KnowMyAgents
identity & governance for AI agents

KNOWMYAGENTS · PASSPORTS FOR AGENTS · RULEBOOKS SCORECARD · EPISODE 5 COMPANION

The Rulebooks Scorecard

Nine rulebooks, from the EU AI Act to EC-Council's ADG framework: what each demands of an organisation running AI agents, how KnowMyAgents scores against every requirement, honestly, and a blank sheet to score your own estate the same way.

PrimeThoughts

PrimeThoughts Innovation Private Limited · knowmyagents.com · Edition 1 · 20 September 2026 · Scores are for KnowMyAgents kit version 0.9.6 as verified on a live installation on 15 September 2026, not for planned features. Regulatory dates are as at publication; the EU AI Act dates follow the Digital Omnibus in force since 27 July 2026. Re-issued when the kit or a rulebook changes.

The rulebooks, and the dates that bite

RULEBOOK	STATUS AND DATES	WHAT IT DEMANDS THAT TOUCHES AGENTS
EU AI Act (Regulation 2024/1689) as amended by the Digital Omnibus on AI	In force 1 Aug 2024. Omnibus in force 27 Jul 2026: Annex III high-risk obligations deferred to 2 Dec 2027, Annex I embedded systems to 2 Aug 2028. Art. 50 transparency applies since 2 Aug 2026. Fines up to 7% of global turnover.	Art. 12 automatic record-keeping and traceability; Art. 14 human oversight with the ability to intervene and stop; Art. 26 deployer duties: monitoring, assigned human oversight, logs kept at least six months; Art. 15 accuracy, robustness, cybersecurity
EC-Council ADG AI Governance Framework	Launched 29 May 2026; free and open; advisory board from Citi, JPMorgan Chase, Microsoft, KPMG, Deloitte, NTT Data, GE Healthcare, GlobalLogic, Prudential, Salesforce; aligned to EU AI Act, ISO 42001, NIST AI RMF, OWASP, MITRE ATLAS	Three pillars (Adopt, Defend, Govern); nine governance surfaces; twelve Minimum Controls MC-1 to MC-12 with evidence artefacts; nine deployment overlays; three autonomy tiers HITL, HOTL, HOOTL; four harm classes; four-phase roadmap
NIST AI RMF 1.0 and Generative AI Profile (NIST AI 600-1)	Jan 2023; profile Jul 2024; voluntary	Govern, Map, Measure, Manage; inventory; accountability roles; incident response
ISO/IEC 42001:2023	Dec 2023; certifiable	AI management system: roles, resources and inventory, impact assessment, lifecycle and logging, third parties
OWASP Top 10 for Agentic Applications 2026	Released 9 Dec 2025 by the OWASP Gen AI Security Project	Goal hijack, tool misuse, identity and privilege abuse, agentic supply chain, unexpected code execution, memory and context poisoning, insecure inter-agent communication, cascading failures, human-agent trust exploitation, rogue agents
NIST SP 800-63-4; eIDAS 2.0 and the EU Digital Identity Wallet	SP 800-63-4 final Jul 2025; eIDAS 2.0 in force May 2024, wallets due 2026	Issued credentials, holder binding, selective disclosure, revocation; the EU wallet uses the credential format KnowMyAgents issues
GDPR; DPDP Act 2023 and DPDP Rules 2025 (India)	GDPR in force; DPDP Rules notified 13 to 14 Nov 2025, phased to May 2027	Records of processing, security of processing, purpose limitation, accountability
DORA and NIS2 (EU)	DORA applies since Jan 2025; NIS2 transposition from Oct 2024	ICT logging and integrity, third-party access control, incident reporting
Singapore IMDA Model AI Governance Framework for Agentic AI	Launched 22 Jan 2026 at the World Economic Forum; updated 20 May 2026	Bound risks upfront; humans meaningfully accountable; technical controls and processes; end-user responsibility

Different books, one demand. Name it: an inventory with an owner. Bound it: a purpose, a limit, a time. Prove it: a record nobody can quietly edit. And be able to stop it.

KnowMyAgents scored against each rulebook

Method. Each requirement is scored 0 to 3 for KnowMyAgents version 0.9.6 as verified on the test machine on 15 September, not as planned: 3 KnowMyAgents produces the primary evidence · 2 substantial contribution, another control completes it · 1 indirect or partial · 0 outside KnowMyAgents' scope. Two totals per rulebook: **all** requirements, and **in scope**, meaning requirements where an identity, authorisation, attribution, logging, monitoring, revocation or third-party-trust control is the expected mitigation. Rows marked • are in scope.

47% ADG · all twelve Minimum Controls	**67%** ADG · in-scope controls	**100%** ADG · Identity surface	**61%** ADG · six agentic overlays
52% / 75% EU AI Act · all / in scope	**58% / 67%** NIST AI RMF · all / in scope	**54% / 72%** ISO 42001 · all / in scope	**53% / 81%** OWASP agentic · all / in scope
89% Identity standards (800-63-4, eIDAS 2.0)	**50% / 75%** GDPR and DPDP · all / in scope	**58% / 67%** DORA and NIS2 · all / in scope	**67%** Singapore agentic framework

ADG · nine governance surfaces

SURFACE (ADG DEFINITION)	WHAT KNOWMYAGENTS PROVIDES	SCORE
• **Identity** · credentials, service accounts, delegated authority, agent identity, secrets	Issued per-agent passport; human owner and operator on the passport; private key sealed in the wallet; fresh proof per call; delegation through missions and the actor chain; did:web issuer resolution; revocation; co-existence with the enterprise directory via ID-JAG	3
• **Tools** · APIs, plugins, actions, code execution, MCP capabilities	Every tool call passes the door; capabilities per passport; discovered places inventoried; MCP routes governed; exit overlay for all outbound calls. Missing: per-tool trust tiering and per-invocation argument policy	2
• **Telemetry** · logs, traces, evaluations, replay data	Hash-chained event log, five events per call, actor chain, export, console summaries. Missing: evaluation and fairness metrics (not KnowMyAgents' role)	3
• **Safety Layer** · guardrails, policy engines, semantic filters, classifiers, circuit breakers	Deterministic policy engine at the door; budget circuit breakers (pause, downgrade, ask a human); observe and enforce modes. Missing: semantic filters and content classifiers	2
• **Orchestration** · planners, workflow graphs, multi-step flows, agent routing	Sub-agent spawn under a parent passport; mission scope and stop-at-budget. Missing: planner-level controls	1
Prompt · system prompts, agent instructions	Persona fingerprint: a hash of the instructions, stable across runs, changes when the prompt changes. Detects, does not gate	1
Model · foundation models, routers, versions	Model family on the passport; per-model metering; downgrade at budget threshold	1
Context · retrieval, memory, hidden context	Not governed	0
Learning Loop · training sources, alignment, feedback	Not governed	0
Surfaces total 13 / 27 · in-scope five surfaces 11 / 15		**48% · 73%**

ADG · twelve Minimum Controls, with the evidence artefact each asks for

CONTROL	ADG EVIDENCE ARTEFACT	WHAT KNOWMYAGENTS PRODUCES	SCORE
• **MC-1 AI System Inventory**	Published inventory with named owner per system, reviewed quarterly	Discovery of every agent that touches the door, including shadows; named human owner on every passport; inventory synced to the console and the family tree per program. Quarterly review cadence is the customer's	3
MC-2 Risk Classification	Documented classification by sensitivity, autonomy, exposure, harm, criticality	Assurance tiers 0, 1, 2 classify external exposure; capabilities and missions bound autonomy. No harm or criticality taxonomy	1
• **MC-3 Separation of Duties**	RACI matrix per system; no single function holds deploy, validate, approve	Owner and operator are distinct fields; the console orders and the engine executes; the passport office issues and the engine verifies; the engine cannot mint, the console cannot connect in. No RACI document produced	2
MC-4 Pre-Production Evaluation	Signed evaluation report across four harm classes	Not KnowMyAgents' role	0
• **MC-5 Change Control**	Change log with approvals for prompts, tools, models, sources	Persona fingerprint changes when instructions change and is logged per call; policy and passport changes are logged and hot-reloaded. Detects and records changes; does not gate them	2
MC-6 Context Policy	Published context policy per system: provenance, retention, access, trust ordering	Zones and the exit overlay define what leaves; retention clocks on the log. Nothing on retrieval provenance or trust ordering	1
• **MC-7 Tool & MCP Register**	Published register with per-tool risk assessment, trust tiering, invocation controls	Doors and tool stand-ins are registered; discovered places inventoried; MCP routed through the door; capabilities per passport gate invocation. Missing: per-tool risk assessment and trust tiers	2
• **MC-8 Runtime Monitoring**	Active monitoring for abuse, drift, leakage, unsafe actions, bias; thresholds and SLAs	Every call verified and logged; budget thresholds with pause, downgrade, ask a human; detections (unnamed callers, invalid proofs, teachable challenges); console alerts. Missing: content-level leakage and bias detection	2
• **MC-9 AI Incident Response**	Playbook with replayable evidence capture; annual tabletop	Replayable evidence: hash-chained log with actor chain, exportable; one-step revocation; mission pause. Playbook and tabletop are the customer's	2
MC-10 Periodic Governance Review	Review records, exception dispositions, board reporting for high-risk	Console reporting per agent, program and tenant feeds a review. No review or exception workflow	1
MC-11 Fairness & Bias Evaluation	Fairness evaluation report	Not KnowMyAgents' role	0
• **MC-12 Shared Responsibility Documentation**	Signed responsibility matrix; vendor due diligence	Accepted-issuer lists and tiers define trust in partner agents; console boundary defines what the vendor sees. No contract artefacts	1
Controls total 17 / 36 · in-scope seven controls 14 / 21			**47% · 67%**

ADG · deployment overlays, autonomy tiers, harm classes

ADG ELEMENT	ADG ASKS FOR	WHAT KNOWMYAGENTS PROVIDES	SCORE
• Agent overlay	Business mission, action boundaries, liability model, segregation of duties	Missions with scope and budget; capabilities; human owner as the liable principal; owner and operator separated	3

• Agentic Hardening overlay	Authority bounds, action isolation, kill switches, forensic replay, semantic firewalls	Authority bounds, kill switch (revocation, pause), forensic replay (chained log): yes. Action isolation: partial via the exit. Semantic firewall: no	2
• Tools & MCP overlay	Tool discovery, capability registration, trust tiering, per-invocation policy	Discovery and registration: yes. Trust tiering and per-invocation argument policy: no	2
• Multi-Agent Interop overlay	Agent-to-agent trust, value alignment, collusion detection, cascading prevention	Agent-to-agent trust (teachable challenge, did:web, tiers, agent cards): yes. Collusion detection: no. Cascading prevention: budgets and revocation, partial	2
• LLMOps overlay	Versioning, evaluation, rollback, cost control, performance	Cost control per agent and mission only	1
• Agentic Orchestration overlay	Planner layer, retries, subtask decomposition, stop conditions	Stop conditions through budgets and missions only	1
Context, Pre/Post-training, Multi-Modal overlays	Memory, provenance, training governance, cross-model risk	Outside scope	0
Six agentic overlays 11 / 18			**61%**
Autonomy tiers HITL, HOTL, HOOTL	Controls tied to the tier an agent runs at	Observe mode with human approval at thresholds evidences HOTL; missions with bounds evidence HOOTL; the passport can record which tier a system is declared at. KnowMyAgents makes the tier visible and enforceable at the door	`supports all three`
Harm classes	Technical, societal, operational, systemic	Technical (credential theft, replay, exfiltration path): primary. Systemic (agent-to-agent trust, cascading spend): primary. Operational (attribution of failures): partial. Societal: none	`2 of 4 primary`

EU AI Act

ARTICLE	REQUIREMENT	WHAT KNOWMYAGENTS PROVIDES	SCORE
• Art. 12	Automatic record-keeping; traceability of the system's functioning across its lifetime	Five chained events per call, automatically, with agent, human principal, verdict, budget and outcome	3
• Art. 14	Human oversight: understand, monitor, intervene, stop	Observe and enforce modes; ask-a-human at budget thresholds; one-step revocation; console per-agent view. No output-level review UI	2
• Art. 15	Accuracy, robustness, cybersecurity, including against manipulation	Credential replay defeated; keys never exposed; fail-closed default. Nothing on model accuracy	2
• Art. 26	Deployer duties: monitor, assign human oversight, keep logs at least six months	Monitoring and assigned owner: yes. Note: the kit's default retention is 90 days for calls; set retention to at least 180 days, or schedule exports, to meet the six-month floor	2
Art. 9	Risk management system	Evidence input only	1
Art. 13	Transparency and instructions to deployers	Passport and log are self-describing; not a provider document	1
Art. 50	Transparency to natural persons	Outside scope	0
Total 11 / 21 · in scope 9 / 12			**52% · 75%**

NIST AI RMF

FUNCTION	REQUIREMENT	WHAT KNOWMYAGENTS PROVIDES	SCORE
• Govern	Roles, accountability, policies, culture	Named owner per agent; policy at the door; separation of issuer, verifier and console	2
• Map	Inventory, context, third parties, intended use	Discovery and inventory; missions as intended use; partner agents through tiers and accepted issuers	2
Measure	Metrics, testing, evaluation, monitoring	Runtime metrics and spend per agent; no evaluation or testing	1
• Manage	Respond, prioritise, decommission, incident handling	Revocation, pause, budget actions, exportable evidence	2
Total 7 / 12 · in scope 6 / 9			**58% · 67%**

ISO/IEC 42001 Annex A

CONTROL AREA	REQUIREMENT	WHAT KNOWMYAGENTS PROVIDES	SCORE
• A.3 Internal organisation	Roles, responsibilities, reporting of concerns	Owner and operator on every passport; actor chain	2
• A.4 Resources	Inventory of AI systems, data, tools, computing	Agent inventory; discovered places; per-provider metering	2
A.5 Impact assessment	Assess impacts on individuals and society	Outside scope	0
• A.6 Lifecycle	Objectives, responsible design, verification, deployment, operation and monitoring, event logging	Event logging and operational monitoring: primary. Design and verification: no	3 on logging, 1 elsewhere; counted 2
A.7 Data	Data provenance, quality, preparation	Data residency only	1
A.8 Information for interested parties	Documentation, incident communication	Exportable evidence	1
• A.9 Responsible use	Objectives and intended use; processes for responsible use	Missions bound intended use; policy enforces it	2
• A.10 Third parties and customers	Allocating responsibilities with suppliers and customers	Accepted issuers, tiers, agent cards; console boundary	2
Total 13 / 24 · in scope 13 / 18			**54% · 72%**

OWASP Top 10 for Agentic Applications 2026

ENTRY	RISK	WHAT KNOWMYAGENTS PROVIDES	SCORE
ASI01	Agent goal hijack	Scope and budget limit the blast radius; no content inspection	1
• ASI02	Tool misuse and exploitation	Capabilities per passport; every tool call through the door; logged	2
• ASI03	Identity and privilege abuse	Per-agent identity, sealed keys, fresh proof per call, delegation chain, revocation	3

ASI04	Agentic supply chain vulnerabilities	Trust in partner agents by issuer and tier; nothing on code or model supply	1
ASI05	Unexpected code execution	Outside scope	0
ASI06	Memory and context poisoning	Outside scope	0
• ASI07	Insecure inter-agent communication	Teachable challenge, signed presentations, audience binding, did:web resolution, agent cards	3
• ASI08	Cascading failures	Budgets, mission pause, revocation, one-call-behind metering	2
• ASI09	Human-agent trust exploitation	Attribution with the human first; no content controls	1
• ASI10	Rogue agents	Shadow discovery, tamper-evident record, revocation, enforce mode	3
Total 16 / 30 · in scope 13 / 16 (entry names to be checked against the published PDF)			**53% · 81%**

Identity standards · NIST SP 800-63-4 and eIDAS 2.0

REQUIREMENT	WHAT KNOWMYAGENTS PROVIDES	SCORE
• Issued credential from a verifiable issuer	SD-JWT VC issued under a did:web issuer, the format the EU Digital Identity Wallet uses	3
• Holder binding and proof of possession	cnf key with key-binding JWT over nonce, audience, time and hash	3
• Selective disclosure	Salted disclosures with decoys	3
• Revocation and status	Bitstring status list	3
• Issuer discovery and federation	did:web resolution; issuer allowlists; assurance tiers	3
• Identity proofing of the human principal	Delegated to the enterprise directory (ID-JAG); KnowMyAgents does not proof people	1
Total 16 / 18		**89%**

GDPR and DPDP

REQUIREMENT	WHAT KNOWMYAGENTS PROVIDES	SCORE
• Records of processing (who, what, why)	Per-call record with agent, human principal, mission and outcome; no data categories recorded	2
• Security of processing	Sealed keys, replay-proof presentation, single exit, integrity of the log	2
• Purpose limitation	Missions and capabilities bind purpose per agent	2
• Minimisation and residency	Content never leaves the data plane; console receives summaries only; selective disclosure on the passport	3
Impact assessment (DPIA)	Evidence input only	1
Data principal rights	Outside scope	0
Total 10 / 18 (rights excluded: 9 / 12) · in scope 9 / 12		**50% · 75%**

DORA and NIS2

REQUIREMENT	WHAT KNOWMYAGENTS PROVIDES	SCORE

• Logging integrity and retention	Hash-chained, exportable; retention configurable (see the six-month finding)	3
• Access control for machine identities and third parties	Per-agent identity; partner agents by issuer and tier; revocation	2
• Detection and incident reporting	Detections logged and summarised; no reporting workflow	1
ICT third-party risk register	Accepted issuers and discovered places are inputs; not a register	1
Total 7 / 12 · in scope 6 / 9		**58% · 67%**

Singapore Model AI Governance Framework for Agentic AI

DIMENSION	WHAT KNOWMYAGENTS PROVIDES	SCORE
• Assess and bound risks upfront	Missions, capabilities, budgets and tiers bound each agent before it acts; no risk assessment method	2
• Make humans meaningfully accountable	Human principal on every passport and first on every log line; ask-a-human at thresholds	3
• Technical controls and processes	Identity, proof per call, policy, revocation, tamper-evident log; multi-agent and third-party agent guidance from the May 2026 update is directly met by tiers and agent cards	2
Enable end-user responsibility	Transparency to end users is the deployer's; KnowMyAgents gives the evidence	1
Total 8 / 12		**67%**

Your own scorecard

Score your estate the same way, one requirement at a time: **3** we produce the primary evidence · **2** substantial, another control completes it · **1** indirect or partial · **0** not in place. Name the evidence an auditor could hold, and the owner. Total each rulebook as a percentage of the maximum, and separately for the rows an identity, authorisation and evidence layer is meant to answer.

REQUIREMENT	0	1	2	3	EVIDENCE WE CAN SHOW	OWNER · DATE
EU AI ACT						
Art. 12 record-keeping	□	□	□	□		
Art. 14 human oversight	□	□	□	□		
Art. 15 robustness and cybersecurity	□	□	□	□		
Art. 26 deployer duties incl. six-month logs	□	□	□	□		
Art. 9 risk management	□	□	□	□		
Art. 13 transparency to deployers	□	□	□	□		
Art. 50 transparency to persons	□	□	□	□		
EC-COUNCIL ADG, TWELVE MINIMUM CONTROLS						
MC-1 AI system inventory	□	□	□	□		
MC-2 risk classification	□	□	□	□		
MC-3 separation of duties	□	□	□	□		
MC-4 pre-production evaluation	□	□	□	□		
MC-5 change control	□	□	□	□		
MC-6 context policy	□	□	□	□		
MC-7 tool and MCP register	□	□	□	□		
MC-8 runtime monitoring	□	□	□	□		
MC-9 AI incident response	□	□	□	□		
MC-10 periodic governance review	□	□	□	□		
MC-11 fairness and bias evaluation	□	□	□	□		
MC-12 shared responsibility documentation	□	□	□	□		
ADG, NINE GOVERNANCE SURFACES						
Model	□	□	□	□		
Prompt	□	□	□	□		
Context	□	□	□	□		
Tools	□	□	□	□		
Orchestration	□	□	□	□		
Identity	□	□	□	□		
Safety layer	□	□	□	□		
Telemetry	□	□	□	□		
Learning loop	□	□	□	□		
NIST AI RMF						

Govern	□	□	□	□
Map	□	□	□	□
Measure	□	□	□	□
Manage	□	□	□	□
ISO/IEC 42001 ANNEX A				
A.3 internal organisation	□	□	□	□
A.4 resources	□	□	□	□
A.5 impact assessment	□	□	□	□
A.6 lifecycle	□	□	□	□
A.7 data	□	□	□	□
A.8 information for interested parties	□	□	□	□
A.9 responsible use	□	□	□	□
A.10 third parties and customers	□	□	□	□
OWASP TOP 10 FOR AGENTIC APPLICATIONS 2026				
ASI01 goal hijack	□	□	□	□
ASI02 tool misuse	□	□	□	□
ASI03 identity and privilege abuse	□	□	□	□
ASI04 agentic supply chain	□	□	□	□
ASI05 unexpected code execution	□	□	□	□
ASI06 memory and context poisoning	□	□	□	□
ASI07 insecure inter-agent communication	□	□	□	□
ASI08 cascading failures	□	□	□	□
ASI09 human-agent trust exploitation	□	□	□	□
ASI10 rogue agents	□	□	□	□
IDENTITY STANDARDS (NIST SP 800-63-4, EIDAS 2.0)				
Issued credential from a verifiable issuer	□	□	□	□
Holder binding and proof of possession	□	□	□	□
Selective disclosure	□	□	□	□
Revocation and status	□	□	□	□
Issuer discovery and federation	□	□	□	□
Identity proofing of the human principal	□	□	□	□
GDPR AND DPDP				
Records of processing	□	□	□	□
Security of processing	□	□	□	□
Purpose limitation	□	□	□	□
Minimisation and residency	□	□	□	□
Impact assessment	□	□	□	□
Data principal rights	□	□	□	□

DORA AND NIS2				
Logging integrity and retention	□	□	□	□
Access control for machine identities and third parties	□	□	□	□
Detection and incident reporting	□	□	□	□
ICT third-party risk register	□	□	□	□
SINGAPORE MGF FOR AGENTIC AI				
Assess and bound risks upfront	□	□	□	□
Make humans meaningfully accountable	□	□	□	□
Technical controls and processes	□	□	□	□
Enable end-user responsibility	□	□	□	□

Bands, per rulebook: under 40 percent, travelling without papers; 40 to 69 percent, papers for some, not all; 70 percent and above, passport-ready, now prove it.

Method note: each requirement is scored 0 to 3 for kit version 0.9.6; "in scope" counts the requirements where an identity, authorisation, attribution, logging, monitoring, revocation or third-party-trust control is the expected mitigation. KnowMyAgents holds no certification against any of these frameworks; the mapping is self-assessed and published so that it can be challenged. "Tamper-evident" is used deliberately, never "immutable". Corrections to hello@primethoughts.com.