Artificial Intelligence (AI) Security and Risk Policy

Purpose

This policy defines Morehouse College's standards for the secure, ethical, and responsible use of Artificial Intelligence (AI) technologies across all institutional units. It is designed to:

- Promote effective use of AI for operational purposes
- Ensure AI systems are deployed and used securely and in compliance with applicable laws
- Minimize risk to institutional data, reputation, and ethical standards
- Support alignment with Morehouse's IT Security, Data Classification, and Acceptable Use Policies

Scope

This policy governs institutional data, systems, and risk management. Academic use of AI in teaching, learning, and research is guided by the Morehouse College Academic AI Guidelines.

The AI Governance Committee will focus on matters related to this policy. The AI Governance Committee serves in an advisory capacity to the CIO and Provost on matters of risk, compliance, and institutional AI use.

This policy applies to the use, management, or development of AI systems at Morehouse College, including:

- All Morehouse employees, faculty, students, contractors, and vendors whose use of AI tools involves Morehouse data, systems, or resources, whether the tool is institutionally provided.
- Any use of AI or machine learning tools, including public AI (e.g., ChatGPT) and internal or third-party AI systems
- All devices, accounts, networks, or platforms that store or process data related to AI activity
- Cloud-based, open-source, and proprietary AI technologies used in connection with institutional data or systems

Relationship to Academic and Operational Use

This policy governs institutional risk, data, and systems.

Academic use of AI in teaching, learning, and scholarship is guided by the Morehouse College Guidelines for the Academic Use of Generative AI, which preserve faculty autonomy within a risk-aware framework.

Operational guidance for staff use of AI tools is provided in the AI Usage Guidelines for Staff, which translate this policy into day-to-day practice.

Definitions

Artificial Intelligence (AI)	Systems capable of performing tasks that traditionally require human intelligence.

Generative AI	AI capable of creating new content (e.g., text, images, code) based on prompts.
Public AI	Vendor-owned systems available to general users where input/output data may be collected externally.
Private AI	AI systems developed or hosted within Morehouse-controlled environments.
Sensitive Data	Data classified as confidential or restricted under Morehouse's Data Classification Policy.
Responsible AI:	The ethical application of AI with consideration for fairness, transparency, and privacy.

Governance Framework

An AI Governance Committee, co-led by the CIO and Provost or a designee, will:	
1.	Evaluate proposed AI tools for security and compliance
2.	Approve exceptions to policy (e.g., for research)
3.	Review risk, legal exposure, and ethical impacts of AI use
4.	Coordinate training and monitor emerging AI risks
5.	Maintain documentation of authorized AI systems and approved use cases

Principles of Responsible AI Use

All AI activities must align with the following principles:

- Privacy
- Fairness
- Transparency
- Security
- Accountability
- Reliability

Acceptable Use of AI

Users must disclose the use of AI in their outputs where applicable. AI-generated content must be labeled and reviewed by a human before dissemination. Faculty, staff, and students must refer to the Morehouse College Academic AI Guidelines for teaching, learning, and research-specific guidance.

Permitted Uses (with Authorization):

- Research, data analysis, and student support functions
- Administrative automation and analytics
- Teaching enhancement tools (outside of classroom policy scope)
- Institutional communications and digital marketing (monitored)

Prohibited Use of AI

- Uploading or processing sensitive data in public AI tools
- Allowing AI to drive disciplinary decisions, admissions outcomes, or evaluations without meaningful human oversight.
- Deploying AI-generated content as human-authored work without disclosure
- AI use that misrepresents Morehouse College or violates copyright/IP rights
- Incorporating AI-generated code into production systems without review and approval

Security Controls

Additional safeguards include firewall egress controls, API registration requirements, and DLP enforcement for AI tool usage. Unauthorized AI API calls or browser extensions are prohibited.

AI systems may only be used with data classified as Public or Internal under Morehouse's Data Classification Policy, unless explicit authorization is obtained from the Data Owner and CIO. Confidential and Restricted data must not be used with public or unapproved AI systems.

Data Confidentiality	**Data Integrity**	**Data Resiliency**	**IT Controls**
- AES-256 encryption for sensitive data - Formal approval from data owners for AI training data - Prohibit sharing sensitive data with public AI tools	- Verify quality of training data - Label AI-generated content - Audit datasets for tampering	- Weekly backups - Quarterly testing of RTO/RPO	- MFA required for AI access - Regular access and audit logs - Intrusion detection systems - Secure key management and rotation

Public vs. Private AI Usage

Public AI	**Private AI**
Must not be used for institutional or sensitive data (e.g., OpenAI, Gemini)	Must be approved by IT and subject to institutional control and monitoring

Training and Awareness

Completion of an AI Security & Ethics course (s) will be tracked through the Learning Management System and reported to HR. All users must complete an AI Security & Ethics course provided by the Office of Information Technology and Human Resources. The objective of such course is to support faculty, staff, and students in their use of AI tools to accomplish essential work happening at Morehouse College.

All employees and students must complete training on:

- Responsible AI use
- Data protection and privacy
- Security risks and safeguards
- Institutional expectations and updates

While the focus of training will be on enablement, failure to complete required training may result in suspension of AI access or disciplinary action after a three-step escalation process (reminder, supervisor notice, HR referral).

Certain roles, such as system developers, data stewards, or AI researchers, may require supplemental role-based training prior to being granted access to advanced AI systems.

Completion of required training may be tied to continued access to approved AI systems.

Legal and Regulatory Compliance

This policy supports compliance with:

- FERPA, HIPAA, GLBA
- GDPR (as applicable)
- CPRA, PIPEDA, and Copyright Law
- Internal data governance and IT security requirements

Monitoring and Enforcement

For institutional data and official work, only approved AI tools may be used. Individuals may explore AI tools for learning and experimentation, provided no sensitive or institutional data is used. The list of approved tools, along with permitted data classifications for each, will be published and maintained by the Office of Information Technology. Use of unapproved or experimental AI tools must be reviewed by the AI Governance Committee.

- Monitoring may be conducted for compliance and risk assessment
- Violations may result in disciplinary action, up to and including termination
- Legal escalation may occur for data breaches or unlawful use

Approved AI Tools

Any suspected misuse of AI tools, including unauthorized data use, biased or harmful outputs, or data breaches involving AI systems, must be reported immediately to the IT Security Office. These incidents will be documented, investigated, and addressed in accordance with Morehouse College's Incident Response Policy. AI tools used in critical decisions must maintain an audit trail of inputs, outputs, and decisions for at least 12 months.

An Approved AI Tools List will be maintained by the Office of Information Technology and published on the College intranet. Faculty, staff, and students are responsible for consulting this list before adopting or using any AI technology in connection with College systems or institutional data."

AI Misuse and Incident Reporting

All AI systems used in high-impact decisions (e.g., admissions, financial aid, disciplinary action) must undergo a documented review process to ensure outputs are explainable and free of systemic bias. A summary of model behavior and testing results must be reviewed and endorsed by the AI Governance Committee before deployment.

Explainability and Bias Review

Each AI system must maintain a model card documenting its purpose, data sources, limitations, and approval status. Ongoing bias monitoring and third-party audits may be required for high-impact systems.

All AI systems used in high-impact decisions, such as admissions, financial aid, student conduct, hiring, or disciplinary processes, must undergo a formal review to ensure they are both explainable and free from systemic bias. Specifically:

- AI tools must provide clear documentation describing how their outputs are generated and what data sources are used.
- A bias and fairness assessment must be conducted prior to deployment of any AI tool used in decision-making processes that affect individuals or protected groups.
- All reviews must be documented and submitted to the AI Governance Committee for approval.
- Post-deployment, such systems must be monitored regularly for continued fairness, accuracy, and transparency.

Non-Compliance

Violations of this policy will be treated like other allegations of wrongdoing at Morehouse College Allegations of misconduct will be adjudicated according to established procedures. Sanctions for non-compliance may include, but are not limited to, one or more of the following:

1. Disciplinary action according to applicable Morehouse College policies.
2. Termination of employment.
3. Legal action according to applicable laws and contractual agreements.

Related Policies

- Data Security and Classification Policy
- Acceptable Use of Technology Policy
- IT Governance Policy
- Information Security Policy

Document Control and Versioning

This policy is maintained within the College's compliance management platform. Version control, effective dates, ownership, and review cycles are automatically tracked and auditable through that system. The Chief Information Officer (CIO) and the Office of Compliance serve as the approval authorities for this policy. Any revisions or exceptions to this policy are recorded within the compliance platform, ensuring traceability and compliance with institutional and accreditation requirements.

APPENDICES – Artificial Intelligence (AI) Policy

Appendix A: AI Risk & Governance Frameworks

This appendix identifies the frameworks, standards, and assessment tools that guide Morehouse College's implementation, evaluation, and monitoring of Artificial Intelligence (AI) systems. These resources ensure alignment with institutional policies, accreditation requirements, and applicable legal and ethical standards.

1. Core AI Governance Frameworks

Framework	Publisher/Authority	Purpose/Use Case
NIST AI Risk Management Framework (RMF) 1.0	National Institute of Standards and Technology (NIST)	Establishes a risk-based approach for identifying, assessing, and mitigating AI risks, emphasizing governance, transparency, and accountability.
ISO/IEC 42001:2023 (AI Management Systems)	International Organization for Standardization (ISO)	Defines management system requirements for developing, deploying, and maintaining trustworthy AI systems.
OECD AI Principles	Organisation for Economic Co-operation and Development	Provides international guidance on fairness, transparency, accountability, and human oversight.
EU AI Act (Reference)	European Union	Offers a global regulatory benchmark on AI risk tiers and human-in-the-loop obligations.

2. Security and Compliance Alignment

Framework	Purpose / Alignment
ISO 27001 + Annex for AI Controls	Ensures supporting infrastructure meets internationally recognized information security standards.
SOC 2 (Type II) + AI Mapping	Provides third-party assurance for vendors handling institutional or student data.
CSA STAR Certification	Evaluates security and privacy maturity for cloud-based AI tools used within the College environment.
NIST SP 800-53 / 800-171 Crosswalk	Aligns AI-related controls with federal data protection baselines, especially for grant-funded or research data.

3. Ethical and Fairness Auditing Tools

Category	Examples	Purpose
Independent Auditors	Credo AI, Holistic AI, PwC Responsible AI, Deloitte Trustworthy AI	Conduct external reviews for bias, fairness, documentation completeness, and 'responsible AI' scoring.
Research Frameworks	MITRE ATLAS, Responsible AI Institute (RAI), Partnership on AI	Provide open evaluation rubrics and self-assessment tools for academic and research AI systems.
Academic Benchmarks	Stanford HAI Policy Toolkit, AI Ethics Guidelines Global Inventory	Offer comparative metrics for AI ethics, governance, and institutional maturity.

4. Institutional AI Governance Maturity Model

Morehouse College evaluates AI systems using six governance domains, each scored from 1 (low maturity) to 5 (high maturity):

1. **Governance Structure** – Defined roles, documented approvals, and oversight reviews.
2. **Data Management** – Data classification, consent, and minimization.
3. **Transparency & Explainability** – Model cards, audit trails, and disclosure.
4. **Fairness & Bias Control** – Pre- and post-deployment bias assessments.
5. **Security & Resilience** – NIST/ISO control alignment and monitoring.
6. **Lifecycle & Decommissioning** – Documented retraining, review, and secure retirement.

5. Continuous Improvement and Audit Integration

- Annual self-assessment against NIST AI RMF and ISO/IEC 42001 standards.
- Evidence maintained within Drata (policies, mappings, training records).
- Regular reviews by the AI Governance Committee and Information Security Office.
- Results summarized in the annual AI Governance Report.

6. References

- NIST AI RMF v1.0 (2023)
- ISO/IEC 42001:2023
- OECD AI Principles (2019)
- CSA STAR Certification Program
- MITRE ATLAS
- Responsible AI Institute (RAI)
- Partnership on AI Guidelines

Appendix B: AI Lifecycle Management and Decommissioning

This appendix defines the required lifecycle stages and controls for all Artificial Intelligence (AI) systems developed, procured, or operated by Morehouse College. It ensures that every AI asset—from concept to retirement—is managed securely, ethically, and in compliance with institutional, legal, and accreditation standards.

1. Lifecycle Phases

Phase	Objectives & Required Activities	Responsible Parties
Initiation & Use Case Definition	Document purpose, stakeholders, and outcomes. Perform NIST AI RMF risk assessment and confirm data classification.	Project Sponsor / AI Governance Committee
Design & Data Preparation	Validate data sources, apply minimization/anonymization, and complete model card with lineage and limitations.	Data Owner / Developer / IRB
Development & Testing	Use isolated environments with encryption, perform bias and vulnerability testing.	Developers / IT Security
Approval & Deployment	Submit package for AI Governance Committee approval, register model/API, enable audit logging.	CIO / AI Governance Committee
Operations & Monitoring	Monitor drift and bias quarterly, revalidate data annually, document retraining and incidents in Drata.	System Owner / IT Security
Review & Continuous Improvement	Conduct annual reviews per ISO 42001 and NIST AI RMF metrics, update governance practices.	AI Governance Committee
Decommissioning & Archival	Securely retire or delete models, retain metadata for five years for audit.	CIO / Data Owner / IT Security

2. Security and Compliance Requirements

- Adhere to Information Security, Data Classification, and Incident Response policies.
- Encryption (AES-256) and MFA required for all AI environments.
- API endpoints must be documented and approved.
- High-impact AI systems must undergo annual bias re-testing.

3. Documentation and Audit Evidence

Model cards, approvals, and logs must be stored in Drata or approved repositories. Decommissioning records must include:

1. System name & version
2. Owner and custodian
3. Reason for decommissioning
4. Date of destruction/archival confirmation

Audit evidence will be reviewed annually by the Office of Information Security.

4. Cross-References

- Appendix A – AI Risk & Governance Frameworks
- Data Security and Classification Policy
- Information Security Policy • Incident Response Policy
- Acceptable Use of Technology Policy