

Developing an AI Governance Capability Maturity Model for Responsible and Trustworthy Artificial Intelligence

Firdaus Malek

Independent Researcher
Vadodara, Gujarat, India

Harshwardhansinh Chauhan

Assistant Professor
Drs. Kiran & Pallavi Patel Global University (KPGU), Vadodara
Vadodara, Gujarat, India

Dr. Rocky Upadhyay

Associate Professor & Head of Department,
Department of Computer science and engineering
Drs. Kiran & Pallavi Patel Global University (KPGU), Vadodara

Abstract: - The rapid rise of Artificial Intelligence (AI) in Governments, enterprises, healthcare, finance, manufacturing, and education has increased demand for mechanisms that govern AI systems, making AI accountable, transparent, and trustworthy while complying with emerging regulatory requirements. Some organizations have started developing AI governance policies to meet these requirements. However, there is currently no formal mechanism that assesses governance capabilities, maturity levels, and therefore the areas that require improvement. Most AI maturity models emphasize AI technologies, digital transformation, and readiness while addressing governance issues such as ethics, accountability, transparency, fairness, privacy, security, explainability, and regulatory compliance at a superficial level.

This paper proposes an AI Governance Capability Maturity Model (AI-GCMM) that evaluates and improves AI governance capabilities at different levels and areas, including strategic governance, risk management, ethical concerns, data governance, model life cycle management, human governance, security and privacy, model explainability, regulatory compliance, and monitoring and evaluation. The model uses a five-level maturity framework – Initial, Managed, Defined, Quantitatively Managed, and Optimizing – which measures the ability of organizations to govern AI systems based on selected governance areas and assesses their overall governance maturity.

The model introduces capability indicators that enable organizations to evaluate their ability, identify areas of improvement, take corrective actions, and ultimately move from a lower to a higher maturity level. In addition, this paper evaluates the proposed framework using experts' judgments and compares it to existing AI governance and capability maturity models, demonstrating that the AI-GCMM is a comprehensive and practical framework that enhances AI governance while promoting transparency and accountability.

Finally, the paper concludes that the proposed framework advances research and practice in responsible and trustworthy AI by contributing to the development of a formal mechanism that bridges the theoretical and practical gaps in assessing and improving organizational AI governance maturity. It also provides researchers and policymakers with a decision-making tool that guides the evaluation process of selecting and developing AI governance policies, practices, and processes that promote responsible AI at various organizational levels and stages of AI maturity in different industries.

Keywords: Artificial Intelligence Governance, Responsible AI, Trustworthy AI, AI Governance Capability Maturity Model (AI-GCMM), AI Capability Maturity Model, AI Ethics, AI Risk Management, Explainable AI (XAI), AI Transparency, AI Accountability, AI Compliance, AI Lifecycle Governance, AI Audit, Human-Centered AI, Organizational AI Maturity, AI Policy and Regulation, Data Governance, AI Assurance, Ethical AI, Continuous Governance Improvement.

1. Introduction

Artificial Intelligence (AI) has become a transformative technology that is reshaping organizational operations, decision-making, and service delivery across both public and private sectors. Advances in machine learning, deep learning, and generative AI have accelerated the adoption of intelligent systems in healthcare, finance, manufacturing, education, transportation, government, and numerous other domains. Organizations increasingly rely on AI to automate complex processes, improve operational efficiency, support strategic decision-making, and generate new business value. As AI systems become more autonomous and influential, however, organizations are simultaneously confronted with growing concerns regarding transparency, accountability, fairness, explainability, privacy, security, safety, and regulatory compliance. These concerns have elevated AI governance from a technical consideration to a strategic organizational priority.

The rapid evolution of AI has prompted governments, regulatory authorities, and international standardization bodies to establish governance principles intended to promote responsible and trustworthy AI. Prominent initiatives—including the OECD AI Principles, the NIST AI Risk Management Framework (AI RMF 1.0), ISO/IEC 42001, and the European Union AI Act—provide internationally recognized guidance for managing AI risks, improving transparency, strengthening accountability, and ensuring that AI systems remain aligned with legal and societal

expectations. Collectively, these initiatives represent significant progress in defining what constitutes responsible AI governance and provide organizations with valuable guidance for developing governance policies and risk management practices. Nevertheless, translating governance principles into effective organizational practice remains a considerable challenge.

Existing AI governance research has similarly expanded rapidly over recent years. Current studies have investigated governance principles, ethical AI, trustworthy AI, AI safety, governance frameworks, risk management, regulatory compliance, and organizational governance structures. Systematic literature reviews further demonstrate that AI governance has evolved into a multidisciplinary research domain that integrates computer science, management, information systems, public policy, ethics, and law. Despite these developments, much of the existing literature remains normative in nature, emphasizing **what** organizations should govern rather than **how** organizations can systematically develop, assess, and continuously improve their governance capability. Consequently, organizations may establish governance policies or adopt recognized governance standards without possessing systematic mechanisms to assess, benchmark, and continuously improve governance capability throughout the AI lifecycle.

A similar limitation exists within the maturity model literature. Organizational maturity models have been widely adopted to evaluate process capability, digital transformation, information technology governance, cybersecurity, knowledge management, and organizational performance. These models demonstrate that organizational capability develops progressively through structured organizational processes and continuous capability development. Recent AI-related maturity models have primarily focused on AI adoption, digital readiness, analytics capability, or technological implementation. While these approaches provide useful mechanisms for evaluating technological advancement, they provide comparatively limited attention to organizational AI governance capability. As AI systems increasingly influence critical organizational decisions, assessing governance capability has become equally important as assessing technological capability.

The absence of a capability-oriented governance assessment framework creates both theoretical and practical challenges. From a theoretical perspective, existing AI governance research lacks an integrated model that combines governance principles, organizational capability theory, maturity assessment, and continuous governance improvement within a unified conceptual framework. From a practical perspective, organizations require structured methods to evaluate governance readiness, identify governance capability gaps, prioritize governance improvement initiatives, and monitor governance performance over time. Without such mechanisms, organizations may struggle to operationalize responsible AI despite complying with governance principles or regulatory requirements.

To address these limitations, this study proposes the **AI Governance Capability Maturity Model (AIGCMM)**. Developed using a Design Science Research (DSR) approach, the AIGCMM conceptualizes AI governance as a multidimensional organizational capability that evolves through

progressive stages of maturity. The proposed framework integrates governance capability dimensions, maturity levels, capability assessment, and continuous improvement into a unified capability assessment framework for evaluating AI governance capability. Rather than replacing existing governance standards or regulatory frameworks, the AIGCMM complements them by providing organizations with a practical mechanism for assessing how effectively governance principles are institutionalized and continuously improved.

This study makes four primary contributions to AI governance research. **First**, it reconceptualizes AI governance from a principle-based perspective to an capability-oriented perspective, emphasizing governance implementation rather than governance intention. **Second**, it introduces a structured maturity model that enables organizations to evaluate governance capability systematically across multiple governance dimensions. **Third**, it integrates capability assessment with continuous organizational improvement, allowing governance maturity to be monitored and strengthened over time. Finally, the model provides a practical decision-support framework that can assist organizations, regulators, auditors, and policymakers in evaluating governance readiness for responsible and trustworthy AI deployment.

The remainder of this paper is organized as follows. Section II reviews the existing literature on AI governance, organizational capability, and maturity models and identifies the research gap addressed by this study. Section III describes the Design Science Research methodology adopted for developing the proposed framework. Section IV presents the AI Governance Capability Maturity Model (AIGCMM), including its governance dimensions, maturity levels, assessment framework, and continuous improvement process. Section V discusses the theoretical and practical contributions of the proposed model. Finally, Sections VI and VII present the research limitations, future research directions, and conclusions.

II. Literature Review and Research Gap

2.1 Evolution of AI Governance

The rapid advancement of Artificial Intelligence (AI) has fundamentally transformed the way organizations create value, make decisions, and deliver services. As AI systems become increasingly autonomous and embedded within critical organizational processes, governance has emerged as a central concern for researchers, practitioners, and policymakers. Initially, AI governance was largely discussed from an ethical perspective, emphasizing fairness, transparency, accountability, privacy, and human rights. More recently, however, AI governance has evolved into a multidisciplinary research domain that encompasses organizational governance, risk management, regulatory compliance, technical assurance, and strategic management.

This evolution has been driven by the increasing complexity of AI systems and the recognition that governance challenges extend beyond algorithmic performance. Modern AI systems influence high-impact decisions in healthcare, finance, criminal justice, public administration,

manufacturing, and critical infrastructure, where failures may generate legal, financial, operational, and societal consequences. Consequently, governments and international organizations have introduced governance frameworks that seek to ensure AI systems remain responsible, trustworthy, transparent, safe, and aligned with societal values. International initiatives such as the OECD AI Principles, the NIST AI Risk Management Framework (AI RMF 1.0), ISO/IEC 42001, and the European Union AI Act collectively demonstrate the global transition from ethical AI discussions toward operational AI governance.

Recent systematic literature reviews further indicate that AI governance research has expanded rapidly during the past five years. Contemporary studies investigate governance mechanisms, organizational governance structures, AI safety, trustworthy AI, human oversight, regulatory compliance, and governance implementation across diverse organizational settings. Despite this rapid growth, the literature remains fragmented across multiple disciplines, with relatively limited integration between governance theory, organizational capability, and governance assessment. As a result, organizations often receive extensive guidance regarding governance principles but comparatively little support for evaluating how effectively those principles are embedded within organizational practice.

2.2 AI Governance Frameworks: Current State of Knowledge

Existing AI governance frameworks have significantly advanced responsible AI by defining governance principles, organizational responsibilities, and regulatory expectations. Rather than focusing exclusively on technical controls, these frameworks recognize that effective AI governance requires leadership commitment, organizational accountability, risk management, transparency, human oversight, and continuous monitoring throughout the AI lifecycle.

International standards and governance initiatives have played an essential role in standardizing governance expectations. The OECD AI Principles establish widely accepted values for trustworthy AI, while the NIST AI RMF provides structured guidance for identifying, assessing, and managing AI-related risks. Similarly, ISO/IEC 42001 introduces organizational management system requirements specifically designed for AI governance, reflecting the increasing institutionalization of governance practices within organizations.

Parallel to these developments, academic research has proposed numerous conceptual frameworks for AI governance. Existing studies investigate governance architectures, organizational governance mechanisms, governance ecosystems, responsible AI implementation, and governance best practices. Several systematic literature reviews conclude that governance research increasingly recognizes AI governance as an organizational challenge rather than solely a technological issue. These studies consistently emphasize the importance of integrating governance structures, organizational processes, technical safeguards, and stakeholder engagement to support responsible AI deployment.

Nevertheless, existing governance frameworks predominantly prescribe **what organizations should govern** rather than **how organizations should evaluate governance capability**. Most frameworks describe governance principles, governance processes, or regulatory obligations but provide limited mechanisms for assessing governance maturity, benchmarking governance capability, or supporting continuous governance improvement. Consequently, organizations often lack practical tools for determining whether governance practices have been effectively institutionalized across the AI lifecycle.

2.3 Organizational Capability and Maturity Models

Capability maturity models (CMMs) have been widely adopted to assess organizational capability and support continuous improvement across domains such as software engineering, information technology governance, cybersecurity, digital transformation, and organizational process management. These models share a common premise that organizational capability develops progressively through structured learning, standardized processes, performance evaluation, and continuous improvement.

Building upon this foundation, researchers have recently proposed maturity models for AI adoption, analytics capability, digital transformation, and organizational AI readiness. These models provide valuable approaches for evaluating technological implementation, organizational preparedness, and digital capability. However, their primary focus remains technological adoption rather than governance capability. Governance-related dimensions are frequently incorporated only as supporting organizational factors instead of representing the central objective of maturity assessment.

The literature therefore reveals an important distinction between **AI capability** and **AI governance capability**. An organization may demonstrate advanced AI deployment while simultaneously possessing relatively immature governance processes. Conversely, organizations with well-established governance structures may still require systematic mechanisms for evaluating governance performance, identifying capability gaps, and guiding continuous improvement. Existing maturity models provide only limited support for addressing these governance-specific challenges, highlighting the need for a dedicated AI governance capability maturity framework that evaluates governance as an organizational capability rather than a technological outcome.

2.4 Critical Research Gap

The synthesis of the reviewed literature reveals four significant research gaps.

First, contemporary AI governance research has successfully established governance principles, ethical guidelines, regulatory frameworks, and governance architectures; however, comparatively

limited attention has been devoted to evaluating organizational AI governance capability. Existing studies primarily define governance expectations without providing structured approaches for measuring governance maturity.

Second, existing maturity models predominantly assess AI adoption, digital transformation, analytics capability, or technological readiness. Although these models contribute substantially to organizational capability research, they rarely position AI governance capability as the primary assessment objective.

Third, current research often treats AI governance, trustworthy AI, AI safety, organizational capability, and maturity assessment as independent research streams. Limited research integrates these perspectives into a unified framework capable of supporting governance assessment and continuous organizational improvement.

Finally, recent systematic literature reviews consistently identify the need for practical governance assessment methodologies that enable organizations to operationalize responsible AI beyond policy development and regulatory compliance. While governance standards define desired governance characteristics, organizations continue to lack practical mechanisms for evaluating governance readiness, identifying capability deficiencies, and systematically improving governance performance over time.

Accordingly, this study addresses these limitations by proposing the **AI Governance Capability Maturity Model (AIGCMM)**. The proposed framework integrates organizational capability theory, AI governance principles, maturity assessment, and continuous improvement into a unified conceptual model that enables organizations to evaluate, benchmark, and strengthen AI governance capability across the entire AI lifecycle.

Table 1. Critical Synthesis of Existing Literature and Positioning of the AIGCMM

Research Domain	Primary Focus	Strengths	Remaining Limitation	How the AIGCMM Addresses the Gap
AI Governance Frameworks	Governance principles and structures	Establish governance guidance	Limited capability assessment	Capability-oriented governance model
Trustworthy AI	Ethical and trustworthy AI	Defines trustworthy AI characteristics	Limited organizational operationalization	Governance capability implementation

Research Domain	Primary Focus	Strengths	Remaining Limitation	How the AIGCMM Addresses the Gap
AI Safety	Risk mitigation and safe AI deployment	Improves AI reliability	Limited maturity perspective	Safety integrated within governance maturity
Organizational Capability	Capability development	Continuous organizational improvement	Not AI-governance specific	AI governance capability construct
AI Maturity Models	AI adoption and digital maturity	Structured capability progression	Focus on technology rather than governance	Governance-specific maturity assessment
International Standards	Governance requirements and compliance	Standardized governance expectations	No maturity evaluation mechanism	Governance readiness assessment
Proposed AIGCMM	Organizational AI governance capability	Integrates governance, maturity, assessment, and continuous improvement	—	Unified capability-oriented governance framework

III. Research Methodology

3.1 Research Design

This study adopts the **Design Science Research (DSR)** paradigm to develop the proposed **AI Governance Capability Maturity Model (AIGCMM)**. DSR is an established research methodology for creating innovative artefacts that address complex organizational problems while simultaneously contributing to theoretical knowledge. Unlike explanatory research that seeks to describe or predict phenomena, DSR focuses on designing and evaluating artefacts capable of solving identified practical and research problems. Given that the objective of this study is to develop a capability-oriented maturity model for organizational AI governance, DSR provides an appropriate methodological foundation.

The development of the AIGCMM follows the fundamental principles of DSR by integrating knowledge from AI governance, organizational capability theory, maturity model research, and international AI governance standards into a unified conceptual framework. Rather than proposing a new governance principle or regulatory framework, the study designs a practical assessment artefact that enables organizations to evaluate governance capability, identify improvement priorities, and support continuous governance enhancement.

3.2 Research Process

The research was conducted through four sequential stages, ensuring that the proposed framework is grounded in both theory and practice.

Stage 1: Problem Identification

The first stage identified the central research problem through a critical review of recent AI governance literature. Although existing studies provide comprehensive guidance on AI governance principles, ethical AI, trustworthy AI, and regulatory compliance, they provide comparatively limited support for assessing organizational governance capability. Simultaneously, existing maturity models primarily evaluate technological capability, AI adoption, or digital transformation rather than governance capability. This analysis established the need for a capability-oriented AI governance maturity model.

Stage 2: Knowledge Synthesis

The second stage synthesized evidence from four complementary knowledge domains:

- AI governance literature;
- organizational capability theory;
- capability maturity model research; and
- international AI governance standards and frameworks.

The synthesis identified recurring governance concepts, organizational governance requirements, capability dimensions, governance processes, and continuous improvement mechanisms. These insights formed the theoretical foundation of the proposed model.

Stage 3: Artefact Development

The third stage involved the design of the **AI Governance Capability Maturity Model (AIGCMM)**. The resulting artefact integrates eight governance capability dimensions, five progressive maturity levels, a multidimensional capability assessment framework, and a continuous improvement cycle into a unified governance capability model. The framework was

designed to support organizational governance assessment while remaining sufficiently flexible for adaptation across different organizational contexts.

Stage 4: Conceptual Evaluation

The final stage evaluated the internal consistency and theoretical coherence of the proposed framework. The governance dimensions, maturity levels, assessment process, and continuous improvement cycle were critically examined to ensure conceptual alignment with existing AI governance research and organizational capability theory. Although empirical validation is beyond the scope of the present study, the proposed model establishes a theoretical foundation for future expert evaluation, case studies, and organizational implementation.

The outcome of this research process is the proposed AI Governance Capability Maturity Model (AIGCMM), presented in Section IV.

3.3 Development Principles of the AIGCMM

The development of the proposed framework was guided by five design principles derived from the literature.

Principle 1 – Capability Orientation: AI governance is conceptualized as an organizational capability rather than solely a compliance or policy function.

Principle 2 – Multidimensional Assessment: Governance capability should be evaluated across multiple organizational dimensions rather than through a single maturity indicator.

Principle 3 – Progressive Capability Development: Organizational governance capability evolves incrementally through structured maturity stages that reflect increasing institutionalization of governance practices.

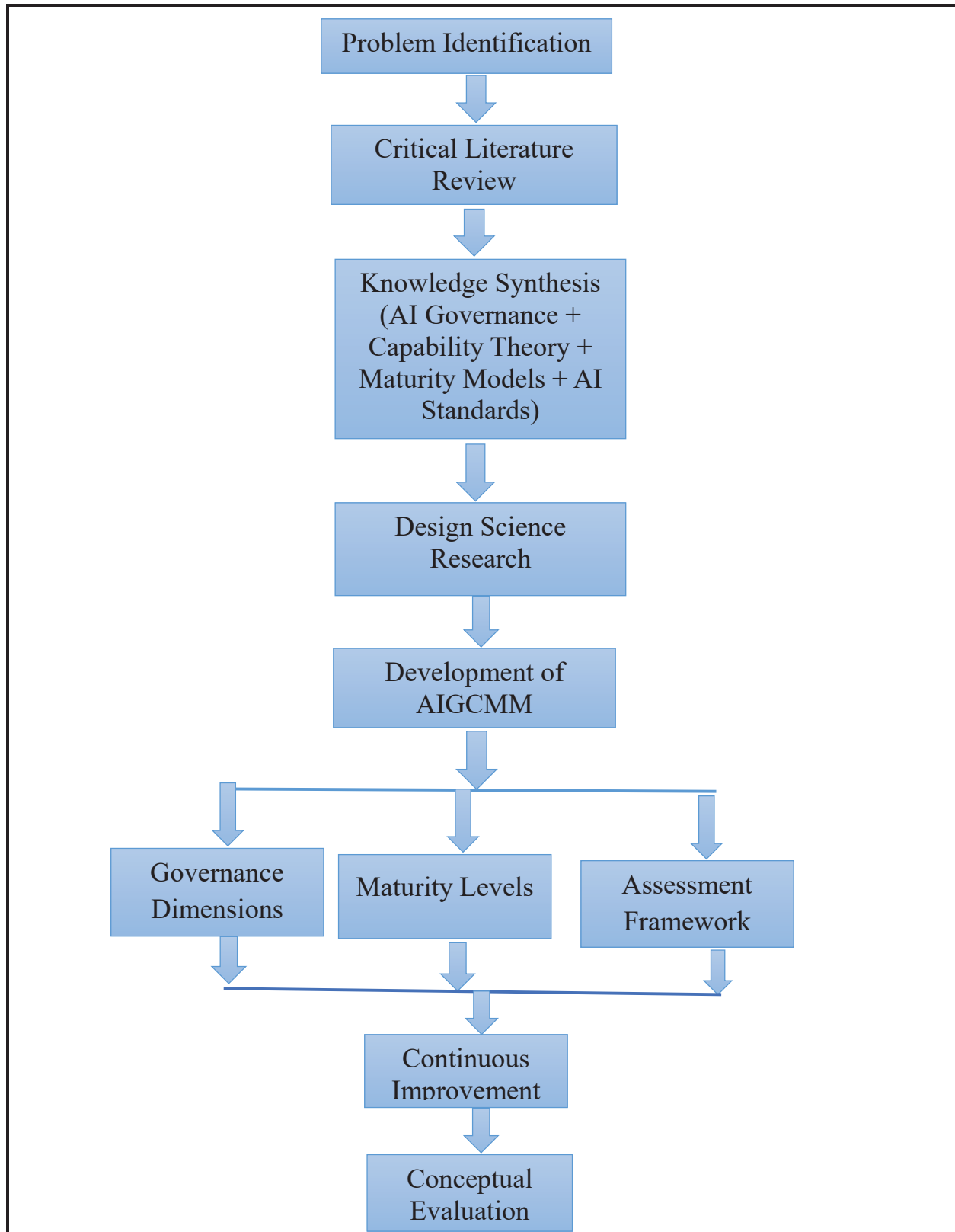
Principle 4 – Continuous Improvement: Governance capability should be continuously monitored, evaluated, and refined to respond to technological change, regulatory evolution, and organizational learning.

Principle 5 – Practical Applicability: The maturity model should provide actionable guidance that supports governance assessment, benchmarking, and capability improvement across diverse organizational environments.

These principles informed both the architecture and implementation logic of the AIGCMM and ensure that the proposed framework is theoretically grounded while remaining applicable in practice.

3.4 Research Methodology Overview

Figure 1. Research Methodology



3.5 Methodological Contribution

Beyond the development of the AIGCMM, this study demonstrates how Design Science Research can be applied to AI governance research by integrating organizational capability theory, governance standards, and maturity assessment within a single artefact development process. This methodological approach contributes to AI governance research by providing a systematic pathway for transforming governance principles into an operational governance capability assessment framework. Furthermore, it establishes a foundation for future empirical studies aimed at validating and refining the proposed maturity model across diverse organizational contexts.

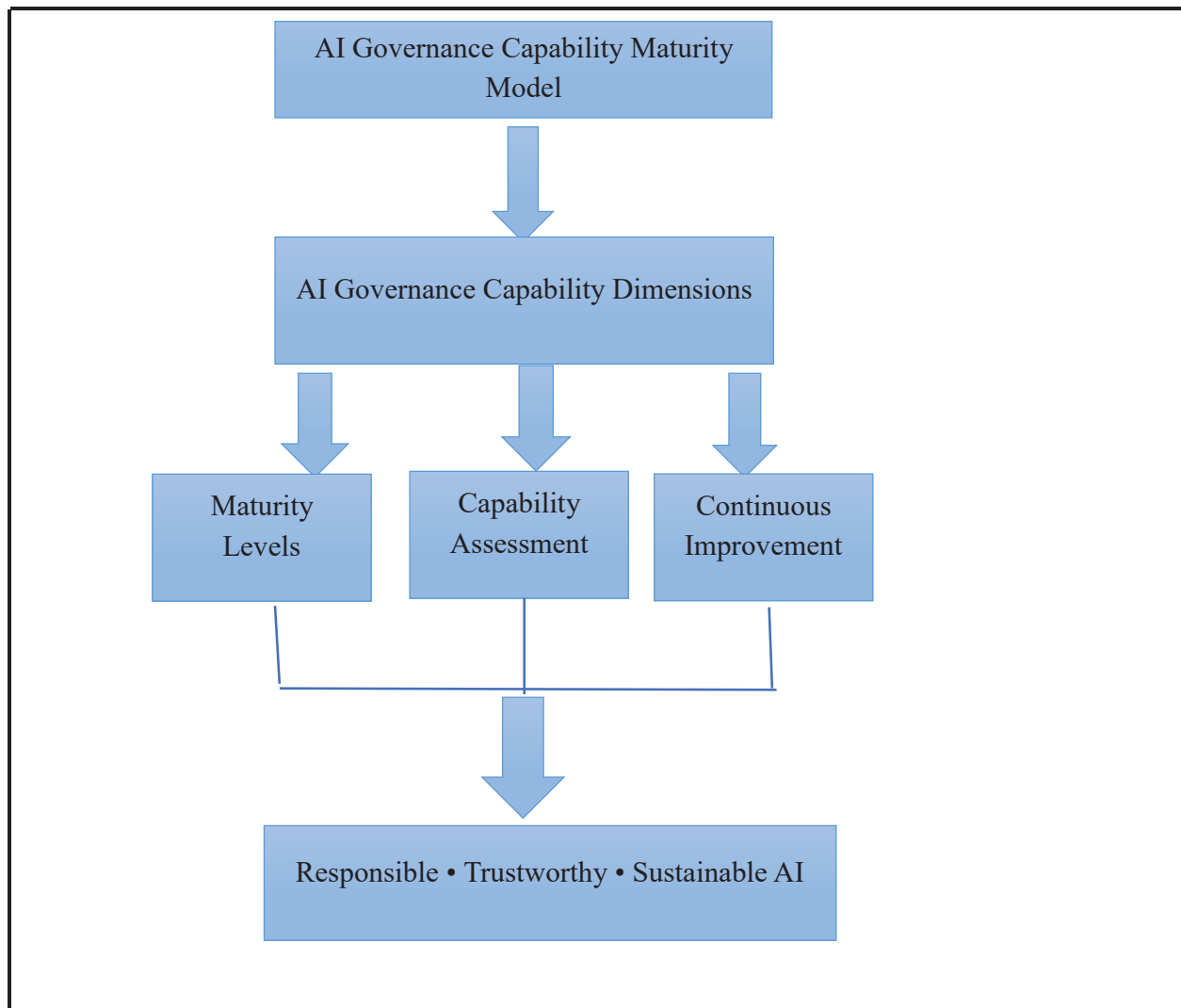
IV. Proposed AI Governance Capability Maturity Model (AIGCMM)

4.1 Overview of the Proposed Framework

The AI Governance Capability Maturity Model (AIGCMM) is proposed as a capability-oriented framework for assessing and improving organizational AI governance. Building upon AI governance research, organizational capability theory, maturity model research, and international governance standards, the model provides a structured approach for evaluating an organization's ability to establish, operationalize, monitor, and continuously improve AI governance throughout the AI lifecycle.

Unlike existing governance frameworks that primarily define governance principles or regulatory requirements, the AIGCMM evaluates **organizational governance capability**. The model integrates four complementary components: (1) AI Governance Capability Dimensions, (2) Governance Maturity Levels, (3) a Capability Assessment Framework, and (4) a Continuous Improvement Cycle. Together, these components enable organizations to benchmark governance maturity, identify capability gaps, and prioritize governance improvement initiatives.

Figure 2. Architecture of the Proposed AIGCMM



4.2 AI Governance Capability

This study conceptualizes **AI Governance Capability (AIGC)** as the organizational ability to establish, operationalize, monitor, evaluate, and continuously improve governance mechanisms that ensure AI systems remain responsible, trustworthy, transparent, accountable, safe, and aligned with organizational, regulatory, and societal expectations throughout their lifecycle.

Unlike governance policies or compliance activities, AI governance capability reflects an organization's capacity to consistently implement governance practices across strategic, organizational, technical, and operational contexts. It therefore represents a dynamic organizational capability that evolves through institutional learning, standardized governance

processes, performance evaluation, and continuous improvement. This conceptualization provides the theoretical foundation for the proposed AIGCMM.

4.3 AI Governance Capability Dimensions

The AIGCMM comprises **eight governance capability dimensions** representing the core organizational capabilities required to achieve responsible and trustworthy AI governance. The dimensions were derived through the synthesis of AI governance literature, organizational capability theory, maturity model research, and international AI governance standards.

1. Strategic Governance and Leadership

Evaluates executive commitment, governance vision, strategic alignment, accountability structures, and resource allocation. Leadership establishes the organizational direction necessary for effective AI governance.

2. Governance Policy, Risk, and Compliance

Assesses governance policies, AI risk management, regulatory compliance, internal controls, and governance accountability to ensure governance activities remain systematic and aligned with evolving regulatory requirements.

3. Data, Model, and Technical Governance

Measures the organization's capability to govern data, AI models, technical infrastructure, cybersecurity, documentation, validation, and lifecycle management to support reliable and trustworthy AI systems.

4. Transparency, Explainability, and Accountability

Evaluates governance mechanisms that promote explainability, traceability, documentation, accountability, and transparent decision-making, enabling stakeholders to understand and evaluate AI-supported decisions.

5. Human Oversight and Ethical Decision-Making

Assesses ethical governance, human supervision, bias mitigation, escalation mechanisms, and responsible decision-making to ensure appropriate human control over AI systems.

6. Organizational Capability and Governance Culture

Examines governance awareness, workforce capability, governance training, cross-functional collaboration, and organizational culture that support the institutionalization of responsible AI practices.

7. Performance Monitoring and Continuous Improvement

Measures the capability to monitor governance performance, conduct governance reviews, evaluate effectiveness, and implement continuous governance improvement based on organizational learning.

8. Stakeholder Trust and Societal Responsibility

Evaluates stakeholder engagement, governance transparency, societal responsibility, and public trust to ensure AI governance supports both organizational objectives and broader societal expectations.

Table 2. AI Governance Capability Dimensions

Dimension	Primary Focus
Strategic Governance & Leadership	Leadership and governance direction
Governance Policy, Risk & Compliance	Governance control and regulatory alignment
Data, Model, and Technical Governance	Data, AI systems, and technical assurance
Transparency, Explainability & Accountability	Transparent and accountable AI
Human Oversight & Ethical Decision-Making	Ethical governance and human control
Organizational Capability & Governance Culture	Governance readiness and culture
Performance Monitoring & Continuous Improvement	Governance evaluation and improvement
Stakeholder Trust & Societal Responsibility	Trust, transparency, and societal value

4.4 AI Governance Capability Maturity Levels

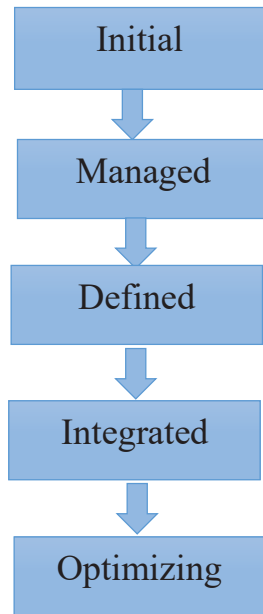
The AIGCMM adopts a five-level maturity structure to represent the progressive development of organizational AI governance capability. The maturity levels describe how governance evolves from ad hoc and reactive practices to an adaptive, continuously improving organizational capability. Progression through the levels reflects increasing institutionalization of governance processes, stronger leadership commitment, greater organizational integration, and more systematic governance evaluation.

Unlike compliance-based approaches, maturity advancement is determined by the balanced development of all governance capability dimensions rather than excellence in a single area. Organizations are expected to demonstrate continuous improvement across strategic, organizational, technical, and operational governance capabilities to achieve higher levels of maturity.

Table 3. AI Governance Capability Maturity Levels

Level	Maturity Stage	Characteristics
Level 1	Initial	Governance activities are informal, inconsistent, and largely reactive. AI governance depends on individual initiatives with limited organizational coordination.
Level 2	Managed	Basic governance policies, responsibilities, and risk management practices are established. Governance is implemented within individual projects but remains fragmented across the organization.
Level 3	Defined	Governance processes are standardized, documented, and consistently applied. Governance roles, policies, and technical controls are integrated into organizational AI development and deployment practices.
Level 4	Integrated	AI governance is embedded within strategic decision-making, enterprise risk management, and organizational operations. Governance performance is regularly monitored using defined metrics and continuous oversight mechanisms.
Level 5	Optimizing	Governance operates as an adaptive organizational capability supported by continuous learning, performance evaluation, stakeholder feedback, and ongoing governance improvement in response to technological and regulatory change.

Figure 3. Progression of AI Governance Capability



The proposed maturity model emphasizes that governance capability develops incrementally rather than instantaneously. As organizations progress through the maturity levels, governance evolves from isolated compliance activities toward an integrated capability that supports responsible, trustworthy, and sustainable AI deployment.

4.5 Governance Capability Assessment Framework

The AIGCMM provides a multidimensional assessment framework that enables organizations to evaluate governance capability across the eight proposed governance dimensions. Rather than producing a single maturity score, the framework encourages organizations to assess each governance dimension independently, enabling the identification of capability strengths, weaknesses, and improvement priorities.

Assessment results may be used for governance benchmarking, internal audits, capability development, and strategic planning. Organizations can periodically repeat the assessment to monitor governance progress and evaluate the effectiveness of improvement initiatives.

Table 4. Governance Capability Assessment Matrix

Governance Dimension	Assessment Criteria	Score (1–5)
Strategic Governance & Leadership	Executive leadership, governance strategy, accountability	
Governance Policy, Risk & Compliance	Policies, risk management, regulatory compliance	
Data, Model & Technical Governance	Data quality, security, model governance	
Transparency, Explainability & Accountability	Explainability, documentation, accountability	
Human Oversight & Ethical Decision-Making	Human review, ethics, bias management	
Organizational Capability & Governance Culture	Training, awareness, governance culture	
Performance Monitoring & Continuous Improvement	KPIs, audits, continuous improvement	
Stakeholder Trust & Societal Responsibility	Stakeholder engagement, transparency, public trust	

Following the assessment, organizations can compare capability scores across the eight governance dimensions to identify strengths, capability gaps, and areas requiring improvement. Dimensions with lower maturity scores indicate governance capabilities that require priority attention, while higher-scoring dimensions represent established organizational strengths. The assessment results support evidence-based decision-making by enabling organizations to prioritize governance improvement initiatives, allocate resources effectively, and monitor capability development over time. Because AI governance capability is multidimensional, the overall maturity of an organization should be determined through balanced performance across all governance dimensions rather than excellence in a single capability area.

4.5.1 Governance Capability Scoring Methodology

To support consistent application of the proposed AI Governance Capability Maturity Model (AIGCMM), governance capability is assessed across the eight governance capability dimensions

presented in Table 4. Each dimension is evaluated using a five-point maturity scale, where **Level 1** represents the lowest level of governance capability and **Level 5** represents the highest level.

The assessment should be conducted using organizational evidence, including governance policies, documented procedures, governance committees, technical documentation, audit reports, performance indicators, training records, and other relevant governance artefacts. Based on this evidence, an assessor assigns a maturity score between **1** and **5** for each governance dimension.

The overall organizational governance maturity is calculated by averaging the maturity scores across all eight dimensions:

$$\text{Overall Governance Maturity} = \frac{\sum_{i=1}^8 S_i}{8}$$

where S_i represents the maturity score assigned to the i th governance capability dimension.

In the absence of established empirical evidence supporting differential weighting, all governance capability dimensions are assigned equal importance in the conceptual model. Future empirical studies may investigate alternative weighting schemes for specific organizational contexts.

The resulting average score is interpreted using the maturity thresholds presented in Table 5.

Table 5. Overall Governance Maturity Classification

Average Score	Overall Maturity Level	Interpretation
1.00–1.99	Level 1 – Initial	Governance is informal and largely reactive.
2.00–2.99	Level 2 – Managed	Basic governance practices are established but inconsistently applied.
3.00–3.99	Level 3 – Defined	Governance processes are standardized and implemented across the organization.
4.00–4.49	Level 4 – Integrated	Governance is embedded within organizational strategy and operational decision-making.
4.50–5.00	Level 5 – Optimizing	Governance functions as a continuously improving organizational capability.

Although the overall maturity score provides a useful summary of governance capability, organizations should not rely solely on the aggregate result. Individual dimension scores should also be examined to identify capability strengths and weaknesses, as deficiencies in a single

governance dimension may significantly affect overall governance effectiveness. Consequently, governance improvement should prioritize dimensions with lower maturity scores while maintaining continuous improvement across all governance capabilities.

4.6 Governance Improvement Roadmap

The AIGCMM is intended not only as an assessment framework but also as a governance improvement tool. Following the maturity assessment, organizations should develop targeted improvement initiatives based on capability gaps identified across the eight governance dimensions. Rather than attempting to achieve the highest maturity level simultaneously across all dimensions, organizations should adopt a phased improvement approach that aligns governance priorities with organizational objectives, available resources, regulatory requirements, and AI deployment risks.

Progression through the maturity levels should be supported by measurable governance improvements. Organizations at lower maturity levels should first establish governance structures, policies, and accountability mechanisms before introducing advanced governance monitoring and optimization practices. As governance capability develops, emphasis should gradually shift from governance implementation to governance integration, organizational learning, and continuous improvement.

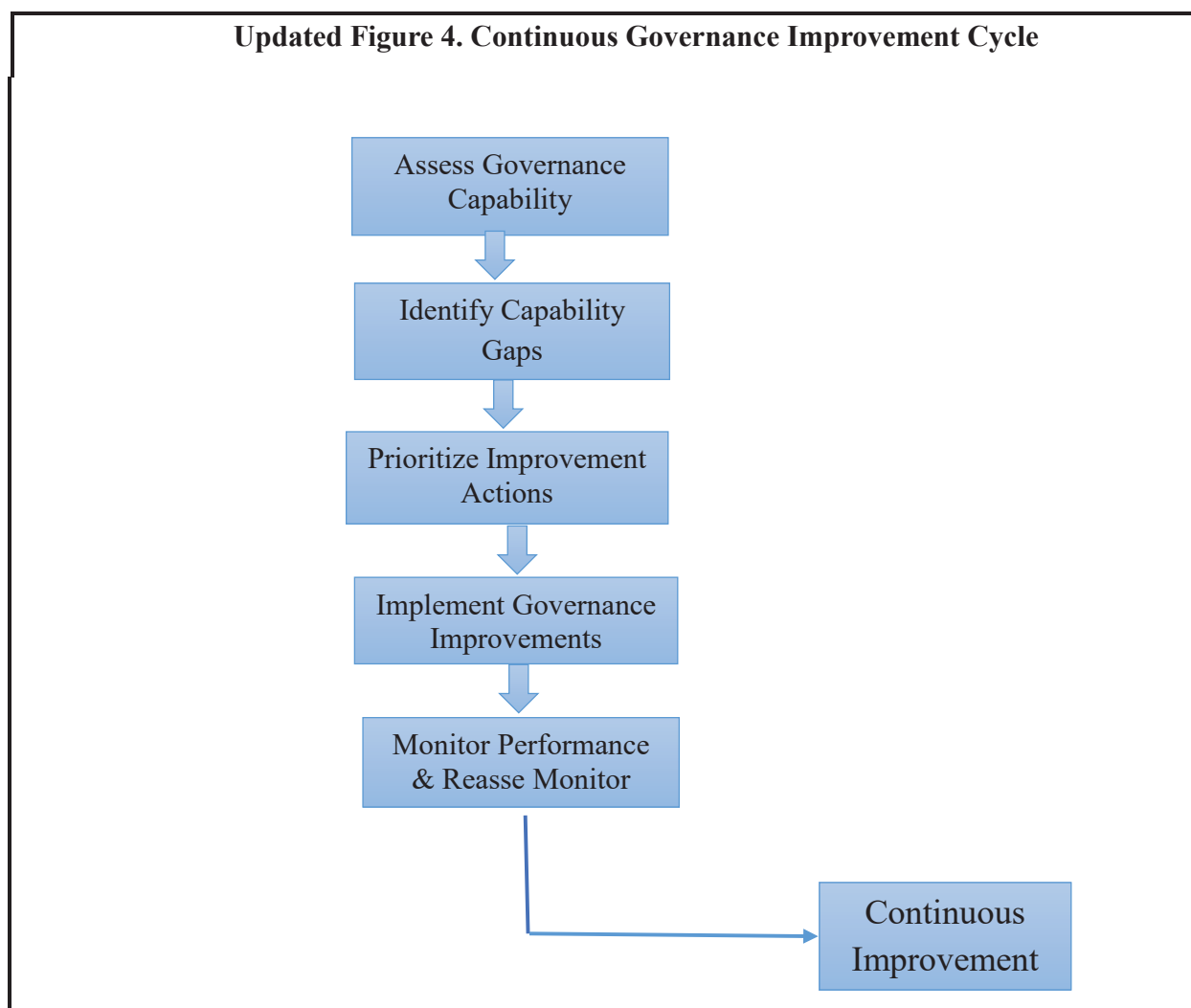
Table 6. Governance Improvement Roadmap

Current Maturity Level	Primary Improvement Objective	Typical Governance Actions
Level 1 -Initial	Establish governance foundation	Define governance roles, leadership commitment, and basic governance policies.
Level 2-Managed	Standardize governance practices	Implement documented governance processes, risk management, and compliance procedures.
Level 3 -Defined	Integrate governance across the organization	Embed governance into AI development, operational processes, and organizational decision-making.
Level 4-Integrated	Measure governance performance	Establish governance KPIs, conduct audits, monitor governance effectiveness, and support evidence-based decision-making.

Current Maturity Level	Primary Improvement Objective	Typical Governance Actions
Level 5 – Optimizing	Sustain continuous improvement	Continuously refine governance through organizational learning, stakeholder feedback, technological advances, and regulatory updates.

The roadmap recognizes that governance capability develops incrementally. Advancement to higher maturity levels should be based on demonstrated organizational capability rather than isolated governance initiatives. Organizations are therefore encouraged to periodically reassess governance maturity, monitor progress, and update improvement priorities in response to changing organizational and regulatory environments.

Updated Figure 4. Continuous Governance Improvement Cycle



Collectively, the governance dimensions, maturity levels, assessment framework, scoring methodology, and continuous improvement roadmap establish an integrated capability-oriented framework for evaluating and strengthening organizational AI governance. The following section discusses the theoretical, practical, and managerial implications of the proposed AIGCMM.

V. Discussion

5.1 Theoretical Contributions

This study contributes to the AI governance literature by conceptualizing **AI governance as an organizational capability** rather than solely as a collection of governance principles, ethical guidelines, or regulatory requirements. While existing research has significantly advanced responsible AI through governance frameworks, standards, and policy recommendations, comparatively limited attention has been given to understanding an organization's capability to operationalize and continuously improve these governance practices. By introducing the **AI Governance Capability Maturity Model (AIGCMM)**, this study extends existing AI governance research from principle-based governance toward capability-oriented governance assessment.

A second contribution is the integration of four complementary research streams—AI governance, organizational capability theory, capability maturity models, and international AI governance standards—into a unified conceptual framework. Rather than treating these domains independently, the AIGCMM demonstrates how they collectively support the systematic evaluation of governance capability. This integrated perspective provides a stronger theoretical foundation for understanding governance maturity as an evolving organizational capability.

Finally, the study contributes to maturity model research by proposing a governance-specific maturity model. Existing maturity models primarily assess technological adoption, digital transformation, or AI readiness, whereas the AIGCMM focuses on governance capability. This distinction broadens the application of maturity models within AI governance research and establishes a conceptual basis for future empirical investigation.

5.2 Practical Implications

Beyond its theoretical contribution, the AIGCMM provides practical value for organizations seeking to implement responsible and trustworthy AI. The framework enables organizations to systematically assess governance capability, identify governance weaknesses, prioritize improvement initiatives, and monitor governance development over time. Instead of relying solely on regulatory compliance or policy documentation, organizations can evaluate how effectively

governance practices are embedded within leadership, organizational processes, technical operations, and continuous improvement activities.

The framework is also relevant for regulators, auditors, and policymakers who require structured mechanisms for evaluating organizational governance readiness. By providing a common assessment structure, the AIGCMM supports governance benchmarking, internal governance reviews, and strategic planning across diverse organizational contexts. Its multidimensional design further allows organizations to identify specific governance capability dimensions requiring improvement rather than relying on a single maturity indicator. By supporting governance assessment, benchmarking, and continuous capability improvement, the AIGCMM provides organizations with a practical mechanism for operationalizing responsible and trustworthy AI.

5.3 Research Implications

The proposed AIGCMM establishes a foundation for future research on capability-oriented AI governance by extending existing governance frameworks beyond principle-based guidance toward systematic capability assessment. The framework provides researchers with a structured basis for investigating how organizational governance capability influences the responsible development, deployment, and management of AI systems.

The multidimensional structure of the AIGCMM also creates opportunities for empirical research. Future research may develop quantitative assessment instruments, validate the proposed governance dimensions through expert evaluation, examine organizational governance maturity using case studies or surveys, and investigate the relationship between governance capability and organizational outcomes such as regulatory compliance, stakeholder trust, AI risk management, and organizational performance. Longitudinal research may further explore how governance capability evolves over time and how continuous governance improvement contributes to sustainable AI adoption.

By establishing AI governance capability as a measurable organizational construct, the proposed framework provides a theoretical foundation for future empirical validation, comparative studies across industries, and the development of governance benchmarking tools.

5.4 Comparison with Existing Approaches

Existing AI governance frameworks and international standards have established important principles for responsible AI, including transparency, accountability, fairness, safety, and regulatory compliance. However, these approaches primarily define governance expectations rather than providing mechanisms for evaluating organizational governance capability. Similarly, AI maturity models generally focus on AI adoption, technological capability, or digital transformation instead of governance maturity.

The AIGCMM complements these existing approaches by shifting the emphasis from **"what should be governed"** to **"how well governance is implemented and continuously improved."** Rather than replacing governance standards, the model provides a capability-oriented assessment framework that enables organizations to operationalize governance principles through structured capability assessment and maturity progression.

Table 5 provides a high-level conceptual comparison based on the primary focus and intended purpose of representative governance frameworks, standards, and maturity models.

Table 5. Comparison of the AIGCMM with Existing Approaches

Comparison Criterion	OECD AI Principles	NIST AI RMF	ISO/IEC 42001	Existing Maturity Models	Proposed AIGCMM
Governance Principles	✓	✓	✓	Partial	✓
Capability Assessment	✗	Partial	Partial	✓	✓
Maturity Levels	✗	✗	✗	✓	✓
Continuous Improvement	Partial	✓	✓	Partial	✓
Governance Readiness Assessment	✗	✗	Partial	Partial	✓

5.5 Managerial Implications

For organizational leaders, AI governance has become a strategic capability rather than solely a compliance requirement. The AIGCMM provides managers with a structured decision-support framework for evaluating organizational governance readiness, prioritizing governance investments, and monitoring organizational governance capability over time. The maturity model also supports continuous governance monitoring, enabling organizations to adapt governance practices as AI technologies, regulatory requirements, and stakeholder expectations evolve. Consequently, the framework encourages organizations to embed governance within strategic planning rather than treating governance as a reactive compliance activity.

Collectively, the theoretical, practical, research, and managerial implications demonstrate that the proposed AIGCMM extends existing AI governance research beyond principle-based guidance by introducing a capability-oriented maturity framework. By integrating governance capability dimensions, maturity assessment, and continuous improvement within a single conceptual model, the AIGCMM provides a structured foundation for both organizational implementation and future empirical research on responsible and trustworthy AI governance.

VI. Limitations and Future Research

Although the proposed AI Governance Capability Maturity Model (AIGCMM) advances capability-oriented AI governance research, several limitations should be acknowledged.

First, the AIGCMM is a conceptual framework developed using Design Science Research and a synthesis of existing literature. While the model is theoretically grounded, it has not yet been empirically validated in organizational settings. Future research should evaluate the reliability, usability, and applicability of the framework through expert evaluation, organizational case studies, survey-based assessments, and longitudinal studies.

Second, the proposed framework is intended to provide a generic assessment model applicable across diverse organizational contexts. However, governance requirements vary according to industry sectors, regulatory environments, organizational size, AI application domains, and risk profiles. Future research should investigate sector-specific adaptations of the AIGCMM for industries such as healthcare, finance, manufacturing, public administration, and critical infrastructure.

Third, the maturity model currently assumes that all governance capability dimensions contribute equally to overall governance maturity. Although this assumption supports simplicity and practical implementation, future studies may develop weighted assessment approaches that reflect the relative importance of individual governance dimensions under different organizational contexts and AI risk levels.

Finally, the present study focuses on conceptual model development rather than quantitative measurement. Future research should develop validated measurement scales, assessment instruments, decision-support tools, and software-based governance assessment platforms to facilitate consistent organizational implementation and benchmarking.

Overall, the proposed AIGCMM establishes a theoretical foundation for future empirical research while providing a structured framework for assessing and improving organizational AI governance capability. Continued validation and refinement will further strengthen its practical applicability and support the advancement of responsible, trustworthy, and sustainable AI governance.

VII. Conclusion

Artificial Intelligence is transforming organizational decision-making, business operations, and public services, making effective AI governance an essential organizational capability. Although existing governance frameworks, ethical principles, and international standards provide valuable guidance for responsible AI, organizations continue to face challenges in systematically assessing and improving their governance capability. This study addressed this challenge by proposing the AI Governance Capability Maturity Model (AIGCMM), a capability-oriented framework that extends AI governance beyond compliance and principle-based guidance.

Developed through a Design Science Research approach, the AIGCMM integrates eight governance capability dimensions, five maturity levels, a structured assessment methodology, and a continuous improvement mechanism into a unified conceptual framework. By conceptualizing AI governance as an organizational capability, the model enables organizations to evaluate governance readiness, identify capability gaps, prioritize governance improvements, and monitor governance development over time.

The proposed framework contributes to AI governance research in several important ways. First, it introduces a capability-oriented perspective that complements existing AI governance principles and standards. Second, it extends maturity model research by providing a governance-specific assessment framework rather than focusing solely on AI adoption or technological capability. Third, it offers a practical foundation for governance benchmarking and continuous capability improvement that can support organizations, regulators, policymakers, and researchers.

While empirical validation remains an important direction for future research, the proposed AIGCMM provides a robust conceptual foundation for advancing capability-oriented AI governance. As AI technologies continue to evolve, organizations will require governance approaches that are adaptive, measurable, and continuously improving. The proposed AIGCMM represents an important step toward achieving this objective by providing a structured, capability-oriented framework for assessing, improving, and sustaining responsible, trustworthy, and sustainable AI governance.

Reference

- [1] OECD, "OECD Principles on Artificial Intelligence," OECD Publishing, Paris, France, May 2019. [Online]. Available: <https://oecd.ai/en/ai-principles>. [Accessed: Jul. 2026].

- [2] National Institute of Standards and Technology (NIST), "Artificial Intelligence Risk Management Framework (AI RMF 1.0)," NIST AI 100-1, Gaithersburg, MD, USA, Jan. 2023.
- [3] *Information Technology—Artificial Intelligence—Management System*, ISO/IEC Standard 42001:2023, ISO/IEC, Geneva, Switzerland, 2023.
- [4] European Parliament and Council of the European Union, "Regulation laying down harmonised rules on Artificial Intelligence (Artificial Intelligence Act)," *Official Journal of the European Union*, 2024.
- [5] UNESCO, "Recommendation on the Ethics of Artificial Intelligence," Paris, France, Tech. Rep. SHS/BIO/REC-AI, 2021.
- [6] High-Level Expert Group on Artificial Intelligence, "Ethics Guidelines for Trustworthy AI," European Commission, Brussels, Belgium, Apr. 2019.
- [7] L. Floridi and J. Cowls, "A unified framework of five principles for AI in society," *Harvard Data Science Review*, vol. 1, no. 1, Jul. 2019.
- [8] D. Leslie, "Understanding artificial intelligence ethics and safety: A guide for the responsible design and implementation of AI systems in the public sector," The Alan Turing Institute, London, U.K., Tech. Rep., 2019.
- [9] B. Mittelstadt, P. Allo, M. Taddeo, S. Wachter, and L. Floridi, "The ethics of algorithms: Mapping the debate," *Big Data & Society*, vol. 3, no. 2, pp. 1–21, Dec. 2016.
- [10] S. Russell, *Human Compatible: Artificial Intelligence and the Problem of Control*. New York, NY, USA: Viking, 2019.
- [11] J. Morley, L. Floridi, L. Kinsey, and A. Elhalal, "From what to how: An initial review of publicly available AI ethics tools, methods and research to translate principles into practices," *Science and Engineering Ethics*, vol. 26, no. 4, pp. 2141–2168, Aug. 2020.
- [12] M. Veale and F. Z. Borgesius, "Demystifying the Draft EU Artificial Intelligence Act," *Computer Law Review International*, vol. 22, no. 4, pp. 97–112, Aug. 2021.
- [13] R. Ashok, S. Narayanan, and R. Rajan, "Artificial intelligence governance: A systematic literature review," *IEEE Access*, vol. 12, pp. 23345–23372, 2024.
- [14] A. Jobin, M. Ienca, and E. Vayena, "The global landscape of AI ethics guidelines," *Nature Machine Intelligence*, vol. 1, no. 9, pp. 389–399, Sep. 2019.
- [15] M. Brundage *et al.*, "Toward trustworthy AI development: Mechanisms for supporting verifiable claims," arXiv:2004.07213, Apr. 2020.

- [16] World Economic Forum, "Empowering AI Leadership: AI Governance Toolkit," World Economic Forum, Geneva, Switzerland, Tech. Rep., 2023.
- [17] D. J. Teece, G. Pisano, and A. Shuen, "Dynamic capabilities and strategic management," *Strategic Management Journal*, vol. 18, no. 7, pp. 509–533, Aug. 1997.
- [18] D. J. Teece, "Explicating dynamic capabilities: The nature and microfoundations of sustainable enterprise performance," *Strategic Management Journal*, vol. 28, no. 13, pp. 1319–1350, Dec. 2007.
- [19] R. M. Grant, "The resource-based theory of competitive advantage: Implications for strategy formulation," *California Management Review*, vol. 33, no. 3, pp. 114–135, Apr. 1991.
- [20] C. E. Helfat and M. A. Peteraf, "The dynamic resource-based view: Capability lifecycles," *Strategic Management Journal*, vol. 24, no. 10, pp. 997–1010, Oct. 2003.
- [21] R. Amit and P. J. H. Schoemaker, "Strategic assets and organizational rent," *Strategic Management Journal*, vol. 14, no. 1, pp. 33–46, Jan. 1993.
- [22] A. R. Hevner, S. T. March, J. Park, and S. Ram, "Design science in information systems research," *MIS Quarterly*, vol. 28, no. 1, pp. 75–105, Mar. 2004.
- [23] K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information systems research," *Journal of Management Information Systems*, vol. 24, no. 3, pp. 45–77, Dec. 2007.
- [24] J. Venable, J. Pries-Heje, and R. Baskerville, "A comprehensive framework for evaluation in design science research," in *Proc. 7th Int. Conf. Design Science Research in Information Systems (DESRIST)*, Las Vegas, NV, USA, 2012, pp. 423–438.
- [25] CMMI Product Team, *CMMI for Development, Version 2.0*. Pittsburgh, PA, USA: Software Engineering Institute, Carnegie Mellon University, 2018.
- [26] M. C. Paulk, B. Curtis, M. B. Chrissis, and C. V. Weber, "Capability Maturity Model for Software," Software Engineering Institute, Carnegie Mellon University, Pittsburgh, PA, USA, Tech. Rep. CMU/SEI-93-TR-024, Feb. 1993.
- [27] A. Van Looy, A. De Backer, G. Poels, and M. Snoeck, "Choosing the right business process maturity model," *Information & Management*, vol. 50, no. 7, pp. 466–488, Nov. 2013.
- [28] R. Maier, *Knowledge Management Systems: Information and Communication Technologies for Knowledge Management*, 3rd ed. Berlin, Germany: Springer, 2007.
- [29] Gartner, "AI Maturity Model," Gartner Research, Stamford, CT, USA, Tech. Rep., 2023.

- [30] Deloitte, "State of AI in the Enterprise," Deloitte Insights, Tech. Rep., 2023.
- [31] IBM Institute for Business Value, "The Enterprise Guide to AI Governance," IBM Corporation, Armonk, NY, USA, Tech. Rep., 2023.
- [32] Microsoft, "Responsible AI Standard, Version 2," Microsoft Corporation, Redmond, WA, USA, Tech. Rep., Jun. 2022.
- [33] Google, "Responsible AI Practices," Google Research, Mountain View, CA, USA, Tech. Rep., 2023.
- [34] D. Gunning and D. Aha, "DARPA's Explainable Artificial Intelligence Program," *AI Magazine*, vol. 40, no. 2, pp. 44–58, 2019.
- [35] S. Wachter, B. Mittelstadt, and C. Russell, "Counterfactual explanations without opening the black box: Automated decisions and the GDPR," *Harvard Journal of Law & Technology*, vol. 31, no. 2, pp. 841–887, 2018.
- [36] F. Doshi-Velez and B. Kim, "Towards a rigorous science of interpretable machine learning," arXiv:1702.08608, Feb. 2017.
- [37] R. Guidotti, A. Monreale, S. Ruggieri, F. Turini, F. Giannotti, and D. Pedreschi, "A survey of methods for explaining black box models," *ACM Computing Surveys*, vol. 51, no. 5, pp. 1–42, Jan. 2019.
- [38] M. T. Ribeiro, S. Singh, and C. Guestrin, "'Why should I trust you?': Explaining the predictions of any classifier," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowledge Discovery and Data Mining (KDD '16)*, San Francisco, CA, USA, 2016, pp. 1135–1144.