ISACA

Leveraging COBIT for Effective AI System Governance

Governance

CONTENTS

ABSTRACT

With artificial intelligence (AI) becoming an essential driver of innovation and efficiency across industries, organizations face mounting pressure to govern these systems responsibly. This white paper explores the role of the COBIT® framework as a robust, adaptable solution for effective AI governance and management. Traditionally employed for information and technology (I&T) governance, COBIT is uniquely positioned to address the distinct challenges AI systems introduce, including issues related to ethics, accountability, transparency, and compliance. The framework's structured approach provides a holistic, life cycle-based model that guides organizations on how to align AI initiatives with strategic business objectives, optimize resource allocation, and mitigate AI-specific risk.

The Role of AI in Modern Enterprises

As AI has become increasingly integral to modern enterprises, the need for robust governance and management of AI has grown. Many organizations use machine learning (ML), deep learning (DL), and generative AI to streamline operations through automation, improve decision making, and enhance customer and user experience. Whether automating user or customer service processes or generating predictive analytics, AI systems offer new opportunities for efficiency and innovation.

However, deploying AI also presents unique challenges that demand careful governance. Unlike traditional IT systems, AI introduces ethical concerns around bias, fairness, and transparency, which must be managed across multiple jurisdictions and industries. As AI rapidly advances, the pace of innovation often surpasses the ability to fully anticipate its impacts and implications, leaving AI systems vulnerable to threats like manipulation, data poisoning, cyberattacks, and security breaches.

Unlike traditional IT systems, AI introduces ethical concerns around bias, fairness, and transparency, which must be managed across multiple jurisdictions and industries. As AI rapidly advances, the pace of innovation often surpasses the ability to fully anticipate its impacts and implications, leaving AI systems vulnerable to threats like manipulation, data poisoning, cyberattacks, and security breaches.

In addition, management must also be aware that AI features are embedded in commercial-off-the-shelf software applications and commonly used search engines that are likely already in use in the enterprise. Lack of clear ownership of AI governance exacerbates these issues, often resulting in fragmented oversight and difficulty aligning AI systems with an organization's overall business strategy.

AI Systems and Strategic Importance

AI systems encompass several key technologies, each with unique capabilities that contribute strategically across industries. ML, for example, involves training algorithms on data to recognize patterns and make predictions. In enterprises, ML can be used for demand forecasting, fraud detection, and targeted marketing, ultimately leading to improved decision making.

DL, a subset of ML, uses neural networks with many layers to analyze vast and complex datasets. DL enables advanced image recognition, natural language processing, and autonomous systems, all of which have a significant impact on diagnostics, analysis, and workflow automation.

In addition, generative AI models, including generative adversarial networks (GANs) and large language models (LLMs), create realistic text, images, or simulations that are similar to the input data provided. Businesses often use generative AI for content creation, personalized experiences, and rapid prototyping.

AI's strategic significance can vary by industry. In healthcare, AI can enhance diagnostics, enable personalized treatment plans, and accelerate drug discovery, all of which improve patient outcomes and reduce costs. In the finance industry, AI can automate trading, enhance fraud detection capabilities, and transform financial operations to make them faster, safer, and more data-driven. AI can also perform predictive maintenance to optimize production lines, enhance quality control, and manage supply chain logistics for manufacturing organizations. In addition, AI can support the retail industry by personalizing customer interactions, managing inventory, predicting demand, driving sales, and improving the customer experience.

Given AI's potential to autonomously influence operations, manage sensitive data, and impact consumer experiences at scale, specialized governance frameworks are vital. Effective AI governance ensures that AI systems align with organizational and societal ethics, protect data privacy, and maintain compliance, all while managing the unique risk associated with these autonomous systems.

Risk and Challenges of AI Systems

To achieve AI's benefits, organizations must navigate its risk and challenges. First, AI bias in models can emerge from the use of data that may reflect systemic or historic inequalities or human biases. Left unchecked, biased AI can lead to discriminatory practices that have legal repercussions, especially in sensitive applications like hiring or lending. For example, in the case of *Saas v. Major, Lindsey & Africa LLC*,[1] the plaintiff alleged the company's algorithmic bias resulted in her being screened out of job opportunities. There was also a case in France where algorithms designed to detect people most likely to commit welfare fraud led to allegations of discrimination against people with low incomes.[2] Mitigating this risk requires deliberate bias testing and regular audits of training data and model behavior.

In addition, many AI models, particularly in DL, function as "black boxes," which makes it difficult to understand how AI decisions are made. This lack of transparency poses challenges to accountability, eroding trust among users, consumers, and stakeholders. Employing explainable AI techniques is critical to help reveal model logic, enable more responsible AI use, and comply with evolving regulatory guidance. Article 50 of the EU Artificial Intelligence Act provides guidance on the transparency expectations of providers and deployers of AI systems.[3]

Recital 93 also requires disclosures upon request to people when they interact with high-risk AI systems and when AI is used in decision making (e.g., safety).[4]

Many AI models, particularly in DL, function as "black boxes," which makes it difficult to understand how AI decisions are made. This lack of transparency poses challenges to accountability, eroding trust among users, consumers, and stakeholders. Employing explainable AI techniques is critical to help reveal model logic, enable more responsible AI use, and comply with evolving regulatory guidance.

Further, noncompliance with data protection and privacy regulations and legislation (such as the EU General Data Protection Regulation [GDPR] or California Consumer Privacy Act [CCPA]) can lead to legal penalties and reputational damage. For example, AI users can intentionally or unintentionally compromise trade secrets and other intellectual property through the use of ChatGPT.[5] Organizations must implement privacy-preserving techniques, such as data anonymization and differential privacy, to maintain compliance.

AI systems are vulnerable to more complex cyberattacks, including adversarial attacks that intentionally manipulate model inputs to produce erroneous outputs. For instance, altering a few pixels in an image could deceive an AI model and potentially misclassify critical information. This form of adversarial attack could conceivably result in a stop sign being perceived as a speed limit sign in a deep neural network.[6] Also, model-inversion attacks pose the risk of sensitive data being extracted from trained models. Such risk makes it essential for organizations to adopt holistic governance and management practices related to AI systems.

1 *Saas v. Major, Lindsey & Africa, LLC*, "Memorandum Opinion," United States District Court, D. Maryland, 10 May 2024, https://scholar.google.com/scholar_case?case=7736063661252278773
2 Meaker, M.; "Algorithms Policed Welfare Systems for Years. Now They're Under Fire for Bias," *Wired*, 16 October 2024, https://www.wired.com/story/algorithms-policed-welfare-systems-for-years-now-theyre-under-fire-for-bias/
3 EU Artificial Intelligence Act, "Article 50: Transparency Obligations for Providers and Deployers of Certain AI Systems," https://artificialintelligenceact.eu/article/50/
4 EU Artificial Intelligence Act, "Recital 93," https://artificialintelligenceact.eu/recital/93/
5 ISACA, "The Promise and Peril of the AI Revolution: Managing Risk," 12 September 2023, https://www.isaca.org/resources/white-papers/2023/the-promise-and-peril-of-the-ai-revolution
6 Elsayed, G.; Mozer, M.; "Images Altered to Trick Machine Vision Can Influence Humans Too," Google DeepMind, 2 January 2024, https://deepmind.google/discover/blog/images-altered-to-trick-machine-vision-can-influence-humans-too/

Governance Guidance Examples

Several AI governance frameworks, standards, and laws have emerged to address ethical use, security, accountability, and fairness, including:

- The Organisation for Economic Co-operation and Development (OECD) AI Principles emphasize human rights, democratic values, transparency, and accountability in AI applications.[7] These principles guide organizations to ensure AI is used responsibly and ethically. By aligning AI initiatives with these principles, businesses can gain public trust and avoid potential ethical and legal pitfalls. Transparent AI systems that minimize bias enhance customer satisfaction and brand loyalty, while those that fail to meet these expectations could lead to reputational damage, legal consequences, or regulatory scrutiny. However, without enforceable measures, adherence relies only on voluntary organizational commitment.
- In the United States, the National Institute of Standards and Technology (NIST) AI Risk Management Framework (AI RMF) focuses on managing cyberrisk associated with AI.[8] NIST's framework provides organizations with a structured approach to developing AI systems that are secure, resilient, and compliant with cybersecurity best practices. Adopting the NIST AI RMF reduces an organization's risk of cyberattacks or data breaches impacting its AI systems, which could otherwise result in financial losses, reputational damage, regulatory noncompliance, and customer distrust.
- ISO/IEC 42001:2023 is a global standard that centers around creating an Artificial Intelligence Management System (AIMS) and includes specific requirements for risk assessment, transparency, and data governance. The standard emphasizes the need for structured risk management and encourages organizations to analyze and document their AI systems' societal, environmental, and data privacy impacts.[9] Furthermore, ISO/IEC 42001 supports a culture of continuous improvement, requiring organizations to regularly review their AI processes to stay aligned with industry best practices and regulatory expectations. By adopting this standard, organizations can build stakeholder trust, streamline compliance, and proactively manage potential AI-related risk, thus positioning themselves as leaders in responsible AI usage.
- The EU AI Act is a comprehensive regulatory framework proposed by the European Union that categorizes AI systems based on risk levels and implements strict requirements for acceptable use.[10] Organizations using AI applications classified as high risk, such as those used in healthcare or finance, must implement rigorous compliance processes to meet the Act's stringent requirements. Noncompliance could result in fines, similar to the penalties issued under the GDPR.

By proactively aligning with this guidance, enterprises can mitigate the risk of penalties, enhance their reputations as trustworthy organizations, and appeal to a broader customer base concerned with data privacy and ethical AI usage. While it offers a structured approach for ethical and responsible AI, it may lack flexibility for fast-evolving technologies.

By proactively aligning with this guidance, enterprises can mitigate the risk of penalties, enhance their reputations as trustworthy organizations, and appeal to a broader customer base concerned with data privacy and ethical AI usage.

Gaps in Coverage

While these examples offer a strong foundation for AI governance, significant gaps remain. The limited coordination between countries has resulted in inconsistent regulations and border enforcement. For example, the EU AI Act emphasizes stringent controls and risk-management practices for high-risk AI systems. Other regions may have more lenient standards or lack comprehensive AI governance altogether. These disparities can create challenges for organizations operating in multiple regions, as they must navigate a complex web of regulations that may conflict or lack interoperability. Inconsistent regulatory approaches can produce uncertainty, complicate compliance efforts, and

7 Corba, J.; Plonk, A.; et al.; "Evolving with innovation: The 2024 OECD AI Principles update," OECD.AI, 20 May 2024, https://oecd.ai/en/wonk/evolving-with-innovation-the-2024-oecd-ai-principles-update
8 NIST, "AI Risk Management Framework," 2024, https://www.nist.gov/itl/ai-risk-management-framework
9 Wright, D.; "AI Governance Blueprint: ISO 42001 & NIST AI RMF," Techstrong.ai, 1 August 2024, https://techstrong.ai/articles/ai-governance-blueprint-iso-42001-nist-ai-rmf/
10 European Commission, "European approach to artificial intelligence," 18 December 2024, https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence

increase operational costs for businesses that need to tailor their AI governance strategies to meet varying international standards.

The limited coordination between countries has resulted in inconsistent regulations and border enforcement.

Certain industries, such as healthcare, have made substantial progress in AI governance, driven by the critical need to ensure patient safety, data privacy, and ethical decision making in AI-driven medical applications.[11]

Sectors like finance and education lag behind, lacking comprehensive governance models that address the risk AI poses in these fields. In education, for example, AI is being employed for personalized learning and administrative efficiencies but governance standards to ensure fairness, transparency, and bias mitigation are still in their infancy.[12] This sectoral divide leaves potential risk unaddressed, particularly concerning ethical use, bias, and data security.

The most critical gap, however, lies within organizations themselves, where there is often a lack of clear ownership of AI governance. Boards and executive management must understand AI benefits and risk to help ensure accountable and responsible roles are clearly defined to avoid fragmented oversight, which can hinder an organization's ability to fully understand the broader impact of AI on its operations.

Boards and executive management must understand AI benefits and risk to help ensure accountable and responsible roles are clearly defined to avoid fragmented oversight, which can hinder an organization's ability to fully understand the broader impact of AI on its operations.

Without clear ownership, key stakeholders, including IT leaders, compliance officers, and business executives, may fail to align AI systems with the organization's policies and overall strategy. This misalignment can result in disjointed implementations, inefficiencies, and an unclear understanding of the risk and opportunities associated with AI deployment.

Overview of the COBIT Framework

As AI technology continues to evolve, so will the need for adaptive governance frameworks. The COBIT framework can address the gaps by offering a holistic approach to AI governance. Its primary goal is to help organizations align IT with business objectives, manage risk, and ensure the optimal use of resources.

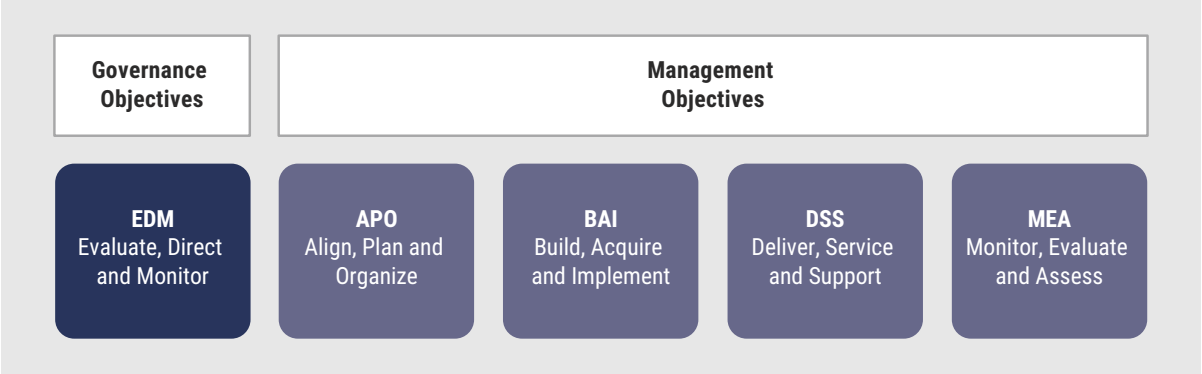
COBIT achieves this by offering a comprehensive framework for the governance and management of I&T. It emphasizes accountability and responsibility to help organizations optimize the value of their I&T investments to reduce the associated risk.

Domains and Objectives

COBIT is structured into five key domains (**figure 1**). These domains include 40 objectives designed to manage IT-related activities, ranging from strategic planning to day-to-day operations and performance monitoring.

11 Price II, W.; "Risks and Remedies for Artificial Intelligence in Health Care," Brookings, 14 November 2019, https://www.brookings.edu/articles/risks-and-remedies-for-artificial-intelligence-in-health-care/
12 Q.ai; "Applications of Artificial Intelligence Across Various Industries," Forbes, 6 January 2023, https://www.forbes.com/sites/qai/2023/01/06/applications-of-artificial-intelligence/

FIGURE 1: COBIT Domains



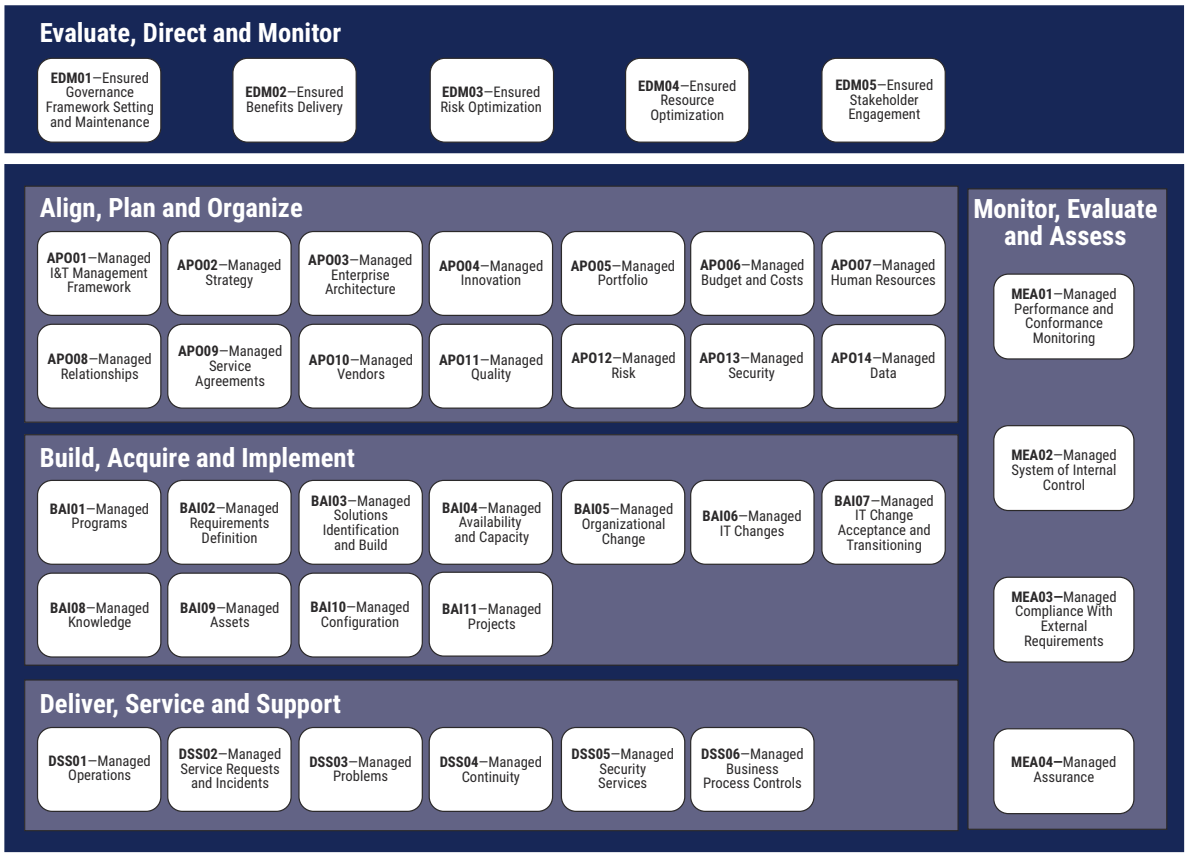
The five COBIT domains are:

- **Evaluate, Direct and Monitor (EDM)**—This domain is primarily focused on governance activities at the highest level. In the EDM domain, COBIT ensures that technology solutions and related I&T initiatives align with the organization's strategic objectives. Within the EDM domain, leadership sets the I&T objectives, goals, and strategy to ensure that all technological investments and efforts directly contribute to business success. Regular monitoring of performance is a critical component of this domain, helping leadership assess whether the technology is performing as expected and delivering the intended value. This continuous alignment between the technology and business outcomes ensures that the organization maintains a clear strategic direction supported by effective, trusted governance for technology management.
- **Align, Plan and Organize (APO)**—The APO domain deals with the strategic planning of technology solutions and initiatives. It focuses on developing IT policies, risk management strategies, and overall resource planning to ensure that technology investments are structured to deliver maximum value. APO objectives help organizations design systems that are well-planned, aligned with business goals, and capable of supporting long-term objectives. By focusing on both short-term operational needs and long-term strategic goals, APO objectives ensure that systems are implemented with a clear vision for their role in the organization's success.
- **Build, Acquire and Implement (BAI)**—In this domain, COBIT addresses the practical aspects of developing and deploying systems and technologies. This includes any systems that are built, acquired, or integrated into the business. The BAI objectives ensure effective and efficient development and implementation of these systems, minimizing the risk associated with new technology deployments. The BAI domain covers the full spectrum of activities needed to ensure that technology solutions meet performance requirements and are deployed in a manner that minimizes disruption to ongoing operations.
- **Deliver, Service and Support (DSS)**—The DSS domain focuses on the ongoing operations of technology and systems once they are deployed. The DSS objectives govern the delivery of IT services, ensuring that they meet the agreed-upon service levels, including security, system continuity, and performance. The DSS objectives also ensure that services are supported by robust security frameworks and continuity plans to minimize downtime and address potential disruptions.
- **Monitor, Evaluate and Assess (MEA)**—This domain emphasizes the importance of continuous improvement within the organization's I&T governance framework. It focuses on evaluating the effectiveness of systems and governance structures, ensuring that they continue to meet organizational goals and perform effectively. Through regular performance assessments and governance evaluations, the MEA objectives help organizations identify areas for improvement, adapt to changing business needs, and ensure that the technology remains aligned with the organization's evolving strategic objectives. This ongoing feedback loop is essential for maintaining a dynamic and resilient I&T governance system.[13]

Figure 2 illustrates the 40 objectives organized into the five domains.

13 ISACA, *COBIT® 2019 Framework: Introduction and Methodology*, USA, 2018, https://store.isaca.org/s/store#/store/browse/detail/a2S4w000004Ko5dEAC

FIGURE 2: COBIT Core Model



Source: ISACA, *COBIT® 2019 Framework: Introduction and Methodology*, USA, 2018

Components of a Governance System

Each enterprise needs to establish, tailor, and maintain a multifaceted governance system to satisfy governance and management objectives. COBIT defines seven component types that interact to align I&T initiatives with business objectives:

1. Processes
2. Organizational Structures
3. Information
4. People, Skills, and Competencies
5. Principles, Policies, and Procedures
6. Culture, Ethics, and Behavior
7. Services, Infrastructure, and Applications

Each plays a critical role in supporting specific governance and management objectives, working together to ensure that systems are effectively managed across their life cycles.

AI, as part of the Services, Infrastructure, and Applications component, represents a key technological advancement that requires alignment with broader governance components. However, enterprises often develop services, applications, and infrastructure enabled by AI without focusing on all the governance components (e.g., policies, information flow, culture, ethics and behaviors).

AI technologies, including ML models, decision-making algorithms, and intelligent systems, naturally fit into this component. They are infrastructure elements providing critical services or applications designed to enable automation, decision making, and insights. However, their successful integration and use require strong governance

mechanisms, as software is rarely static and software owners routinely push out updates with additional features that should be reevaluated.

However, other COBIT governance components have a supporting role in the governance of AI. These AI technologies cannot function effectively in isolation. Governance components that play essential roles in enabling, supporting, and regulating the use of AI technologies include:

- **Processes**—Structured workflows and methods are necessary to guide the development, deployment, and management of AI technologies. For example, a robust AI life cycle management process ensures these systems are maintained, updated, and aligned with business goals.
- **Organizational Structures**—Governance structures define who is responsible for AI oversight, including ethical decision making, risk management, and accountability for outcomes.
- **Information**—AI relies heavily on data to operate effectively. Proper governance ensures data quality, privacy, security, and appropriate usage, avoiding biases and inaccuracies that could compromise AI functionality.
- **People, Skills, and Competencies**—Skilled professionals are critical for designing, deploying, and maintaining AI. Governance ensures ongoing development and alignment of expertise with organizational needs.
- **Principles, Policies, and Procedures**—Clear policies are essential to guide the ethical use of AI. For example, organizations must establish guidelines to address data privacy, algorithmic transparency, and AI-driven decision-making accountability.
- **Culture, Ethics, and Behavior**—This component governs how individuals and teams interact with AI systems. This governance component fosters a culture of trust, ethical awareness, and responsible usage, ensuring that technology aligns with organizational values.

Today, organizations often focus heavily on developing cutting-edge technologies like AI while neglecting the necessary governance ingredients that ensure these technologies are used responsibly and effectively.

Today, organizations often focus heavily on developing cutting-edge technologies like AI while neglecting the necessary governance ingredients that ensure these technologies are used responsibly and effectively.

This imbalance can lead to:

- Ethical concerns such as bias in AI models
- Inefficiencies due to unclear processes, unclear requirements, or lack of requisite expertise
- Resistance to adoption due to poor organizational culture alignment

A balanced governance approach ensures that AI-enabled services, applications, and infrastructure contribute effectively to business objectives while adhering to ethical standards and organizational policies. By considering all governance components, an enterprise can mitigate risk and unlock AI's full potential.

All of the components shown in **figure 3** are interdependent and must be considered collectively as organizations progress through the phases of IT development.[14]

FIGURE 3: COBIT Components of a Governance System

Source: ISACA, *COBIT® 2019 Framework: Introduction and Methodology*, USA, 2018

14 ISACA, *COBIT® 2019 Framework*

Governance System Design Considerations

In addition to the seven main components, the COBIT framework introduces the goals cascade as a key design element for determining the initial scope of the governance system. It is embedded in the 11 design factors to consider when selecting the most appropriate governance and management objectives and ensuring that IT-related activities are closely aligned with the overall enterprise goals and objectives. Considering enterprise goals and the application of the goals cascade is a key step in the governance system design workflow.

The cascade starts with "Stakeholder Drivers and Needs," which identifies stakeholder expectations, market demands, and other external factors. These drivers define what the business needs to achieve to remain competitive and effective.

The business needs are translated into "Enterprise Goals." COBIT defines 13 generic enterprise goals that span financial, customer, internal, and growth dimensions. These goals include improving customer satisfaction, optimizing business costs, and ensuring compliance with regulations.

"Alignment Goals" relate to the larger enterprise goals and provide the primary link to the governance or management objective. This linkage helps ensure that I&T initiatives contribute to the success of the business. Alignment goals help define how systems should be managed, ensuring that technology investments are aligned with the organization's business objectives.

Finally, 40 "Governance and Management Objectives" guide the enterprise in the design and implementation of an effective governance system over I&T processes.

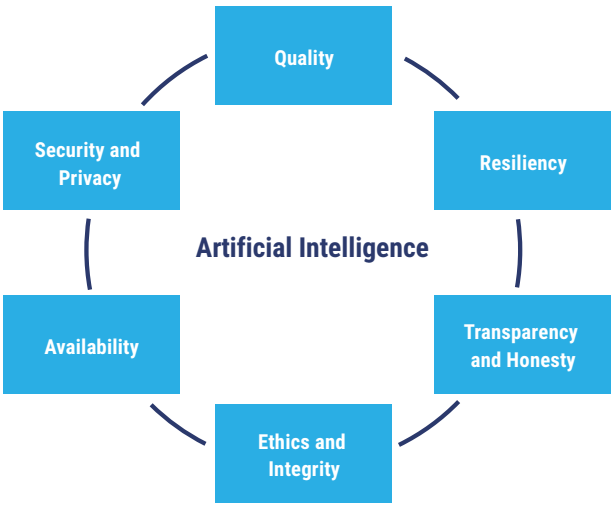
These steps ensure that the efforts of IT and governance teams are focused on activities that directly support the organization's overall strategy. By following this approach, companies can prioritize their I&T governance activities and allocate resources to the processes that will have the most significant impact on achieving stakeholder needs.[15] These steps can be applied to design a coordinated governance approach to managing AI.

The Principles of Trustworthy AI

Trustworthy AI encompasses principles that help mitigate risk associated with AI systems to foster stakeholder confidence—such as transparency, accountability, fairness, privacy, reliability, safety, resilience, and security.[16] Applying COBIT's governance and management objectives to AI across the organization establishes a comprehensive framework that spans from high-level governance to operational processes and systems. This approach ensures that AI-related initiatives are managed end-to-end, aligning strategic goals with operational execution while embedding the principles of trustworthy AI at every level. By addressing organizational governance, policymaking, risk management, and compliance at the top and cascading these principles down through well-defined processes, roles, and systems, COBIT helps integrate AI effectively into the organization's ecosystem.

Tailoring COBIT's governance and management objectives to AI goals can help the enterprise design a structured governance approach to managing AI-related risk while ensuring that the initiatives align with the elements of trustworthy AI shown in **figure 4**.

FIGURE 4: Elements of Trustworthy AI



Source: ISACA, "Using the Digital Trust Ecosystem Framework to Achieve Trustworthy AI," 30 April 2024, https://www.isaca.org/resources/white-papers/2024/using-dtef-to-achieve-trustworthy-ai

15 ISACA, *COBIT® 2019 Framework*
16 NIST, "AI Risk Management Framework"

Data quality is important to ensure datasets contain accurate information and bias is identified and mitigated. The organization must ensure there are established quality criteria and that periodic assessments are performed. Safety in AI systems is also tied in with AI trustworthiness because it protects users and the environment from potential harm that might arise from AI system operations.

- APO11 (Managed Quality) supports the need for documentation to outline the quality assurance and safety practices followed for AI systems, including the data-related aspects.
- BAI03 (Managed Solutions Identification and Build) ensures that data used to build AI models is high quality, relevant, and secure.

Resiliency is the ability of an AI system to adapt to unexpected changes or disruptions, mitigate the likelihood of failure, and respond appropriately after a failure, which helps maintain operational integrity. Resilient AI systems can withstand disturbances such as data shifts or cyberattacks.

- AI systems often require frequent updates, from model changes to algorithm adjustments. BAI06 (Managed IT Changes) supports structured change management processes, ensuring that AI system modifications are well-coordinated to minimize disruption and align with business priorities.

Transparency and honesty in AI mean that the system's processes, decisions, and data sources are clearly documented and understandable to users and stakeholders. This is foundational to explainable AI (XAI), and enterprises will be accountable for providing evidence that supports it. Accountability will also drive designated individuals or teams to be responsible for AI system governance and outcomes, including addressing any issues that arise.

- EDM01 (Ensured Governance Framework Setting and Maintenance) requires the establishment and enforcement of a governance framework that clarifies roles and maintains organizational accountability. This objective aligns with the need to designate AI oversight, ensuring that decision-making and accountability structures are established and visible across teams, especially given AI's complex interdependencies and data requirements.

Ethics and integrity in AI align closely with fairness in AI systems. A key goal is to prevent biases and ensure equitable treatment across different user groups.

Ethics and integrity in AI align closely with fairness in AI systems. A key goal is to prevent biases and ensure equitable treatment across different user groups.

This involves mitigating data or algorithmic biases that could lead to discriminatory outcomes.

- EDM01 (Ensured Governance Framework Setting and Maintenance) can align the ethical development, deployment, and use of AI technologies with the enterprise's direction, goals, and objectives. Organizations should ensure that AI systems are designed and utilized in ways that uphold societal values, respect the natural environment, and address the interests and rights of internal and external stakeholders. This includes fostering transparency, fairness, accountability, and inclusivity in AI processes to mitigate potential biases, protect privacy, and promote trust in AI-driven decision making.
- Ethical considerations in AI also align with COBIT's APO01 (Managed I&T Management Framework), which mandates that technology strategies (including AI) adhere to ethical standards. EDM01 (Ensured Governance Framework Setting and Maintenance) and EDM02 (Ensured Benefit Delivery) also support this by considering the potential impact of unethical use and data processing on society and the natural environment, as well as monitoring the outcomes of AI systems to ensure they align with ethical use cases and organizational values.

Security in AI involves implementing safeguards to protect the AI system and its data from unauthorized access and attacks. AI systems need robust security controls to prevent the exploitation of vulnerabilities and ensure data integrity.

- Given AI's susceptibility to data breaches and adversarial attacks, DSS05 (Managed Security Services) aids in enforcing security protocols, while BAI09 (Managed Assets) provides guidelines for securing the infrastructure housing AI systems. Together, these objectives establish secure environments for data storage, model training, and operations.

Privacy ensures that AI systems protect personal information and comply with relevant data protection regulations. This principle is especially critical for AI systems handling sensitive or identifiable information.

- It is important to have an effective data management strategy that outlines the roles and responsibilities for metadata management and the data quality strategy, which help protect intellectual property and sensitive data as outlined by the practices and activities in APO14 (Managed Data).

The COBIT objectives related to effective risk management, regulatory compliance, stakeholder engagement, and monitoring for effective internal controls provide horizontal support across all the principles of trustworthy AI.

- **Risk**—EDM03 (Ensured Risk Optimization) mandates that organizations identify, assess, and manage risk. APO12 (Managed Risk) further supports this by establishing structured risk assessments for AI systems, guiding risk mitigation strategies. Both objectives ensure AI risk management processes align with overall organizational appetite and remain within its risk tolerance level, helping to mitigate AI-specific risk.
- **Compliance**—MEA03 (Managed Compliance with External Requirements), APO14 (Managed Data), and APO11 (Managed Quality) provide clear guidance on regulatory adherence, data quality, and privacy. These objectives ensure that data governance policies support compliance across jurisdictions, particularly in managing data quality and privacy concerns central to AI ethics.
- **Stakeholder Expectations**—EDM05 (Ensured Stakeholder Engagement) and APO04 (Managed Innovation) emphasize aligning AI initiatives with stakeholder expectations. These objectives ensure that AI governance is responsive to user needs—from internal departments to external customers—fostering trust and promoting alignment with business goals.
- **Monitoring**—MEA01 (Managed Performance and Conformance Monitoring) practices can be facilitated by setting monitoring targets for AI performance. Similarly, MEA02 (Managed System of Internal Control) ensures internal controls such as security and compliance checks are embedded, allowing for real-time assessments and corrective actions as needed.

For a comprehensive selection of applicable governance and management objectives, the *COBIT 2019 Design Guide: Designing an Information and Technology Governance Solution* and Design Toolkit[17] should be used to account for AI use within the organization. These design factors include the enterprise strategy, enterprise goals, risk profile, IT-related issues, threat landscape, compliance requirement, role of IT, sourcing model for IT, IT implementation methods, technology adoption strategy, and enterprise size.

COBIT Throughout the AI Life Cycle

COBIT can be applied to AI technology governance and management by integrating its principles into each stage of the AI life cycle—from design and development to deployment, operations, and monitoring.[18] For example, an enterprise looking to build an in-house AI system can use COBIT's EDM domain to ensure that AI initiatives are aligned with the organization's overall strategy in the design phase. The principles in this domain can help decision-makers evaluate AI's potential impact on the business, direct resources accordingly, and monitor performance against strategic objectives. Organizations can also use these principles to ensure their AI systems adhere to principles of ethical use, fairness, and transparency, which are essential from the start.

During the development phase, the APO domain helps organizations define clear management structures and establish the processes necessary to build secure and trustworthy AI systems. For organizations purchasing AI systems, these principles are just as important. This phase involves setting clear policies for data governance,

17 ISACA, *COBIT 2019 Design Guide: Designing and Information and Technology Governance Solution*, 2019, https://store.isaca.org/s/store#/store/browse/detail/a2S4w000004Ko9bEAC; Design Toolkit may be downloaded at https://www.isaca.org/-/media/files/isacadp/feature/downloads/c/cobit-2019-design-toolkit_tkt_eng_0222.zip

18 IT Modernization Centers of Excellence, "Understanding and Managing the AI Lifecycle," GSA, https://coe.gsa.gov/coe/ai-guide-for-government/understanding-managing-ai-lifecycle/

security, and compliance; planning for the resource needs of the AI initiative; and ensuring that bias, security vulnerabilities, and compliance challenges are addressed.

In the deployment phase, COBIT's BAI domain can be leveraged to address how AI systems are integrated into the business. Regardless of whether the organization buys or builds the AI system, this domain ensures that it is implemented securely and efficiently and that proper testing and validation processes are in place. This domain also covers aspects of change management, making sure that new AI solutions are introduced without disrupting existing systems or creating unforeseen risk.

Once an AI system is deployed, the DSS domain covers the day-to-day operations and ongoing support of the AI solution. DSS principles ensure that AI systems perform efficiently, service levels are maintained, and any issues are promptly addressed. This domain also ensures the security and continuity of AI systems to protect them from cyberthreats and other risk while in operation.

DSS principles ensure that AI systems perform efficiently, service levels are maintained, and any issues are promptly addressed.

Finally, in the monitoring phase, COBIT's MEA domain supports the continuous evaluation of AI systems to ensure they meet performance, compliance, and risk-management objectives. This domain emphasizes the importance of assessing AI outcomes, measuring performance against business goals, and identifying areas for improvement. MEA principles also play a critical role in ensuring AI systems comply with evolving regulations and industry standards. Through continuous monitoring and assessment, organizations can refine AI models and processes to ensure long-term success and alignment with both internal goals and external requirements.

Example Use Case

Consider a global e-commerce company that is planning to implement an AI-driven customer service system to enhance customer satisfaction and reduce operational costs. The company's leadership has identified key stakeholder needs for faster response times, higher customer engagement, and more personalized interactions. However, the organization also recognizes the risk associated with AI (such as potential bias in customer interactions), regulatory compliance (especially regarding customer data), and the need for ongoing support to maintain AI system performance.

To ensure the success of the AI initiative, the company applies COBIT principles to translate stakeholder needs into actionable I&T objectives. The primary drivers are customer demands for faster and more efficient service, management's need to reduce operational costs, and compliance with international data privacy regulations (such as GDPR). Stakeholders also expect ethical and bias-free AI interactions.

Based on these drivers, the enterprise identifies its goals, which include improving customer-oriented services, optimizing business continuity and availability, and maintaining compliance with laws and regulations. Its goals are mapped to specific I&T objectives that directly support the enterprise's business priorities.

The next step is translating and aligning the enterprise goals into IT-specific goals. For this AI project, the goals include ensuring high performance and reliability of AI systems, optimizing the customer experience through personalization, and adhering to data security and privacy regulations.

Finally, the goals are further broken down into specific governance and management objectives that will guide the AI system's implementation, monitoring, and continuous improvement.

Once the priorities are determined, the company leverages COBIT's objectives to structure the governance of its AI-driven customer service system. The objectives begin by defining specific governance and management processes to guide the development, deployment, and continuous monitoring of the AI system. For example, a process is set up to test and validate the AI model to ensure that its outputs (customer interactions) are accurate, bias-free, and aligned with the company's standards for customer service (e.g., BAI03.06). This process is linked

to practices, such as regular retraining of the AI model, to account for evolving customer needs or language variations across global markets.

The company also implements data governance policies that ensure customer information is processed securely and complies with GDPR and similar regulations. The data used to train the AI model is regularly audited for quality, accuracy, and fairness to avoid biased interactions that could harm the customer experience (e.g., MEA04.06). In addition, the company establishes a set of policies and procedures to govern the day-to-day use of the AI system. These include policies on how all AI-driven customer interactions should be logged and procedures for how the AI system should be updated or modified.

Further, the company creates a governance structure that involves multiple stakeholders, such as a board-level AI steering committee with representation from the global company's IT managers, compliance and privacy officers, customer service leaders, and data scientists (e.g., EDM05.01). These roles are mapped through a RACI chart, ensuring each decision-making entity knows its responsibility regarding the AI system. In this case, IT managers are responsible for ensuring the technical functionality of the system, while compliance/privacy officers monitor adherence to data privacy regulations. Customer service leaders provide feedback on AI system performance based on user experiences. Continuous training programs are implemented to keep all of the teams updated on new AI trends, governance requirements, ethical concerns, and the latest developments in AI (e.g., APO04.04).

The organization fosters a culture that prioritizes responsible AI use and ethical customer interactions. Leadership emphasizes transparency in how the AI system makes decisions, ensuring that customers can understand how and why certain outcomes (such as responses to queries) are generated. The enterprise encourages ethical behavior through training programs that help employees understand the implications of AI's impact on customer interactions.

Finally, the company invests in a robust cloud infrastructure to support the AI system (e.g., APO03). This ensures that the system can scale to meet the demands of global operations while maintaining security and performance (e.g., BAI04). Regular maintenance and updates are scheduled to ensure the AI application remains aligned with business needs and continues to deliver reliable service to customers.

Challenges and Considerations

Organizations may face several challenges when integrating AI into their governance and management practices, primarily due to the complexity, rapid evolution, and integration demands of AI. The lack of clear ownership of AI governance often results in fragmented oversight and difficulty aligning AI systems with the overall business strategy. Traditional governance frameworks can be difficult to fully implement in the context of AI, especially as AI systems often operate in silos, disconnected from traditional IT infrastructures.

The lack of clear ownership of AI governance often results in fragmented oversight and difficulty aligning AI systems with the overall business strategy.

Additionally, AI brings unique risk, such as bias and ethical concerns, which require more specialized governance models. Integrating this risk into COBIT's broader risk management framework can be complex and may necessitate significant customization. The risk profile of the enterprise needs to be considered, which requires a high-level analysis to identify relevant risk scenarios. There also needs to be an assessment of the impact and likelihood of each scenario, given current risk mitigation controls, to effectively rate the risk. Example AI risk categories the enterprise should consider can be found in **figure 5**.

FIGURE 5: AI Risk Categories

Risk Category	Example Risk Scenario
Ethical Usage Risk	**Bias and Discrimination**—AI systems unintentionally amplify biases in training data, leading to unfair treatment of specific groups (e.g., biased hiring recommendations). **Unintended Consequences**—AI systems learn and act in ways not anticipated by developers, potentially leading to harmful outcomes.
Policy and Governance Risk	**Lack of Accountability**—Undefined ownership of AI decisions leads to no clear accountability when errors or harm occur. **Noncompliance with Global Standards**—AI systems fail to meet cross-border regulatory requirements (e.g., GDPR, EU AI Act), leading to fines or operational delays.
Technology and Infrastructure Risk	**Algorithmic Failures**—AI models produce incorrect predictions or classifications due to errors in design, data, or deployment. **Adversarial Attacks**—Hackers manipulate AI inputs (e.g., images or text) to exploit weaknesses and produce desired outcomes.
Operational and Organizational Risk	**Overreliance on AI**—Employees rely excessively on AI outputs without critical oversight, leading to blind trust in flawed systems. **Failure to Align with Strategic Goals**—AI investments focus on trendy technologies rather than solving actual business problems.
Emerging and Strategic Risk	**Misuse of Generative AI**—Generative AI creates harmful content (e.g., explicit images or extremist propaganda), impacting brand reputation. **Legal Precedents Against AI**—New lawsuits or legal rulings limit how organizations can deploy or utilize AI systems.

Another AI challenge is obtaining security stakeholder buy-in. Since AI projects typically involve specialized technical teams, nontechnical stakeholders may not fully understand the AI governance requirements. This makes it difficult to achieve cross-functional collaboration and acquire the specialized AI expertise needed in the governance team.

Moreover, AI governance needs active involvement from senior leadership and solid measures of the effectiveness of AI governance processes to ensure strategic alignment with business goals. Without executive support, AI governance initiatives may lack the necessary resources and attention, resulting in fragmented efforts that undermine the success of AI integration.

To overcome these challenges, organizations should prioritize obtaining executive support and emphasize aligning I&T and AI initiatives with business objectives. This ensures that AI governance efforts receive the necessary resources and focus from senior leadership. Additionally, fostering cross-functional collaboration between departments like IT, data science, compliance, and legal is essential to create a unified enterprise governance framework that addresses AI-specific risk.

Finally, as AI technologies evolve rapidly, continuous education is crucial. Organizations must ensure that both technical and nontechnical teams stay updated on emerging trends and risk, enabling them to adapt governance structures within COBIT to effectively manage new AI challenges as they arise.

Benefits of Using COBIT for AI Technology

By using COBIT, enterprises can ensure that AI initiatives align with their broader organizational strategies by helping to establish a governance system that aligns AI objectives directly to business goals. Through COBIT's principles, AI projects are designed with a clear focus on strategic alignment, ensuring that all activities, from

development to deployment, are intended to serve an organization's core mission and align with its broader strategy. This alignment helps create and demonstrate measurable value and enables enterprises to achieve specific, quantifiable outcomes, such as increased operational efficiency, positive return on investment (ROI), enhanced decision making, or improved customer engagement.[19]

COBIT also emphasizes accountability, ensuring that each AI initiative has clear ownership and a roadmap for achieving defined business outcomes. As a result, organizations are better equipped to drive real, tangible benefits from their AI investments, making them more competitive in their industry.

AI technologies bring unique risk but COBIT provides a robust framework for identifying, assessing, and mitigating this risk throughout the AI life cycle. COBIT mandates a thorough risk assessment that considers factors such as data integrity, model bias, and regulatory compliance. This approach extends into the deployment phase when continuous monitoring for potential threats or security vulnerabilities becomes crucial.

COBIT provides a robust framework for identifying, assessing, and mitigating this risk throughout the AI life cycle.

By implementing COBIT's risk management guidelines, organizations can systematically address and reduce risk, ensuring that AI systems remain reliable, compliant, and ethically sound. Furthermore, COBIT's focus on performance and conformance monitoring evaluation and assessment ensures that any emerging risk is promptly identified and mitigated to minimize potential disruptions to the organization.

AI systems require various resources, such as high-quality data, skilled personnel, and a robust technology infrastructure. COBIT assists organizations in managing these resources efficiently to avoid waste and maximize the ROI.

COBIT helps allocate data effectively, ensuring that only relevant and high-quality data is used in AI models. It also provides guidelines for recruiting and retaining talent with the specialized skills needed for AI development, from data scientists to machine learning engineers. By ensuring that infrastructure and technology investments align with both current and future demands, COBIT promotes scalability. This forward-thinking approach helps organizations avoid unnecessary costs and prepares them to scale AI initiatives as the organization's needs evolve, all while maintaining control over expenditures.

Furthermore, COBIT's governance framework emphasizes continuous improvement, which is essential for evolving AI systems that need to adapt to changing business environments and technological advancements. By embedding ongoing monitoring and assessment practices, COBIT ensures that AI systems are continually evaluated against performance metrics, compliance standards, and evolving business requirements. These practices enable organizations to identify performance gaps, adapt to emerging trends, and integrate feedback into the AI system's life cycle.

COBIT also encourages a feedback loop that includes stakeholders from various departments, allowing a holistic view of how AI systems impact the organization. This continuous improvement cycle ensures that AI systems remain relevant, effective, and aligned with an organization's strategy over time. Additionally, COBIT's governance principles guide organizations in documenting these improvements, fostering a culture of transparency and accountability that enhances trust in AI systems across all levels of the organization.

19 ISACA, "Artificial Intelligence Governance Brief," https://www.isaca.org/-/media/files/isacadp/project/isaca/resources/ebooks/ai-governance-brief_1124.pdf

Conclusion

Leveraging COBIT for AI governance and management provides organizations with a structured, comprehensive approach to aligning AI systems with their strategic objectives, ensuring responsible and ethical AI practices, and mitigating unique AI-related risk. By integrating COBIT's governance principles across the AI life cycle—from strategic alignment and risk management to continuous improvement—organizations can better address the complexities of AI deployments. The COBIT framework not only facilitates resource optimization and enhances accountability but also builds a foundation of trust in AI initiatives.

Through COBIT, organizations are better equipped to navigate the challenges of AI adoption in a fast-evolving landscape while driving sustained value and fostering resilience in AI-enabled operations.

Acknowledgments

Lead Developer

Meghan Maneval

CISM, CRISC

AI and Risk Strategist, The Risk Optimist, USA

Expert Reviewers

J. Winston Hayden

CISA, CISM, CGEIT, CRISC, CDPSE

South Africa

Terence Law

CISA, CISSP, CFA, CPA

Certified Banker

Hong Kong

Dina Numan

CRISC, CDPSE, COBIT 2019 Lead Assessor, ISO 20000 LA, ITIL

Jordan

Zachy Olorunojowon

CISA, CISM, CGEIT, CET

Executive Director, Digital Health Strategic Initiatives BC Ministry of Health, Canada

Dirk Steuperaert

CISA, CGEIT, CRISC

Belgium

Board of Directors

John De Santis, Chair

Former Chairman and Chief Executive Officer, HyTrust, Inc., USA

Niel Harper, Vice-Chair

CISA, CRISC, CDPSE, CISSP, NACD.DC

Chief Information Security Officer and Data Protection Officer, Doodle, Former Chief Information Security Officer, United Nations Office for Project Services (UNOPS), Germany

Stephen Gilfus

Managing Director, Oversight Ventures LLC, Chairman, Gilfus Education Group and Founder, Blackboard Inc., USA

Gabriela Hernandez-Cardoso

NACD.DC

Former President and CEO, GE Mexico, Independent Board Member, Mexico

Jason Lau

CISA, CISM, CGEIT, CRISC, CDPSE, CIPM, CIPP/E, CIPT, CISSP, FIP, HCISPP

Chief Information Security Officer, Crypto.com, Singapore

Massimo Migliuolo

Independent Board Member, Malaysia

Jamie Norton

CISA, CISM, CGEIT, CIPM, CISSP

Partner, McGrathNicol, Australia

Maureen O'Connell

NACD.DC

Board Chair, Acacia Research (NASDAQ), Former Chief Financial Officer and Chief Administration Officer, Scholastic, Inc., USA

Erik Prusch

Chief Executive Officer, ISACA, USA

Asaf Weisberg

CISA, CISM, CGEIT, CRISC, CSX-P, CDPSE

Chief Executive Officer, introSight Ltd., Israel

Pamela Nigro

ISACA Board Chair 2022-2023

CISA, CGEIT, CRISC, CDPSE, CRMA

Vice President, Security, Medecision, USA

Tracey Dedrick

ISACA Board Chair, 2020-2021

Former Executive Vice President and Head of Enterprise Risk Management, Santander Holdings, USA

Brennan P. Baybeck

ISACA Board Chair, 2019-2020

CISA, CISM, CRISC, CISSP

Senior Vice President and Chief Information Security Officer for Customer Services, Oracle Corporation, USA

About ISACA

ISACA® (www.isaca.org) is a global community advancing individuals and organizations in their pursuit of digital trust. For more than 50 years, ISACA has equipped individuals and enterprises with the knowledge, credentials, education, training and community to progress their careers, transform their organizations, and build a more trusted and ethical digital world. ISACA is a global professional association and learning organization that leverages the expertise of its 180,000+ members who work in digital trust fields such as information security, governance, assurance, risk, privacy and quality. It has a presence in 188 countries, including 225 chapters worldwide. Through the ISACA Foundation, ISACA supports IT education and career pathways for underresourced and underrepresented populations.

DISCLAIMER

ISACA has designed and created *Leveraging COBIT for Effective AI System Governance* (the "Work") primarily as an educational resource for professionals. ISACA makes no claim that use of any of the Work will assure a successful outcome. The Work should not be considered inclusive of all proper information, procedures and tests or exclusive of other information, procedures and tests that are reasonably directed to obtaining the same results. In determining the propriety of any specific information, procedure or test, professionals should apply their own professional judgment to the specific circumstances presented by the particular systems or information technology environment.

RESERVATION OF RIGHTS

ISACA
1700 E. Golf Road, Suite 400
Schaumburg, IL 60173, USA

Phone: +1.847.660.5505

Fax: +1.847.253.1755

Support: support.isaca.org

Website: www.isaca.org

Participate in the ISACA Online Forums:
https://engage.isaca.org/onlineforums

X: www.x.com/ISACANews

LinkedIn:
www.linkedin.com/company/isaca

Facebook:
www.facebook.com/ISACAGlobal

Instagram:
www.instagram.com/isacanews/