IIA Chicago

Generative AI and Internal Audit

Key risks and considerations

EY

Building a better working world

Contents

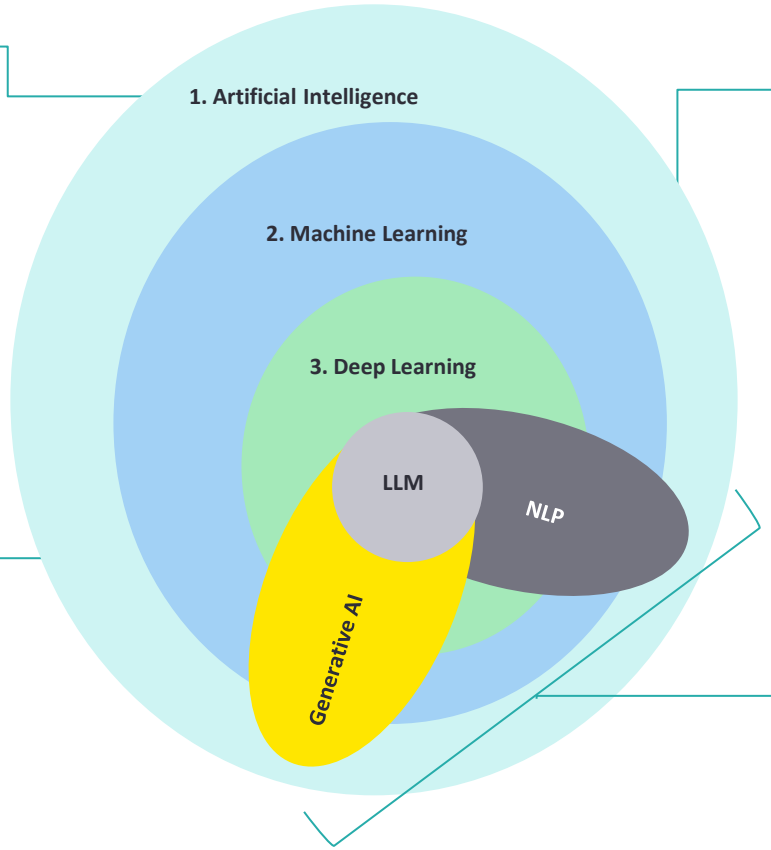
1. Artificial intelligence terminology

Artificial Intelligence and it's evolution

Artificial Intelligence "is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable."[1]

- **Artificial Intelligence (AI)** originated circa 1950s.
- A branch of computer science concerned with the theory and development of computer **systems able to perform tasks that normally require human intelligence.**

- **Machine Learning (ML) is a subset of AI** that originated circa 1960s.
- Development of **systems that learn and make predictions and decisions without explicit command,** by using algorithms and statistics to analyze and draw inferences from data patterns.



- **Deep Learning (DL) is a subset of AI and ML** that originated circa 1970s.
- The process of **using artificial neural networks, meant to simulate behavior of the human brain,** to recognize intricate patterns in data and perform complex tasks.

- The field of AI is constantly evolving, and there are **many more subsets and types of AI that integrate ML and DL principles to achieve distinct objectives.**
- **Natural Language Processing (NLP), Generative AI, and Large Language Models (LLMs)** are a few examples discussed further on the next slide.

[1]: John McCarthy – "What is Artificial Intelligence" – November 12, 2007.

How ChatGPT fits into the Artificial Intelligence landscape

Various AI disciplines were integrated to create the popular LLM; ChatGPT. AI adoption is being accelerated by LLMs and broader Generative AI. Compared to existing AI/ML models, LLMs lower the technical thresholds for using AI and make democratized AI possible due to its versatility.

NLP

- Natural Language Processing (NLP) is a branch of Artificial Intelligence (AI) that focuses on **understanding the meaning of text data (i.e., human language).**
- NLP is a distinct subset of AI yet interconnected to Machine Learning and Deep Learning.

\+

Gen AI

- Generative AI is a type of Artificial Intelligence that leverages Machine Learning and Deep Learning principles and techniques **to generate new content**, including text, images, music, and decisions.

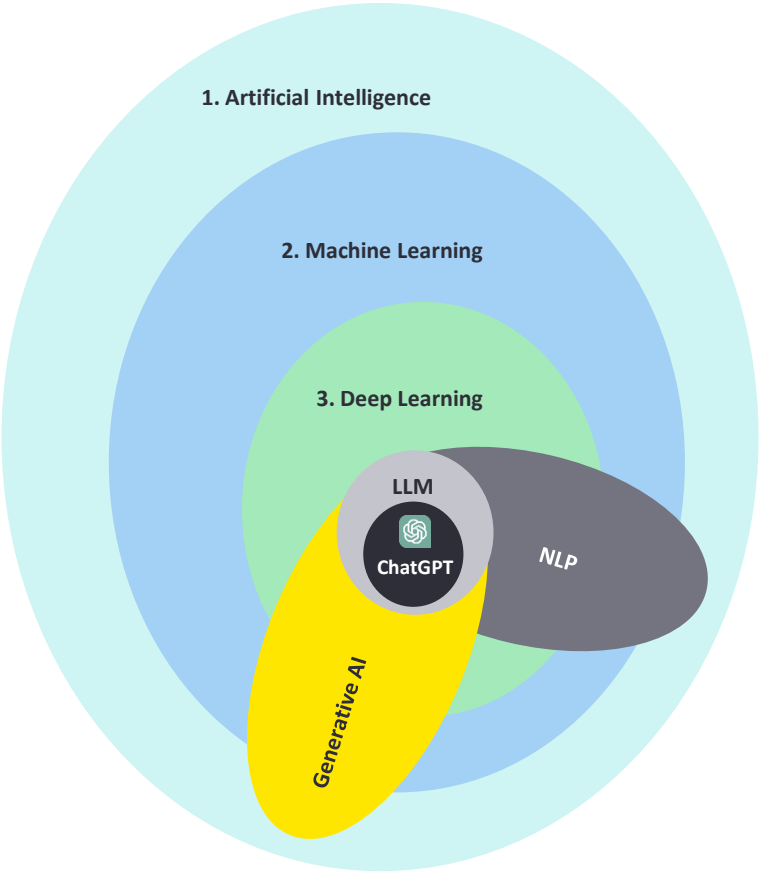
=

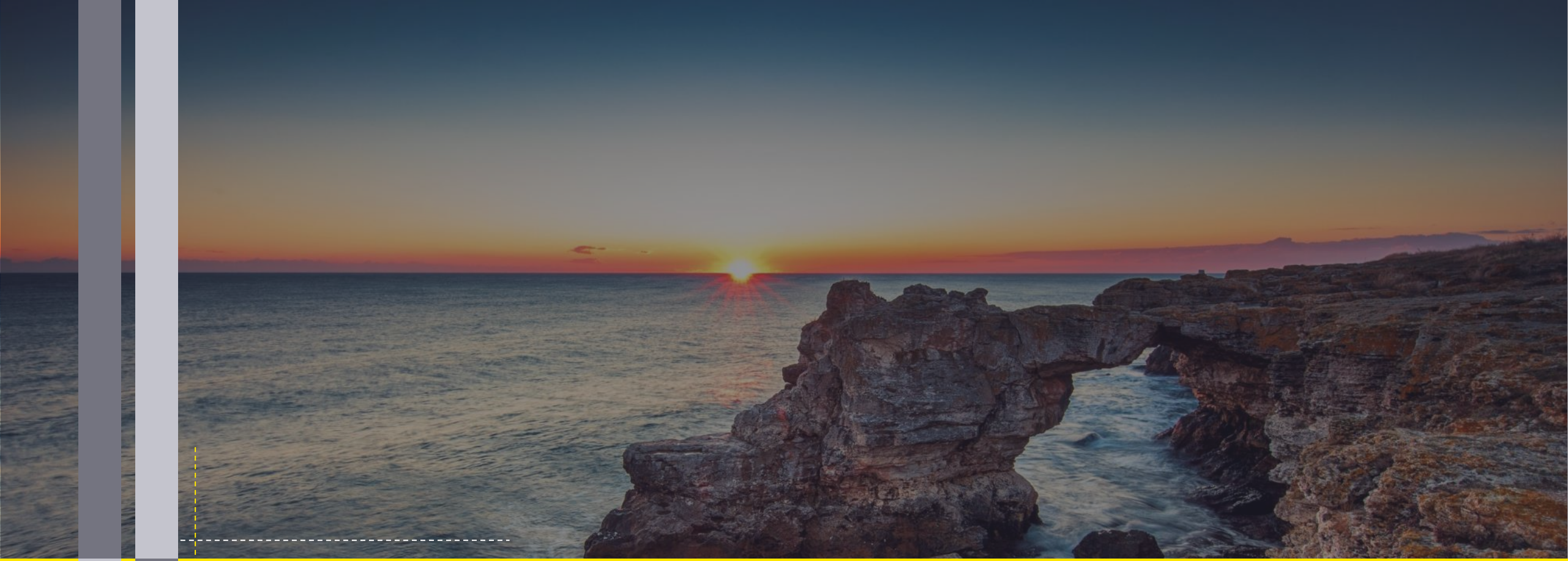
LLM

- Large Language Models (LLMs) **integrate Natural Language Processing (NLP) and Generative AI to create models that understand and generate human-like text** based on the patterns they have learned from vast amounts of training data.

ChatGPT

- **ChatGPT is an example of an LLM** launched by OpenAI in November 2022 that **recorded 1M users in a week from launch.**
- It is powered by Microsoft Azure; Microsoft invested $1B in OpenAI in 2019 and $10B in 2023.
- It builds upon GPT-3 family, **a Large Language Model (LLM) that uses a massive amount of data to understand and generate humanlike text.**
- Other market solutions include Google's Bard and Meta's LlaMa.





2. Generative AI risks and ownership considerations

LLMs introduce new risks in addition to amplifying existing ones

The Evolving AI Risk Taxonomy

Carryover risks from classical AI models (e.g., classical NLP models)

Model Risk

Lack of transparency and explainability leads to concerns about accountability and potential biases

Legal/Compliance Risk

New and more stringent **regulatory and compliance concerns** raised by AI models

Data Risk

Training data limitations and biases potentially lead to unreliable outcomes

Technology Risk

Technology limitations resulting from increased computational needs of AI models

Operational Risk

Business continuity for AI must be established to mitigate the risk of disruption to critical operations

+

Heightened or amplified risks introduced by LLMs

Model Risk

Hallucination resulting in plausible-sounding but inaccurate responses

Legal/Compliance Risk

Toxic information can be introduced by training data

Data/Technology Risk

Data transfer required to finetune LLMs increase **data hosting, sharing, retention, and security** risk

Data/Technology Risk

Model fine-tuning may access or share sensitive data, increasing **data privacy and PII Data** risk

Regulatory/Compliance Risk

Potential violation of **laws and regulations**

Third-party Risk

Most LLMs are provided by and rely on **third parties**

Legal/Copyright Risk

LLMs carry **copyright** risks if copyrighted material is reproduced without authorization

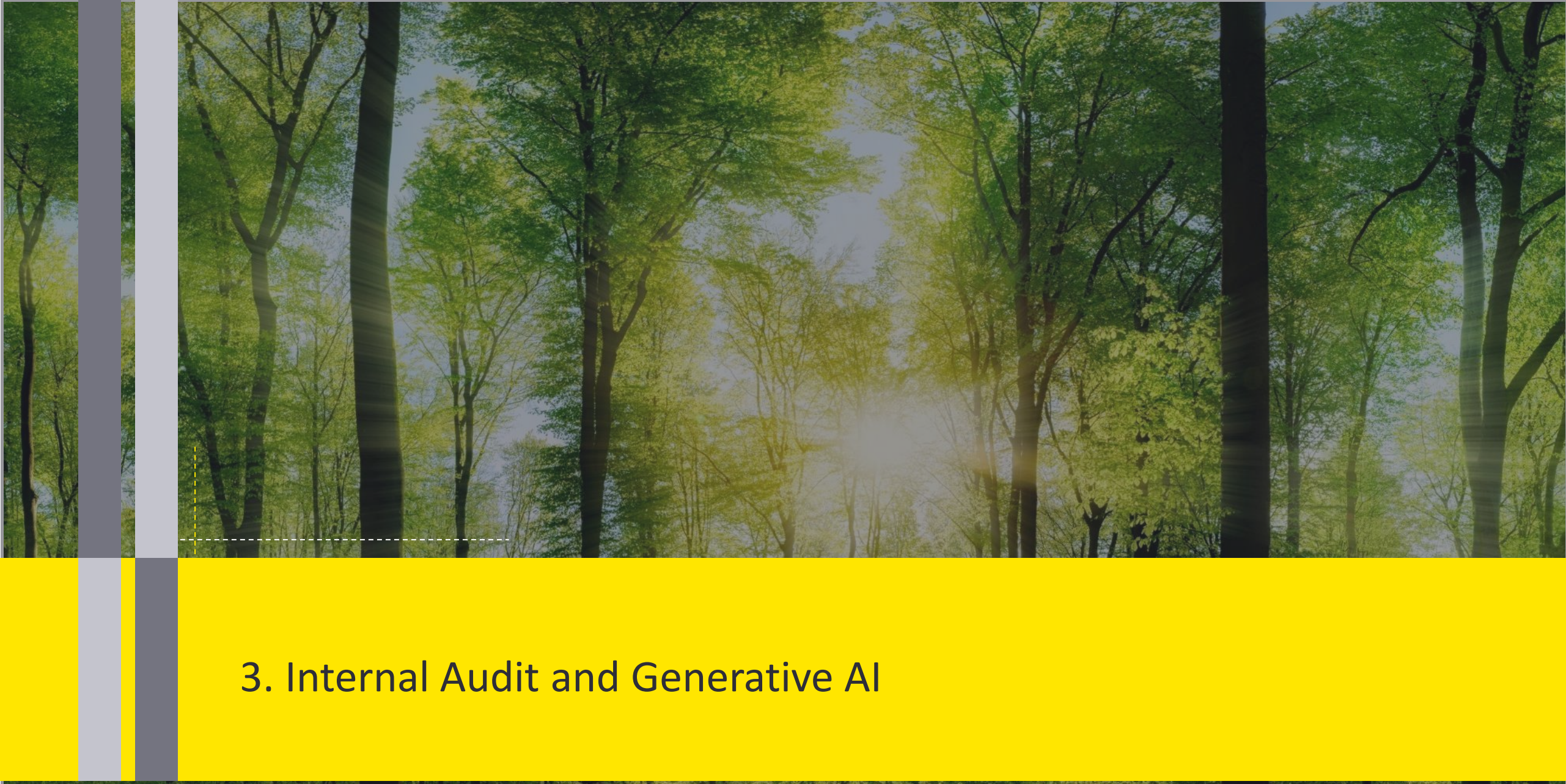
Reputational Risk

All the above risks may lead to **reputational damages**

LLMs require cross functional ownership to mitigate risk

		Owner												
		Developer				MRM			CDO		CTO/CISO	TPRM	Legal	Compliance
	Key Risks	Data Quality	Design/ Controls	Model Testing	Monitoring & Reporting	Concept Assessment	Model Testing	Compliance Assessment	Infrastructure	Mitigation Controls	Infrastructure	Risk Assessment	Legal Review	Compliance Review
1	Bias/Fairness													
2	Hallucination													
3	Toxic Content													
4	Privacy/PII													
5	Reputation Risk													
6	Legal/Reg													
7	Business Continuity													
8	Explainability													
9	Hosting/Retention/Sharing													
10	Data Capabilities													
11	Tech Capability													
12	Data Security													
13	Cyber Security													
14	Third Party Risks													
15	Copyright													

- Developer: key participants across the model life cycle serve as the primary owners of risk leveraging supportive organizations & gating bodies
- MRM: provides adequate challenge to the development team across the model risk component
- CDO: works to enable data capability and controls to mitigate data risks
- CTO: enables technology infrastructure and capability to mitigate technology and data risk
- Legal: advises developers, validators and users on all legal considerations
- Compliance: verifies compliance with new and existing regulatory requirements
- TPRM: addresses all third-party risk considerations



3. Internal Audit and Generative AI

Internal Audit's role in evaluating Generative AI: rethinking risk assessments

Given the heighted interest and rapid proliferation of LLMs at many organizations**, Internal Audit (IA) functions should assess its core capabilities to effectively assess how their organizations are addressing the heightened risks of LLMs.**

Internal Audit Core Capabilities

	Immediate	Near-term	Long-term
Risk Assessment & Audit Planning	✓ Seat at the table' for AI governance: ✓ Understand maturity of AI adoption ✓ Assess ownership and capabilities to address AI risks ✓ Effective challenge on risk mitigation across 1st and 2nd lines ✓ Establishing risk metrics and scoring methodology	✓ Review/leverage enterprise AI applicability/impact assessment ✓ Map AI / LLM models to auditable entities to assess risk exposure and coverage requirements ✓ Refresh audit universe and annual planning to cover associated risks (against the Enterprise inventory of all AI / LLM models) ✓ Documenting official Gen AI business objectives	✓ Continuous monitoring on AI / LLM adoption ✓ Refresh risk assessment on real-time and/or quarterly basis

Quarterly reporting Audit Committee and Board

Internal Audit's role in evaluating Generative AI: evolving our approach to audit

Given the heighted interest and rapid proliferation of LLMs at many organizations**, Internal Audit (IA) functions should assess its core capabilities to effectively assess how their organizations are addressing the heightened risks of LLMs.**

Internal Audit Core Capabilities

	Immediate	Near-term	Long-term
Audit Execution and Delivery	✓ Enterprise AI Readiness Assessment ✓ Develop standard audit programs to assess design and operating effectiveness of AI / LLM controls ✓ Perform skill-assessment surrounding AI / LLMs ✓ Identify opportunities for subject matter expertise and training ✓ Tailor/execute audit program to address specific / emerging risk (e.g., end-to-end, thematic, LLM specific audits) ✓ Model input/output documentation	✓ Governance review of AI adoption, including: ✓ Policies/Procedures ✓ Roles/Responsibilities ✓ Risk mitigation ✓ Effective challenge ✓ Management and Board Reporting ✓ Scoring methodology application review ✓ Conduct Training/awareness: ✓ AI and related concept definitions ✓ Enterprise response to AI / LLMs ✓ Key risk & controls ✓ Testing strategies	✓ Enhance audit programs based on maturity of AI / LLM use ✓ Assess significant risk areas for enhanced continuous monitoring and assurance coverage ✓ Continuous monitoring/back testing ✓ Uplift resourcing using skills-assessment (hiring, consultants, etc.) ✓ Ongoing training/awareness to audit teams

Quarterly reporting Audit Committee and Board

Gauging the readiness of our enterprise

Given the heighted interest and rapid proliferation of LLMs at many organizations**, Internal Audit (IA) functions should assess its core capabilities to effectively assess how their organizations are addressing the heightened risks of LLMs.**

Internal Audit Core Capabilities

	Gen AI ecosystem assessment	Governance, risk, and controls	Operating Model
Enterprise readiness assessment	✓ Gen AI LLM selection and adoption ✓ Third party LLM review and assessment ✓ Use case selection and prioritization framework (Innovation CoE) ✓ Standard operating guidelines ✓ Proof of concept evaluation	✓ Overall governance, policies and procedures review ✓ AI development process assessment, and testing guidelines ✓ Cybersecurity and IT controls review ✓ Business Continuity ✓ Change and incident management ✓ Compliance with Laws, Rules and Regulations	✓ Engagement model for businesses to adopt AI and Gen AI ✓ Defining roles & responsibilities ✓ Overall awareness for the workforce ✓ Training and upskilling based on roles & responsibilities

Quarterly reporting Audit Committee and Board

4. Application of Generative AI for audit

Generative AI use cases for Internal Audit

Information processing

Process and synthesize information like human beings and in exceptional speed

1 Knowledge management

Organize and analyse documents, regulations, policies, procedures, guidelines, education and audit work papers

2 Workpaper quality review

Review audit workpapers for quality issues (language, formatting, and alignment with organizational standards)

3 Process, risk and control diagnostics

Review, assess and rationalize Process, Risk, and Controls descriptions and connectivity

Information retrieval

Retrieves specific relevant information with contextual understanding

4 Testing Automation

Summarizing complaints/issues to detect and tag UDAAP / Sales Practices risk and categorize risk topics

5 Issue management

Automatically review, interpret and map issues to risk themes, and establish linkages to Product Risk Classification taxonomy and impacted regulation

6 Interpretation and summarization of unstructured data

Ingest, interpret and summarize Governance, regulatory, contract (underwriting, vendor etc.) documents to allow users to search contextually and summarize on-demand

New content generation

Generate new content by recognizing patterns across multiple sources

7 Business intelligence automation

Utilize natural language user prompts to automatically generate intuitive and insightful visualizations, charts, and summaries to present the analyzed data. Generate education and awareness materials for areas of non-compliance

8 CAE/Audit committee summary reports

Generate summary reports for CAE and audit committees using Gen AI to interpret testing outcomes and existing monitoring routines

9 Risk, control, issue and procedure generation

Automate generation of risk, control, issue and procedure documentation given a regulation/policy, to ensure consistency, clarity, reproducibility. Recommend remediation activities and draft remediation plans

4. Appendix

AI actors in AI lifecycle stages

Key dimensions	Application context	Data and input	AI model	AI model	Task and output	Application context	People and planet
Lifecycle stage	**Plan and design**	**Collect and process data**	**Build and use model**	**Verify and validate**	**Deploy and use**	**Operate and monitor**	**Use or impacted by**
Test, evaluation, verification, and validation (TEVV)	TEVV includes audit and impact assessment	TEVV includes internal and external validation	TEVV includes model testing	TEVV includes model testing	TEVV includes integration, compliance testing and validation	TEVV includes audit and impact assessment	TEVV includes audit and impact assessment
Activities	Articulate and document the system's concept and objectives, underlying assumptions, and context in light of legal and regulatory requirements and ethical considerations	Gather, validate and clean data and document the metadata and characteristics of the dataset, in light of objectives, legal and ethical considerations	Create or select algorithms; train models	Verify and validate, calibrate and interpret model output	Pilot, check compatibility with legacy systems, verify regulatory compliance, manage organizational change and evaluate user experience	Operate the AI system and continuously assess its recommendations and impacts (both intended and unintended) in light of objectives, legal and regulatory requirements and ethical considerations	Use system/technology; monitor and assess impacts; seek mitigation of impacts, advocate for rights
Representative actors	System operators; end users; domain experts; AI designers; impact assessors; TEVV experts; product managers; compliance experts; auditors; governance experts; organizational management; C-suite executives; impacted individuals/communities; evaluators	Data scientists; data engineers; data providers; domain experts; socio-cultural analysts; human factors experts; TEVV experts	Modelers; model engineers; data scientists; developers; domain experts; with consultation of socio-cultural analysts familiar with the application context and TEVV experts		System integrators; developers; system engineers; software engineers; domain experts; procurement experts; third-party suppliers; C-suite executives; with consultation of human factors experts, socio-cultural analysts, governance experts, TEVV experts	System operators, end users and practitioners; domain experts; AI designers; impact assessors; TEVV experts; system funders; product managers; compliance experts; auditors; governance experts; organizational management; impacted individuals/communities; evaluators	End users; operators and practitioners; impacted individuals/communities; general public; policy makers; standards organizations; trade associations; advocacy groups; environmental groups; civil society organizations; researchers

Define risk mitigation for new/heightened risk brought in LLM (1/3)

Risk Type	Challenges	Roles	Responsibilities
Model Risk	**Hallucination and model accuracy**	CDO	• Identify and assess inaccuracies and falsification in training data
		Users	• Understand the LLM's limitations • Ask LLM the same questions in multiple ways to form an "ensemble" model in order to improve the model accuracy • Request LLM for reflection of the output until converge and/or request more details of the output such as sources, references and rationale as applicable
		Developers/ MRM	• Conduct the model outcome testing (e.g., sensitivity, scenario, stress testing) to identify model limitations • Provide prompt guidance to improve the accuracy of the model output and integrate users' feedback on a timely basis
	Explainability	Developer/MRM	• Actively follow latest develop to identify third-party explainability tools and methodologies to provide insight into the model output
		Users	• Use prompt engineering to design prompts from multiple angles in order to gain a more comprehensive views • Request LLM to explain the rationale of the output and ask multiple times for convergence
Third-party Risk	**All LLMs are provided by third-party**	TPRM	• Review third-party systems to understand LLM capabilities and limitations. • Work with model risk, CDO, compliance, legal, cyber to perform comprehensive risk assessment of vendor solutions with a focus on data protection
Conduct/ Compliance Risk	**Bias/Fairness**	Compliance/ MRM	• Define ethical risk (e.g., bias/unfairness) management standard (definition, detection, remediation, monitoring) • In conjunction with Model Risk team or other utility functions to 1) review LLM models for any ethical risks and define mitigation controls 2) perform ongoing testing of LLM to identify instances of control failure
		Developers/Users	• Understand the model limitations that could result in any ethical concerns; Consider Ethical Principles for AI • Ensure that the data preparation process for fine tuning does not introduce additional bias.
		MRM	• Install a tollgate process to identify the potential ethical risks at the initial stage of model development • Per MRM guidance, as applicable, perform independent reviews to identify bias/unfairness through a wide variety of techniques (e.g., change gender to assess the impact on model performance, identify the association of race with certain activity/occupation to detect serotypes)
	Toxic information	Compliance/ MRM/Developer/Users/CDO	• (Compliance) Define/enhance the standard for toxic information and identify and (CDO) filter potential toxic data in training materials • Developer and MRM perform testing to identify the scenarios potentially trigger toxic output, and come up with a remediation or mitigation controls and ongoing monitoring plan

Define risk mitigation for new/heightened risk brought in LLM (2/3)

Risk Type	Challenges	Roles	Responsibilities
Data/ Technology Risk	Data privacy and PII data	CDO	• Review/update data classification standard and examine the implementation for those data fed to LLM • Implement data privacy assessment/PII scrubbing covering data for fine-tuning, in-context learning, prompt and completion/generated output • Establish guard rails for accessing and handling sensitive data in fine-tuning, in-context learning, prompt and completion/generated output, throughout generative AI POC/production lifecycle (developer, MRM, operation, DevOps, reporting, user, audit, etc.)
		Compliance/Regulations	• Map LLM data risks to existing regulations to update standards and meanwhile develop mitigations • Closely monitor the new regulations of LLM in relation to data privacy and automated decision making in all jurisdictions
	Data hosting, sharing and security	TPRM/CICO	• Review/update existing procedure and standard for data in transit and data in rest (e.g., encryption, key management) • Review/update vendor agreement on • Data retention, event logging and data accessing by vendor personal for system purpose • Data sharing, i.e., use bank's data for other purposes by vendor or other parties implicitly • Review/update vendor platform configurations to identify potential data breaches
		CDO	• Review/update existing data architecture and data flow as well as governance and controls to minimize the unintended violation of relevant policies
	Data capability	CTO/CDO	• Assess and establish the infrastructure requirements for handling heightened data risks due to GenAI use across privacy, compliance and model risk
	Technology capability	CTO	• Provide the appropriate platform for fine-tuning training data storage, processing and data pipeline • Provision needed coding libraries from vendor/open source • Provide the infrastructure to support model fine-tuning, embedding, query, analysis • Provide the technology capability for users to log and send effective feedback to developers to enhance the model performance (which is important give that the output of GAI is mostly related to natural language, hard to define objective metrics and rely on human users to monitor the performance and provide subjective feedback)
		TPRM	• Provide vendor integration and secure environment for experimentation • Assess technology architecture on vendor platform for the developmental environment • Understand the chain of LLM capabilities (e.g., vendor may be licensing GAI capacities from multiple vendors) • Pay close attention to the vendors' self-identified limitations, restrictions and terms and conditions for use • Understand and evaluate vendor's preventive controls for LLM (e.g., toxic data detection)

Define risk mitigation for new/heightened risk brought in LLM (3/3)

Risk Type	Challenges	Roles	Responsibilities
Legal/ Regulatory Risk	**Lawsuit, reg penalty and Copyright**	Legal	• At inception, evaluate if there is legal concerns in the LLM use cases • Track the latest development in output copyright of Generative AI • Provide timely update to all stakeholders, management team and board on law changes
		Reg policy	• At inception, evaluate if there is regulatory concerns in the LLM use cases • Track the latest development in the regulations of Generative AI use as well as data protection • Provide timely update to all stakeholders, management team and board on regulation changes
		Developer	• Raise potential legal/regulatory concerns to GCO or regulatory policy office before and during the model development and post-production run • For any control failure during the LLM development/deployment and production run, consider the potential legal/regulatory consequence, consulting GCO or regulatory policy office
Cyber Risk	**Cyber attack and adversarial attack**	Developer/ MRM/Users	• Perform regular internal and vendor code scan, testing for vulnerabilities, and following secure coding standards • Design prompt to test the boundary or specific prompt approaches to identify the scenario leading to jailbreakers • Log all the prompt history for ex-post analysis • Develop and implement automated monitoring tools to detect suspicious prompt behaviors
		Cyber	• Protection should be developed for all type of training data, prompt history, output, trained LLM model, event log, training history, performance data
Operational risk	**Business continuity**	Developer	• For important use cases, establish a business continuity plan including fall-back solutions, challenger models and include a breach protocol/champion challenger workflow in on-going monitoring to trigger fall-back solution
Reputational Risk	**Linked to all other risks**	All functions	• Ensure timely escalate of incidences of data breach, compliance, legal issues

EY | Building a better working world

EY exists to build a better working world, helping 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.

2307-4276195
ED None

ey.com