Azpirant TECHNOLOGIES

PRISM

Practical Risk-based Inspection System for Machine Intelligence

AI AUDIT FRAMEWORK

A Practitioner's Guide to Auditing AI Systems

Version 1.0 | January 2026

Aligned with: **EU AI Act | ISO/IEC 42001 | NIST AI RMF | DORA**

Prepare by Prabh Nair

Contents

Part 1: Framework Foundation

1.1 What This Framework Is (and Isn't)

PRISM is built for auditors who need to assess AI systems in real-world environments. It's not an academic exercise or a compliance checklist generator.

What PRISM Is:

- A risk-based methodology for evaluating AI systems across their lifecycle
- Practical procedures that work whether you're auditing a startup's ML model or an enterprise AI platform
- Evidence-focused — every control has specific artifacts to collect and tests to perform
- Regulatory-aware — mapped to EU AI Act, ISO 42001, NIST AI RMF, but not dependent on any single standard
- Scalable — from a single high-risk model to an organization-wide AI audit

What PRISM Is Not:

- A certification scheme (though it supports ISO 42001 audits)
- A replacement for professional judgment — you still need to think
- A guarantee of AI safety — no audit framework can promise that
- An ethics evaluation tool — that requires different expertise

The Core Premise

AI audit is fundamentally different from traditional IT audit. Here's why:

Traditional IT Audit	AI Audit
Deterministic systems — same input = same output	Probabilistic systems — same input may yield different outputs
Logic is explicit in code	Logic is learned from data (often opaque)
Change control is event-based	Models drift continuously even without code changes
Testing is binary (pass/fail)	Testing is statistical (accuracy, precision, recall)
Security = prevent unauthorized access	Security = prevent unauthorized access + adversarial attacks + data poisoning
Documentation describes what system does	Documentation must explain why system behaves as it does

This means auditors must adapt their techniques. PRISM provides that adaptation.

1.2 Audit Types and When to Use Them

Audit Type	Purpose	Typical Scope	Duration	Output
Pre-Deployment Review	Assess readiness before go-live	Single AI system	1-2 weeks	Go/No-Go recommendation with conditions
Operational Audit	Evaluate controls on live systems	Single system or portfolio	2-4 weeks	Findings report with remediation plan
Compliance Audit	Verify regulatory conformity	All in-scope systems	4-8 weeks	Compliance status + gap analysis
Incident Investigation	Root cause analysis post-incident	Affected system(s)	1-3 weeks	Incident report + recommendations
Third-Party Assessment	Evaluate vendor AI systems	Vendor product/service	2-3 weeks	Risk assessment + contract recommendations
Certification Audit	ISO 42001 or equivalent	Organization-wide AIMS	6-12 weeks	Certification recommendation

Selecting the Right Audit Type

Decision matrix for audit type selection:

Trigger	Recommended Audit Type	Minimum Scope
New AI system ready for deployment	Pre-Deployment Review	All Domain 3 + 4 controls
Annual audit cycle	Operational Audit	Risk-based selection of systems
EU AI Act compliance deadline	Compliance Audit	All high-risk AI systems
AI system produced harmful output	Incident Investigation	Affected system + related systems
New AI vendor under consideration	Third-Party Assessment	Vendor's AI offering
ISO 42001 certification required	Certification Audit	Full AIMS scope
Board/regulator request	Depends on request	As specified

1.3 Risk-Based Audit Scoping

Not all AI systems require the same audit intensity. PRISM uses a risk-based approach to determine audit depth.

Step 1: System Risk Classification

Use this matrix to classify each AI system:

Risk Factor	Low (1)	Medium (2)	High (3)	Weight
Decision Impact	Informational only	Influences human decisions	Autonomous decisions affecting people	25%
Data Sensitivity	Public/aggregated data	Internal business data	Personal/sensitive data	20%
Affected Population	Internal users only	Limited external users	Large public population	20%
Reversibility	Easily reversible	Reversible with effort	Irreversible or high-cost reversal	15%
Regulatory Scope	Not regulated	General regulations apply	AI-specific regulations apply	20%

Weighted Risk Score = Σ (Factor Score × Weight)

Score Range	Risk Level	Audit Intensity	Control Coverage
2.5 - 3.0	Critical	Full audit — all controls tested	100% of PRISM controls
2.0 - 2.4	High	Comprehensive audit — key controls tested in depth	All controls, sampling for operational
1.5 - 1.9	Medium	Focused audit — risk areas prioritized	Governance + highest-risk domains
1.0 - 1.4	Low	Limited review — spot checks only	Governance + data controls only

Step 2: Determine Sample Size

For organizations with multiple AI systems:

Population Size	Critical Risk	High Risk	Medium Risk	Low Risk
1-5 systems	100%	100%	100%	50%
6-15 systems	100%	100%	60%	30%
16-50 systems	100%	70%	40%	20%
51-100 systems	100%	50%	25%	10%
100+ systems	100%	30%	15%	5%

Note: All Critical risk systems must be audited regardless of population size.

Part 2: The Six Audit Domains

PRISM organizes AI audit into six domains. Each domain contains control objectives, specific audit procedures, evidence requirements, and testing techniques.

Domain	Focus Area	# Controls	Key Question
1. Governance	Who is accountable? What are the rules?	8	Is there clear ownership and policy?
2. Risk Management	What can go wrong? How is it managed?	10	Are AI-specific risks identified and treated?
3. Data Integrity	Is the data trustworthy?	12	Can we trust what the model learned from?
4. Model Lifecycle	Is the model built and maintained properly?	14	Is the model reliable and well-documented?
5. Operational Controls	Is the system running safely?	10	Is the system monitored and protected?
6. Compliance & Ethics	Are we meeting obligations?	8	Are regulatory and ethical requirements met?

DOMAIN 1: GOVERNANCE & ACCOUNTABILITY

This domain assesses whether the organization has established clear ownership, policies, and structures for AI governance. Without this foundation, technical controls are meaningless.

Control Objectives

ID	Control Objective	Risk Addressed
G-01	AI governance policy is documented, approved, and communicated	Unclear rules lead to inconsistent practices
G-02	Roles and responsibilities for AI are defined and assigned	No accountability = no ownership of problems
G-03	AI governance committee/function exists with appropriate authority	Decisions made without proper oversight
G-04	AI system inventory is complete and current	Unknown systems = unmanaged risk
G-05	AI strategy aligns with organizational risk appetite	AI investments misaligned with business tolerance
G-06	Resource allocation for AI governance is adequate	Under-resourced governance fails
G-07	AI competency requirements are defined and training provided	Unqualified personnel making AI decisions
G-08	Board/executive oversight of AI is established	Strategic risks not visible to leadership

Audit Procedures

G-01: AI Governance Policy

Evidence to collect:

- AI governance policy document
- Policy approval records (board minutes, sign-off)
- Communication records (emails, intranet announcements, training materials)
- Policy acknowledgment records from staff

Tests to perform:

1. Verify policy exists and has been approved within the last 24 months
2. Check policy covers: scope definition, risk classification, approval requirements, roles, prohibited uses
3. Interview 3-5 staff in different roles — ask them to describe AI policy requirements
4. Verify policy is accessible (not buried in SharePoint)
5. Check policy references current regulations (EU AI Act, applicable standards)

Red flags:

- Policy hasn't been updated since before EU AI Act (pre-2024)
- Staff cannot articulate basic policy requirements
- Policy exists but no evidence of enforcement

G-02: Roles and Responsibilities

Evidence to collect:

- Organization chart showing AI governance roles
- Job descriptions for AI-related positions
- RACI matrix for AI activities
- System ownership register

Tests to perform:

6. For 5 AI systems in inventory, verify each has a named business owner and technical owner
7. Interview owners — confirm they understand their responsibilities
8. Check for segregation of duties (developer ≠ approver ≠ deployer)
9. Verify escalation paths are documented and known

Red flags:

- Systems with no assigned owner
- Owners unaware of their accountability
- Same person develops, tests, and approves AI systems

G-03: Governance Committee/Function

Evidence to collect:

- Committee charter/terms of reference
- Meeting minutes from last 12 months
- Membership list with roles
- Decision log

Tests to perform:

10. Verify committee meets at required frequency (at least quarterly for high-risk)
11. Review 3-5 meeting minutes — check for substantive discussion, decisions, action items
12. Confirm committee has authority to approve/reject AI deployments
13. Check membership includes appropriate expertise (technical, legal, business, ethics)
14. Verify committee reviews high-risk AI before deployment

Red flags:

- Committee exists on paper but doesn't meet
- Meetings are rubber-stamp exercises with no real discussion
- No technical expertise on committee
- High-risk AI deployed without committee review

G-04: AI System Inventory

Evidence to collect:

- AI system inventory/register
- Inventory update procedures
- Last inventory reconciliation

Tests to perform:

15. Obtain current inventory — check completeness of required fields
16. Select 5 systems from inventory — verify they exist and details are accurate
17. Identify 3 AI use cases through interviews/observation — verify they're in inventory
18. Check procurement records for AI purchases — verify all appear in inventory
19. Review cloud billing for ML services (SageMaker, Azure ML, etc.) — cross-reference with inventory

Red flags:

- Inventory hasn't been updated in >6 months
- Systems found in production that aren't in inventory
- ChatGPT/Copilot usage not tracked
- No process to add new systems to inventory

DOMAIN 2: RISK MANAGEMENT

This domain assesses whether AI-specific risks are systematically identified, assessed, treated, and monitored. Generic enterprise risk management is insufficient for AI.

Control Objectives

ID	Control Objective	Risk Addressed
R-01	AI risk assessment methodology is defined and applied	Ad-hoc risk identification misses critical risks
R-02	AI systems are classified by risk level	All systems treated equally = over/under governance
R-03	Risk treatment plans exist for identified risks	Risks identified but not addressed
R-04	Bias and fairness risks are specifically assessed	Discrimination risk not evaluated
R-05	Third-party AI risks are evaluated	Vendor AI creates unmanaged exposure
R-06	Fundamental rights impact is assessed for high-risk AI	Human rights violations
R-07	Residual risk is formally accepted by appropriate authority	Unacknowledged risk exposure
R-08	Risk register is maintained and current	No visibility into aggregate AI risk
R-09	Emerging AI risks are monitored	New threats not identified
R-10	Risk assessments are periodically refreshed	Stale assessments don't reflect current state

Audit Procedures

R-01: Risk Assessment Methodology

Evidence to collect:

- AI risk assessment procedure/methodology document
- Risk assessment templates
- Completed risk assessments (sample of 5)

Tests to perform:

20. Verify methodology addresses AI-specific risks (bias, drift, adversarial attacks, opacity)
21. Check methodology considers: impact, likelihood, data sensitivity, autonomy level
22. Review 5 completed assessments — verify methodology was actually followed
23. Interview assessors — confirm they understand and can apply the methodology
24. Verify assessments are performed before deployment (not after)

Red flags:

- Using generic IT risk methodology without AI adaptations
- Risk assessments completed post-deployment
- Assessments don't consider bias, fairness, or explainability
- Copy-paste assessments with no system-specific analysis

R-04: Bias and Fairness Assessment

Evidence to collect:

- Bias testing methodology/procedure
- Bias test results for high-risk systems
- Fairness metrics definitions
- Mitigation actions taken

Tests to perform:

25. Verify bias testing is required for all Tier 1/2 systems
26. Review bias test results for 3 high-risk systems
27. Check protected characteristics tested (gender, age, race, disability, etc.)
28. Verify fairness metrics are appropriate for use case (demographic parity, equalized odds, etc.)
29. Confirm mitigation actions were implemented where bias detected
30. Check ongoing bias monitoring is in place (not just pre-deployment testing)

Red flags:

- No bias testing performed on decision-making systems
- Testing done on training data only, not production data
- Bias detected but no remediation
- Fairness metrics not defined or inappropriate for use case

DOMAIN 3: DATA INTEGRITY

This domain assesses whether the data used to train and operate AI systems is trustworthy. Garbage in = garbage out. This is often the highest-risk area.

Control Objectives

ID	Control Objective	Risk Addressed
D-01	Training data sources are documented and validated	Model trained on inappropriate/biased data
D-02	Data quality standards are defined and enforced	Low-quality data degrades model performance
D-03	Data lineage is tracked from source to model	Cannot trace decisions back to data
D-04	Personal data processing has legal basis	Privacy violations, regulatory breach
D-05	Consent management is implemented where required	Using data without proper consent
D-06	Data labeling quality is assured	Incorrect labels = incorrect learning
D-07	Data version control is maintained	Cannot reproduce training conditions
D-08	Synthetic/augmented data use is controlled	Synthetic data introduces artifacts
D-09	Data retention and deletion policies are enforced	Retaining data beyond legal limits
D-10	Data access controls are implemented	Unauthorized access to training data
D-11	Data drift is monitored	Production data differs from training data
D-12	Data documentation (data sheets) is maintained	No understanding of data characteristics

Audit Procedures

D-01: Training Data Validation

Evidence to collect:

- Data source documentation
- Data acquisition agreements/licenses
- Data validation reports
- Data sheet (if available)

Tests to perform:

31. For 3 high-risk models, trace training data back to original sources
32. Verify data sources are appropriate for the use case (representative, relevant)
33. Check for proper licensing/permission to use data for ML training
34. Review data validation checks performed before training
35. Verify no prohibited data sources (e.g., scraped without permission, biased historical data)

Red flags:

- Training data sources unknown or undocumented
- Using publicly scraped data without license verification
- Historical data with known biases used without correction
- No validation performed on training data quality

D-06: Data Labeling Quality

Evidence to collect:

- Labeling procedures/guidelines
- Labeler qualifications/training records
- Inter-annotator agreement metrics
- Label quality review results

Tests to perform:

36. Review labeling guidelines — check for clarity and completeness
37. Verify labelers were qualified and trained for the task
38. Check inter-annotator agreement scores (should be >80% for most tasks)
39. Review sample of labels for accuracy (request 50 samples with ground truth)
40. Verify label disputes were resolved with documented process

Red flags:

- Labels created by unqualified/untrained personnel
- No inter-annotator agreement measured
- Single labeler with no quality review
- Labeling guidelines vague or nonexistent

DOMAIN 4: MODEL LIFECYCLE

This domain assesses whether AI models are developed, tested, deployed, and maintained with appropriate rigor. This is the technical heart of AI audit.

Control Objectives

ID	Control Objective	Risk Addressed
M-01	Model development follows documented methodology	Inconsistent, unreproducible development
M-02	Model requirements are defined and traceable	Model doesn't meet business needs
M-03	Model architecture decisions are documented and justified	Inappropriate model for problem
M-04	Model training is reproducible	Cannot recreate or validate model
M-05	Model testing is comprehensive and documented	Untested failure modes
M-06	Model validation is independent from development	Developer blind spots not caught
M-07	Model performance metrics are defined and measured	No objective measure of quality
M-08	Model explainability is assessed and documented	Black box decisions
M-09	Model versioning and change control is implemented	Lost track of which model is in production
M-10	Model deployment follows defined process	Ad-hoc deployments bypass controls
M-11	Model documentation (model cards) is maintained	No understanding of model behavior
M-12	Model retraining triggers and procedures are defined	Model degrades without update
M-13	Model decommissioning process exists	Obsolete models remain in use
M-14	Transfer learning and foundation models are controlled	Inheriting unknown biases/risks

Audit Procedures

M-05: Model Testing

Evidence to collect:

- Test strategy/plan documents
- Test case documentation
- Test execution results
- Test data documentation
- Edge case and failure mode testing results

Tests to perform:

41. Review test strategy — verify it covers: functional, performance, bias, adversarial, edge cases
42. Check test data is representative and separate from training data
43. Review test results — verify pass/fail criteria were defined in advance
44. Confirm negative testing performed (what happens with bad inputs?)
45. Verify testing includes real-world scenarios, not just clean lab data
46. Check for adversarial testing (attempts to fool the model)

Red flags:

- Testing only on same data distribution as training
- No edge case or failure mode testing

- Pass/fail criteria defined after seeing results
- No adversarial testing on security-relevant systems

M-08: Model Explainability

Evidence to collect:

- Explainability requirements for system
- Explainability method documentation
- Sample explanations generated
- User comprehension testing results

Tests to perform:

47. Verify explainability requirements are defined based on use case and risk level
48. Review explainability method used (SHAP, LIME, attention weights, etc.)
49. Request explanation for 5 sample predictions — assess quality and usefulness
50. For high-risk systems, verify explanations are understandable to affected individuals
51. Check explanations are logged and available for audit/appeal

Red flags:

- High-risk decisions with no explanation capability
- Explanations are technical jargon incomprehensible to users
- Explanations not logged or available after the fact
- Using opaque model where explainability is legally required

DOMAIN 5: OPERATIONAL CONTROLS

This domain assesses whether AI systems are securely operated, monitored, and protected in production. Where things actually go wrong.

Control Objectives

ID	Control Objective	Risk Addressed
O-01	Human oversight mechanisms are implemented	Autonomous harmful decisions
O-02	Performance monitoring is continuous	Degradation not detected
O-03	Model drift detection is implemented	Model becomes inaccurate over time
O-04	Anomaly detection identifies unusual behavior	Attacks or failures not detected
O-05	Incident response procedures include AI-specific scenarios	Unprepared for AI failures
O-06	Business continuity plans address AI dependencies	AI failure disrupts operations
O-07	Security controls protect against adversarial attacks	Model manipulation
O-08	Access controls limit who can modify AI systems	Unauthorized changes
O-09	Audit logging captures AI decisions and changes	No evidence trail
O-10	Fallback mechanisms exist for AI failure	No alternative when AI fails

Audit Procedures

O-01: Human Oversight

Evidence to collect:

- Human oversight design documentation
- Override mechanism documentation
- Override usage logs
- Escalation procedures

Tests to perform:

52. For high-risk systems, verify human oversight model is defined (in-the-loop, on-the-loop, over-the-loop)
53. Test override mechanism — can a human actually stop/reverse a decision?
54. Review override logs — are overrides happening? Too many = system unreliable. Zero = rubber stamp?
55. Interview human reviewers — do they understand when to override? Are they empowered to do so?
56. Check response time — can humans intervene before harm occurs?

Red flags:

- No override mechanism for consequential decisions
- Override exists but never used (suggests reviewers are rubber-stamping)
- Humans cannot intervene in real-time for time-critical decisions
- Reviewers not trained on when/how to override

O-03: Model Drift Detection

Evidence to collect:

- Drift detection methodology
- Drift monitoring dashboards/reports
- Drift alert configurations
- Retraining trigger criteria

Tests to perform:

57. Verify drift monitoring is implemented for all production models
58. Review drift metrics tracked (data drift, concept drift, prediction drift)
59. Check alert thresholds are defined and appropriate
60. Review drift alerts from past 6 months — verify they were investigated and resolved
61. Confirm retraining is triggered when drift exceeds thresholds

Red flags:

- No drift monitoring on production models
- Alerts configured but no one reviewing them
- Model in production >12 months with no drift analysis
- Drift detected but model not retrained

DOMAIN 6: COMPLIANCE & ETHICS

This domain assesses whether regulatory and ethical obligations are met. The most consequential domain for avoiding fines and reputational damage.

Control Objectives

ID	Control Objective	Risk Addressed
C-01	Regulatory requirements are identified and tracked	Non-compliance with applicable laws
C-02	Conformity assessments are completed for high-risk AI	EU AI Act non-compliance
C-03	Transparency obligations are met	User deception, regulatory breach
C-04	Technical documentation meets regulatory requirements	Insufficient evidence for compliance
C-05	Incident reporting procedures meet regulatory timelines	Late or missed notifications
C-06	AI ethics principles are defined and operationalized	Ethical violations
C-07	Prohibited AI uses are identified and prevented	Deploying banned AI systems
C-08	Records are maintained for regulatory demonstration	Cannot prove compliance

Audit Procedures

C-02: Conformity Assessment (EU AI Act)

Evidence to collect:

- High-risk AI classification documentation
- Conformity assessment reports
- Technical documentation per Annex IV
- Quality management system documentation
- CE marking records
- EU database registration

Tests to perform:

62. Identify all AI systems that may fall under Annex III high-risk categories
63. Verify conformity assessment was performed (internal or notified body as required)
64. Review technical documentation against Annex IV requirements
65. Check quality management system addresses EU AI Act requirements
66. Verify CE marking applied correctly
67. Confirm registration in EU database for applicable systems

Red flags:

- High-risk systems deployed without conformity assessment
- Technical documentation doesn't meet Annex IV requirements
- No quality management system for AI
- Systems not registered in EU database when required

C-03: Transparency Obligations

Evidence to collect:

- User notification mechanisms (chatbot disclosures, etc.)
- Transparency documentation for affected individuals
- Explanation provision mechanisms
- Deepfake/synthetic content labeling

Tests to perform:

68. For chatbots/virtual assistants: verify users are informed they're interacting with AI
69. For emotion recognition systems: verify users are informed
70. For synthetic content generation: verify outputs are labeled as AI-generated
71. For high-risk systems: verify affected individuals can request explanations
72. Test user interfaces — is disclosure clear and prominent?

Red flags:

- Chatbots pretending to be human
- Synthetic content not labeled
- No mechanism for affected individuals to request explanations
- Transparency notices buried in fine print

Part 3: Audit Execution

3.1 Evidence Collection Techniques

Technique	When to Use	Strength	Limitation
Document Review	Policy, procedure, documentation assessment	Efficient, comprehensive	May not reflect actual practice
Interview	Understanding processes, culture, awareness	Reveals nuance and context	Subjective, may be rehearsed
Observation	Process execution, control operation	Shows actual practice	Hawthorne effect, point-in-time
Walkthrough	End-to-end process understanding	Identifies gaps and handoffs	Time-consuming, may be scripted
Re-performance	Testing control effectiveness	Direct evidence of operation	Sample may not be representative
Data Analysis	Large-scale pattern detection	Objective, comprehensive	Requires technical capability
Technical Testing	Security, performance, bias assessment	Definitive technical evidence	Requires specialized skills
System Inspection	Configuration, code, model review	Direct verification	Point-in-time, requires access

3.2 AI-Specific Testing Procedures

Beyond traditional audit techniques, AI audit requires specialized testing:

Bias Testing

Test Type	Description	Tools/Methods
Demographic Parity	Equal positive prediction rates across groups	Compare prediction rates by protected attribute
Equalized Odds	Equal true positive and false positive rates	Confusion matrices by group
Individual Fairness	Similar individuals receive similar predictions	Distance metrics on feature space
Calibration	Predictions are well-calibrated across groups	Reliability diagrams by group
Disparate Impact	Selection rates across groups (80% rule)	Calculate selection rate ratios

Robustness Testing

Test Type	Description	Tools/Methods
Adversarial Inputs	Model behavior with malicious inputs	FGSM, PGD, C&W attacks
Out-of-Distribution	Model behavior on unusual inputs	Generate OOD samples, measure confidence
Boundary Testing	Behavior at decision boundaries	Systematic boundary exploration
Stress Testing	Performance under load	Load testing with concurrent requests
Input Fuzzing	Random/malformed input handling	Automated fuzz testing

Performance Testing

Metric	Definition	Acceptable Range (typical)
Accuracy	Correct predictions / total predictions	>90% (varies by use case)
Precision	True positives / predicted positives	>85% for high-stakes
Recall	True positives / actual positives	>85% for critical safety
F1 Score	Harmonic mean of precision and recall	>80%
AUC-ROC	Area under ROC curve	>0.85
Latency	Response time	<100ms for real-time
Throughput	Predictions per second	Depends on requirements

3.3 Finding Classification

Findings must be classified consistently to enable prioritization and tracking:

Severity	Definition	Examples	Remediation Timeline
Critical	Immediate risk of significant harm; regulatory breach; fundamental control failure	High-risk AI deployed without assessment; bias causing discrimination; no human oversight on autonomous decisions	Immediate (stop activity)
High	Significant control weakness; likely regulatory finding; material risk exposure	Incomplete risk assessments; no drift monitoring; inadequate documentation	30 days
Medium	Control gap that could lead to issues; best practice deviation	Policy not updated; training gaps; incomplete testing	90 days
Low	Minor weakness; improvement opportunity; efficiency gain	Documentation formatting; process optimization	180 days
Observation	Not a deficiency; suggestion for enhancement	Industry practices; emerging guidance	Optional

Finding Documentation Requirements

Each finding must include:

- Finding ID: Unique identifier
- Title: Brief descriptive title
- Domain: Which PRISM domain (G, R, D, M, O, C)
- Control Reference: Specific control objective
- Severity: Critical/High/Medium/Low/Observation
- Condition: What was found (factual description)
- Criteria: What should be in place (policy, standard, regulation)
- Cause: Why the gap exists (root cause)
- Effect: What could go wrong (risk/impact)
- Recommendation: What to do about it
- Management Response: Owner, action plan, target date
- Evidence Reference: Workpaper reference

3.4 Reporting Standards

Executive Summary Structure

Every AI audit report should begin with a 1-2 page executive summary:

- Audit Scope: What was audited, what was excluded
- Overall Assessment: Summary opinion/rating
- Key Findings: Top 3-5 most significant findings
- Immediate Actions Required: Critical items needing immediate attention
- Positive Observations: What's working well
- Next Steps: Recommended follow-up activities

Rating Scale

Rating	Definition	Criteria
Effective	Controls are well-designed and operating effectively	No Critical/High findings; <3 Medium findings
Needs Improvement	Controls exist but have gaps	No Critical; ≤2 High findings; any number of Medium
Ineffective	Significant control weaknesses	Any Critical finding OR >2 High findings
Not Assessed	Insufficient evidence to conclude	Access limitations, scope exclusions

Report Distribution

Audience	Content	Format
Board/Audit Committee	Executive summary, overall rating, critical findings only	1-2 pages, high-level
Executive Management	Executive summary + all High/Critical findings + trends	5-10 pages
AI Governance Committee	Full report with all findings	Complete report
System Owners	Findings relevant to their systems	Extracted findings
Regulators (if required)	As specified by regulation	Per regulatory format

Part 4: Workpaper Templates

The following templates support PRISM audit execution. Adapt as needed for your organization.

WP-01: Audit Planning Checklist

#	Planning Step	Completed	Notes
1	Obtain AI system inventory from management	☐	
2	Classify systems by risk level	☐	
3	Determine sample size based on risk	☐	
4	Review prior audit findings and status	☐	
5	Identify regulatory requirements applicable	☐	
6	Request preliminary documentation	☐	
7	Schedule interviews with key personnel	☐	
8	Arrange system access for testing	☐	
9	Confirm audit timeline and milestones	☐	
10	Communicate audit scope to stakeholders	☐	

WP-02: System Assessment Summary

Field	Response
System Name	
System ID	
Business Owner	
Technical Owner	
System Description	
AI/ML Technology	
Risk Classification	☐ Critical ☐ High ☐ Medium ☐ Low
Regulatory Scope	☐ EU AI Act High-Risk ☐ DORA ☐ Other: _____
Data Types Processed	
Decision Impact	
Human Oversight Model	☐ In-the-loop ☐ On-the-loop ☐ Over-the-loop ☐ None
Last Assessment Date	
Key Concerns Identified	

WP-03: Control Testing Workpaper

Field	Response
Control ID	
Control Objective	
System(s) Tested	
Test Procedure Performed	
Sample Size / Selection Method	
Evidence Reviewed	
Test Results	
Exceptions Noted	
Root Cause of Exceptions	
Control Assessment	☐ Effective ☐ Partially Effective ☐ Ineffective
Finding Reference (if applicable)	
Auditor	
Date	
Reviewer	
Review Date	

WP-04: Finding Documentation

Field	Response
Finding ID	
Finding Title	
Domain	☐ Governance ☐ Risk ☐ Data ☐ Model ☐ Operations ☐ Compliance
Control Reference	
Severity	☐ Critical ☐ High ☐ Medium ☐ Low ☐ Observation
Condition (What was found)	
Criteria (What should be)	
Cause (Why the gap exists)	
Effect (Risk/Impact)	
Recommendation	
Evidence Reference	
Management Response	
Action Owner	
Target Remediation Date	
Auditor	
Date	

WP-05: Interview Record

Field	Response
Interviewee Name	
Title/Role	
Date/Time	
Location/Format	☐ In-person ☐ Video ☐ Phone
Interviewer(s)	
Topics Covered	
Key Points Discussed	
Documents Provided/Promised	
Follow-up Items	
Interviewee Confirmation	☐ Notes shared ☐ Confirmed accurate ☐ Corrections made

WP-06: Bias Testing Results

Field	Response
System Name	
Model Version	
Test Date	
Test Dataset Description	
Protected Attributes Tested	
Fairness Metrics Used	
Overall Accuracy	
Accuracy by Group	
Selection Rate by Group	
Disparate Impact Ratio	
Statistical Significance	
Bias Detected?	☐ Yes ☐ No
Bias Description (if yes)	
Mitigation Implemented	
Residual Bias Accepted?	☐ Yes (by: _____) ☐ No ☐ N/A
Tester	
Reviewer	

Quick Reference: PRISM Control Summary

Complete list of all 62 control objectives for reference:

Domain 1: Governance (8 controls)

ID	Control Objective
G-01	AI governance policy is documented, approved, and communicated
G-02	Roles and responsibilities for AI are defined and assigned
G-03	AI governance committee/function exists with appropriate authority
G-04	AI system inventory is complete and current
G-05	AI strategy aligns with organizational risk appetite
G-06	Resource allocation for AI governance is adequate
G-07	AI competency requirements are defined and training provided
G-08	Board/executive oversight of AI is established

Domain 2: Risk Management (10 controls)

ID	Control Objective
R-01	AI risk assessment methodology is defined and applied
R-02	AI systems are classified by risk level
R-03	Risk treatment plans exist for identified risks
R-04	Bias and fairness risks are specifically assessed
R-05	Third-party AI risks are evaluated
R-06	Fundamental rights impact is assessed for high-risk AI
R-07	Residual risk is formally accepted by appropriate authority
R-08	Risk register is maintained and current
R-09	Emerging AI risks are monitored
R-10	Risk assessments are periodically refreshed

Domain 3: Data Integrity (12 controls)

ID	Control Objective
D-01	Training data sources are documented and validated
D-02	Data quality standards are defined and enforced
D-03	Data lineage is tracked from source to model
D-04	Personal data processing has legal basis
D-05	Consent management is implemented where required
D-06	Data labeling quality is assured
D-07	Data version control is maintained
D-08	Synthetic/augmented data use is controlled
D-09	Data retention and deletion policies are enforced
D-10	Data access controls are implemented
D-11	Data drift is monitored

D-12	Data documentation (data sheets) is maintained

Domain 4: Model Lifecycle (14 controls)

ID	Control Objective
M-01	Model development follows documented methodology
M-02	Model requirements are defined and traceable
M-03	Model architecture decisions are documented and justified
M-04	Model training is reproducible
M-05	Model testing is comprehensive and documented
M-06	Model validation is independent from development
M-07	Model performance metrics are defined and measured
M-08	Model explainability is assessed and documented
M-09	Model versioning and change control is implemented
M-10	Model deployment follows defined process
M-11	Model documentation (model cards) is maintained
M-12	Model retraining triggers and procedures are defined
M-13	Model decommissioning process exists
M-14	Transfer learning and foundation models are controlled

Domain 5: Operational Controls (10 controls)

ID	Control Objective
O-01	Human oversight mechanisms are implemented
O-02	Performance monitoring is continuous
O-03	Model drift detection is implemented
O-04	Anomaly detection identifies unusual behavior
O-05	Incident response procedures include AI-specific scenarios
O-06	Business continuity plans address AI dependencies
O-07	Security controls protect against adversarial attacks
O-08	Access controls limit who can modify AI systems
O-09	Audit logging captures AI decisions and changes
O-10	Fallback mechanisms exist for AI failure

Domain 6: Compliance & Ethics (8 controls)

ID	Control Objective
C-01	Regulatory requirements are identified and tracked
C-02	Conformity assessments are completed for high-risk AI
C-03	Transparency obligations are met
C-04	Technical documentation meets regulatory requirements
C-05	Incident reporting procedures meet regulatory timelines

C-06	AI ethics principles are defined and operationalized
C-07	Prohibited AI uses are identified and prevented
C-08	Records are maintained for regulatory demonstration

— End of PRISM AI Audit Framework —

Version 1.0 | For professional use | Adapt to organizational context