity and Robustness

Validity and Robustness

Ensuring that AI systems work as intended and are resilient across the range of contexts to which they are likely to be exposed is critical to building trust. This is a particular challenge for widely deployed generative AI systems, since they may be used in a broad range of contexts and thus may have greater exposure to misuse and attacks. This flexibility is a key advantage of the technology, but requires rigorous measures and testing to be put in place to avoid misuse and unintended consequences.

Developers of generative AI systems would:

- **use a wide variety of testing methods** across a spectrum of tasks and contexts, including adversarial testing (e.g., red-teaming), to measure performance and identify vulnerabilities.

Developers, deployers, and operators of generative AI systems would:

- **employ appropriate cybersecurity measures to prevent or identify adversarial attacks** on the system (e.g., data poisoning).

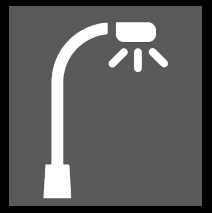
?

Are **Validity and Robustness Lampposts**, **Guardrails** or **Both**?

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

4. METHODS OF AI GOVERNANCE

#6 Accountability





Accountability

Accountability

Generative AI systems are powerful tools with broad and complex risk profiles. While internal governance mechanisms are important to any organization developing, deploying, or operating AI systems, particular care needs to be taken with generative AI systems to ensure that a comprehensive and multifaceted risk management process is followed, and that employees across the AI value chain understand their role in this process.

Developers, deployers, and operators of generative AI systems would:

- **ensure that multiple lines of defence** are in place to secure the safety of their system, such as ensuring that both internal and external (independent) audits of their system are undertaken before and after the system is put into operation; and
- **develop policies, procedures, and training** to ensure that roles and responsibilities are clearly defined, and that staff are familiar with their duties and the organization's risk management practices.

Is **Accountability** a **Lamppost**, **Guardrail** or **Both**?

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

5. A PROGRAM TO GOVERN AI

Effectively steering towards mission success

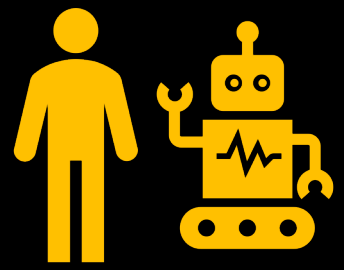


Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

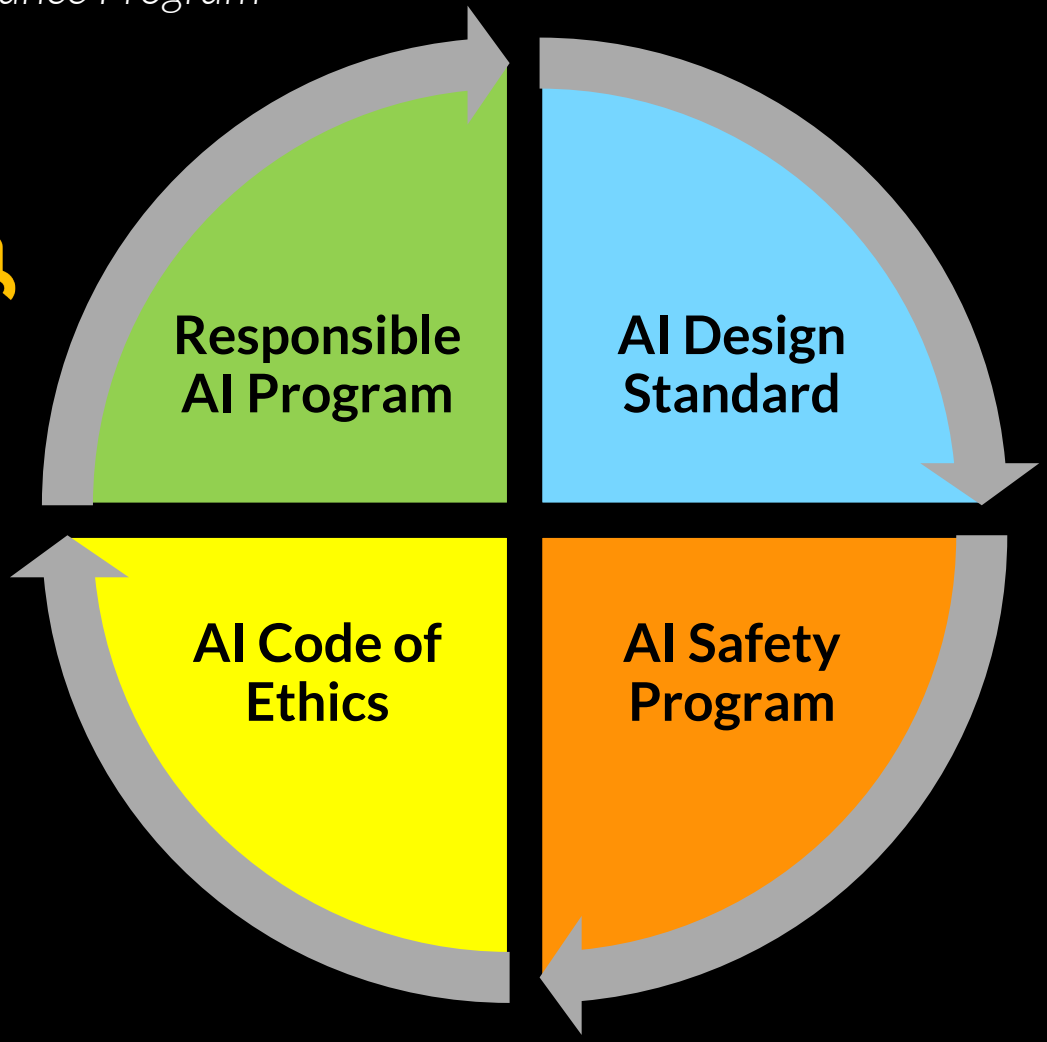
LEAN COMPLIANCE

5. A PROGRAM TO GOVERN AI

Elements of an AI Governance Program



Creating Your Path





Engineering Your Path



Guiding Your Path



Guarding Your Path

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

5. A PROGRAM TO GOVERN AI

Elements of an AI Governance Program



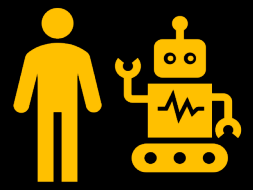
Guiding Your Path

AI CODE OF ETHICS

Purpose: To establish guiding principles and values for the ethical development and use of AI within the organization.

Scope:
- All AI-related activities, from research and development to deployment and maintenance
- All employees, contractors, and partners involved in AI projects

Approach:
- Develop a set of ethical principles aligned with organizational values
- Create clear guidelines for ethical decision-making in AI development and use
- Establish a process for ethical review of AI projects
- Provide regular ethics training for AI teams



Creating Your Path

RESPONSIBLE AI PROGRAM

Purpose: Ensure AI systems are developed and used ethically, fairly, and transparently. Considers the broader societal impact of AI and aim to mitigate potential harms while maximizing benefits.

Scope:
- All stages of the AI lifecycle, from conception to retirement
- All departments and teams involved in AI projects
- All products or services where AI is used (internal or third-party)

Approach:
- Establish clear roles and responsibilities for AI governance
- Develop processes for impact assessments and risk management
- Create mechanisms for stakeholder engagement and feedback
- Implement transparency measures for AI decision-making
- Establish metrics for measuring and reporting on responsible AI practices

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

5. A PROGRAM TO GOVERN AI

Elements of an AI Governance Program



Engineering Your Path

AI DESIGN STANDARD

Purpose: AI design standards are a set of guidelines, principles, and best practices that govern the design, development, and deployment of artificial intelligence (AI) systems.

Scope:
- All phases of AI system development, from conceptualization to implementation
- All AI projects and products or services within the organization

Approach:
- Define technical standards for AI system architecture and development
- Establish best practices for data management, model training, and testing
- Incorporate ethical considerations into the design process
- Create templates and checklists for AI system documentation
- Implement a review process to ensure adherence to the design standard



Guarding Your Path

AI SAFETY PROGRAM

Purpose: Prevent AI systems from causing harm, often focusing on technical safeguards and preventing unintended consequences.

Scope:
- All AI systems developed or used by the organization
- Both technical and operational aspects of AI safety

Approach:
- Conduct comprehensive risk assessments for AI systems
- Develop safety protocols and contingency plans
- Implement robust testing and validation procedures
- Establish monitoring systems for deployed AI
- Create incident response and recovery plans
- Provide ongoing safety training for AI teams and users

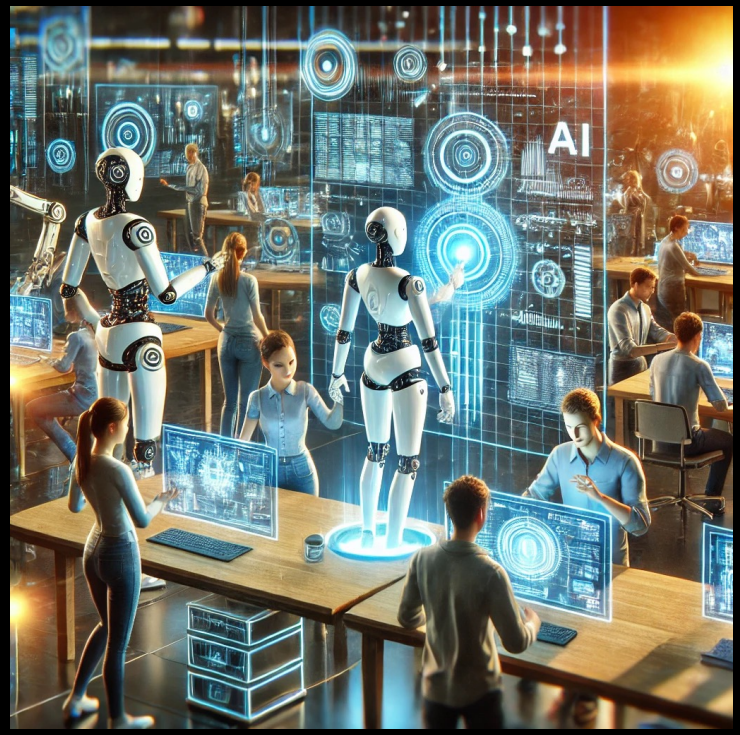
Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

5. A PROGRAM TO GOVERN AI

AI Governance: Steering Towards Mission Success in the Presence of Uncertainty Using Guardrails and Lampposts



AI GOVERNANCE PROGRAM



MISSION SUCCESS

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

Q&A



Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

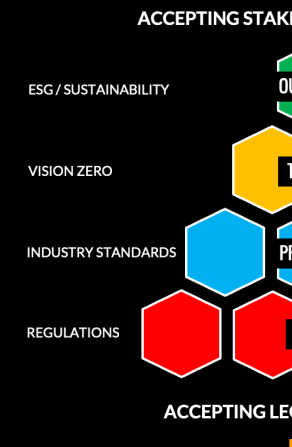
Thank You!

Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved

2. THE PROBLEM WITH AI REGULATION

The AI regulatory terrain is complex





Copyright, © 2024 Lean Compliance Consulting, Inc.—All rights reserved