OFFICE of
information
technology
services

August 2025

ARTIFICIAL INTELLIGENCE (AI)

Governance Policy, Standard, and Guideline

Table of contents

1.0 Introduction

1.1 Overview

Based on the surge of interest in artificial intelligence (AI) technologies, ITS created an AI framework to deliver a structured pathway for responsibly harnessing AI across Idaho state government. Grounded in eight foundational principles, the framework balances ethical rigor with practical implementation. This ensures AI capabilities create real public value for citizens and adhere to the highest standards of transparency, accountability, privacy, and fairness. More importantly, this approach places Idahoans at the center, with AI serving as a tool to enhance human potential rather than replace it.

This section operationalizes the AI framework for State of Idaho users, providing specific information on AI technologies and prescribing measures to establish and reinforce the framework at ITS. It also aligns directly with the National Institute of Standards and Technology AI Risk Management Framework (NIST AI RMF) and adopts its core functions: GOVERN, MAP, MEASURE, and MANAGE. AI is an emerging technology, and it is critically important for all personnel to be aware of and adhere to ITS' AI usage requirements.

1.2 Purpose

The purpose of this document is to ensure State Agencies have formal, documented policies and procedures to facilitate the implementation of AI technologies.

1.3 Scope

This document applies to all State Data, supported information systems, activities, and assets owned, leased, controlled, or used by Idaho agencies, their agents, contractors, or other business partners on behalf of the agency. These policies apply to all State of Idaho employees, contractors, sub-contractors, and their respective facilities supporting official State business operations, wherever State data is stored or processed, including any third-party contracted by ITS to handle, process, transmit, store, or dispose of State data.

This standard does not supersede any other applicable law or stricter agency directive or existing labor management agreement in effect as of the effective date of this policy.

ITS reserves the right to revoke, change, or supplement this document at any time without prior notice. Such changes must be effective immediately upon approval by management, unless otherwise stated.

2.0 AI Policies

2.1 (P.ITS-01) AI Governance Policy

Purpose

The purpose of the policy is to establish a framework for artificial intelligence (AI) within State of Idaho agencies and departments. It defines risk classification methodologies, oversight responsibilities, and implementation requirements to enable innovation and ensure appropriate safeguards for privacy, security, and fairness.

Scope

The policy applies to the use of all AI tools by all Idaho State personnel, contractors, and affiliates (personnel) conducting state business.

Policy

The policy defines the structures that enable Idaho to govern AI implementation responsibly across varying system complexities, agency and department sizes, and use cases. The model balances comprehensive oversight with operational flexibility, allowing agencies and departments to fulfill their missions according to established principles. ITS must:

(a) Establish the following:
  1. AI Executive Committee
  2. Ethics Advisory Committee
  3. Technical Review Board
  4. AI Innovation Team

(b) Define the roles and responsibilities for:
  1. The AI Executive Committee
  2. The Ethics Advisory Committee
  3. The Technical Review Board
  4. The AI Innovation Team
  5. Information Owners
  6. Agencies and Departments
  7. IT Liaisons
  8. State Personnel

(c) Risk Classification Model:
  1. Define a risk classification model incorporating risk factors and risk tiers
  2. Classify all AI implementations according to a multi-factor risk classification model that aligns with existing ITS policy Security Classification (RA-02)

(d) AI System Governance:
  1. Designate specific approval authorities that align with the risk classification tiers
  2. Align AI System Governance with the NIST AI RMF
  3. Document requirements based off risk classifications
  4. Follow a structured AI implementation process throughout its lifecycle

(e) Define Generative AI (GenAI) requirements

(f) Define AI privacy and security requirements

(g) Define documentation requirements for all AI systems

(h) Define AI procurement requirements

(i) Define transparency and fairness requirements

(j) Define AI system monitoring and maintenance requirements

(k) Include AI in the incident response plans

(l) Define training and awareness requirements

(m) Adopt a shared responsibility model for AI governance: see Appendix A

(n) Standard Alignment: AI Governance must comply with the (S.ITS-01) AI Governance Standard

3.0 AI Standards

3.1 (S.ITS-01) AI Governance Standard

(P.ITS-01) AI Usage Policy requires ITS to establish consistent implementation requirements that enable agencies and departments to realize the benefits of AI and manage potential risks. These include:

(a) Responsibilities:

1. **AI Executive Committee:** The AI Executive Committee is established under the Idaho Technology Authority and shall be responsible for:
   - i. Setting statewide priorities for AI implementation
   - ii. Reviewing high-risk implementations
   - iii. Advocating for shared resources for AI initiatives
   - iv. Ensuring alignment with Idaho's broader technology strategy
   - v. Reviewing proposed changes to standard governance processes
   - vi. Coordinating Idaho's AI governance approach with relevant standards and guidelines
2. **Ethics Advisory Committee:** The Ethics Advisory Committee is established under the Information Technology Leadership Council and shall be responsible for:
   - i. Advising on ethical risks, fairness concerns, and demographic impact
   - ii. Providing specialized consultation on high-risk consultations
   - iii. Reviewing privacy and fairness assessments for medium and high-risk systems
   - iv. Developing ethical guidelines for AI development and use
3. **Technical Review Board:** The Technical Review Board shall be responsible for:
   - i. Advising on technical issues related to AI proposals and implementations
   - ii. Conducting technical reviews of medium and high-risk systems
   - iii. Evaluating model development, platform standards, risk assessments, and digital infrastructure
   - iv. Providing guidance on implementing appropriate technical safeguards
   - v. Reviewing security controls and mitigation strategies
4. **AI Innovation Team:** The AI Innovation Team shall be responsible for:
   - i. Providing implementation support to agencies and departments
   - ii. Facilitating communities of practice around AI governance
   - iii. Developing documentation standards and templates
   - iv. Maintaining the state's AI system inventory
   - v. Coordinating testing and evaluation of emerging AI technologies
   - vi. Developing and distributing standardized implementation tools
   - vii. Supporting agency and department-level IT Liaisons
5. **Information Owners:** Information owners shall:
   - i. Validate decisions regarding assignment of security controls and access privileges
   - ii. Execute formal information-sharing agreements with other agencies and departments prior to exchanging AI-generated information
   - iii. Perform periodic risk classification reviews based on changes in sensitivity, value, and impact to the agency or department
   - iv. Classify AI systems based on the "high water mark" (highest impact level) of the system's associated information
   - v. Ensure appropriate documentation and safeguards are implemented based on risk tier
6. **Agencies and Departments:** Agencies and departments shall:
   - i. Establish policies and procedures for managing risk classification within the agency or department

ii. Ensure that information belonging to different classification levels is appropriately protected

iii. Observe and maintain the appropriate security for classification levels assigned by another agency or department's information owner

iv. Provide training to information owners and handlers on this policy

v. Designate internal IT Liaisons for implementation coordination

vi. Ensure all AI systems are disposed of in accordance with established policies and regulations

vii. Maintain an inventory of AI systems used within the agency or department

viii. Implement appropriate incident response procedures for AI-related incidents

7. **IT Liaisons**: IT Liaisons designated by each agency and department shall:

i. Serve as primary point of contact for agency and department AI implementations

ii. Coordinate risk assessment and approval processes

iii. Maintain agency and department-level AI documentation and inventory

iv. Facilitate agency and department compliance with this policy

v. Participate in cross agency and department AI communities of practice

vi. Coordinate AI training for agency and department personnel

vii. Liaise with the AI Innovation Team

8. **State Personnel**: All state personnel shall:

i. Confirm the AI tool is on your agency's approved list - contact your IT liaison if unsure.

ii. Create an account specifically used for state business

iii. Follow all AI policies, standards, and guidelines

(b) **Risk Classification Model**: All AI implementations must be classified according to a multi-factor risk classification model that aligns with existing ITS policy Security Classification (RA-02). This model evaluates AI systems across six dimensions and rates each dimension based on a four-point scale ("Low", "Medium", "High", and "Very High.").

1. Risk Classification Factors:

i. **Personal Data Sensitivity**: Assesses the nature of the data the system uses and maps directly to ITS policy Security Classification (RA-02) levels:

(A) Level 1 (Unrestricted) data

(B) Level 2 (Limited) data

(C) Level 3 (Restricted) data

(D) Level 4 (Critical) data

ii. **Decision Impact**: Assesses how system outputs affect individuals, considering FIPS-199 impact levels (low, medium, high) as referenced in ITS policy Security Classification (RA-02).

iii. **Autonomy Level**: Evaluates human oversight involvement, with fully autonomous systems carrying higher risk than those with human validation.

iv. **Transparency**: Measures how understandable the system's logic and outcomes are to non-technical stakeholders, with "black box" models scoring higher.

v. **Scope and Scale**: Considers system reach, from limited pilots to enterprise-wide deployments impacting thousands.

vi. **Novelty and Complexity**: Evaluates whether the system uses well-established methods or introduces untested approaches with potential unforeseen risks.

2. **Risk Tiers**: The individual dimension ratings described in the previous section are weighted and combined to produce a total risk score. This score places systems into one of three governance tiers aligned with ITS policy Security Classification (RA-02) levels:

- **Tier 1: Low Risk**: These systems generally process non-sensitive information, have limited decision impact, maintain human oversight, and have a narrow scope of operation. Their overall risk profile typically aligns with Classification Levels 1-2 ("Unrestricted" or "Limited") from (RA-02).

i. **Tier 2: Medium Risk**: These systems may process some sensitive information, have moderate decision impact, operate with reduced human oversight, or serve a broader user base. Their overall risk profile typically aligns with Classification Level 3 ("Restricted") from (RA-02).

ii. **Tier 3: High Risk**: These systems process highly sensitive information, have significant decision impact, operate with substantial autonomy, or serve large populations. Their overall risk profile typically aligns with Classification Level 4 (“Critical”) from (RA-02).

3. **Risk Classification Requirements**: All AI implementations must be classified according to the risk assessment methodology defined in P.ITS-01. The risk classification must align with ITS policy Security Classification Policy (RA-02). Classification Process:
   i. The completed assessment must use the State of Idaho Risk Classification Tool and be approved by ITS and the agency or department’s designated IT Liaison.
   ii. The assessment must evaluate each of the risk factors identified in P.ITS-01, including GenAI-specific factors when applicable.
   iii. The assessment must include justification for each factor and identification of specific concerns.
   iv. The assessment must document the data classification levels of information processed by the system according to ITS policy Security Classification (RA-02).
   v. The risk classification must be reviewed annually and when significant changes occur to the system’s functionality, data sources, or operational context.
4. **Risk Management Alignment:** The risk assessment and management process must align with the four core functions of the NIST AI RMF:
   i. **Govern**: Establish and implement a governance structure that incorporates AI risk management into agency and department processes.
   ii. **Map**: Identify, analyze, and document context, capabilities, and potential risks of the AI system.
   iii. **Measure**: Assess and track AI risks and impacts using appropriate qualitative and quantitative tools.
   iv. **Manage**: Allocate resources to address and reduce AI risks throughout the system lifecycle.

CSF Core Function	Function Description	Closest AI RMF Function(s)	Rationale
Identify	Develop an organizational understanding to manage cybersecurity risk	Map, Govern	Both involve understanding systems, stakeholders, risk context, and governance structures
Protect	Develop and implement appropriate safeguards to ensure delivery of critical services	Manage, Govern	Focus on risk mitigation strategies and protective measures for AI systems
Detect	Develop and implement appropriate activities to identify the occurrence of a cybersecurity event	Measure, Manage	Relates to detecting anomalous or harmful AI behavior via measurement and monitoring
Respond	Develop and implement appropriate activities to take action regarding a detected cybersecurity incident	Manage	Addresses how to take action on identified risks, including incident response and adaptation
Recover	Develop and implement appropriate activities to maintain plans for resilience and to restore any capabilities or services that were impaired due to a cybersecurity incident	Manage, Govern	Supports continuity and system improvement through governance and resilience planning

5. **GenAI-specific Risk Factors:** For GenAI systems, additional risk factors must be evaluated:
    i. Potential for hallucinations or factual inaccuracies.
    ii. Vulnerability to prompt injection.
    iii. Data poisoning risks.
    iv. Copyright and intellectual property implications.
    v. Potential for misuse or unintended outputs.

(c) AI Data Management:

1. Data classified as:
    i. **Unrestricted Data (Classification Level 1):** May be processed with approved GenAI tools with appropriate oversight.
    ii. **Limited Data (Classification Level 2):** May be processed with approved GenAI tools with enhanced oversight and content review.
    iii. **Restricted Data (Classification Level 3):** May only be processed with securely deployed, enterprise-approved GenAI solutions with stringent safeguards.
    iv. **Critical Data (Classification Level 4):** Must not be processed with GenAI tools without explicit authorization for high-risk deployments, including AI Executive Committee review, and implementation of robust safeguards.
2. AI tools must have access to and use only data sources they need and respects users' rights and privacy.
3. Institute proper data and information management controls, procedures, and processes for data set selection, evaluation, and preparation for use with AI tools.
4. Data used to train AI systems will be screened for biases and corrected where necessary.
5. Data storage, transfer, and processing will follow all regulatory and internal data management standards.
6. Employees should be aware of when the use of an AI tool may result in the creation of a public record that must be retained under Public Records Law Idaho Code §§74-101 through 74-127.

(d) AI System Governance:

1. **Approval Requirements:** Idaho designates specific approval authorities that align with the appropriate level of oversight for each risk tier. The level of review and approval depends on the system's risk classification:
    i. **Tier 1: Low Risk** systems require ITS and agency (or department) IT leader approval. The AI Innovation Team receives notification for inventory purposes only.
    ii. **Tier 2: Medium Risk** systems require ITS and agency (or department) leadership approval with advisory input from the AI Innovation Team and Technical Review Board. Information Owners manage information sharing agreements and security controls as required by (RA-02).
    iii. **Tier 3: High Risk** systems require ITS and agency (or department) leadership approval with mandatory consultation from the Ethics Committee, Technical Review Board, and Executive Committee. Information Owners maintain enhanced oversight of classification and security controls for critical data as required by (RA-02).
2. These authorities exercise responsibility across key lifecycle activities aligned with the NIST AI RMF functions of GOVERN, MAP, MEASURE, and MANAGE. Idaho implements these functions through:
    i. **Initial Planning and Design (GOVERN, MAP):** ITS and agency (or department) sponsors with Information Owner classification guidance.
    ii. **Development and Testing (MAP, MEASURE):** Risk-appropriate authority approves resources with security validation.
    iii. **Technical Evaluation (MEASURE):** Technical Review Board verifies integration and proper data separation.
    iv. **Deployment Decision (MANAGE):** Risk-appropriate authority approves launch with security verification.
    v. **Monitoring and Evaluation (MANAGE):** Implementation teams oversee performance with periodic

reviews.

- **Recommended Documentation Requirements**: Documentation requirements scale with risk classification. All AI systems must maintain basic documentation, including purpose, data sources, and risk classification. Medium and high-risk systems require additional documentation, including privacy impact assessments, security controls, fairness evaluations, and human oversight mechanisms, where appropriate. Agencies and departments can obtain standard templates for required documentation from the AI Innovation Team.

vi. All AI Systems (Tier 1-3)

(A) AI Concept Brief documenting purpose, anticipated benefits, and alignment with agency or department mission.

(B) Risk Classification Assessment results and approval.

(C) Data sources, classification levels, and information handling procedures.

(D) System performance metrics and evaluation criteria.

(E) User guidance and support materials.

(F) AI Fact Sheet with standardized system information.

(G) Human oversight mechanisms and procedures.

(H) Specific responsibility assignments using the Shared Responsibility model.

(I) Any deviations from the standard responsibility framework with justification.

(J) Contact information for all roles.

(K) Approval signatures from both Agency (or Department) leadership and ITS.

vii. **Medium-Risk Systems (Tier 2)**: Systems classified as medium-risk must maintain all requirements from Section (d.3.i) plus:

(A) Privacy Impact Assessment addressing data collection, use, storage, and sharing.

(B) Security controls implementation, documenting technical and administrative safeguards.

(C) Fairness evaluation results demonstrating testing across demographic groups.

(D) Monitoring plan identifying performance indicators and review schedules.

(E) Vendor assessment documentation (for third-party systems).

viii. **High-Risk Systems (Tier 3)**: Systems classified as high-risk must maintain all requirements from Sections (d.3.i) and (d.3.ii) plus:

(A) Detailed model documentation including architecture, training methodology, and limitations.

(B) Enhanced security assessment documenting specialized AI security controls.

(C) Risk mitigation plan addressing identified concerns.

(D) Incident response procedures for AI-specific failure modes.

3. **Implementation Lifecycle:** AI implementations must follow a structured governance process throughout their lifecycle:

- **Ideation and Concept Development**: Agencies and departments submit solution vetting requests through standard intake processes. For AI capabilities, the intake system gathers additional information for the AI Concept Brief. The Information Owner validates classification decisions and governance teams conduct preliminary risk screening.

i. **Project Proposal and Risk Assessment:** ITS conducts a formal risk assessment in collaboration with the Information Owner, documenting classification levels, security controls, and data separation methods.

ii. **Implementation and Deployment:** Standard deployment processes incorporate additional verification of AI-specific controls. Governance and security teams validate privacy, fairness, and transparency requirements alongside classification-appropriate security controls.

4. **Monitoring and Continuous Improvement:** Standard monitoring processes incorporate AI-specific metrics from the Concept Brief. Information Owners conduct periodic classification reviews as required by ITS policy Security Classification (RA-02), with regular reassessment of risk classifications.

(e) **GenAI Requirements**: GenAI systems must adhere to additional requirements:

1. General Use Requirements:
   i. **ITS and Agency (or Department) Approval**: ITS, Agencies (or Departments, where appropriate) must explicitly approve specific GenAI tools for different use cases based on data classification levels:
      (A) **Unrestricted Data (Classification Level 1)**: May be processed with approved GenAI tools with appropriate oversight.
      (B) **Limited Data (Classification Level 2)**: May be processed with approved GenAI tools with enhanced oversight and content review.
      (C) **Restricted Data (Classification Level 3):** May only be processed with securely deployed, enterprise-approved GenAI solutions with stringent safeguards.
      (D) **Critical Data (Classification Level 4)**: Must not be processed with GenAI tools without explicit authorization for high-risk deployments, including AI Executive Committee review, and implementation of robust safeguards.
   ii. **Human Oversight**: All GenAI outputs used for official purposes must undergo human review before finalization or distribution.
   iii. **Content Identification:** AI-generated content must be clearly identified as such when distributed, both internally and externally.
   iv. **Tool Registration:** All GenAI tools must be registered in the agency or department's software inventory and the AI Innovation Team's AI system inventory.
   v. **Output Logging**: Agencies and Departments must maintain logs of significant AI-generated content, including prompts used and human review status.
2. **Acceptable Use Parameters**: Agencies and departments must establish clear parameters for appropriate and inappropriate use of GenAI. All GenAI applications must maintain human review of outputs used for official purposes.
3. **Acceptable Use Cases**: GenAI tools may be used for the following approved use cases:
   i. Content drafting and editing with appropriate human review, including first drafts of routine communications, summarization of non-sensitive documents, format conversion of public information, and language translation of appropriate content.
   ii. Research assistance with verified information, such as analysis of public documentation, suggestion of relevant policies or resources, or explaining complex topics for internal reference.
   iii. Code assistance and documentation, limited to generating code examples for public-facing applications, documenting existing code, and suggesting improvements to non-sensitive code.
   iv. Workflow optimization, including process documentation, meeting summarization (for non-sensitive meetings), and task organization and prioritization.
   v. Customer service enhancement, such as development of chatbot responses for public inquiries, creation of FAQ content, and generation of information materials.
4. **Prohibited Use Cases:** GenAI tools must not be used for:
   i. Fully autonomous decision-making affecting individual rights, benefits, or services.
   ii. Processing or generating content involved data classified as "Critical" under ITS policy Security Classification Policy (RA-02).
   iii. Generating content for official state communications without human review and verification.
   iv. Creating or processing legal documents, contracts, or official opinions without legal review.
   v. Unauthorized disclosure of sensitive information, including personally identifiable information, protected health information, financial account information, law enforcement or security information, or information exempt from public disclosure.
   vi. Impersonation of state officials or misrepresentation of agency or department positions.
   vii. Generation of deceptive content or deepfakes.
5. **Technical Safeguards:** All GenAI implementations must implement:
   i. Content filtering mechanisms to prevent harmful, biased, or inappropriate outputs.
   ii. Prompt management practices, including documented prompt libraries for approved uses

and monitoring and review of prompt effectiveness (e.g., prompt audit logging).

iii. Output review procedures for public-facing content.

iv. Logging and audit capabilities for all system interactions.

v. Security controls appropriate to the data classification level.

vi. Input validation to address potential prompt injection and other manipulation.

6. **Implementation Safeguards**: GenAI implementations must include content filtering, transparency measures, technical protections against misuse, and monitoring of outputs for quality and appropriateness.

7. **Transparency Requirements**: GenAI use must be transparent to both internal users and citizens:

8. **Public Disclosure**: When GenAI contributes to public-facing content, this must be disclosed in a clear and appropriate manner. Disclaimer language reads as follows:

> *"This (chatbot, content, system, etc.) uses artificial intelligence ("AI") for general information only and should not be a substitute for obtaining legal advice from a licensed attorney. Users should verify information before acting on it. The State of Idaho is not responsible for any errors or inaccuracies that may arise from the use of AI-generated content. Do not share sensitive information, such may be subject to the Idaho Public Records Act."*

9. **Internal Attribution**: Internal documents created or substantially assisted by GenAI must clearly indicate this in the document.

10. **Decision Support Documentation**: When GenAI is used to support decision-making, this must be documented with information about the system used and how outputs were validated.

(f) Privacy and Security:

1. **Privacy Requirements**: AI systems must adhere to Idaho's privacy framework. Systems processing personal information must include privacy impact assessments and implement appropriate safeguards throughout the AI lifecycle.

2. **Security Requirements**: AI systems require safeguards that address both their technical architecture and operational behavior. In accordance with ITS policy Security Classification (RA-02) and (S.MP-01c), Idaho mandates a set of baseline controls for all AI implementations, with additional requirements for systems operating at medium- or high-risk. Over time, these controls will evolve to include protection against emerging threats such as model poisoning, prompt injection, and inference manipulation—security considerations unique to contemporary AI implementations:

i. A Secure by Design Approach to AI Security:

- Security for AI systems builds on existing enterprise cybersecurity best practices, standards, and programs. Rather than creating entirely new processes, agencies and departments shall extend their current efforts to address AI-specific concerns. For all AI systems, agencies and departments must work with ITS to implement:

1. Access controls for system administration and operation.
2. Integrity verification for deployed models.
3. Input validation and sanitization.
4. Basic monitoring of system access and operations.
5. Integration with enterprise security infrastructure.

(B) Security Requirements by Risk Tier:

1. For Tier 1 (Low-Risk) Systems:

I. Apply standard IT security controls.

II. Implement basic access controls.

III. Validate user inputs.

IV. Maintain system logs.

2. For Tier 2 (Medium-Risk) Systems: All Tier 1 measures, plus:

I. Enhance access controls with stronger authentication.

II. Implement more robust monitoring.

III. Conduct periodic security testing.
IV. Develop specific incident response procedures.

3. For Tier 3 (High-Risk) Systems: All Tier 2 measures, plus:
I. Implement advanced security controls.
II. Conduct regular penetration testing.
III. Develop comprehensive monitoring and alerting.
IV. Establish specialized incident response plans.

3. **Third-Party Requirements**: When acquiring AI capabilities from vendors, agencies and departments remain responsible for verifying compliance state security and privacy standards and implementing compensating controls where necessary. Specific vendor assessment procedures and contracting requirements are detailed in (SA-04).

(g) **AI Procurement**: All AI systems (including pilots or free tools) must be reviewed and approved via ITS processes. Agencies and departments shall adhere to specific guidelines when procuring AI systems or services:

1. **Pre-Procurement Vendor Assessment**: Before procurement, agencies and departments must:
i. Evaluate vendor AI governance and risk management practices.
ii. Verify adherence to relevant security and privacy standards.
iii. Assess model development methodologies and testing procedures.
iv. Review known limitations, biases, or ethical concerns.
v. Verify vendor compliance with state security and privacy standards.
vi. Obtain documentation for model development and training methods, data sources and privacy protections, security testing procedures and results, and known limitations or biases.
vii. Assess alignment with state AI ethical principles.
viii. Review vendor incident history and remediation practices.
ix. Complete the standardized Solution Security Questionnaire.

2. **Contractual Requirements**: AI procurement contracts must include:
i. Specific security and privacy requirements aligned with system risk classification.
ii. Clear delineation of data ownership and usage rights.
iii. Performance metrics and quality standards.
iv. Incident reporting and response obligations.
v. Compliance verification and audit provisions.
vi. Requirements for ongoing support and updates.
vii. Monitoring and reporting obligations.
viii. Audit rights for system review and compliance verification.
ix. Warranty provisions for system performance and security.
x. Requirements for ongoing updates and maintenance.
xi. Transparency requirements regarding system capabilities and limitations.

3. **Ongoing Vendor Management**: After procurement, agencies and departments must:
i. Monitor vendor compliance with contractual requirements.
ii. Document and track system updates and changes.
iii. Partner with ITS to periodically reassess system risk classification.
iv. Maintain appropriate vendor management documentation.
v. Regular review of vendor performance and compliance.
vi. Reassessment of risk classification, where appropriate.
vii. Validation of continued security compliance.
viii. Monitoring for reported issues or incidents with the vendor’s AI systems.

(h) Transparency and Fairness:

1. Transparency Requirements:
i. AI systems must provide appropriate transparency to both users and those affected by system decisions:

- **AI Use Disclosure:** Clear notification when AI systems are being used, explanation of system capabilities and limitations, and attribution of AI-generated content.
- **Explanation Capabilities:** Documentation of system purpose and operation, plan language descriptions of how decisions are made, and explanation of key factors influencing outcomes.

2. Fairness Testing Requirements:
    i. To ensure AI systems serve all Idahoans, without bias, agencies and departments must:

- Identify relevant demographic categories for fairness analysis based on system purpose and affected population, historical patterns of disparate impact in similar contexts, and legal and regulatory requirements.
- Establish appropriate fairness metrics, including statistical parity across groups where appropriate, equal error rates across population segments, and consistency of decisions for similar cases.
- Document fairness testing methodologies and results, including testing datasets used and their representatives, statistical methods applied, and limitations of the analysis.
- Implement mitigation strategies for identified biases, including model adjustments or constraints, procedural safeguards and human review, and monitoring mechanisms for ongoing assessment.

(i) AI System Monitoring and Maintenance:

1. **Continuous Monitoring Requirements:** Agencies and departments must implement continuous monitoring of AI systems based on risk tier:
    i. **Tier 1 Systems:** Quarterly performance review, annual risk reassessment, and operational metrics tracking.
    ii. **Tier 2 Systems:** Monthly performance review, semi-annual risk reassessment, regular bias and fairness testing, and security and access control verification.
    iii. **Tier 3 Systems:** Weekly performance monitoring, quarterly risk reassessment, comprehensive bias and fairness testing, proactive security testing.
2. **Maintenance Procedures:** All AI systems require documented maintenance procedures:
    i. **Version Control:** Documentation of all system changes, testing requirements before implementation, and approvals appropriate to risk tier.
    ii. **Performance Optimization:** Regular model retraining or updating, data quality verification, and effectiveness evaluation.
    iii. **Documentation Updates:** Maintenance of current system documentation, user guidance revisions, and training material updates.

(j) **Incident Response:** Agencies and departments must build upon existing incident response processes to include relevant AI-specific incident types, such as hallucinated outputs, misclassification of inputs, inference bias, or inappropriate content generation.

(k) Training and Awareness:

1. **Required Training Programs:** Agencies and departments must implement AI training programs:
    i. **Executive-level Training:** AI governance principles, risk management approach, and strategic considerations.
    ii. **IT Liaison Training:** Implementation procedures, risk assessment methodology, documentation requirements, monitoring, and oversight responsibilities.
    iii. **User-Level Training:** Appropriate AI use cases, security and privacy considerations, oversight responsibilities, and incident reporting procedures.
    iv. **Technical Implementation Training:** Security controls implementation, monitoring techniques, bias detection and mitigation, and system maintenance procedures.

- **Specialized Training:** Additional specialized training is recommended for specific roles:
    v. **High-Risk System Operators:** Advanced risk management, fairness and bias mitigation, adversarial testing, and ethical considerations.
    vi. **AI Procurement Specialists:** Vendor assessment procedures, contractual requirements, and ongoing vendor management.

vii. **GenAI Users**: Prompt engineering best practices, content review procedures, proper attribution, and disclosure.

(l) **AI Shared Responsibility Model**: The State of Idaho adopts a shared responsibility model for AI governance to clearly delineate the roles and responsibilities between ITS and individual agencies or departments:

1. **ITS as the Foundation:** ITS serves as the central hub for AI governance, providing the foundation that agencies and departments can build upon. ITS shall be responsible for:
    - i. Establishing the statewide AI governance framework, policies, and standards.
    - ii. Developing risk assessment methodologies and assigning risk scores.
    - iii. Approving high-risk implementations in coordination with AI Executive Committee review.
    - iv. Maintaining the state's AI system inventory.
    - v. Coordinating enterprise-wide AI initiatives and shared services.
    - vi. Developing standardized templates and assessment tools.
    - vii. Providing technical consultation and advisory services.
2. **Agencies and Departments as the Implementers**: Agencies and Departments shall be responsible for:
    - i. Partnering with ITS to support risk assessments for their proposed AI implementations.
    - ii. Developing agency or department-specific use cases and implementation plans.
    - iii. Ensuring compliance with ITS policies and standards.
    - iv. Maintaining appropriate documentation for AI systems.
    - v. Implementing required security and privacy controls in partnership with ITS.
    - vi. Conducting ongoing monitoring and assessment of agency and department AI systems.
    - vii. Managing vendor relationships for agency and department-specific AI implementations.
    - viii. Ensuring proper training of agency and department personnel on responsible AI use.
3. **Shared Responsibility Implementation**: The shared responsibility model defines how ITS, agencies and departments work together to implement AI governance. Appendix A outlines specific information on governance, risk management, implementation, and privacy and security activities, including who leads each activity and who assists (or provides input).

4.0 AI Guidelines

4.1 (G.ITS-01) AI Governance Guideline

This guideline assists agencies and departments in implementing AI systems in accordance with P.ITS-01 (AI Policies) and S.ITS-01 (AI Standards). It provides recommended approaches for completing risk assessments, documenting systems, implementing technical safeguards, and managing AI throughout its lifecycle.

By following consistent implementation practices, agencies and departments can more effectively manage risks associated with AI and at the same time, realize the benefits these technologies offer. This guideline complements existing security, privacy, and information management frameworks to address the unique considerations introduced by AI systems.

4.2 AI Governance and Organizational Structure

Establishing AI Governance: To implement effective AI governance, agencies and departments should:

- **Designate Leadership and Accountability**: Appoint an IT Liaison responsible for overseeing the agency or department’s AI implementations, establish clear roles and responsibilities for AI oversight and risk management, and create cross-functional teams including IT, business, legal, security, and privacy stakeholders.
- **Develop Agency and Department-Specific Policies**: Adapt Idaho’s AI Framework to agency and department-specific contexts, document acceptable use cases and prohibited applications, and establish approval workflows aligned with risk tiers.
- **Implement the “AI Ambassador” Model**: Identify enthusiastic and knowledgeable staff to serve as AI ambassadors within their teams, provide ambassadors with enhanced training on responsible AI use, and leverage ambassadors to spread knowledge and best practices across the organization.

Organizational Readiness Assessment: Before implementing AI systems, agencies and departments should assess their readiness:

- **Skills and Knowledge Evaluation:** Inventory existing AI-related skills and experience, identify training needs and knowledge gaps, and develop plans for building necessary capabilities.
- **Technical Infrastructure Assessment:** Evaluate existing infrastructure to support AI implementations, identify necessary upgrades or enhancements, and ensure compliance with security and accessibility requirements.
- **Cultural Readiness:** Assess organizational attitude towards AI adoption, identify potential resistance or concerns, and develop change management strategies.

4.3 Risk Assessment Methodology

Risk assessment doesn't have to be complex to be effective. The goal is to identify potential concerns early so you can address them appropriately.

A Straightforward Risk Assessment Process: Risk assessment for AI should be practical and proportionate. Follow these steps for an effective approach:

- **Bring the right people together:** Effective risk management isn't a solo activity. Include people who understand the business purpose, the data involved, the technical aspects, and the potential impacts on users. For small implementations, this might just be two to three people; for larger ones, you might need broader representation.
- **Understand what you’re assessing:** Before evaluating risks, make sure everyone has a shared

understanding of what the AI system will do, what data it will use, and how it will be deployed. This common foundation helps ensure you're all assessing the same thing.

- **Evaluate the key risk factors:** Consider each of the risk factors identified in the policy, but don't get bogged down in excessive detail. Focus on understanding:
- What kind of data will the system use, and how sensitive is it?
- How will the system directly affect decisions about individuals?
- What level of human oversight will exist in practice?
- How understandable will the system's operations be to users and stakeholders?
- How many people will the system affect, and for how long?
- How novel or complex is the technology being used?
- **Document your thinking:** Work with ITS to capture your risk ratings and the reasoning behind them. Brief explanations for each rating help others understand your perspective and provide valuable context for future reviews. For more information about how ITS evaluates risk scoring, you can speak with an ITS governance representative about the risk scoring tool.
- **Partner with ITS to support its work to determine the overall risk tier:** Based on your evaluation, partner with ITS to support its work to determine whether the system falls into Tier 1 (Low-Risk), Tier 2 (Medium-Risk), or Tier 3 (High-Risk). When in doubt, consider rounding up to the highest tier, especially for systems that affect individual rights or benefits.
- **Plan for reassessment:** Risk isn't static. Schedule reviews based on the risk tier--more frequent for higher-risk systems--and be prepared to reassess whenever significant changes occur to the system or its context.

4.3.1 Making Sense of Information Classification

Your AI system's risk tier should align sensibly with the classification level of the information it processes. Here's how to approach this alignment:

For Unrestricted Information (Classification Level 1):
These systems typically fall into tier one if they have limited decision impact and good human oversight. However, even systems using only public information can move to higher risk tiers if they have significant decision impact or limited oversight.

For Limited Information (Classification Level 2):
Depending upon other factors, these systems may fall into Tier 1 or Tier 2. Consider the decision impact and autonomy level carefully—systems with moderate impact or reduced oversight should generally be Tier 2.

For Restricted Information (Classification Level 3):
These systems will typically be at least Tier 2 (Medium-Risk) and maybe Tier 3 if they have significant decision impact or substantial autonomy. The sensitive nature of the data requires additional safeguards.

For Critical Information (Classification Level 4):
These systems will almost always be Tier 3 (High-Risk) and require the highest level of oversight and safeguards. The critical nature of the data combined with AI capabilities presents significant risks that must be carefully managed.

4.3.2 Learning from Examples

Risk assessments become clearer with examples. Here are some typical scenarios and their likely classifications.

4.3.2.1 Likely Tier 1 (Low-Risk) Examples:

- A document summarization tool for internal use processing only public documents.
- The translation tool for public website content with human review.
- An internal knowledge search tool using agency or department-specific information.

The common thread: These systems use non-sensitive information, have minimal direct impact on individuals, and maintain solid human oversight.

4.3.2.2 Likely Tier 2 (Medium-Risk) Examples:

- A system predicting maintenance needs for critical infrastructure.
- A recommendation engine suggesting services to citizens.
- A pattern recognition system supporting eligibility decisions.

What places these in Tier 2: They have more significant potential impacts, may process some sensitive information, or have reduced (but still present) human oversight.

4.3.2.3 Likely Tier 3 (High-Risk) Examples:

- A system determining benefit eligibility.
- An AI processing health information for treatment recommendations.
- A predictive system directly affecting resource allocation to individuals.

The key factors: These systems process highly sensitive information, have direct and significant impacts on individuals, or operate with substantial autonomy in critical domains.

4.4 System Documentation

Documentation shouldn't be busywork; it should serve a practical purpose in helping you implement and manage AI effectively.

4.4.1 Creating a Useful AI Concept Brief

The AI Concept brief is your starting point for any AI implementation. It is a short, strategic document that captures the essential “what” and “why” before diving into detailed planning.

A good concept brief answers these key questions:

What problem are we trying to solve?

- Clearly state the business challenge or opportunity that AI could address. Be specific about the current pain points or opportunities that motivated this initiative.

Why is AI a good approach for this?

- Explain why AI capabilities are well-suited for this particular problem, as opposed to traditional software or manual processes. What specific AI capabilities (like natural language processing, image recognition, or predictive analytics) make it appropriate?

Who will this serve, and how?

- Identify the primary users or beneficiaries and how the AI system will improve their experience or outcomes. This helps maintain focus on the real people who will interact with or be affected by the system.

What data will we need?

- Outline the major data sources that will be needed, noting any potential challenges with data availability, quality, or sensitivity. This helps identify potential hurdles early.

How will humans stay in the loop?

- Describe how human oversight and intervention will be maintained. Be realistic about the level of human involvement that will occur in practice.

The Concept Brief shouldn't be lengthy, typically two to three pages is sufficient. Its purpose is to align stakeholders on the basic direction before investing in detailed planning.

4.4.2 Creating an Informative AI Fact Sheet

Once your AI system is better defined, an AI Fact Sheet improves transparency and provides a standardized way to document essential information about the system.

A good fact sheet includes:

System basics:

- Name and version of the system
- Primary purpose and functions
- Owner and responsible parties
- Risk classification tier

Technical Information:

- AI technologies used
- Performance metrics
- Limitations and constraints
- Data types used

Governance Information:

- Approval status
- Review schedules
- Human oversight mechanisms
- Feedback channels

Keep your Fact Sheet concise, typically one to two pages, but make sure it covers the essential information someone would need to understand what the system does and how it's governed.

4.4.3 Documentation That Works

Effective documentation shares these characteristics:

- **It's concise** -- Focus on what people need to know, not exhaustive detail for its own sake. If it's too long, it won't be read or maintained.
- **It's accessible** – Write for both technical and non-technical audiences, avoiding unnecessary jargon. A good test: Could someone new to the project understand it?
- **It's maintained** -- Documentation that's out of date is worse than useless, it's misleading. Build in periodic review and updates, particularly after significant changes.
- **It's purposeful** -- Each document should serve a clear purpose and objective. Ask yourself: Who needs this information and why?

For GenAI systems, be sure to also document:

- **Prompt management** -- How are prompts developed, tested, and approved?
- **Content filtering** -- What measures prevent inappropriate outputs?
- **Review workflows** -- How are outputs reviewed before use?
- **Attribution procedures** -- How is AI-generated content identified?

Remember, the goal of documentation isn't compliance for its own sake, it's creating a shared understanding that enables effective governance and use of AI technology.

4.5 Using Third-Party AI Tools

Determining Appropriate Use: When considering whether to use third-party tools like ChatGPT, Claude, or Copilot Chat, Idaho state employees should evaluate:

- **Purpose:** Is its intended use aligned with job responsibilities and agency (or department) mission?
- **Data Sensitivity:** What classification level of data would be involved?
    - Classification Level 1 (Unrestricted) data: Generally acceptable to use with appropriate safeguards.
    - Classification Level 2 (Limited) data: Requires caution and additional review.
    - Classification Level 3 (Restricted) data: Should only use agency- and department-approved secure AI tools.
    - Classification Level 4 (Critical) data: Not appropriate for third-party AI tools.
- **Agency (or Department) Policy:** Has your agency (or department) approved the specific tool for this purpose?
- **Human Review:** Will outputs be reviewed before use in official communications?
- **Transparency:** Will AI involvement be disclosed appropriately?

4.5.1 Decision Framework

Use this simplified decision framework to determine if your intended use is appropriate:

- Is the information public?
  If NO → Do not use public third-party AI tools
  If YES → Continue evaluation
- Is the tool properly approved by your agency (or department)?
  If NO → Request approval or use an approved alternative
  If YES → Continue evaluation
- Will you review the output before using it officially?
  If NO → Reconsider use or implement review process
  If YES → Continue evaluation
- Will the use be disclosed appropriately?
  If NO → Implement proper disclosure
  If YES → Proceed with appropriate safeguards

4.5.2 Appropriate Use Examples

Third-party GenAI tools like Claude, Copilot Chat, or ChatGPT may be appropriate for:

- **Document Summarization:** Summarizing public meeting minutes, published reports, or other public documents for internal reference. Example: A policy analyst asks ChatGPT to summarize key points from a 50-page public report on infrastructure needs. The analyst reviews the summary for accuracy before using it in an internal briefing.

- **Content Drafting**: Creating first drafts of routine communications that will undergo human review and editing. Example: An administrative assistant uses Copilot Chat to help draft a first version of a public announcement about an upcoming community event. The draft is edited and approved by the communications team before distribution.
- **Format Conversion**: Reformatting public information for different communication channels. Example: A web content manager asked Claude to help convert technical documentation into a more accessible FAQ format. The result is verified for accuracy before publishing.
- **Research Assistance**: Finding and organizing publicly available information. Example: A researcher uses Claude to help identify relevant studies on a topic, then verifies the information through official sources before including it in a report.
- **Code Examples**: Generating basic code samples for non-sensitive applications. Example: A developer asks ChatGPT for example code to parse CSV files in Python. The developer reviews and tests the code before implementing it.
- **Meeting Summarization**: Creating summaries of non-sensitive meetings or public hearings. Example: An administrator uses Copilot Chat to summarize points from a recorded public meeting, then reviews the summary for accuracy before distributing to staff.
- **Language Translation**: Translating public documents or communications. Example: A communications specialist uses ChatGPT to create an initial translation of public materials into Spanish, then has the translation verified by a fluent speaker before publication.

4.5.3 Inappropriate Use Examples

Third-party GenAI tools should **not** be used for:

- **Sensitive Data Processing**: Uploading or discussing confidential, restricted, or critical information. Example: Pasting internal financial projections or personnel information into Claude to summarize or analyze would be inappropriate.
- **Decision-making**: Relying on AI to make or directly inform decisions affecting individuals. Example: Using ChatGPT to evaluate eligibility for a state program or to determine resource allocation without substantial human review would be inappropriate.
- **Legal or Policy Guidance**: Generating authoritative legal opinions or policy interpretations. Example: Asking Copilot Chat to interpret state laws, draft legally binding documents, or create official policy statements without legal review would be inappropriate.
- **Unreviewed Public Communications**: Publishing AI-generated content without human verification. Example: Directly publishing AI-generated newsletters, announcements, or social media posts without review would be inappropriate.
- **Sharing Personal Information**: Inputting any personally identifiable information about citizens or employees. Example: Uploading citizen correspondence with personal details for analysis or response drafting would be inappropriate.
- **Security Sensitive Information**: Sharing information about security systems, vulnerabilities, or access controls. Example: Asking Claude to review or suggest improvements to network security configurations would be inappropriate.
- **Administrative Credentials**: Sharing login information or system access details. Example: Including system access instructions with usernames or password patterns would be inappropriate.

4.5.4 Best Practices for Using Third-Party AI Tools

When using approved third-party AI tools, follow these best practices:

- **Be transparent**: Clearly identify when content has been generated or assisted by AI. Refer to public disclosure notice for vetted disclaimer language that ensures transparency.

- **Verify information**: Always fact-check AI-generated content against reliable sources. Best practice: Cross-reference information from official state resources, authoritative sources, or subject matter experts.
- **Review Thoroughly**: Carefully review all AI-generated content before using it officially. Best practice: Look for factual errors, inappropriate phrasing, biased language, or incomplete information.
- **Maintain Ownership**: Remember that you are responsible for all content used in your official capacity. Best practice: Treat AI as a drafting tool, not a replacement for your professional judgment and accountability.
- **Respect Boundaries**: Do not share sensitive, confidential, or private information. Best practice: Only use publicly available information that would be appropriate to share in an open forum.
- **Use Professional Prompts**: Write clear, professional prompts that specify the parameters of your request. Best practice: "Please help me draft a public announcement for the upcoming road maintenance on State Highway 55. Include the dates July 1-15, 2025, working hours of 9 AM to 3 PM, and note that one lane will remain open with flaggers present."
- **Keep Records**: Maintain appropriate documentation of significant AI use. Best practice: Save your prompts and the AI-generated responses for important work products.
- **Use Agency (or Department) Templates:** Whenever possible, direct AI to conform to existing agency or department templates and style guides. Best practice: "Please follow the State of Idaho press release format as described in the following template …"
- **Apply Content Warnings**: Add disclaimers to preliminary AI-generated materials. Best practice: Adding "DRAFT – AI ASSISTED – REQUIRES REVIEW" to documents during development.

4.6 Common Questions and Answers

Q: Can I use ChatGPT, Copilot, or Claude to summarize my meeting notes?
A: Yes, if the meeting notes contain only public or Classification Level 1 information. You should review the summary for accuracy and appropriate characterization of the discussion. Do not use these tools to summarize meetings that contain sensitive, confidential, or personal information.

Q: Can I ask AI tools to draft emails for me?
A: Yes, for routine communications that that don't contain sensitive information. Always review and edit AI-drafted emails before sending them.

Q: How do I know if information is appropriate to share with a third-party AI tool?
A: Consider whether the information would be appropriate to share publicly on the Internet. If you would not be comfortable sharing it on a public website, it is not appropriate for third-party AI tools. Refer to (RA-02) for information classification guidance.

Q: Does my agency (or department) need to approve specific AI tools before I use them?
A: Yes, each agency (or department) should maintain a list of approved AI tools and their permitted uses. If you wish to use a tool that is not yet approved, consult with your agency or department's IT Liaison or submit a software vetting request, where appropriate.

Q: Do I need to disclose that I used AI to help create content?
A: Yes, transparency is important. Appropriate disclosure might be as simple as noting "AI-assisted" or "Draft generated with AI assistance and reviewed by [name]" on the document.

Q: Can I use AI tools to help improve my writing style?
A: Yes, AI tools can be helpful for improving writing clarity and style for public information. Always review the suggestions and make your own judgment about what changes to accept.

Q: How should I report problems with AI-generated content?

A: If you encounter biased, inappropriate, or concerning outputs from AI tools, report the issue to your agency or department’s IT Liaison and document the prompt and response.

4.7 Privacy and Security Implementation

A Practical Approach to Privacy:
Protecting privacy in AI systems doesn’t require reinventing the wheel. Instead, build upon existing privacy practices with a focused approach to AI-specific challenges.

Start with understanding your data flows:
Begin by mapping how personal information moves through your AI system. Where does it come from? Where is it stored? Who can access it? How long is it kept? This basic understanding forms the foundation for effective privacy protection.

The goal isn't to create complex data flow diagrams, but rather to gain practical knowledge of how data moves so you can identify where privacy protections are needed.

Apply data minimization thoughtfully:
Ask three practical questions about each piece of personal information: Do we really need this to accomplish our goal? Could we use anonymized or aggregated data instead? How long do we need to keep this information?

Less data means less risk. Only collect what you need, only keep it as long as necessary, and only share it when required.

Make privacy part of your governance process:
Designate clear ownership for privacy within your team. This could be as simple as assigning privacy responsibilities to your IT Liaison or involving your agency or department's existing privacy officer.

Build simple checkpoints into your AI development process to ensure privacy is considered before moving forward with implementation.

For GenAI, focus on specific risks:
Pay special attention to these key risks with GenAI: inadvertent disclosure of sensitive information and outputs; generation of content that appears to be about real individuals; collection of unnecessary personal information through prompts.

Implement basic guardrails like content filtering and prompt restrictions to mitigate these risks.

4.7.1 Secure-by-Design

Security for AI systems works best as an ongoing process, not a one-time checklist.

Integrate with what you already have:
Connect your AI security efforts with your existing security program. Your current security team already has expertise that applies to AI, leverage it rather than creating separate processes.

Use your existing security monitoring, vulnerability management, and incident response processes, extending them to cover AI-specific concerns.

Focus on these key steps:

- **Identify AI-specific risks:** Beyond traditional IT security concerns, consider risks like prompt injection, data poisoning, and model manipulation.

- **Apply targeted controls**: Implement controls that address the specific risks of your AI system. For low-risk systems, standard IT security controls may be sufficient. For higher-risk systems, add AI-specific protections.
- **Test before deploying**: Test security controls before going live. For low-risk systems, basic functional testing may be enough. For higher-risk systems, consider more thorough testing, including adversarial testing when appropriate.
- **Monitor continuously**: Watch for unusual patterns in both system behavior and user interactions. Set alerts for anomalies that might indicate security issues.
- **Improve based on experience**: Use what you learn from operations and any incidents to continually improve your security approach.

For GenAI, pay special attention to these areas: implement content filtering appropriate to your use case; establish clear boundaries and prompts to prevent manipulation; monitor outputs for unexpected or inappropriate content; create a simple process for users to report concerns.

This process-based approach ensures security evolves alongside your AI implementation, providing appropriate protection without unnecessary complexity.

4.8 Fairness and Accessibility

4.8.1 Fairness Testing Methodology

To ensure AI systems serve all Idahoans, agencies and departments should:

- **Identify relevant demographic categories for fairness analysis based on** system purpose and affected population; legal and regulatory requirements; and community feedback and concerns.
- **Establish appropriate fairness metrics** (e.g., statistical parity across groups where appropriate; equal error rates across population segments; consistency of decisions for similar cases; and representation in training and testing data).
- **Test for both direct and indirect bias** (e.g., direct discrimination based on protected characteristics; amplification of existing societal biases; and underrepresentation in training data).
- **Document fairness testing methodologies and results** (e.g., testing datasets used and their representativeness; statistical methods applied; limitations of the analysis; and ongoing monitoring plans).
- **Implement mitigation strategies for identified biases** (e.g., model adjustments or constraints; procedural safeguards and human review; monitoring mechanisms for ongoing assessment; and feedback channels for affected individuals).

4.8.2 Accessibility considerations

To ensure AI systems are accessible to all users, agencies and departments should:

- **Apply universal design principles from the outset** (e.g., design interfaces that work with assistive technologies; ensure AI-generated content meets accessibility standards; test with users having diverse abilities; and follow WCAG guidelines).
- **Provide alternative access methods** (e.g., non-AI alternatives for critical services; multiple interaction channels (voice, text, etc.); human assistance options when needed; and offline alternatives).
- **Consider cognitive accessibility** (e.g., use plain language in user interfaces; provide explanations appropriate to various levels of technical understanding; test comprehension with diverse user groups; and offer simplified versions where appropriate).
- **Ensure equitable service** (e.g., address potential digital divide issues; consider varied levels of technological literacy; provide appropriate accommodations; and verify equitable performance

across user groups).

4.9 GenAI Implementation

GenAI offers exciting possibilities but requires thoughtful implementation. Here's how to make the most of these technologies and manage their unique challenges.

4.9.1 Practical Prompt Engineering

Prompts are the key to getting good results from GenAI. Think of prompt engineering as a skill that improves with practice and refinement.

Start with clear objectives. Before writing prompts, be clear about what you're trying to accomplish. Are you summarizing content? Drafting a response? Analyzing information? The more specific your goal, the better your prompt can be.

Build prompts iteratively. If you are just getting started with this technology, good prompts rarely emerge fully formed on the first try. Start with a basic version, see what you get, and refine based on results. Keep notes on what works and what doesn't.

Include the right context. Providing background information helps the AI understand what you need. Include relevant facts, constraints, or examples that illustrate what you're looking for.

Set clear parameters. Specify the format, length, tone, and style you want. For example, "Draft a brief update for citizens about the new permitting process. Keep it under 250 words, use plain language, and include the three key deadlines."

Maintain a prompt library. Create a collection of effective prompts for common tasks. This saves time and ensures consistency. Share these within your agency or department to spread best practices.

Good prompt example:
"Summarize the attached meeting minutes from the March [insert agency/department/committee] meeting. Focus on the decisions made, action items assigned, and upcoming deadlines. Format the summary with clear headings and keep it under one page."

4.9.2 Managing Content Effectively

GenAI creates content quickly, which requires thoughtful management processes.

Establish clear review workflows. Define who reviews what content and how thoroughly, based on its use and risk. Low-risk internal content might need only a quick check, while citizen-facing content requires more thorough review.

A simple workflow might look like:

- AI generates an initial draft
- Subject matter expert reviews for factual accuracy
- Communications staff reviews for style and clarity
- Final approver signs off before publication

Implement appropriate content filtering. Content filtering helps prevent inappropriate outputs. Most AI platforms have built-in filtering. Make sure it's enabled and configured appropriately for your use case.

Create feedback loops. Establish a simple way for users to report problematic content. Use this feedback to improve prompts and review processes.

Be transparent about AI use. When sharing content externally, be clear when AI has been involved in its creation. This builds trust and sets appropriate expectations. A simple note like "Draft assisted by AI and reviewed by [name]" is often sufficient.

Document your approach. Maintain basic documentation about how you're using GenAI, including:

- Approved use cases
- Review procedures
- Attribution requirements
- Responsibility assignments
- Learning and Improving

<u>GenAI is still evolving. Approach it with a learning mindset.</u>

Start small and expand gradually. Begin with lower-risk use cases to build experience and confidence before tackling more complex applications.

Share what works. Participate in the state's AI community of practice to share experiences and learn from others.

Revisit and refine. Periodically review your use of GenAI to identify improvements and address any emerging concerns.

Remember that GenAI is a tool to augment human capabilities, not replace human judgment. The most successful implementations maintain appropriate human oversight and leverage AI's unique strengths.

4.9.3 Real-World Implementation Examples

Nothing illustrates effective AI governance like actual examples. Here are three hypothetical but realistic scenarios showing how agencies and departments can implement AI systems at different risk levels.

<u>Public Information Chatbot (Tier 1)</u>

<u>The Scenario</u>:
A small regulatory agency wanted to implement a chatbot to answer common questions about licensing requirements. The goal was to reduce call volume and improve citizen access to information.

<u>The Approach</u>:
The agency took a thoughtful but streamlined approach to this low-risk implementation:

First, they discussed the concept with ITS and submitted a Solution Vetting request, later confirming through the process it would likely be a Tier 1 system since it would only access and share public information with significant human oversight.

They designated their communications director as the IT Liaison and assembled a small team, including a program specialist and an IT staff member. This team partnered with ITS to document the system's purpose, data sources, and oversight approach in a simple Concept Brief.

Rather than trying to build a comprehensive solution immediately, they started with the 20 most common questions, creating a prompt library with approved responses. The IT staff implemented basic content filtering and established a weekly review process for chatbot interactions.

Key Governance Elements:

- Clear ownership: The communications director was accountable for content accuracy
  - Appropriate oversight: Regular reviews of all chatbot interactions
- Manageable scope: Starting small allowed for learning and adjustment
- Transparency: Clear labeling of the chatbot as AI-assisted with human review

This approach allowed the agency to successfully implement a valuable service without unnecessary complexity in governance.

Regulatory Document Analysis Tool (Tier 2)

The Scenario:
A medium-sized agency needed to analyze thousands of regulatory documents to identify inconsistencies and suggest improvements. The system would process sensitive information and make recommendations that would influence regulatory updates.

The Approach:
The agency recognized this as a Tier 2 implementation due to the sensitivity of the information and potential impact on regulations.

They formed a cross-functional team including legal, IT, and program staff, and worked with ITS early in the planning process. Together, they conducted a thorough risk assessment as part of the solution vetting process, documenting key concerns and mitigation strategies.

They established a phased implementation plan:

- First, the system analyzed a small batch of documents with extensive human verification
- After confirming accuracy, they expanded to more document types
- They maintained a “human in the loop” model, with staff reviewing all recommendations before use

The agency documented responsibility assignments using a shared responsibility model, clearly identifying who would review outputs, make decisions based on the analysis, and monitor system performance.

Key Governance Elements:

- Collaborative planning: Early involvement of ITS and diverse agency perspectives
- Incremental rollout: Starting small and scaling up based on performance
- Shared responsibilities: Clear documentation of who does what
- Ongoing review: Regular assessment of system outputs and performance

This methodical approach enabled the agency to harness AI capabilities and maintain appropriate governance and oversight.

Benefits Eligibility Support System (Tier 3)

The Scenario:
A large agency implemented an AI system to support decisions about benefits eligibility for vulnerable populations. The system processed confidential personal information and directly affected individuals’ access to essential services.

The Approach:
The agency correctly identified this as a Tier 3 implementation requiring the highest level of governance and oversight.

They engaged extensively with ITS, the Technical Review Board, and the Ethics Advisory Committee from the beginning, working through a comprehensive governance process. The AI Executive Committee also reviewed the implementation plan.

The agency established multiple layers of oversight:

- Technical team monitoring system performance and security
- Program specialists reviewing individual recommendations
- Management team assessing broader patterns and impacts
- Independent testers evaluating for potential bias

They created a detailed shared responsibility matrix documenting responsibilities across teams, with clear escalation paths for concerns. The agency leadership remained directly accountable for system outcomes.

Perhaps most importantly, they established a robust appeals process ensuring individuals could request human review of any decisions influenced by the system.

Key Governance Elements:

- **Executive oversight**: High-level accountability and engagement
- **Multiple review layers:** Different perspectives checking different aspects
- **Clear documentation:** Comprehensive but usable documentation of responsibilities
- **Human safety net:** Appeals process ensuring recourse for affected individuals

This comprehensive governance approach allowed the agency to leverage AI for a high-impact application and maintain appropriate safeguards for the vulnerable populations it served.

Common Success Factors
Across all three examples, several factors contributed to successful implementation:

- **Appropriate scaling of governance:** Each agency aligned governance complexity with actual risk
- **Clear ownership:** Specific individuals were accountable for system performance and outputs
- **Incremental approach:** All started with limited scope before expanding
- **Learning mindset:** Each agency viewed implementation as a learning process
- **Collaboration:** Effective partnerships within agencies and with ITS

4.10 Making Shared Responsibility Work in Practice

Shared responsibility isn't just a framework; it's about people working together effectively. This section provides practical guidance on making the model work in your agency or department.

4.10.1 Practical Steps for Making This Work

Start with clear designation of roles:
Identify your agency or department’s IT Liaison who will serve as the primary contact for AI initiatives. This person doesn't need to be an AI expert but should be good at coordination and have enough authority to get things done.

For significant AI implementations, consider forming a small committee with representatives from business units, IT, legal, and privacy and security teams. Keep it small enough to be effective, usually three to five people is ideal.

Establish communication channels with ITS:
Regular check-ins prevent problems before they start. For low-risk systems, quarterly updates may be sufficient. For higher-risk systems, more frequent communication helps ensure alignment.

Join the statewide AI community of practice to share experiences and learn from other agencies and departments. This forum helps surface common challenges and solutions.

Develop right-sized processes:
Start simple and build as needed. For your first AI implementation, focus on the basics: document the purpose, assess risks, and establish appropriate oversight.

Use templates from ITS to streamline documentation. These templates are designed to capture essential information without creating unnecessary bureaucracy.

4.11 Continuous Improvement and Evaluation

AI systems aren’t “set it and forget it” technologies. They require ongoing attention to perform well and remain aligned with your agency or department's needs. Think of continuous improvement as part of responsible AI ownership, not an additional burden.

4.11.1 Tracking what matters

The key to effective AI system management is knowing what to measure and why. Rather than tracking everything possible, focus on metrics that tell you whether your system is meeting its intended purpose and operating safely.

System performance metrics help you understand whether the AI is performing effectively. Track accuracy rates for systems making predictions or classifications, response times for user-facing applications, and error rates that might indicate problems. The specific metrics depend on your system's purpose. For example, a document summarization tool needs different measurement mechanisms than a chatbot.

User experience indicators reveal how well the system serves its intended audience. User satisfaction surveys, adoption rates across different user groups, and the volume of support requests can all signal whether the system is working as intended. Pay particular attention to accessibility metrics that show whether all users can effectively use the system.

Risk management indicators help you spot potential problems before they become serious issues. Track security incidents, privacy compliance measures, and fairness indicators across different demographic groups. For GenAI systems, monitor for inappropriate outputs or content that might need filtering.

Business value measures connect your AI system to organizational goals. Document efficiency improvements, cost savings, service quality enhancements, and process improvements. These metrics help justify continued investment and identify opportunities for expansion.

The goal isn't to create extensive dashboards but to establish a reasonable set of indicators that help you understand system health and value.

4.11.2 Creating Sustainable Review Cycles

Effective AI governance requires regular review, but the frequency and depth should match the system's risk level and complexity.

Daily and weekly operational oversight focuses on immediate system health. Monitor for outages, unusual patterns in user interactions, or performance degradation. This level of review typically involves technical staff and doesn't require extensive documentation, just awareness of system status and quick response to problems.

Monthly and quarterly tactical reviews step back to look at broader patterns. Analyze trends in performance metrics, user feedback schemes, and any emerging issues. This is when you might adjust prompts for GenAI systems, update training materials, or make minor configuration changes. Document these reviews briefly to track what you've learned and what you've changed.

Annual strategic assessments provide the opportunity for comprehensive evaluation. Reassess the system's risk classification, validate that it still serves its intended purpose, evaluate alignment with current agency or department priorities, and plan major enhancements. This deeper review should involve stakeholders beyond the technical team and result in documented decisions about the system's future.

The key is making review cycles sustainable. Don't create review processes so complex that they become burdensome or get skipped. Better to have simple, consistent reviews than elaborate processes that fall by the wayside.

4.11.3 Building a Culture of Learning

The most successful AI implementations treat continuous improvement as an opportunity to learn and adapt rather than a compliance obligation.

Make feedback easy to provide and act upon. Create simple ways for users to report issues, suggest improvements, or share insights about system performance. This might be as straightforward as an e-mail address for feedback or a brief monthly survey. More important than the specific mechanism is ensuring feedback gets reviewed and, when appropriate, acted upon.

Learn from both successes and problems. Document what works well so you can replicate successful approaches in other contexts. When problems arise, focus on understanding root causes and preventing similar issues rather than just fixing immediate symptoms.

Share knowledge across your organization and with other agencies and departments. Participate in the state's AI community of practice to learn from others' experiences and contribute your own insights. Internal knowledge sharing helps build AI capabilities across your agency (or department) and prevents others from encountering the same challenges you've already solved.

Stay current with evolving best practices. AI technology and governance practices continue to evolve rapidly. Regularly review guidance from ITS, participate in relevant training opportunities, and stay informed about developments in AI risk management and best practices.

Approach changes thoughtfully. Continuous improvement doesn't mean constant change. Evaluate potential improvements carefully, considering their impact on users, system stability, and resource requirements. Sometimes the best improvement is maintaining stable, reliable service rather than implementing new features.

The goal of continuous improvement is to ensure your AI systems continue to serve their intended purpose effectively and safely. This requires ongoing attention, but it doesn't require perfection or elaborate processes. Focus on building sustainable practices that help you learn, adapt, and improve over time.

5.0 Appendix A: Shared Responsibility Model

5.1 Shared Responsibility Model for AI Governance

P	Primary Responsibility – Who leads each activity
S	Support Role – Who assists or provides input

5.1.1 Governance Activities

Activity	ITS	Agency/Department Leadership	Agency/Department IT Liaison	Agency/Department IT Staff	End Users
Establish statewide AI policies	P	S	S	S	S
Develop agency/department-specific AI procedures	S	P	P	S	S
Maintain AI system inventory	P	S	P	S	S
Define acceptable use parameters	P	S	S	S	S
Monitor policy compliance	P	S	P	S	S

5.1.2 Risk Management Activities

Activity	ITS	Agency/Department Leadership	Agency/Department IT Liaison	Agency/Department IT Staff	End Users
Define risk assessment methodology	P	S	S	S	S
Conduct risk classifications	P	S	S	S	S
Approve Tier 2 implementations	P	S	S	S	S
Approve Tier 3 implementations	P	S	S	S	S
Implement risk mitigations	S	P	P	P	S

5.1.3 Implementation Activities

Activity	ITS	Agency/Department Leadership	Agency/Department IT Liaison	Agency/Department IT Staff	End Users
Define AI use cases	S	P	S	S	P
Select AI solutions	S	P	P	S	S
Deploy AI systems	S	S	P	P	S
Test AI systems	S	S	P	P	S
Conduct use acceptance testing	S	S	P	S	P

5.1.4 Security and Privacy Activities

Activity	ITS	Agency/Department Leadership	Agency/Department IT Liaison	Agency/Department IT Staff	End Users
Establish security standards	P	S	S	S	S
Implement security controls	S	S	P	P	S
Conduct security assessments	P	S	S	P	S
Conduct privacy assessments	P	S	S	P	S
Manage privacy controls	S	S	P	P	S
Respond to security incidents	P	S	P	P	S

6.0 Appendix B: References

- ITS Information Security Manual
- Idaho's AI Advantage: A Framework for Responsible Innovation
- NIST Artificial Intelligence Risk Management Framework (AI RMF 1.0)
- Idaho Code §§74-101 through 74-127

7.0 Appendix C: Revision History

Revision	Date	Description	Revised By
1	August 2025	Initial draft	Elizabeth Knox