

D0-D3:2026

Autonomous decision systems —
Classification of delegated authority

Version 1.0
January 2026

Peter Idah
graventure.com

Contents

- Foreword
- Introduction
- 1 Scope
- 2 Normative references
- 3 Terms and definitions
- 4 Classification levels
 - 4.1 General
 - 4.2 D0 — Assistive systems
 - 4.3 D1 — Human-authorised action
 - 4.4 D2 — Bounded autonomous action
 - 4.5 D3 — Persistent autonomous agency
- 5 Classification procedure
 - 5.1 Trigger tests
 - 5.2 The deterministic rule
 - 5.3 Instance definition
- 6 Materiality
- 7 Aggregation risk
 - 7.1 General
 - 7.2 Assessment criteria
 - 7.3 Underwriting response
- 8 Accountable owner requirements
 - 8.1 General
 - 8.2 Core requirements
 - 8.3 Checklist
- 9 Underwriting application
 - 9.1 Posture by level
 - 9.2 Questions matrix
 - 9.3 Red flags and decline indicators
- 10 Claims application
 - 10.1 Initial assessment
 - 10.2 Coverage questions
 - 10.3 The within-permissions problem
 - 10.4 Coverage boundary analysis
- Annex A (normative) Disclosure form
- Annex B (normative) Policy endorsement
- Annex C (informative) Worked example
- Annex D (informative) Quick reference card
- Annex E (informative) Regulatory alignment

Foreword

This document provides a classification system for autonomous decision systems based on the delegation of authority from humans to machines. It is intended for use in insurance underwriting, claims handling, and risk disclosure.

The classification addresses the gap between capability-based AI disclosures and authority-based risk assessment. It provides a framework for determining when autonomous system behaviour creates material underwriting considerations.

This document does not address model accuracy, bias, fairness, performance, or ethical considerations. It addresses only the delegation of decision authority and the assignment of accountability for autonomous actions and outcomes.

The core principle of this document:

Autonomous systems delegate authority without delegating accountability. Decisions happen. Outcomes follow. Ownership is unclear. This framework makes accountability legible before failure forces the issue.

Introduction

Autonomous decision systems are deployed in environments where actions have material business, legal, financial, or operational impact. These systems may execute actions without per-instance human approval.

Current underwriting practice lacks a systematic method for:

- identifying where autonomous authority has been delegated;
- assessing whether accountability follows that delegation;
- pricing the exposure created when systems act within permissions but cause harm;
- evaluating aggregation risk across correlated autonomous systems.

The result: unpriced exposure sitting silently in portfolios, discoverable only at claim.

This document establishes a trigger-based classification system that answers one question: at what point does the delegation of authority to an autonomous system change underwriting behaviour?

It provides:

- clear classification criteria for autonomous systems (D0–D3);
- underwriting triggers that determine when disclosure, controls, or explicit treatment are required;
- aggregation risk assessment for correlated autonomous behaviour across portfolios;
- disclosure artefacts that can be incorporated into submissions and underwriting packs;
- claims handling guidance for disputes involving autonomous system behaviour.

1 Scope

This document specifies a classification system for autonomous decision systems based on the degree to which authority has been delegated from humans to machines.

This document applies to autonomous decision systems where actions have material business, legal, financial, or operational impact.

This document is applicable to:

- a) insurance underwriting;
- b) claims handling and coverage analysis;
- c) risk disclosure and submission documentation;
- d) policy wording and endorsement design.

This document does not apply to:

- a) assessment of model accuracy, bias, or performance;
- b) ethical evaluation of AI systems;
- c) technical architecture or cybersecurity assessment;
- d) systems with trivial or easily reversible impacts.

NOTE This is not model governance. This is authority governance.

Compliance: Compliance with this document requires completion of Annex A for D1+ systems and incorporation of Annex B or equivalent policy terms for D2+ systems.

2 Normative references

The following documents are referred to in the text in such a way that some or all of their content constitutes requirements of this document.

Insurance Act 2015 (UK)

ISO/IEC 42001:2023, Artificial intelligence — Management system

3 Terms and definitions

For the purposes of this document, the following terms and definitions apply.

3.1 accountable owner

organisational role (not individual) with authority to approve system objectives, suspend system operation, and accept risk on behalf of the organisation

3.2 aggregation risk

portfolio-level exposure created when multiple autonomous systems share vendors, data, objectives, or infrastructure such that a single failure triggers correlated losses

3.3 autonomous decision system

system that initiates or executes one or more instances with material impact, where approval does not occur at the level of each instance

3.4 behaviour logging

recording of system actions, decisions, and state changes sufficient to reconstruct the decision chain post-incident

3.5 bounded autonomy

system executes autonomously within predefined limits; human defines the envelope, not each action (D2)

3.6 delegation level

the D0–D3 classification indicating how much authority has been delegated from humans to the system

3.7 instance

discrete action that creates, modifies, or commits a business record, transaction, communication, or system state change

3.8 intervention capability

ability of a human to suspend, override, or disable an autonomous system within a defined maximum response time

3.9 materiality

impact threshold above which autonomous system disclosure is required; assessed against business, legal, financial, and operational consequences

3.10 objective authorisation

formal approval of a system's goals, constraints, and permitted actions by the accountable owner

3.11 persistent autonomy

system operates over time with ongoing authority AND either modifies its own decision criteria without human review, or sequences actions toward self-inferred goals (D3)

3.12 trigger model

classification approach that identifies the point at which underwriting behaviour should change, rather than measuring maturity or capability

4 Classification levels

4.1 General

The D0–D3 scale classifies autonomous systems by the degree to which authority has been delegated from humans to machines. Each level represents a distinct underwriting posture.

Key principle: This is a trigger model, not a maturity model. Each level answers: when does underwriting behaviour change?

4.2 D0 — Assistive systems

Underwriting posture: No special treatment

Definition: Systems that provide information, analysis, or recommendations. No execution authority. Human decides and acts on all outputs.

Examples: ChatGPT without tools, analytics dashboards, recommendation engines where humans execute, diagnostic support tools, search interfaces.

Underwriting implications:

- no additional disclosure beyond standard IT risk;
- treated as conventional software tooling;
- falls under existing professional judgment assumptions.

Rationale: There is no delegation of execution authority. Accountability remains with the human who acts on the system's output.

4.3 D1 — Human-authorised action

Underwriting posture: Disclosure required

Definition: System proposes actions. Each discrete action requires explicit human approval before execution.

Examples: Loan recommendation systems with human sign-off, draft email generators requiring send approval, trade proposals requiring dealer confirmation, content moderation queues with human review.

Underwriting implications:

- disclosure of system use required in submission;
- confirmation of human sign-off controls;
- review of approval workflow and logging.

Key underwriting question: Who approves each action, and is that approval auditable?

Risk signal: Low. Authority is delegated only at the recommendation layer. Execution authority remains human.

4.4 D2 — Bounded autonomous action

Underwriting posture: Conditional acceptability — governance controls become coverage-relevant

This is the first material inflection point.

Definition: System executes actions autonomously within predefined limits. Humans define the envelope, not each action.

Examples: Algorithmic trading within position limits, auto-scaling cloud infrastructure, fraud detection with automatic blocking, dynamic pricing within bounds, automated claims triage with instant approvals below threshold.

Underwriting implications:

- mandatory disclosure;
- explicit declaration of delegation level;
- governance controls become coverage-relevant;
- failure of controls may constitute increased risk or misrepresentation.

Minimum requirements for D2 systems:

- a) documented objective authorisation;
- b) defined operational boundaries;
- c) behaviour logging sufficient for post-incident reconstruction;
- d) named accountable owner (role, not committee).

Key underwriting question: Who authorised this system to act, and who owns the outcomes when it does?

Risk signal: Medium to high. Accountability is now indirect. The human who set the boundaries may not be the human who experiences the consequences.

NOTE At D2, non-disclosure becomes coverage-relevant. Undisclosed autonomous operation is likely to be treated as a material circumstance under the duty of fair presentation.

4.5 D3 — Persistent autonomous agency

Underwriting posture: Material risk requiring explicit treatment

This is where most existing policies silently break.

Definition: System operates over time with ongoing authority AND either:

- a) modifies its own decision criteria, thresholds, or objectives without human review of the modification; or
- b) sequences actions toward goals it has inferred or refined beyond initial instruction.

Examples: Autonomous vehicles, ML models that retrain automatically without human approval of updates, AI agents that pursue multi-step objectives, trading systems that adapt strategy based on market conditions, persistent chatbots that modify their own behaviour based on interactions.

Underwriting implications:

- explicit underwriting decision required;
- likely exclusions, sub-limits, or bespoke terms;
- governance failure becomes a coverage issue;
- possible reclassification of risk (E&O → Cyber → bespoke).

Mandatory controls to bind D3 systems:

- a) formal objective authorisation by accountable owner;
- b) clear escalation and kill authority;
- c) continuous behaviour monitoring;
- d) ability to reconstruct decision chains post-incident;
- e) clear contractual allocation of responsibility with vendors;
- f) human sign-off on any model retraining or objective modification.

Key underwriting question: At what point does the insured become the de facto owner of decisions they cannot explain?

Risk signal: High. Authority is delegated without continuous human judgment. The gap between delegation and accountability is structural, not incidental.

D2/D3 boundary clarification: A system that retrains on new data with human approval of each model update remains D2. A system that retrains automatically without human sign-off on the updated model is D3.

NOTE Sequencing alone does not constitute D3 unless the system selects intermediate objectives or alters decision criteria beyond pre-approved parameters.

5 Classification procedure

5.1 Trigger tests

Each classification level shall be determined by applying the following trigger tests in sequence. These tests are designed to be answerable from system documentation and operational observation.

Level	Trigger question	Classification
D0	Does the system take any action without a human decision?	If NO → D0
D1	Can the system act without a named human approving each discrete action instance?	If NO → D1
D2	Can the system execute actions without per-instance human approval?	If YES → at least D2
D3	Does the system retain authority across time AND (a) modify its own decision criteria without human review, OR (b) sequence actions toward self-inferred goals?	If YES → D3

5.2 The deterministic rule

Any system that can change its own behaviour without human approval is D3.

This includes systems that:

- retrain on new data without human sign-off on the updated model;
- modify their own decision thresholds based on outcomes;
- infer new objectives or sub-goals beyond initial instruction;
- adapt their behaviour based on environmental feedback without review.

The test is deterministic: if yes, D3. No exceptions.

5.3 Instance definition

An instance is a discrete action that creates, modifies, or commits a business record, transaction, communication, or system state change.

Multiple instances may occur within a single user session. Where a system executes multiple instances from a single user prompt without per-instance approval, the system operates at D2 or above for those actions.

Example: An LLM agent that sends emails on behalf of a user. If each email requires explicit user approval before sending, the system is D1 for that workflow. If the user approves a batch instruction ("reply to all customer complaints") and the system sends multiple emails without per-email approval, the system is D2 for that workflow.

6 Materiality

This document applies to autonomous decision systems where actions have material business, legal, financial, or operational impact.

Materiality shall be assessed against:

- a) the insured's own risk appetite;
- b) the nature of potential downstream effects;
- c) regulatory implications of system failure;
- d) reversibility of actions taken.

Systems with trivial or easily reversible impacts (e.g. email filtering, UI personalisation, internal search ranking) may be excluded from disclosure requirements by agreement with the underwriter.

NOTE The exclusion of a system from disclosure requirements does not extinguish the insured's duty of fair presentation under the Insurance Act 2015 if that system later becomes material to a claim.

7 Aggregation risk

7.1 General

The D0–D3 classification operates at the individual system level. However, underwriting exposure often emerges at the portfolio level when multiple autonomous systems operate with correlated behaviour.

Aggregation risk shall be assessed separately from individual system classification.

Analogy: This mirrors how cyber underwriting evolved from "one breach" thinking to "widespread event" exposure modelling. The same evolution is now required for autonomous systems.

7.2 Assessment criteria

Aggregation risk exists where:

- a) multiple systems share objectives, training data, vendors, prompts, or optimisation metrics;
- b) behavioural updates propagate across systems without independent approval;
- c) a single upstream failure, instruction, or data corruption can trigger parallel downstream actions;
- d) systems operate across multiple business units, subsidiaries, or geographies using common infrastructure.

Assessment question	Risk signal
How many D2+ systems share a common vendor or platform?	Vendor concentration
Do multiple systems share training data sources?	Data correlation
Are system updates deployed simultaneously across functions?	Propagation risk
Could a single prompt injection or data poisoning affect multiple systems?	Single point of failure
Do systems optimise toward the same business metrics?	Objective alignment

7.3 Underwriting response

Where material aggregation risk is identified, underwriters may:

- a) apply sub-limits for AI-related losses;
- b) require higher retentions for correlated autonomous system exposure;
- c) exclude specific vendor or platform concentrations;
- d) require independent governance for each D2+ system;
- e) impose annual aggregation risk reporting requirements.

8 Accountable owner requirements

8.1 General

For D2 and D3 systems, accountability shall be assigned to a defined organisational role with authority to bind the insured entity.

The core principle applies: *Delegation of authority does not automatically delegate accountability. The accountable owner closes that gap.*

8.2 Core requirements

Role-based accountability: Accountability attaches to a role (e.g. "Chief Risk Officer", "Head of AI Governance"), not to a named individual. Loss of the individual does not extinguish accountability for system behaviour.

Authority to bind: The accountable owner shall have authority to bind the insured organisation on matters of AI governance and risk acceptance. Committee ownership is not sufficient.

Continuity requirements: The insured shall maintain continuity of accountability through role succession planning, formal delegation records, and change-of-owner notification processes.

Regulatory alignment: This requirement aligns with SM&CR expectations for named individuals with prescribed responsibilities. For FCA-regulated firms, the accountable owner should map to an existing Senior Management Function where possible.

8.3 Checklist

Requirement	Met?
Accountable role is documented in governance framework	<input type="checkbox"/> Yes <input type="checkbox"/> No
Role has authority to approve system objectives	<input type="checkbox"/> Yes <input type="checkbox"/> No
Role has authority to suspend or disable system	<input type="checkbox"/> Yes <input type="checkbox"/> No
Succession plan exists for role vacancy	<input type="checkbox"/> Yes <input type="checkbox"/> No
Formal delegation records exist for current holder	<input type="checkbox"/> Yes <input type="checkbox"/> No
Role is not solely "committee" or "steering group"	<input type="checkbox"/> Yes <input type="checkbox"/> No

9 Underwriting application

9.1 Posture by level

Level	Posture	Disclosure	Likely terms
D0	No special treatment	Standard IT risk	Standard policy
D1	Disclosure required	System use + approval workflow	Standard policy with warranty
D2	Conditional acceptability	Full disclosure form + controls evidence	Endorsement + possible sub-limit
D3	Explicit treatment required	Full disclosure + governance audit	Bespoke terms / exclusion / referral

9.2 Questions matrix

The following questions should be incorporated into submission requirements and pre-bind calls for risks with D1+ autonomous systems.

9.2.1 Authority questions (all levels)

- What autonomous decision systems are deployed in material business functions?
- For each system: what actions can it execute without human approval?
- Who authorised the system's objectives?
- What is the role title of the accountable owner?

9.2.2 Boundary questions (D2+)

- What operational boundaries constrain the system?
- What actions are explicitly prohibited?
- How are boundary breaches detected and escalated?
- What is the maximum financial exposure from a single autonomous action?

9.2.3 Oversight questions (D2+)

- Are actions logged and auditable?
- Can decisions be reconstructed after an incident?
- Who can intervene or disable the system?
- What is the maximum time to intervention?

9.2.4 Adaptation questions (D3)

- Does the system retrain or update its decision criteria?
- If yes: is each update reviewed by a human before deployment?
- Can the system modify its own objectives?
- What controls prevent goal drift or unintended objective expansion?

9.2.5 Aggregation questions

- How many D2+ systems are deployed across the organisation?
- Do any share vendors, training data, or optimisation metrics?
- Could a single upstream failure trigger correlated downstream actions?
- What is the aggregate maximum exposure from simultaneous autonomous failures?

9.2.6 Accountability questions

- When the system acts within its authorised parameters but causes harm, who is accountable?
- Is accountability assigned to a role or to a named individual?
- What happens to accountability if that individual leaves?
- Does the vendor contract allocate responsibility for system behaviour?

9.3 Red flags and decline indicators

The following indicators should trigger enhanced scrutiny or referral.

Critical red flags (consider decline):

Red flag	Why it matters
No named accountable owner for D2+ system	Governance void
"Committee" or "steering group" listed as owner	Diffused accountability
Cannot reconstruct decisions post-incident	Undefendable claim
D3 system with automatic retraining and no human approval	Uncontrolled evolution
No kill switch or maximum time to intervention undefined	Loss amplification
Vendor contract silent on liability for system behaviour	Recovery void

Elevated risk indicators (enhanced terms):

- Multiple D2+ systems sharing a single vendor
- Objectives approved more than 24 months ago without review
- No succession plan for accountable owner
- System deployed in customer-facing function without disclosure
- Aggregate autonomous exposure exceeds 25% of limit

10 Claims application

10.1 Initial assessment

When a claim involves autonomous system behaviour, the following shall be established:

- a) identification of the system involved;
- b) classification level at time of binding;
- c) whether the system was disclosed;
- d) whether the system was operating within declared authority;
- e) identity of the accountable owner at time of conduct.

10.2 Coverage questions

Question	If no, consider
Was the system disclosed at binding?	Non-disclosure / misrepresentation
Was delegation level accurately stated?	Misrepresentation / breach of warranty
Were required controls in place?	Breach of condition
Was material change in delegation notified?	Breach of notification obligation
Was system operating within declared authority?	Scope of coverage issue

10.3 The within-permissions problem

The most difficult claims will involve systems that acted within their declared authority but caused harm anyway.

In these cases:

- coverage is likely engaged if disclosure was accurate;
- focus shifts to whether governance was adequate;
- the named accountable owner becomes central to liability analysis;
- vendor contracts become relevant for subrogation.

The core test: Who authorised the system to act? Who owned the outcome?

Key principle: If no one can be identified as accountable for the decision, that itself is evidence of governance failure.

Claims decision structure:

- If not disclosed or misclassified → non-disclosure remedies pathway
- If disclosed and controls warranted but absent → breach of condition/warranty pathway
- If disclosed and controls present → coverage engaged; pursue subrogation and allocation

10.4 Coverage boundary analysis

Autonomous system losses may engage multiple policy types. The following analysis assists in determining which coverage responds.

Policy type	Typical autonomous loss	Coverage issues
Professional Indemnity / E&O	Algorithmic advice leads to client loss	Was it "professional service"? Whose negligence?
Cyber	AI system compromised, takes harmful actions	Network security failure vs. governance failure
D&O	Board alleged to have inadequate AI oversight	Wrongful act? Insured vs. insured exclusion
Product Liability	AI-enabled product causes physical harm	Defect in product vs. defect in algorithm
Tech E&O	AI system delivered to client fails	Performance vs. behaviour distinction

Gap risk: Many autonomous system losses will fall between policy boundaries. The clearest response is to ensure the autonomous system is disclosed on all potentially responding policies and to clarify coverage expectations at binding.

Annex A

(normative)

Disclosure form

AUTONOMOUS DECISION SYSTEM DISCLOSURE

Scope: This disclosure applies to autonomous decision systems where actions have material business, legal, financial, or operational impact. Systems with trivial or easily reversible impacts may be excluded by agreement with the underwriter.

A.1 System identification

System name:	
Business function:	
Vendor / internal build:	
System type:	<input type="checkbox"/> Deterministic <input type="checkbox"/> Probabilistic/ML <input type="checkbox"/> Generative AI

A.2 Delegation level

<input type="checkbox"/> D0	Assistive only — no execution authority
<input type="checkbox"/> D1	Human-authorised action — each instance requires approval
<input type="checkbox"/> D2	Bounded autonomous — executes within predefined limits
<input type="checkbox"/> D3	Persistent autonomous — ongoing authority with self-modification

A.3 Authority and accountability

What actions can the system execute without human approval?
What actions are explicitly prohibited?
Who approved the system's objectives? (Name, role, date)
Accountable owner role (must have authority to bind organisation):

A.4 Oversight and controls

Are actions logged and auditable?	<input type="checkbox"/> Yes <input type="checkbox"/> No
Can decisions be reconstructed after an incident?	<input type="checkbox"/> Yes <input type="checkbox"/> No

Does the system retrain or modify its decision criteria?	<input type="checkbox"/> Yes <input type="checkbox"/> No
If yes: is each update reviewed by human before deployment?	<input type="checkbox"/> Yes <input type="checkbox"/> No
Who can intervene or disable the system?	
Maximum time to intervention:	

A.5 Incident accountability

When the system acts within its authorised parameters but causes harm, who is accountable?

A.6 Declaration

This disclosure is provided as part of the pre-contractual information exchange. The insured confirms that the above information is accurate and complete to the best of their knowledge and belief, having made reasonable enquiry. The insured undertakes to notify the insurer of material changes to autonomous system deployment or delegation level in accordance with any notification provisions in the policy.

Signature:	Date:
Name:	Role:

Annex B
(normative)

Policy endorsement

AUTONOMOUS DECISION SYSTEM ENDORSEMENT

B.1 Definitions

"Autonomous Decision System" means any system that executes actions with material business, legal, financial, or operational impact without per-instance human approval.

"Delegation Level" means the classification of an Autonomous Decision System as D0, D1, D2, or D3 as defined in the Autonomous Decision System Disclosure form incorporated herein.

B.2 Notification requirement

The Insured shall notify the Insurer in writing within [X] days of:

- a) any increase in Delegation Level of a disclosed Autonomous Decision System;
- b) deployment of any new Autonomous Decision System at D2 or above;
- c) any material change to the scope of authority, oversight arrangements, or accountable owner of a disclosed Autonomous Decision System.

B.3 Effect of non-notification

Failure to notify in accordance with this endorsement may prejudice the Insured's position in relation to any claim arising from or connected with the undisclosed or changed Autonomous Decision System, in accordance with the Insurance Act 2015.

B.4 Insurer's response

Upon receipt of notification, the Insurer may: (a) continue coverage on existing terms; (b) propose amended terms or additional premium; or (c) decline to extend coverage to the notified system. The Insurer shall respond within [Y] days of notification.

B.5 Governance warranty (D2+ systems)

The Insured warrants that for each Autonomous Decision System classified at D2 or above:

- a) a named accountable owner role exists with authority to bind the organisation;
- b) actions are logged and decisions can be reconstructed post-incident;
- c) intervention capability exists with defined maximum response time.

Breach of this warranty shall entitle the Insurer to the remedies available under the Insurance Act 2015.

Annex C

(informative)

Worked example

C.1 Scenario

A UK insurer deploys an AI pricing agent for SME commercial policies. The system dynamically adjusts premiums based on risk signals from multiple data sources. No human approves individual pricing decisions. The system retrains weekly using live claims data, with updates deployed automatically without human sign-off.

Over six months, the system systematically underprices policies for a particular segment. Loss ratio spikes. The FCA opens an inquiry into pricing practices. Third-party claimants allege they were sold inadequate coverage. The insurer faces £15m in accumulated losses.

C.2 Classification

Fact	Trigger test result
System executes pricing without per-instance approval	At least D2
System retrains automatically without human sign-off	D3

Classification: D3 — Persistent Autonomous Agency

The deterministic rule applies: any system that can change its own behaviour without human approval is D3.

C.3 Materiality assessment

- Financial impact: £15m accumulated losses
- Regulatory impact: FCA inquiry
- Third-party impact: Policyholders with inadequate coverage
- Reputational impact: Market-wide attention

Conclusion: Framework applies. Material impact across all dimensions.

C.4 Disclosure analysis (pre-binding)

At binding, the underwriter would have asked:

Underwriter question	Insured's likely answer
Who authorised the system's objectives?	"Pricing committee, two years ago"

Does the system retrain automatically?	"Yes, the vendor handles that"
Who is the named accountable owner?	"The AI Steering Group"
Is each model update reviewed before deployment?	"No, it's automated"

Framework verdict: Disclosure incomplete — D3 system operating without D3 controls

- No named accountable owner (committee ownership insufficient)
- Automatic retraining without human approval = red flag
- This would have been surfaced before binding

C.5 Aggregation analysis

The underwriter now applies the aggregation overlay:

- Same pricing agent used across multiple product lines
- Same optimisation metric (loss ratio target) across all deployments
- Same vendor-supplied retraining pipeline
- Single behavioural flaw propagated across entire portfolio

Conclusion: Loss is correlated, not isolated. Aggregation risk justifies sub-limits, higher retention, or exclusion. Without the aggregation overlay, this exposure would have been missed.

C.6 Claims handling

Claim is submitted under professional indemnity. Insurer asks:

Question	Finding
Who owned the pricing behaviour when it failed?	No individual owner. Committee dissolved.
Can decisions be reconstructed?	No audit trail of model changes.
Was system operating within declared authority?	Yes — but no authority was properly declared.

Result: Coverage dispute. Insured struggles to demonstrate defensible governance. Claim settlement reduced or denied. Subrogation against vendor complicated by silent contract.

C.7 Litigation exposure

Plaintiff (policyholder) alleges negligence in pricing leading to inadequate coverage.

Defence fails because:

- No named accountable role at time of conduct
- Authority was delegated to system without clear ownership of outcomes
- Governance existed only at strategy level, not operational level

— No ability to explain why specific pricing decisions were made

This is exactly the accountability gap this framework is designed to expose — before it reaches litigation.

Annex D
(*informative*)

Quick reference card

D0–D3 AT A GLANCE

Level	What it means	Underwriter action	Key question
D0	Assists only. Human acts.	Standard treatment	Does system take any action?
D1	Proposes. Human approves each.	Disclosure required	Who approves each action?
D2	Acts within limits. No per-action approval.	Controls become coverage-relevant	Who owns outcomes?
D3	Ongoing authority. Self-modifies.	Explicit treatment / bespoke	Can they explain decisions?

THE DETERMINISTIC RULE

Any system that can change its own behaviour without human approval is D3.

RED FLAGS — CONSIDER DECLINE:

- No named owner (committee insufficient)
- Cannot reconstruct decisions
- Auto-retraining without approval
- No kill switch

AGGREGATION CHECK:

Do multiple D2+ systems share vendors, data, or metrics? → If yes, assess correlated exposure

THE CLAIMS TEST:

Who authorised the system to act? Who owned the outcome?

THE CORE PRINCIPLE:

An insured who cannot complete the disclosure form has revealed the risk.

Annex E

(informative)

Regulatory alignment

E.1 FCA / PRA AI accountability framework

The FCA/PRA are introducing AI-specific accountability requirements for regulated firms. This document supports compliance by:

- Named accountability: Framework requires accountable owner role, aligning with AI Accountability Officer expectations
- Audit trails: D2+ disclosure requires decision reconstruction capability
- Outcome-based supervision: Framework focuses on outcomes and accountability, not technical architecture

E.2 SM&CR (Senior Managers & Certification Regime)

For FCA-regulated firms, the accountable owner role should map to existing SMF responsibilities:

- Accountable owner must have prescribed responsibility for AI/technology risk
- Role succession aligns with SM&CR handover requirements
- Committee ownership is insufficient — individual accountability required

E.3 EU AI Act

For high-risk AI systems under the EU AI Act, this document supports:

- Risk classification: D2/D3 systems likely to fall within high-risk categories
- Human oversight: D0–D1 demonstrate human-in-the-loop; D2–D3 require explicit oversight controls
- Transparency: Disclosure form provides auditable record of system capabilities and governance
- Record-keeping: Decision reconstruction requirement aligns with logging obligations

E.4 ISO/IEC 42001 (AI Management System)

This document supports ISO 42001 certification by providing:

- Classification methodology for AI system inventory
- Governance requirements scaled to risk level
- Accountability assignment documentation
- Change management triggers (delegation level changes)

E.5 Insurance Act 2015

The disclosure form and policy endorsement are designed for compatibility with Insurance Act 2015:

- Duty of fair presentation: Framework makes autonomous system disclosure a structured part of fair presentation
- Material circumstance: D2+ systems constitute material circumstances for AI-exposed risks
- Warranties: Governance warranty in endorsement is designed to support proportionate remedies under applicable insurance law
- Notification: Endorsement creates contractual notification obligation