



Responsible AI & Governance

How to Navigate the Changing Environment

WSAIQ24, Main Stage

JoAnn Stonier, Mastercard Fellow, Data & AI

December 20204



Our Dynamic World



ChatGPT: Optimizing Language Models for Dialogue

We've trained a model called ChatGPT which interacts in a conversational way. The dialogue format makes it possible for ChatGPT to answer follow-up questions, draw in context, challenge incorrect premises, and repeat explanations. ChatGPT is a sibling model to GPT-3.5 which is used to follow an instruction and generate a response.



The Economist

Paranoid nationalists and corruption
The property price paradox
Anticorruption: Oppenheimer's nightmare
Chile, 60 years after the coup
Editorial page 10



THE ENGINEER WEBINAR

Thursday 14th November, 2024
starting at 10:00am GMT

REGISTER NOW

How leading companies are integrating AI into core operations

Sponsored by
[expleo]



AP

LIVE:
PREPARATIONS FOR NOTRE DAME CATHEDRAL'S OFFICIAL REOPENING



SOUTH KOREA DECLARES EMERGENCY MARTIAL LAW

NDTV



TECH

NVIDIA SHARES SPIKE 26% ON HUGE FORECAST BEAT DRIVEN BY A.I. CHIP DEMAND

Nvidia CEO Jensen Huang said the company was seeing "surging demand" for its data center products.



KNOWLEDGE CENTER

Global AI Regulations Roundup

Top Stories of March 2024



2024 AI State Law

VIEW INTERACTIVE MAP



ARTIFICIAL INTELLIGENCE REGULATION

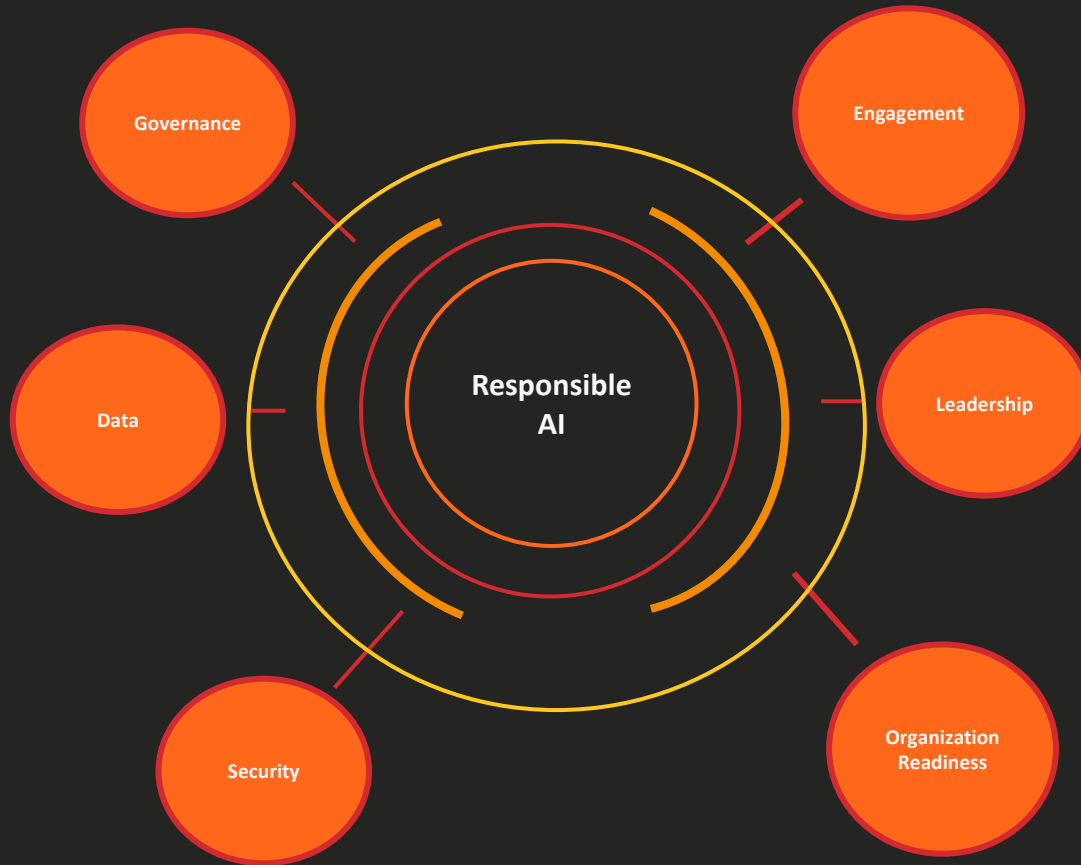
EU 2024/1689

EUROPEAN AI ACT

PUBLISHED IN THE OFFICIAL JOURNAL OF THE EUROPEAN UNION ON 12.07.2024



Elements of Responsible AI





Governance



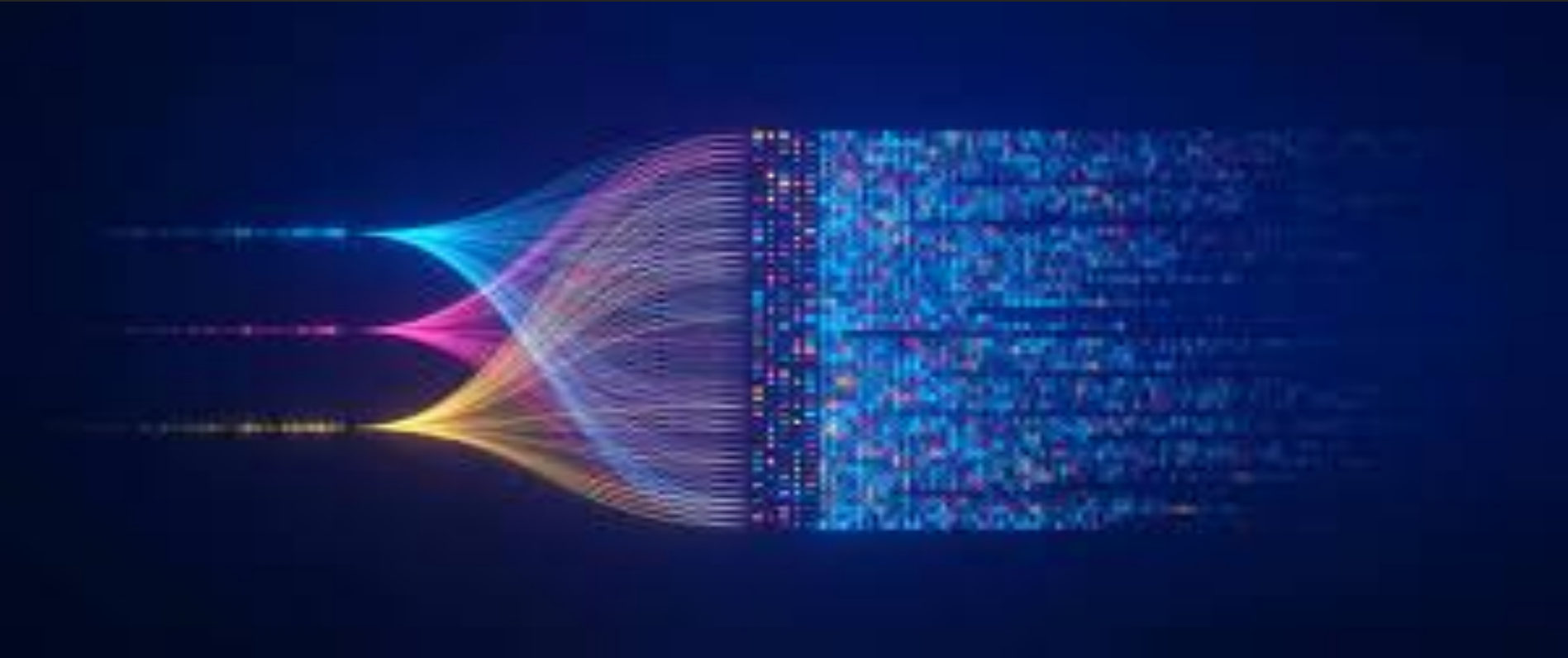


Governance





Data





Data

- Source of Data
- Rights Obtained
- Primary & Secondary Uses
- Ownership Rights and Limitations
- Personal information – privacy compliance
- Sensitive data
- PETs
- Licenses
- Ownership of AI Generated content

**Rights,
Privacy, IP**

**Quality &
Fitness**

**Bias,
Fairness &
Outcomes**

*Requires highly
trained staff*

- Accuracy/ Validity
- Completeness
- Consistency
- Timeliness/ Recency
- Integrity
- Uniqueness
- Fitness for Purpose
- Reliable/Useful
- Represents the population set related to the outcome

- What is the purpose of the project/product/service
- What is the data source – how was it obtained – what gaps does it have naturally – missing geographic or demographic groups that are important to the outcomes
- What other bias – human, storage, training, analytic bias has been introduced via the process and needs to be address
- How are the models observed – observational AI?
- Are the outputs – viewed in light of outcomes to be achieved?
- Training /Red teaming/ reviews/ audits





Security





Security



Attacks are exploiting vulnerabilities at every interaction stage of Gen AI:

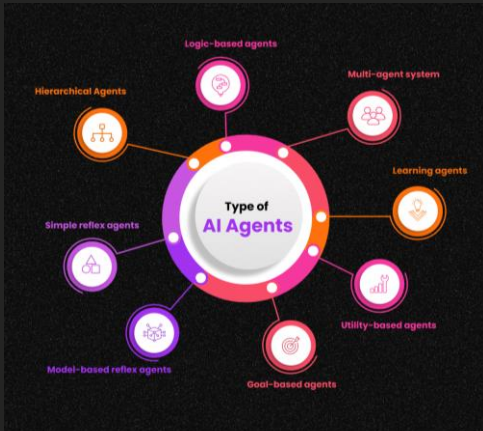
- Prompts
- RAG
- Tool Output
- Model Response



Attackers are motivated to access and steal proprietary data, generate malicious content/code

The introduction of AI Agents adds another layer of complexity – as these autonomous systems can interact with different environments and make independent decisions – creating a new threat landscape

Organizations will need to continue to invest in robust security measures to monitor their infrastructure as well as their partnership eco-systems



Organization Readiness



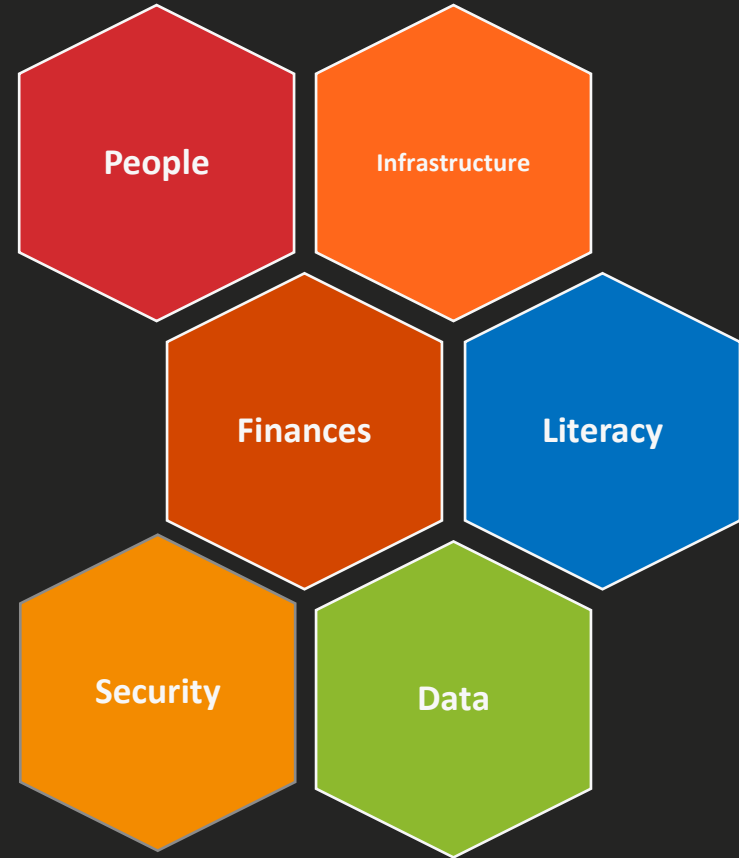


Organization Readiness



Level of Maturity

- 05** Optimized – Enterprise Ready
- 04** Managed – Portfolio of Projects
- 03** Defined – Programmatic Elements
- 02** Repeatable – Teams Evidence Maturity
- 01** Ad Hoc – Project by Project





Leadership

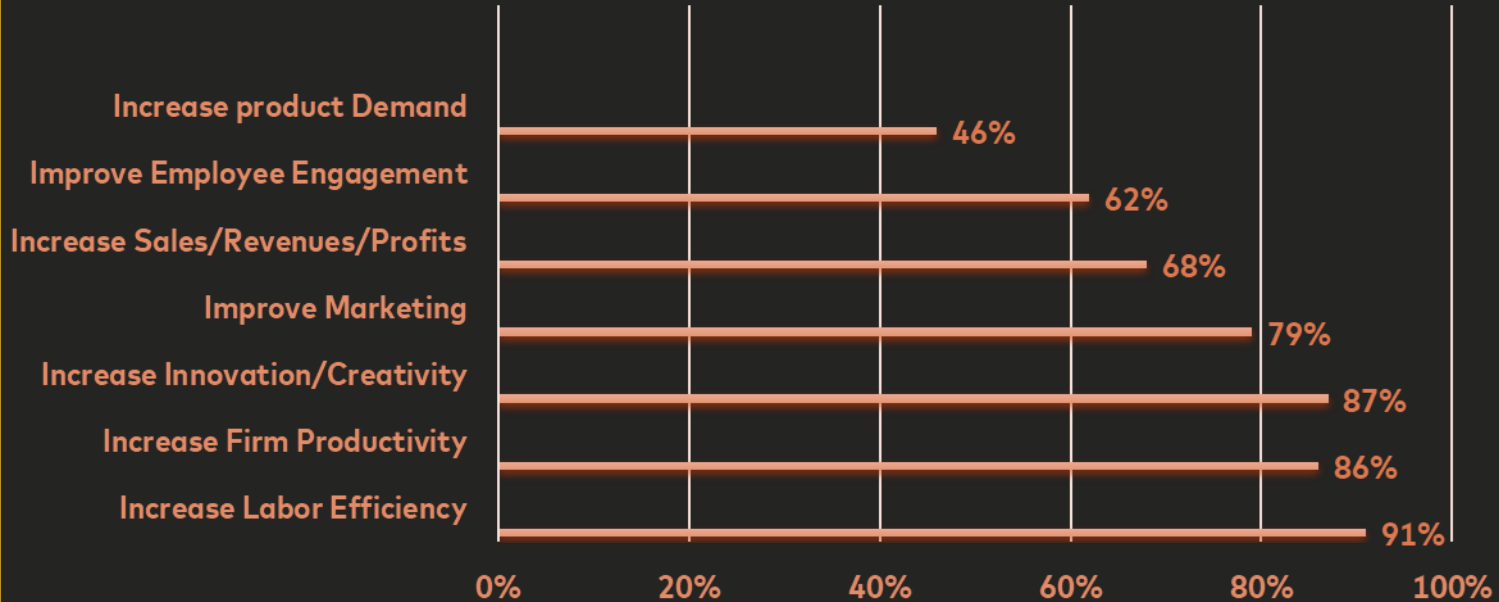




Leadership



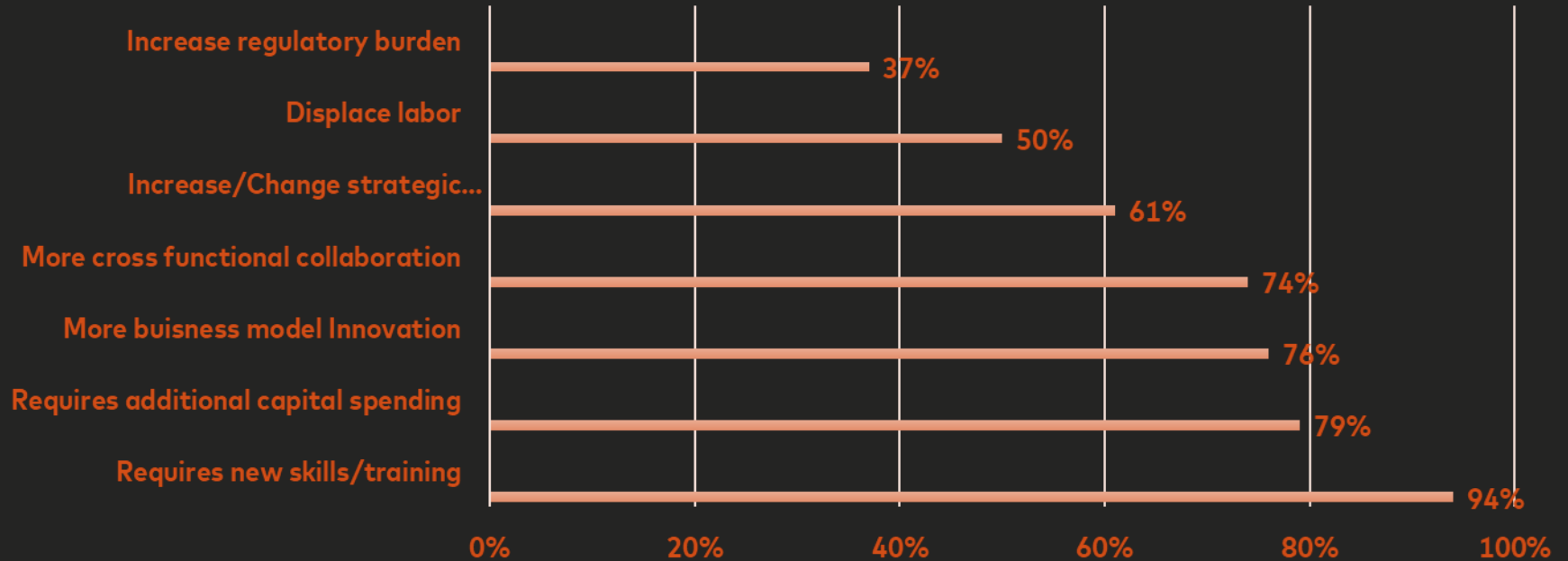
CEO Perspectives on AI – Opportunities



Source: The Conference Board – C-Suite Outlook 2024



CEO Perspective on AI – Risks & Challenges



Source: The Conference Board – C-Suite Outlook 2024

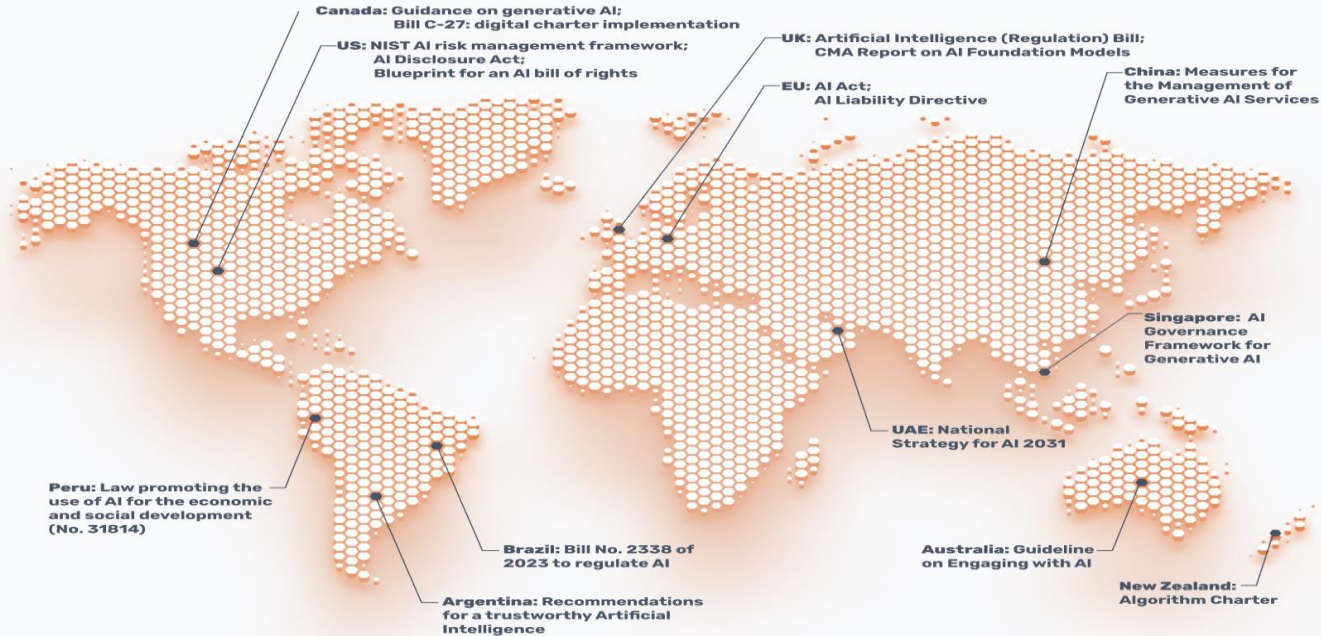


Engagement





1000+ Global AI Regulations





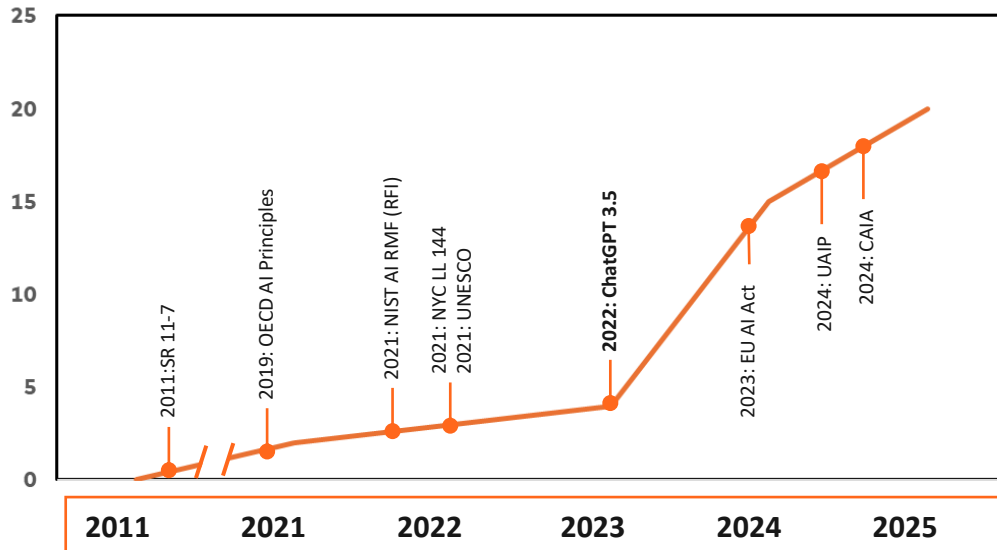
Timeline – Overview of AI Regulations & Initiatives



Select AI Governance events between 2020-2024 include:

1. Frameworks
2. Initiatives
3. Standards efforts underway
4. Passed & pending legislation
5. Take aways

AI Governance Related Events by Year



Year	2020	2021	2022	2023	2024
Count	2	3	4	15	20*



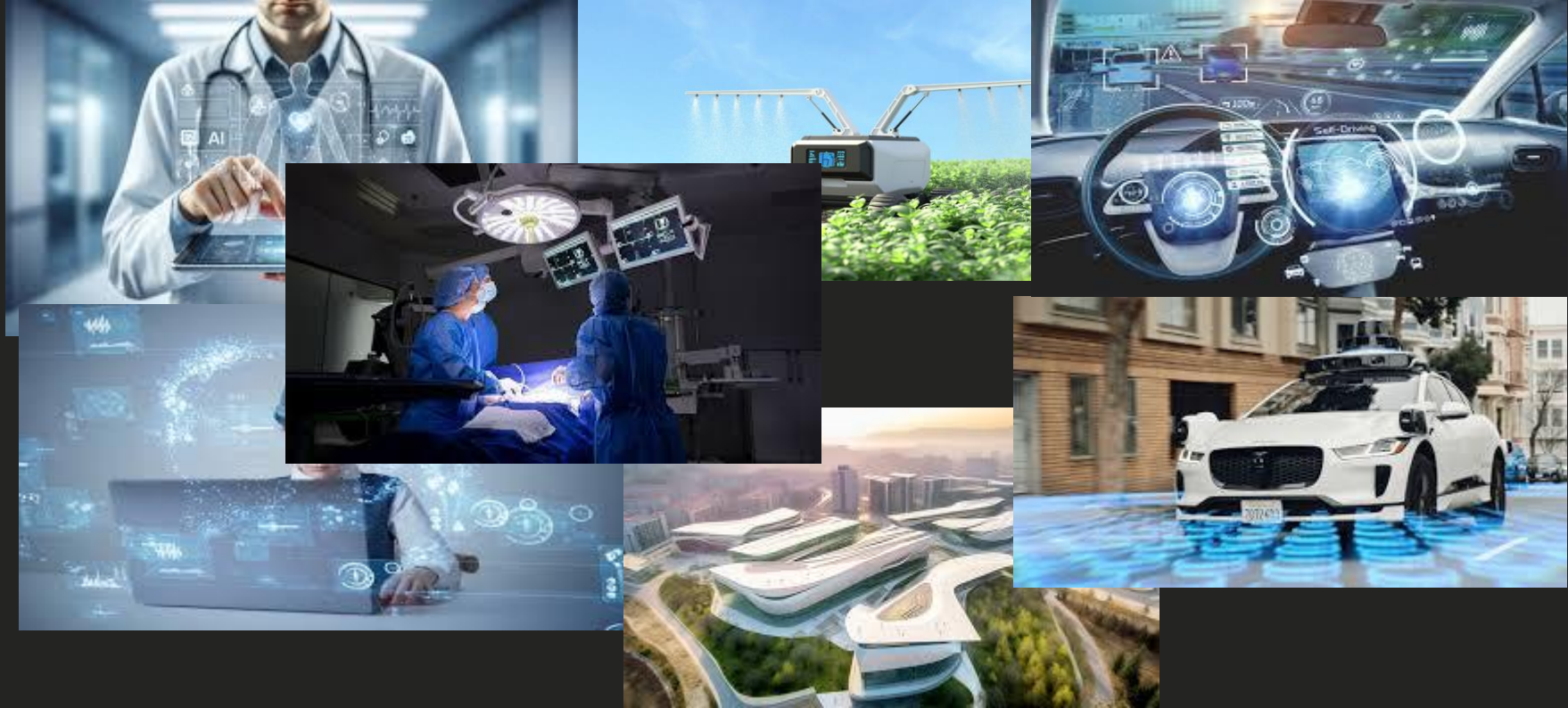
Public-Private Dialogue



- **Public-Private Dialogue assists in building standards and in operationalizing Responsible AI**
- **Forward looking frameworks** (e.g., OECD, NIST AI RMF) are helpful in providing guidance for all types of organizations to follow – enabling innovation across organizations
- Assists in LLMs to build next generation models in a more effective manner
- Assists to define responsibilities across a complicated global innovation eco-system – you can not have responsible AI in a vacuum – you are only as good as your partner organizations
- The more dialogue between public and private enterprise the better likelihood for improved standards that assist in interoperability and better outcomes for individuals, business and governments
- Organizations to watch in 2025 – OECD, NIST, Implementation acts of the EU as well as China and Singapore acts
- Join organizations – International Chamber of Commerce, Data & Trust Alliance, Executive Tech Board among others



Responsible AI – We all Benefit





Thank you!



JoAnn.Stonier@mastercard.com

<https://www.linkedin.com/in/joann-stonier-55>