Generative AI Risk and Governance

Way towards a responsible and trusted AI

January 2024

EY

Building a better working world

Generative AI is reshaping the modern enterprise

Revolutionizing the modern enterprise, reshaping industries and unlocking unparalleled opportunities for innovation

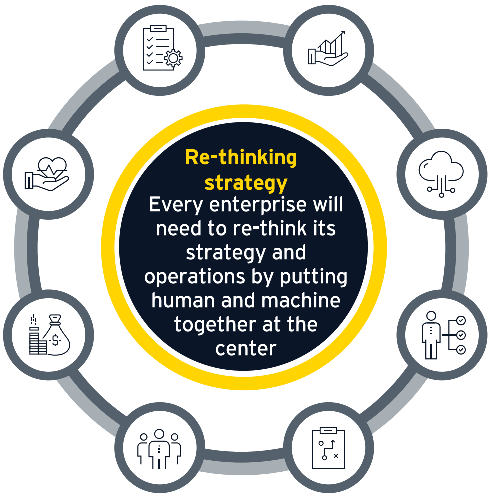
Though generative AI will not replace corporate leadership, it does significantly enable them…

Operations
Assist in generating optimal production schedules, identifying bottlenecks, and suggesting process improvements to enhance operational efficiency

Healthcare
Analyze vast amount of data to detect pattern, identify potential diseases, provide personalized care using virtual nursing assistants, assist in clinical trials, etc.

Finance
Detect fraudulent activities in financial transactions, assess creditworthiness, provide finance advice as well as streamline banking systems using chatbots and virtual assistants, etc..

People
Empower human resource and workforce through virtual assistants and chatbots to assist in generation of accurate, interactive and helpful responses, enable efficient and personalized customer service, etc.



Re-thinking strategy
Every enterprise will need to re-think its strategy and operations by putting human and machine together at the center

Marketing and sales
Analyze emerging market trends, customer needs, and generate marketing campaigns to increase customer satisfaction and sales

Technology
Aid developers in automating repetitive tasks, propose solutions, creation and analysis of code snippets for potential vulnerabilities, automate software product quality and reliability processes

Legal and compliance
Generate compliance reports, enforce version controls, automate process to flag potential risks, and provide ease to manage legal documents as per regulatory compliance

Research and development
Acts as a catalyst in empowering research as a research assistant, assist in swifter identification and summarizing key information from diverse data sources, etc.

Generative AI: speed of adoption

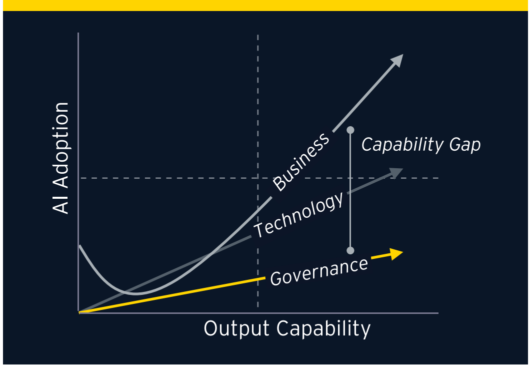
Generative AI adoption is leading innovation boom to unleash unprecedented opportunities across industries

Across industries, organizations are experiencing increasing and unpredictable changes in the business environment due to the exponential increase in the advances of Artificial Intelligence (AI) and its ubiquitous adoption by large organizations, nimble start-ups and the public alike.

Market growth will accelerate…
Global AI market is predicted to snowball in the next few years, reaching a US$190.61 billion market value in 2025[1]

Hence, AI can be the true game-changer…
Given its transformative potential of AI, +90% of consulted companies will increase their investment in AI in the next 18 months[2]

But the growth in AI adoption and advances in technology is only one part of a bigger story…
Holistically speaking, the risk posted to organizations due to the lack of awareness/knowledge of how AI operates, explain ability considerations for complex algorithms and decision frameworks and the widespread, irresolute nature of the impact to business operations, people and the overall business environment is seldom considered by organizations when adopting AI. Many organizations are also unaware of the state of AI adoption and use by service providers that they regularly engage with.



> “The rate of growth of Artificial Intelligence is far outpacing the regulations, and regulatory bodies across the world are taking notice…

[1] *CISION PR Newswire* [2] *EY Study*

Empowering innovation: the building blocks of Generative AI

Components of Generative AI that drive innovation with their ability to create, synthesize, and unlock new possibilities

Generative AI components serve as the foundational building blocks that empower innovation, driving the creation of new and diverse outputs through advanced algorithms and models. These components enable the generation of novel data, empowering businesses to unlock new perspectives, insights, and possibilities in their pursuit of innovation and creativity.

Data readiness

- Depending on the use-case, ascertain high quality diverse datasets as foundation for Generative AI
- Design data ingestion pipeline architecture based on identified training data sources
- Curate and pre-process data prior to modelling to ensure outputs' accuracy and reliability through use of effective data cleaning techniques

Model selection

- Determine Generative AI modality (text, code, image, audio, video, mixed) and design model architecture
- Utilize complex modelling techniques such as deep learning models, probabilistic models, recurrent neural networks, or transformers, etc.
- Ascertain computational infrastructure and quality requirements as part of the modeling process

Training and fine-tuning

- Conduct prompt engineering and document state maintenance guide
- Develop training and model fine-tuning processes to refine and optimize the model's parameters
- Enhance training efficiency by leveraging algorithmic capabilities, ascertaining training protocols and re-training procedure

Inference and evaluation

- Harness trained models and learned patterns for optimized generation of outputs
- Evaluate and measure model's performance against designated test sets and quality considerations
- Leverage model monitoring and performance insights to maintain high quality standards

Artificial Intelligence (AI) evolution has triggered multiple regulations across the world

Canada

- Publication: the Digital Charter Implementation Act, Bill C-27
- Date: June 2022

US

- Biden Executive Order
- Date: Oct 2023

Mexico

- Law for the Ethical Regulation of Artificial Intelligence for the Mexican United States

EU

- EU Parliament voted on draft AI law
- Date June 2023
- Publication: the EU Artificial Intelligence Act (AIA)
- Date: April 2021

UK

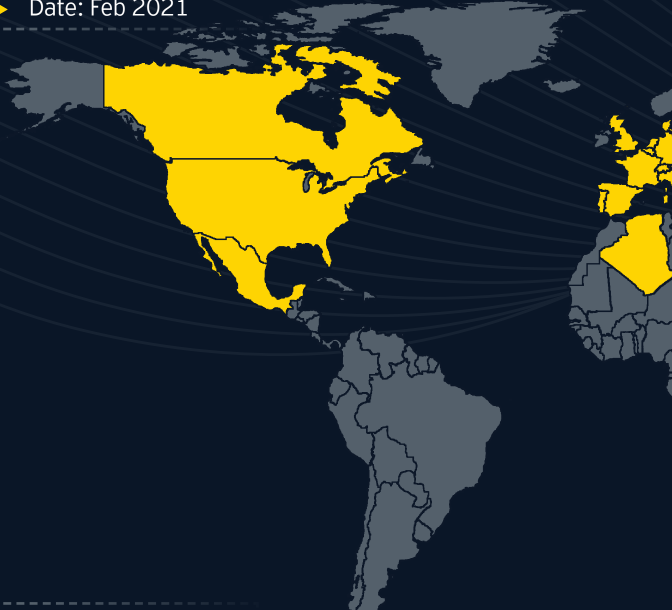
- UK Launches AI regulation roadmap
- Publication: Guidance on AI and data protection;
- Date: July 2020

Germany

- Publication: AI Cloud Service Compliance Criteria Catalogue (AIC4)
- Date: Feb 2021

Saudi Arabia

- SA proposes AI regulation via the new Intellectual Property Law
- Date: May 2023



Scoring Factors

- AI regulations
- Data Regulations (data, cyber and privacy)
- Strategy, roadmap and investment
- Infrastructure and Tooling
- Skill and Education

Algeria

- Presented national strategy on research and innovation in AI
- Date: Jan 2021

South Africa

- Launches "AI for Africa" blueprint in collaboration with other African nations
- Date: Nov 2021

Ethiopia

- Finalizing the preparation of national policy on AI
- Date: June 2023

AI systems are vulnerable to both conventional and net new security attacks

AI should be considered an additional layer with its own unique attack surface instead of being considered part of the application layer as it introduces threats which are not accounted for within conventional security defenses due to the nature of the AI lifecycle.

AI	Data Poisoning	Prompt Injection	AI Supply Chain Attack	Model Inversion	Data Leakage
Application Layer	Cross-Site Scripting	Cross-Site Request Forgery	Remote Code Execution	Remote File Inclusion	XML External Entity
Data Layer	SQL injection	Denial of Service	Data Breaches	Insider Threat	Data Tampering
Host Layer	Malware	Rootkits	Privileged Escalation	Zero Day Exploits	Buffer Overflow
Network Layer	Denial of Service	IP Spoofing	Man in the Middle	ARP Poisoning	DNS Spoofing

Configuration of AI

AI is often incorporated into the application layer, offering various configuration options. However, regardless of how it is oriented, it is important to acknowledge that AI carries its own inherent risks.

Rising global guidelines/regulations on Responsible AI signal urgency

Egypt

- Egypt's National Council for AI announces the launch of "Egyptian Charter for Responsible AI"
- Date: April 2023

UAE

- UAE launches Generative AI guide
- Date: April 2023

Vietnam

- Instructs cross border platforms to use AI and remove toxic content
- Date: June 2023

S Korea

- PIPC publishes guidelines on personal data processing in AI
- Date: Aug 2023

Japan

- Amendment that allows level four automated driving
- Date: April 2023

Thailand

- ETDA proposes three new AI laws
- Date: Sep 2023

Philippines

- University of Philippines released draft set of AI regulations
- Date: July 2023

Sri Lanka

- Announces 1 Billion fund for AI
- Date: Sep 2023

Indonesia

- MCI is drafting ethical guidelines for privacy protection
- Date: Aug 2023

Malaysia

- Considering a new law to label AI generative products either "AI-generated" or "AI-assisted"
- Date: July 2023

Australia

- Royal Commission Report into Robodebt Scheme
- Date: July 2023

New Zealand

- NZ government releases Digital Strategy for Aotearoa
- Date: Sep 2022

Singapore

- Singapore and the EU signed a Digital Partnership
- Date Feb 2023
- Publication: the Model AI Governance Framework
- Date: Jan 2019

© Australian Bureau of Statistics, GeoNames, Microsoft, Navinfo, Open Places, OpenStreetMap, TomTom, Zenrin

Threat landscape with ever evolving use of Gen AI

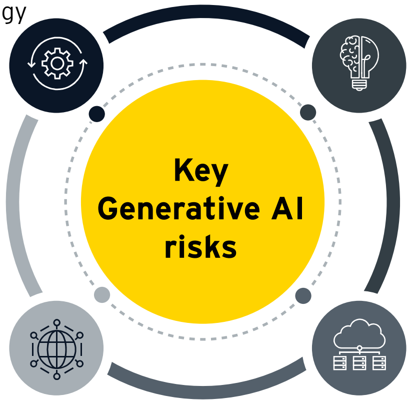
Generative AI risks and considerations that every organization must think deeply about

Design risks

- Ambiguous transparency in third party AI systems
- Poor lab to production resiliency strategy
- Ill-equipped fail-over mechanism for inoperable AI systems
- Inadequate AI system monitoring
- Poor model design for ethic-based content filtering

Performance risks

- Unoptimized model parameter fine-tuning
- Inconsistent quality and monotonous outputs
- Inadequate error and crash handling mechanism
- Inappropriate oversight from in-house IT Teams
- Impaired scalability and output generation capability



Algorithmic risks

- Biases and fairness concerns
- Manipulation Vulnerability
- Compromised model integrity
- Output accuracy and authenticity
- Poisoned AI inputs

Data risks

- Inappropriate data provenance
- Unauthorized data reconstruction
- Poor data quality impairing AI outcomes accuracy
- Improper use and misrepresentation of copyrighted content in AI systems
- Unauthorized access to proprietary or sensitive information

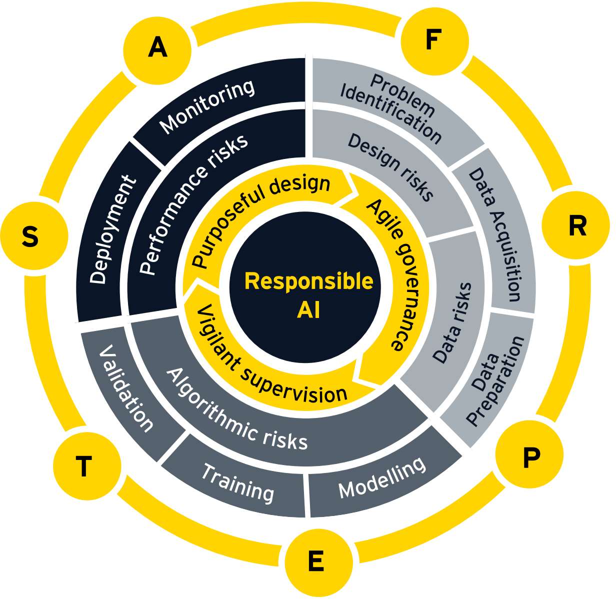
Trusting AI will require expanding the risk and control attributes

The Responsible AI framework developed by EY enables clients to mitigate AI risks while complying with emerging AI regulations. It can evaluate AI risks and build controls across seven trust attributes and four risk categories.

Accountability
there is unambiguous ownership over AI systems and their impacts across the AI development lifecycle.

Sustainability
the design and deployment of AI systems are compatible with the goals of sustaining physical safety, social well-being, and planetary health.

Transparency
appropriate levels of openness regarding the purpose, design, and impact of AI systems is provided so that end users and system designers can understand, evaluate, and correctly employ AI outputs.



Fairness
AI systems are designed with consideration for the need of all impacted stakeholders and to promote inclusiveness and positive societal impact.

Reliability
outcomes of AI systems are aligned with stakeholder expectations and perform at a desired level of precision and consistency, whilst being secured from unauthorized access, corruption, and/or adversarial attack.

Privacy
AI systems are design with consideration to data rights regarding how personal information is collected, stored, and used.

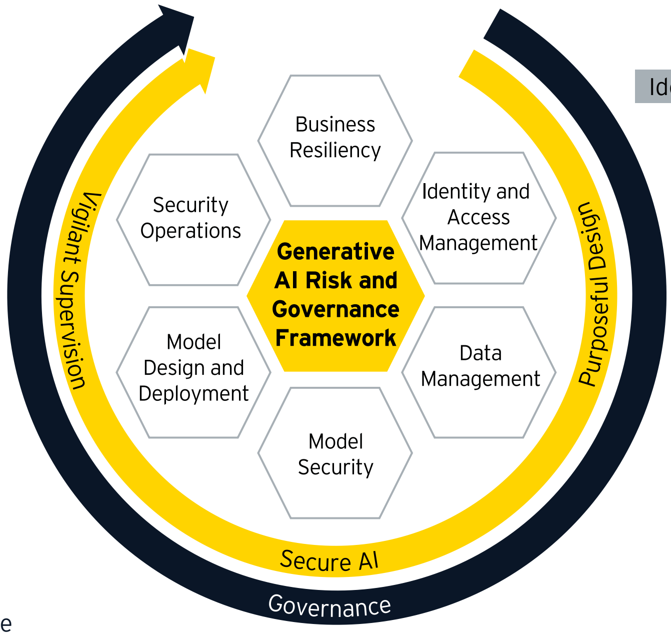
Explainability
appropriate levels of explanation are enabled so that the decision criteria of AI systems can be reasonably understood, challenged, and/or validated by human operators.

EY's Generative AI Risk and Governance Framework focuses on seven key domains to establish robust framework and governance processes that align with industry-leading standards for ethical and responsible use of GenAI

BR
Business Resiliency
- AI Redundancy Architecture
- AI Operational Continuity Management
- AI Disaster Recovery

SO
Security Operations
- GenSec Monitoring
- AI Incident Management
- AI Vulnerability Shield Management
- AI Threat and Vigilance Management

MDD
Model Design and Deployment
- Model Design and Architecture
- AI Capacity Planning
- Model Bias and Fairness
- Secure AI Development
- AI Deployment Assurance



IAM
Identity and Access Management
- Identity Management
- AI AuthSec Management
- Privilege Access Management
- AI Interface Access Management

DM
Data Management
- Data Provenance
- AI Data Privacy
- AI Data Security

MS
Model Security
- AI Input Integrity
- AI Prompt Validation
- AI Output Management
- AI Ethics Validation
- AI Plugin Security
- AI Codebase Control

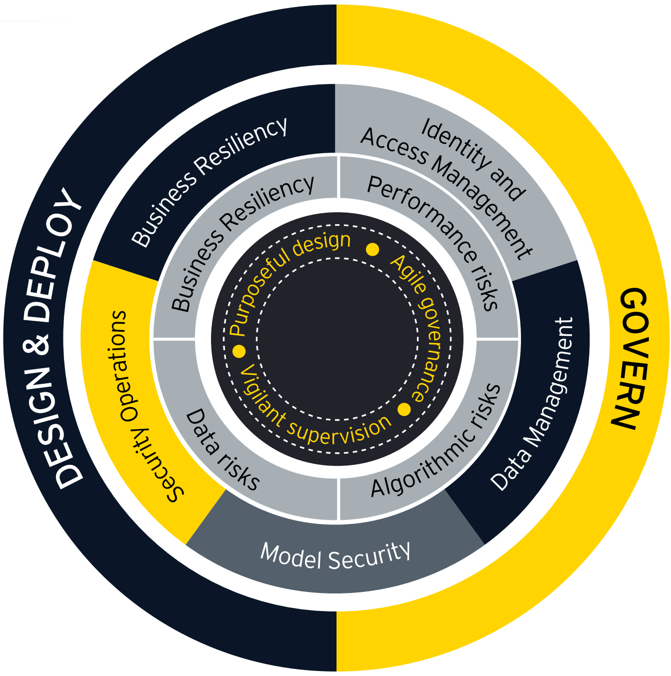
GRC Governance
- Trusted AI/ML
- Governance Guardrails
- Regulatory Compliance
- Third Party Risk
- Training & Awareness

Empowering success with strong governance

leveraging drivers, mitigating risks, and ensuring responsible business practices

EY Generative AI Risk and Governance Framework

A robust risk management framework and governance processes that establish organizational standards for ethical and responsible use of Gen AI based on NIST AI RMF, MITRE ATLAS, OWASP Top 10 for LLM, ENISA, HITRUST, ISO/IEC 23894, ISO 42001, etc.



How EY can help with robust governance for Generative AI

Establishing a trusted AI ecosystem is the first key towards transformation

EY provides clients with a wide range of AI governance-centric services to ensure trust in their Generative AI products and services. EY understands the heightened risks and significance of design, performance, and algorithmic factors in Generative AI. EY has developed and designed the below salient offerings to assist clients in building robust and resilient Generative AI governance.

GenAI Risk & Governance Advisory

- Using EY's Generative AI Risk and Governance Framework, assess organization's existing policies, procedures, security standard documents to determine adequacy of governance processes and controls associated with Generative AI and evaluate implementation effectiveness
- Develop/Update organization's policies, procedures and security standard documents; and design tailored governance processes and controls specific to Generative AI

GenAI LLM Assessment

- Examine the Gen AI LLM by performing a detailed evaluation of its resilience, data integrity, and protective measures against potential threats and vulnerabilities.
- Assess the LLM's ability to effectively handle security challenges, ensuring data accuracy and maintaining robust defences.

NIST AI RMF based Risk Assessment

- Perform risk assessment for the existing Generative AI solution to evaluate controls implemented for AI risk management and review current state to ascertain applicability of NIST AI RMF security and privacy requirements
- Identify potential risks associated with Generative AI solution based on NIST AI RMF guidelines across the four functions (i.e., Govern, Map, Measure and Manage)

HITRUST Assessment

- Perform HITRUST readiness assessment for Generative AI solution and related IT controls based on the latest version of HITRUST Common Security Framework (CSF v11.2.0)
- Conduct HITRUST Certification assessment to demonstrate assurance that the security and operational controls within the AI system have been effectively implemented and maintained.

Ernst & Young LLP

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. EYG member firms do not practice law where prohibited by local laws. For more information about our organization, please visit ey.com.

Ernst & Young LLP is one of the Indian client serving member firms of EYGM Limited. For more information about our organization, please visit www.ey.com/en_in.

Ernst & Young LLP is a Limited Liability Partnership, registered under the Limited Liability Partnership Act, 2008 in India, having its registered office at Ground Floor, Plot No. 67, Institutional Area, Sector - 44, Gurugram - 122 003, Haryana, India.

EYIN2401-022
ED None

This publication contains information in summary form and is therefore intended for general guidance only. It is not intended to be a substitute for detailed research or the exercise of professional judgment. Neither EYGM Limited nor any other member of the global Ernst & Young organization can accept any responsibility for loss occasioned to any person acting or refraining from action as a result of any material in this publication. On any specific matter, reference should be made to the appropriate advisor.

RS2

ey.com/en_in

@EY_India

EY Careers India

EY India

@ey_indiacareers

EY

For more information please contact

Abbas Godhrawala
Partner
Technology Risk
Consulting India
abbas.godhrawala@in.ey.com

Abhijit Kumar
Partner
Technology Risk
Consulting India
abhijit.kumar@in.ey.com

Aditya Iyer
Director
Technology Risk
Consulting India
aditya.iyer@in.ey.com