PASC
PAN-AFRICAN STANDARDS COUNCIL

DIRECTIVE / INSTITUTIONAL EDITION

PASC-AI-01 · ED. 1.0 · NOVEMBER 2026

Responsible, sovereign and demonstrable artificial intelligence

Governance · Rights · Model assurance · Independent evidence

Issuing authority
Pan-African Standards Council (PASC)

Publication status
Council-approved · Published

Publication date
1 November 2026

Institutional architecture

PASC AI 01 defines dedicated governance and assurance requirements for artificial intelligence. OSPCRM v1.0 governs related cybersecurity risk, AI/ML attack vectors, supply-chain exposure, sovereignty and cyber-assurance. PASC AI 01 governs human rights and agency, lawful data use, discriminatory outcomes, reliability, explainability, autonomous action and lifecycle oversight. The standards operate through traceable cross-references and distinct conformity scopes.

1 Scope and applicability

1.1 This standard specifies organizational and system-level requirements for developing, procuring, adapting, deploying, operating, monitoring and retiring AI systems. It covers predictive and generative AI, foundation-model integrations, algorithmic decision support, embodied AI and tool-using or agentic systems.

1.2 It applies to providers and deployers according to their actual control, including government, education, public services, financial services, healthcare, employment, agriculture, industrial operations and consumer systems. The declared conformity scope SHALL specify which legal entities, products, deployments and lifecycle stages are included.

1.3 It addresses harms to individuals, communities, institutions, rights, livelihoods, public trust, physical safety, information integrity and strategic autonomy. Pure cybersecurity-risk assessment and incident-priority requirements remain governed by OSPCRM where applicable.

1.4 This voluntary PASC standard does not supersede applicable law, international human rights obligations, sector regulations, data-protection requirements or legitimately binding contracts. A legal compliance claim SHALL identify its jurisdiction, obligation and evidence separately.

2 Normative language and definitions

2.1 **SHALL** indicates a mandatory requirement for a conformity claim; **SHOULD** indicates a recommendation; **MAY** indicates permission. A justified exception or accepted risk does not convert an unmet SHALL into a passed test.

2.2 *AI system* means a machine-based system that infers outputs such as predictions, content, recommendations or decisions from inputs, with varying levels of autonomy. *Affected person* includes people whose interests, rights, opportunities or safety may be changed by system output, even without direct interaction.

2.3 *High-impact use* means a deployment whose reasonably foreseeable failure, misuse or exclusion could materially impair rights, safety, access to essential services, livelihood, democratic participation or important public interests. It is a PASC assurance trigger, not a legal reclassification of an EU or national regulated category.

2.4 *Meaningful human oversight* requires sufficient information, authority, time and practical ability to question, override, suspend, correct or escalate an AI-mediated result before or after consequential action as appropriate.

2.5 *Demonstrability* means traceable, repeatable and reviewable evidence of an implemented outcome; a written policy or vendor declaration alone is not evidence that the outcome occurred.

3 System inventory, intended purpose and governance

AI.GV.01 The organization SHALL maintain a versioned AI system register covering internally built and externally procured systems. Each entry SHALL identify its intended purpose, prohibited or unsupported uses, accountable owner, provider and deployer roles, model and system versions, data sources, material third parties, affected groups, operating geography and decision/action authority.

AI.GV.02 Management SHALL assign accountability for lawful use, quality, privacy, human-impact review, security, model operation, independent evaluation, incident response and cessation of use. The same individual MAY hold multiple roles only where independence obligations remain satisfied.

AI.GV.03 Before deployment and after a material change, the owner SHALL perform and approve an AI impact and assurance assessment covering context, stakeholders, expected benefits, reasonably foreseeable misuse, uncertainty, potential harm and the availability of non-AI alternatives.

AI.GV.04 The assessment SHALL distinguish high-impact uses from lower-impact uses and set proportionate approval gates, independent evaluation, change thresholds and monitoring intervals. High-impact deployments SHALL receive documented senior authorization and a tested human intervention or withdrawal mechanism before production use.

AI.GV.05 Organizations SHALL establish an accessible channel for affected persons and personnel to raise concerns, report errors, request correction or seek review, with non-retaliation safeguards when applicable.

AI.GV.06 The system owner SHALL maintain a current register of laws, regulator guidance and contractual restrictions relevant to each operating jurisdiction, including cross-border data, sector-specific decisions and vulnerable groups. Voluntary standards SHALL not be presented as legal approvals.

4 Data provenance, rights and stewardship

AI.DA.01 The organization SHALL document the provenance, permitted purposes, acquisition grounds, rights and restrictions, retention periods and material limitations of datasets used to train, tune, retrieve, evaluate or operate in-scope systems, insofar as it controls or can procure that information.

AI.DA.02 Collection and processing SHALL have a documented lawful basis where required; purpose limitation, data minimization, confidentiality, access control and deletion or retention rules SHALL be applied. A public webpage’s accessibility alone SHALL not be treated as conclusive authorization for arbitrary extraction or model training.

AI.DA.03 The organization SHALL evaluate representation, measurement error, missingness, historical bias and foreseeable effects on relevant affected populations, including underrepresented groups and locally used languages. Corrective action SHALL be documented when these factors invalidate intended use.

AI.DA.04 Sensitive data, minors' data, biometric information and confidential institutional records SHALL receive enhanced necessity, authority, access and impact review. Data or prompts from production users SHALL NOT be repurposed for additional model training without the requisite authorization.

AI.DA.05 Where deletion, correction, withdrawal, data-subject rights or confidentiality duties apply, the organization SHALL document feasible response procedures and technical limitations across source stores, retrieval layers, logs, fine-tuning artefacts and service providers. Inability to erase a trained model precisely SHALL be disclosed in the assessment and treated where necessary.

5 Human rights, fairness and social context

AI.HR.01 The impact assessment SHALL identify reasonably foreseeable adverse effects on dignity, autonomy, privacy, equal treatment, access to opportunities, freedoms and effective remedy. It SHALL consider impacts on people indirectly affected by decisions.

AI.HR.02 High-impact systems SHALL be evaluated using relevant disaggregated performance and error measures where lawful, feasible and necessary, including group-specific false-positive/false-negative consequences. The organization SHALL document when direct demographic testing is restricted and use a justified alternative method rather than infer fairness from an overall average.

AI.HR.03 Design and evaluation SHALL consider local language proficiency, dialects, accessibility, cultural context and distributional effects relevant to each deployment. Translation of user interfaces alone SHALL not count as proof of equivalent model performance.

AI.HR.04 Where a system recommends or influences a consequential individual decision, affected persons SHALL receive appropriate notice of AI involvement, meaningful information concerning the basis and limitations of the result, and access to a competent appeal or correction path consistent with applicable law.

AI.HR.05 An organization SHALL NOT deploy an AI use where a mandatory legal prohibition applies. For other materially harmful uses, it SHALL document prevention, refusal, redesign or a specifically authorized residual-risk decision; this is not a waiver of law or of the standard's mandatory safeguards.

6 Model quality, validity and safety

AI.QA.01 The owner SHALL define a written claim of fitness for purpose with measurable performance and safety criteria, relevant populations and conditions, acceptable error boundaries, known limitations and conditions requiring abstention, escalation or shutdown.

AI.QA.02 Pre-deployment evaluation SHALL use data, scenarios and human review appropriate to real operating conditions. Training-set or vendor benchmark scores alone SHALL not establish performance on the declared local use case.

AI.QA.03 Testing SHALL consider calibration, robustness, uncertainty, out-of-distribution inputs, language variation, adversarial manipulation and foreseeable misuse, proportional to risk. Generative systems SHALL additionally test unsupported assertions, unsafe advice, fabricated sources and task-specific error patterns.

AI.QA.04 High-impact systems SHALL document independent challenge or evaluation, including acceptance thresholds, failed cases, remediation, residual-risk judgment and sign-off by an assessor not solely responsible for development or commercial delivery.

AI.QA.05 The organization SHALL monitor for drift and changing error patterns, set trigger thresholds and define a tested pause, fallback or rollback procedure. Changes in model weights, prompts, retrieval collections, tools, external APIs and decision thresholds SHALL be subject to impact-appropriate change review.

AI.QA.06 AI outputs SHALL NOT be described as inherently factual, unbiased, complete or suitable for autonomous consequential decisions without use-case-specific evidence. Human review SHALL itself be tested for ability to detect and correct errors, not assumed effective by the mere presence of a human.

7 Autonomy, agency and control of consequential action

AI.AG.01 A tool-using or agentic system SHALL have an explicit authorization boundary defining the identities, tools, data, resources, destinations, transaction limits, environments and actions it may access.

AI.AG.02 Consequential or irreversible actions SHALL be guarded by approvals proportionate to harm, with an authenticated human decision point where required by the approved risk policy. A generic instruction to “ask first” in a model prompt SHALL not replace system-enforced authorization.

AI.AG.03 The deployment SHALL apply least privilege, tool-output validation, external-content trust separation, rate and spend limits, transaction logging and immediate revocation or kill controls. Security attack handling SHALL additionally conform to OSPCRM when applicable.

AI.AG.04 The owner SHALL test agents against prompt and tool-output manipulation, authority confusion, cascading action errors, hidden instruction conflicts, false tool results, unbounded loops and loss of human control. Tests SHALL exercise real tool permission boundaries or representative constrained replicas.

AI.AG.05 Multi-agent and delegated workflows SHALL preserve traceable responsibility for each action and SHALL prevent unverified hand-off of authority to another agent, supplier service or execution environment.

8 Transparency, traceability and public accountability

AI.TR.01 Information to users and affected persons SHALL describe the system's role, intended use, material limitations, means of obtaining human support and relevant complaint routes, in comprehensible and accessible language.

AI.TR.02 The organization SHALL maintain version-linked records of model lineage, deployment configuration, major datasets or sources, evaluation reports, approvals, human interventions, significant incidents and corrective actions, subject to legitimate privacy and security restrictions.

AI.TR.03 Where synthetic content, impersonation or manipulated media creates material deception risk, the organization SHALL use proportionate disclosure, provenance or detection controls and evaluate their effectiveness. Metadata marking alone SHALL not substitute for deceptive-use prevention.

AI.TR.04 Claims of explainability SHALL describe what is explained, to whom, with what fidelity and limitations. A fluent narrative generated after a decision SHALL not be represented as verified causal explanation without validation.

9 Sovereignty, procurement and third-party assurance

AI.SO.01 The owner SHALL map training, inference, logging, fine-tuning, retrieval, model updates, moderation, identity administration and control-plane decision authority across providers and jurisdictions. It SHALL identify who can alter behaviour, access data, suspend service or extract sensitive information.

AI.SO.02 Procurement SHALL identify rights to documentation, testing, audit evidence, data export, version notice, incident notification, retention/deletion, service exit and continuity. Opaque dependencies SHALL be recorded as assurance limitations, not silently treated as compliant.

AI.SO.03 Critical deployments SHALL test contingency for supplier outage, model withdrawal, policy change, unexpected price or licensing change and loss of access to required data or controls. A credible non-AI or alternative-supplier service path SHOULD be considered.

AI.SO.04 When OSPCRM applies, cyber asset context, P0–P4 decisions, sovereignty, AI-specific cyber threats and supplier cybersecurity requirements SHALL be demonstrated under OSPCRM. Evidence MAY be reused with traceable references; obligations SHALL not be counted as satisfied solely because another certificate exists.

10 Operations, incident handling and continuous assurance

AI.OP.01 The owner SHALL continuously monitor actual use, error rates, complaints, harmful outputs, fairness drift, model changes, security events and human escalation effectiveness, in proportion to impact. Metrics SHALL identify coverage, denominators, populations and observation windows.

AI.OP.02 The organization SHALL establish incident thresholds, containment and notification responsibilities, evidence preservation, human support and restoration criteria. Material harm or loss of control SHALL trigger suspension or mitigation when continued operation cannot meet accepted limits.

AI.OP.03 Root-cause review SHALL distinguish technical malfunction, data defects, governance failure, supplier dependency, misuse, human oversight failure and contextual deployment error. Corrective action SHALL include an owner, due date and effectiveness check.

AI.OP.04 The system SHALL have a documented retirement process addressing replacement, archival, model and data retention, record traceability, vendor revocation, pending appeals and assurance obligations.

11 Conformity, independent assessment and certification design

AI.CO.01 A PASC AI 01 conformity claim SHALL state the exact version and scope, system(s), provider/deployer role, jurisdictions, intended purposes, excluded uses, high-impact determinations, evaluated configurations and material limitations.

AI.CO.02 All applicable SHALL controls within that scope SHALL be demonstrably met. The assessor SHALL review primary evidence and observe or reproduce risk-appropriate system outcomes; policy statements, checklist completion, marketing labels and untested vendor declarations alone are insufficient.

AI.CO.03 For high-impact AI, independent assessment SHALL test representative performance, harmful failure modes, fairness or affected-population impacts, human intervention, version controls, monitoring, complaints and remediation. The assessor SHALL have documented competence and independence from the delivery function.

AI.CO.04 Each assessment report SHALL contain objective, method, sample, results, exceptions, unresolved risk, limitations, assessor identity, decision authority, effective date and reassessment triggers. Failure of a mandatory protection SHALL result in nonconformity, not a positive label with an undisclosed caveat.

AI.CO.05 A PASC AI 01 certificate or attestation SHALL follow a separately authorised scheme with defined surveillance and appeal rules. AI-system conformity, OSPCRM certification and PASC AI 01 certification SHALL be recorded as distinct claims with their own evidence and scopes.

AI.CO.06 Material changes to intended purpose, affected population, model, data rights, autonomous authority, supplier control or observed harm SHALL trigger re-evaluation before continued claims of equivalent assurance.

Annex A — Minimum demonstrability dossier (Normative)

The dossier SHALL contain: (A1) scoped system register and ownership; (A2) intended purpose, prohibited use and affected-person mapping; (A3) legal and data-provenance assessment; (A4) AI impact assessment and high-impact determination; (A5) measurable quality and safety criteria; (A6)

local-context testing and disaggregated results or justified alternative; (A7) autonomy authorization and intervention tests, if applicable; (A8) human notice, complaint and appeal process; (A9) system and model version lineage; (A10) sovereignty and supplier dependency map; (A11) security evidence and OSPCRM reference when applicable; (A12) monitoring and incident records; (A13) independent acceptance report; and (A14) reassessment plan.

Annex B — Representative assurance scenarios (Informative)

B1 Public-benefit eligibility: record where AI enters decision-making; prove applicable rights and legal basis; test group-specific error costs, local languages, notice, human appeal and reversal of an incorrect rejection.

B2 Agricultural advisory in local languages: establish validity across regions, crops and vernacular usage; test critical misinformation, uncertainty communication, escalation to qualified expertise and offline fallback where relevant.

B3 Autonomous administrative agent: show enforced tool permissions, authenticated approvals for publication or payment, resistance to poisoned documents, action audit, revocation and replayed failure tests.

B4 Procured foundation model in healthcare: show procurement evidence and dependency map, prohibited clinical autonomy boundaries, local validation, safety escalation and provider-update reassessment.

Annex C — Interoperability without duplication (Informative)

Concern	Primary PASC AI 01 requirement	OSPCRM retained requirement
AI lifecycle and responsibility	AI.GV, AI.OP	OSPCRM cyber risk governance when relevant
Data rights, fairness, dignity	AI.DA, AI.HR	No attempt to reframe these as cyber priority scores
Model fitness, safety, drift	AI.QA	OSPCRM cyber performance and incident metrics separately
Agent capability and authority	AI.AG behavioural and human-control requirements	§8.3 AI attack vectors and applicable risk methodology
Sovereignty and supplier control	AI.SO AI-specific model, data and platform control	§8.4–8.5 and core contextual clauses
Evidence and assessment	AI.CO separately scoped assurance	§11 OSPCRM certification demonstrability where claimed

Annex D — External reference map (Informative)

1. PASC, *Open Sovereign Predictive Cyber Risk Management*, v1.0, November 2025, https://doi.org/10.5281/zenodo.17642943.
2. ISO/IEC 42001:2023, *Artificial intelligence — Management system*, https://www.iso.org/standard/81230.html. Management-system alignment; not proof of individual-system safety.
3. ISO/IEC 23894:2023, *Artificial intelligence — Guidance on risk management*, https://www.iso.org/standard/77304.html.
4. NIST, *AI Risk Management Framework 1.0*, https://www.nist.gov/itl/ai-risk-management-framework; Govern/Map/Measure/Manage are reference concepts, not an externally mandated certification.
5. European Union, Regulation (EU) 2024/1689, consolidated AI Act text, https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:02024R1689-20260727. Legal scope and commencement dates SHALL be verified per deployment; no universal PASC date is inferred.
6. OWASP, *Top 10 for LLM Applications*, 2025, https://genai.owasp.org/resource/owasp-top-10-for-llm-applications-2025/; and *Top 10 for Agentic Applications*, https://genai.owasp.org/2025/12/09/owasp-top-10-for-agentic-applications-the-benchmark-for-agentic-security-in-the-age-of-autonomous-ai/.

Annex E — French administrative summary (Informative)

Intitulé : Directive PASC AI 01 — Intelligence artificielle responsable, souveraine et démontrable.
Objet : établir un référentiel autonome couvrant la gouvernance du cycle de vie de l’IA, les droits fondamentaux, la licéité des données, l’équité, la validité des modèles, l’autonomie des agents, le contrôle humain, la transparence et l’évaluation indépendante. La cybersécurité demeure régie par OSPCRM lorsque ce dernier est applicable. Une éventuelle certification PASC AI exigera des preuves vérifiables et un dispositif d’évaluation spécifiquement approuvé.