The Decision-Grade PMO: tables

Every table from the book, at full width. Free, and yours to print, mark up, or adapt with attribution. The figures are a separate download.

Nothing here replaces the book. A table is an instrument, and the chapter around it is what tells you how to use it and where it stops being appropriate. Table 1.1 in particular will give you a stage; only Chapter 1 tells you what to do about it.

Two things to carry with you. Bramwold, Stellar and Brackmere are composite cases, and the numbers in them illustrate a pattern rather than reporting measurements from a single programme. Figure 4.1, in the figures download, is an original diagram after a categorisation published by the Canadian Nuclear Safety Commission, the Office for Nuclear Regulation and the US Nuclear Regulatory Commission; the diagram is the author's own and the underlying categorisation is the regulators'.

Companion to the v1.96 edition, 30 August 2026. Master SHA-256 4052653b83bd5021. If your copy of the book disagrees with a table here, the book is the source of truth and this pack is out of date.

Reader	Chapter sequence
Foundation reader (new to AI in PMO)	Ch 1 → Ch 2 → Ch 3 → one of Ch 7–12 → Ch 17 → Ch 21
PMO Director / Head of Project Controls	Ch 1 → Ch 3 → Ch 4 → Ch 6 → Ch 16 → Ch 20 → Ch 21 (about 6 hours)
Digital and Data lead / AI governance partner	Ch 3 → Ch 4 → Ch 5 → Ch 13 → Ch 14 → Ch 19 → Appendix C (about 8 hours)
Head of Project Controls / PMO Manager (full path)	Read in order, Ch 1–21 and Appendices A–F; in regulated contexts follow the Stellar vignettes, in commercial contexts the Bramwold thread (about 20 hours over a fortnight)
Programme controls practitioner (domain specialist)	Ch 2 → your domain chapter (Ch 7, 8, 9, 10, 11, or 12) → Ch 18 → Ch 21
Sceptic	Ch 19 (what goes wrong) → Ch 15 (trust) → Ch 3 and Ch 4 (governance constraints) → Ch 20 (measuring value) → Ch 2 if persuaded
3–4 hour strategic architecture	Preface → Ch 1–4 → Ch 16 → Ch 17 → narrative of Ch 21 (about 3–4 hours)

Reader	Chapter sequence
Minimum viable implementation	Ch 5 → Ch 6 → one relevant domain chapter (Ch 7–12) → Ch 17 → Ch 19

Table 0.1, Reading-route quick reference. This table is the single reading-route map for the book: find your role or your time budget and follow the chapter sequence in that row.

Acronyms at a Glance

Acronym	Meaning	Acronym	Meaning
NC	Nerve Centre: integrated PMO operating model (this book)	**LoA**	Level of Autonomy (0–4): what AI is permitted to do within a service
KDE	Key Data Element: a data point whose quality affects decisions	**FM**	Failure Mode: a specific way an AI capability can fail in a PMO
DAS	Decision-Augmented System: the NC model's AI-integrated architecture	**PMO**	Project Management Office
EVM	Earned Value Management	**CPI**	Cost Performance Index (Earned Value ÷ Actual Cost)
SPI	Schedule Performance Index (Earned Value ÷ Planned Value)	**EAC**	Estimate at Completion
BAC	Budget at Completion	**TCPI**	To-Complete Performance Index (cost)
TSPI	To-Complete Schedule Performance Index	**QRA**	Quantitative Risk Analysis
WBS	Work Breakdown Structure	**CBS**	Cost Breakdown Structure
RACI	Responsible, Accountable, Consulted, Informed	**KPI**	Key Performance Indicator
ROI	Return on Investment	**FTE**	Full-Time Equivalent
PEP	Project Execution Plan	**LFE**	Learning From Experience
AI	Artificial Intelligence (this book: automation, ML, GenAI, agents)	**ML**	Machine Learning

Acronym	Meaning	Acronym	Meaning
GenAI	Generative AI (LLMs and diffusion models)	**RAG**	Retrieval-Augmented Generation (AI grounded in specific documents)
API	Application Programming Interface	**NEC**	New Engineering Contract (NEC3/NEC4 suites)
ONR	Office for Nuclear Regulation (UK)	**SNI**	Sensitive Nuclear Information
NISR	Nuclear Industries Security Regulations 2003	**IPA**	Infrastructure and Projects Authority (UK)
PMI	Project Management Institute	**APM**	Association for Project Management (UK)
PMBOK	Project Management Body of Knowledge (PMI)	**ISO 19650**	Information management in built-asset sector
ISO 42001	AI Management Systems (AIMS) standard	**MSP**	Managing Successful Programmes
P3M3	Portfolio, Programme & Project Management Maturity Model	**CAE**	Claims, Arguments, Evidence
AMLAS	Assurance of Machine Learning for use in Autonomous Systems	**CNSC**	Canadian Nuclear Safety Commission
USNRC	US Nuclear Regulatory Commission	**IAEA**	International Atomic Energy Agency
IET	Institution of Engineering and Technology	**SRO**	Senior Responsible Owner
NISTA	National Infrastructure and Service Transformation Authority (UK), formerly the IPA	**BIM**	Building Information Modelling
EPC	Engineering, Procurement and Construction	**HAZOP**	Hazard and Operability study
P3O	Portfolio, Programme and Project Offices	**BoE**	Basis of Estimate

Dimension	Stage 1: Reactive	Stage 2: Defined	Stage 3: Analytical	Stage 4: Predictive	Your Stage
Data & Cut-offs	Spreadsheets, manual exports; no consistent cut-off.	Some systems-of-record; cut-offs exist but not always enforced.	Defined Data Products with cut-offs, versioning, and ownership.	Automated extract-validate-publish; lineage logged.	☐
Governance & Decisions	Meetings happen; decisions informal or not recorded.	Forums exist; action logs kept, but not always.	Decision rights defined; approvals recorded; evidence links present.	Exception-led rhythm; AI-assisted options papers; full audit trail.	☐

Reporting	Late, bespoke packs; no standard format; debates about numbers.	Standard templates; packs mostly on time; still backward-looking.	Decision-spine packs; exceptions identified; options proposed.	Automated assembly; constrained GenAI drafting; Publication Gates.	☐
Evidence & Audit	No consistent evidence; document hunts when challenged.	Some version control; approvals logged in some areas.	Publication records standard; evidence index for key outputs.	Full Evidence Spine; retrieval on demand; sampling routine.	☐
Technology	Point tools; heavy spreadsheet reliance; no integration.	Scheduling and cost tools in use; limited integration; manual reconciliation.	Governed Data Products; API/extract patterns; access controls.	AI services (automation, ML, GenAI) with gates, logging, safe mode.	☐

Table 1.1, Digital PMO maturity self-assessment. Record one stage per row; the stage recorded most often is your starting stage, and where two tie, take the lower.

Your Stage	**Where to start**	**Chapters most relevant right now**
Stage 1: Reactive	Chapters 4–6	Ch 4 (governance and decision rights), Ch 5 (data readiness), Ch 6 (operating process embedding). Do not add AI capability until these are stable.
Stage 2: Defined	Chapters 4–6, then 7–12	Close the Ch 4–6 gaps first: at Stage 2 the gate and the evidence discipline are not yet in place. Then Ch 7–12 (pick one domain service and Thin-Slice it) and Ch 17 (roadmap and sequencing).
Stage 3: Analytical	Chapters 3, 13 and 14	Ch 3 (AI types and constraints), Ch 13 (architecture), Ch 14 (assurance and audit), Ch 16 (what the predictive state looks like in operation). Apply AI in its constrained role alongside governance already in place.
Stage 4: Predictive	Chapters 15, 18, 20–21	Ch 15 (stakeholder trust), Ch 18 (roles and capability uplift), Ch 20 (value measurement), Ch 21 (decision advantage and sustained operation).
Stage 5: Autonomous	Chapters 19–21	Ch 19 (failure modes, ethics, monitoring), Ch 20 (value measurement and continuous improvement), Ch 21 (sustained decision advantage). Foundation built progressively through Stages 2–4.

Table 1.2, Reading path by starting stage

NC Domain Service	Key Decision Supported	Ch.
Planning & Scheduling	Schedule health, baseline drift, near-term milestone risk, logic integrity	7
Estimating & Cost Modelling	Cost variance, EAC/ETC forecast credibility, estimate validation	8
Risk & Quantitative Risk Analysis	Risk exposure, QRA input quality, emerging risk detection, risk–cost linkage	9
Change & Contract Management	Change exposure, entitlement, contract compliance, commercial position	10
Resource & Capability Management	Demand vs capacity balance, skills gaps, resource conflict resolution	11
Reporting, Performance & Insight	Pack assembly, narrative consistency, KPI monitoring, insight briefs	12

Table 2.1, The Six NC Domain Services: decisions supported and chapter references

Level	What the AI may do	What a human must review before the output is used	Accountable for the resulting decision
LoA 0, No AI	Nothing. The output is produced by a human using conventional tooling. Deterministic automation is classified here as well, and that placement describes the tooling rather than exempting it from control.	The output itself, on the same terms as any other manual product.	The service owner.
LoA 1, Assist	Observe and generate. AI drafts a component, flags a pattern, or summarises a source set. It does not produce the finished output.	The drafted component against its named sources, and the parts of the output the AI did not touch.	The service owner, who completes and publishes.

Level	What the AI may do	What a human must review before the output is used	Accountable for the resulting decision
LoA 2, Recommend	Observe, generate and recommend. AI produces the first draft end to end, including a proposed reading of the data.	Every material claim and number, clause by clause, against the Evidence Index. The recommendation is a proposal, not a finding.	The responsible discipline manager (project controls, planning, cost or commercial), with the Project Manager carrying overall responsibility for the report.
LoA 3, Route	Observe, generate, recommend and route. AI coordinates outputs across services under explicit rules and directs them to named owners. It does not publish.	The combined artefact, the routing decisions taken, and the declared LoA of each contributing service.	The Project Manager for project-level artefacts, or the Portfolio Manager for portfolio synthesis. The PMO Director operates the Gate but does not sign project content.
LoA 4, Publish (restricted)	Execute. AI completes a defined action without per-instance human sign-off, inside pre-declared bounds.	The policy, the bounds and the exception log, on a defined cycle. There is no per-instance review; that is what the rung means, and why it is restricted.	The sponsor, on the policy rather than on the instance. Accountability is not delegated to the system.

Table 3.1, Decision rights by Level of Autonomy: what the AI may do, what a human must review, and who is accountable for the decision that follows

Level	Who may accept, reject or override	Mandatory escalation conditions	Audit evidence required on every output
LoA 0	The service owner, through the normal approval route.	Any use of the output beyond its stated purpose or audience.	Source references, cut-off, reviewer and approver.
LoA 1	The service owner accepts, amends or discards the drafted component without having to justify doing so. The assistance is advisory.	A drafted component that cannot be traced to a named source. Any claim the model introduced that no source supports.	As LoA 0, plus the AI contribution identified in the Evidence Index and the tool, model and version used.

Level	Who may accept, reject or override	Mandatory escalation conditions	Audit evidence required on every output
LoA 2	The named reviewer may accept, amend or reject clause by clause, and must record substantive amendments.	A hallucinated commercial, contractual or safety claim. A material number that cannot be reconciled to a KDE inside its refresh cadence. A disagreement between reviewer and model that the reviewer cannot settle from evidence.	As LoA 1, plus the model draft preserved alongside the reviewer amendments, and a confidence note on any KDE refreshed against a provisional source.
LoA 3	The signing human may reject the combined artefact and return any routed component. Any named owner in the route may refuse a routed item, and must say why.	A routing failure that sent an output to the wrong owner or the wrong audience. A contributing service operating outside its declared LoA. A conflict between services that the routing rules do not resolve.	As LoA 2, plus the routing rule version, the route actually taken, and every refusal or reroute with its reason.
LoA 4	The sponsor may withdraw the authorisation at any time. Any named recipient may challenge an instance and require it to be withdrawn and reviewed.	Any instance outside the pre-declared bounds. Any change in the consequence class of the output. Any breach detected after publication.	Full ex-post traceability for every instance: inputs, rule version, action taken, timestamp, and the standing authorisation it was executed under.

Table 3.2, Control points by Level of Autonomy: override rights, mandatory escalation, and the audit evidence required on every output

Level	Are the actions reversible?	Monitoring	Independent assurance
LoA 0	Fully. Nothing is released without human action.	Normal service performance review.	Standard assurance sampling.
LoA 1	Fully. The human completes and publishes.	The service owner reviews AI-assisted components in the routine cycle.	Sampling at the assurance partner's discretion.
LoA 2	Reversible before publication. After publication, correction follows the normal errata route.	Track the reviewer amendment rate and the Gate failure rate attributable to the model. A falling amendment rate is a signal to test, not a signal to relax.	Assurance samples the model draft against the published output at a defined frequency.

Level	Are the actions reversible?	Monitoring	Independent assurance
LoA 3	Reversible before publication. Routing actions already taken may not be. Design the route so that no irreversible action occurs before sign-off.	Track routing accuracy, refusal rate and time in route, alongside the LoA 2 measures for each contributing service.	Independent verification of the routing rules before first use and after any rule change.
LoA 4	Restricted to actions that are reversible and low in consequence. If an action cannot be undone, it does not belong at this rung.	Continuous monitoring against the declared bounds, with an alert on any out-of-bounds instance and a standing exception report to the sponsor.	Independent verification before authorisation, and scheduled re-verification at a stated interval. The authorisation lapses if the re-verification is not completed.

Table 3.3, Reversibility, monitoring and independent assurance by Level of Autonomy

Key Data Element (KDE)	Quality Standard	Owner / Gate
Schedule activities	Every work-package activity has a WBS code, driving logic and a resource assignment. Milestones are scoped out of the resource test only and keep the driving-logic test, because a milestone has no resources assigned. Hammock and level-of-effort activities keep the resource test but are scoped out of the driving-logic test, and are listed, because their dates are calculated from the activities they span. A level-of-effort activity with its own set duration is not exempt. No open ends on critical work packages. Every activity with negative float is reported, with the cause recorded and a named owner. Negative float is not removed by changing a constraint, a calendar or a logic link without a recorded reason.	Planning Data Owner; validated at cut-off

Key Data Element (KDE)	Quality Standard	Owner / Gate
Cost commitments	CBS allocation complete with no uncommitted spend above threshold. Actuals, commitments, and forecast fields populated and reconciled.	Cost Data Owner; reconciled at month-end
Risk register items	All active risks have probability, impact, owner, mitigation, and a WBS or activity reference for the work they affect. Risks whose subject is not work that this programme plans are scoped out of the reference test: the field carries a scoped-out value rather than a blank, the exemption is listed, and the risk is routed to a named owner at the level it does sit at. These are typically funding, regulatory approvals not yet in the plan, market capacity, workforce availability and interfaces beyond this WBS. A reference above the level the work is planned at is a false link, not a reference. Reviewed and signed off at each cut-off date.	Risk Coordinator; validated before Publication Gate
Change events	Notification date, awareness date, contract reference, status and cost impact populated for every event. Time-bar status is derived from those fields, not asserted here: whether a notice was served in time is a contractual determination and stays with the Commercial Manager.	Commercial Lead; validated against change log at cut-off
Resource demand	Demand mapped to work-package level and reconciled against the current approved schedule baseline. Versioned snapshot at cut-off.	Resource Lead; reconciliation note filed

Table 5.1, Key Data Element (KDE) Quality Criteria Matrix: minimum standards for governance-grade data inputs

Readiness criterion	The question the gate asks	It fails when
Ownership	Is there a named owner for this data, not a team?	The owner is a function, a system, or a vacancy.
Common definitions	Does the measure mean the same thing across every project and supplier that feeds it?	Two projects can both be right and still disagree.
Lineage	Can the value be traced back to the system and the record it came from?	The trail stops at a spreadsheet.
Completeness	Is the field populated where it matters, and is the gap pattern understood?	Missingness is unexplained, or concentrated in one supplier, project type or period.
Timeliness	Is the record refreshed inside the cadence the decision needs?	The refresh cadence is slower than the decision cycle it feeds.
Accuracy	Has the value been checked against the source, rather than against the rule?	The quality check tests presence rather than correctness.
Interoperability	Can this join to the other data it must be read alongside?	The join needs a manual mapping that only one person can perform.
Stable identifiers	Do the keys survive replanning, restructuring and contract change?	Activity or package identifiers are reissued between cycles.
Access permissions	Is access controlled, current, and correct for the audience of the output?	Permissions were set at mobilisation and never re-tested.
Classification	Is the sensitivity marking present and reviewed?	Classification is inherited, assumed, or older than the review interval.
Historical depth	Is there enough consistent history for the intended analysis, counted in delivery cycles rather than months?	The history exists but the definitions changed inside it.
Outcome labels	For any predictive use, is there a recorded answer to learn from?	Outcomes were never captured, or were captured only for the successes.
Known limitations	Are the weaknesses written down where a user of the data will actually see them?	The limitations are known to the data team and to nobody else.

Readiness criterion	The question the gate asks	It fails when
Remediation ownership	For every gap, is there a name and a date?	The gap is recorded as a risk with no owner and no target.

Table 5.2, Data-readiness gate criteria: what is assessed, the question the gate asks, and what constitutes a fail

Step	Purpose	Output	Who
1, Service catalogue mapping	Map NC services to their data inputs, processing needs, and publication requirements	Data Product requirements per service	PMO Lead + Data Architect
2, Define Data Products	Define Data Products with owners, cut-offs, lineage, and access rules (not ad hoc extracts)	Data Product catalogue and ownership register	Data & Tooling Owner
3, Integration patterns	Design stable extract-validate-publish patterns that survive system changes	Integration architecture diagram and runbook	Data Architect + IT
4, Identity, access, audience variants	Control who can see what, by service and confidentiality level, from the outset	Access control matrix and audience variant rules	Security/IT + PMO Lead
5, Evidence Spine (logging)	Log every AI output, decision, and publication to support retrieval under assurance	Evidence log schema and retention policy	Data Architect + Tooling Owner
6, Safe AI retrieval	Design retrieval boundaries for GenAI and agents constrained to approved sources only	Retrieval scope definition and prohibited source list	Data Architect + PMO Lead
7, Environments and release management	Separate development, test, and production; control releases so changes do not break live services	Environment register and release runbook	IT + Tooling Owner

Table 13.1, The seven architecture steps: purpose, output and owner. Use it as the agenda for a first briefing with a data architect or IT function.

Requirement	What it means in practice
Partitioned by period first	The period is the primary dimension. A submission belongs to a period, and the set for that period closes when the cut-off passes. This is what makes period-on-period comparison mechanical rather than reconstructive.

Requirement	What it means in practice
A defined set per project, per period	The PMO knows what should arrive before it arrives. A missing dataset is detectable as an absence rather than discovered later as a hole in a report.
Fixed shape and coding	Structure, field names and coding fixed in advance, so ingestion is deterministic. Where projects genuinely differ, the difference is recorded as a mapping rule under change control, not absorbed by somebody's spreadsheet.
Metadata carried as properties	Project, period, dataset, version, cut-off date, owner and sign-off status, held as properties that can be validated rather than encoded in filenames that decay.
Submission as a controlled act	Loading data is a governed step with an owner and a timestamp, not a file copy. What was submitted, by whom and when is itself part of the Evidence Spine.
Raw data only, plus a structured attestation	Datasets, not packs. The one exception is the discipline attestation: a structured dataset, not free prose, in which the accountable discipline manager records each material movement against defined fields (driver, quantum, evidence reference, corrective action, action owner, confidence). It is authored and signed by the discipline manager, carried in the same submission, under the same cut-off. Discipline narrative for the pack is authored on its own route (Chapter 12), not here. Everything else at this interface is data.

Table 13.2, The Submission Contract: the six requirements that make a handover a controlled interface rather than a filing convention

Dimension	RAG as commonly deployed	DAS
Starting point	Retrieves over a text corpus	Reads the Evidence Spine with KDEs as first-class citizens
Figure origin	Composes figures from retrieved text	Reads them from the authoritative KDE
Traceability	Cites the passage retrieved, not the figure's source of record	Carries a lineage path per figure
Refresh awareness	Has none by default	Enforces it through the KDE refresh cadence
Confidence handling	Carries confidence implicitly in phrasing	Declares it per KDE
Provisional data	Blends provisional and authoritative sources unless the corpus carries the distinction and the retriever filters on it	Flags the provisional component and offers a defensible alternative

Dimension	RAG as commonly deployed	DAS
Audit posture	Reconstructed after the fact	Legible in-flight
Gate cost	Expensive, because defensibility is constructed retrospectively	Cheap, because defensibility is inherent in the answer

Table 13.3, RAG and DAS compared across the eight dimensions that decide whether an answer can be published, and at what cost

Link in the chain	How it fails	What assurance tests
Source system	Stale data. The extract ran on schedule, but the source had not been updated since the previous cycle.	The age of the source record at the cut-off, not the age of the extract.
Extraction	Partial or silently truncated extracts. Records excluded by a filter that nobody remembers setting.	Record counts and control totals reconciled between source and extract, every cycle.
Transformation	Incorrect schedule, cost, risk or work-breakdown mappings. A renamed cost code lands in the wrong bucket and the total still balances.	Mapping versions under change control, with a reconciliation run on any mapping change.
Data-quality controls	Rules that pass because they test presence rather than correctness. A populated field is not a correct field.	A sample of records tested against the source, not against the rule.
Model	The failures the validation requirements in Table 16.2 exist to prevent: drift, miscalibration, and use outside the conditions it was validated for.	The validation record, its date, and whether the conditions it assumed still hold.
Business rules	Thresholds and overrides applied after the model that change the answer without changing the record.	Every post-model rule identified, versioned, and visible in the Evidence Index.
Retrieval and citation	The right answer attached to the wrong source, or a citation to a document that does not contain the claim.	Citations opened and read, on a sample. A reference that is never followed is not evidence.
Access control	The output is correct and the audience is wrong. Classification and audience-variant rules fail quietly.	Distribution lists tested against classification, and permissions re-tested after any role change.
Interface and human review	Misreading of what the screen is saying. Automation bias, where the reviewer agrees because the system proposed it. Overrides made and not recorded.	Reviewer amendment rates, override logs, and a check that a dissenting review is possible and has actually happened.

Link in the chain	How it fails	What assurance tests
Decision and audit record	The decision is taken and the record does not capture what it rested on, so the chain cannot be walked backwards.	A retrieval test: take a published decision at random and reconstruct the chain end to end.

Table 14.1, The decision chain: where an AI-assisted output fails, and what assurance tests at each link

Dimension	The question the plan answers	Where the evidence comes from
Product quality: is the capability any good?		
Accuracy	How well does it do the task, measured how, and against what?	The validation record, Table 16.2; for a drafting service, periodic blind review
Robustness	How does it behave on inputs it was not built for?	Sensitivity and scenario testing
Reliability	Does it give the same answer from the same inputs, and is it available when the cycle needs it?	Reproducibility test at the Gate; service availability record
Explainability	Can a reviewer see why the output says what it says, in terms they can argue with?	Reviewer walkthrough; the Evidence Index
Reproducibility	Can the output be regenerated from the same snapshots at the same cut-off?	The evidence standard, Chapter 5
Security	Who can reach the data, the prompts, the logs and the outputs?	Access control tests, Chapter 13
Privacy	What personal or people data does it touch, and on what basis?	The privacy assessment, Chapter 19
Data dependency	Which KDEs must be healthy for this to work, and what happens when one is not?	The KDE quality matrix, Table 5.1
Resilience to missing or abnormal data	What does it do when an input is absent, late or implausible? Does it say so, or does it fill the gap?	Negative testing; the safe-mode trigger list
Quality in use: is it any good for the people who rely on it?		

Dimension	The question the plan answers	Where the evidence comes from
Decision usefulness	Does the output change what the forum can decide, or only what it can read?	The value hypothesis and scorecard, Chapter 20
User comprehension	Do the people who must act on it read it the way it was meant?	Review of misreadings; audience-variant testing, Chapter 15
Controllability	Can a user stop it, narrow it, or run it differently?	The safe-mode trigger list, Chapter 4
Operational reliability	Does it hold up across a full governance cycle, under real load and real deadlines?	Cycle-level service performance measures
Accessibility	Can it be used by people with different access needs and different levels of technical fluency?	Accessibility check on the output formats
Consequences of error	What is the worst credible outcome of a wrong output reaching a decision, and who carries it?	Consequence classification, Chapter 19
Ability to challenge or override	Can a user disagree, and does disagreement leave a record?	Override log; the contestability route, Chapter 15
Fitness for the governance purpose	Is this the right instrument for the decision it is being used in?	The LoA declaration and the Publication Gate

Table 14.2, The AI Quality Plan: what the plan must answer for a service, and where the evidence for each answer comes from

Stage	What happens	Who owns it	What is recorded
1. Signal	A model, rule or threshold produces an exception flag. The flag states what changed, against what baseline, with what confidence, at what cut-off. It states nothing about what to do.	Service owner	The flag, the model and version that produced it, the data cut-off, and the confidence or interval attached to it.
2. Human review	A named reviewer tests the flag in proportion to its consequence class. A low-consequence flag gets a short check; a governance-facing flag gets the full evidence walk. The reviewer confirms, downgrades or dismisses, and records which.	The named reviewer for the service	The review decision, the evidence consulted, the time taken, and the reason for any dismissal.

Stage	What happens	Who owns it	What is recorded
3. Intervention	A surviving flag is given an intervention with an accountable owner and a date. Monitor is a valid intervention. Noted is not.	The decision owner in the relevant forum	The intervention, its owner, its date, the forum that set it, and the options considered.
4. Outcome	What actually happened is recorded against the intervention at a stated review point, whether or not it is flattering. An intervention with no recorded outcome is an open item, not a closed one.	The intervention owner	The outcome, the date it was assessed, and whether it is attributable to the intervention or to something else.
5. Learning	Two questions are answered separately. Was the signal right? Was the intervention effective? A correct signal followed by an ineffective intervention teaches something different from a false signal followed by wasted effort.	Service owner, with the assurance partner	The evaluation of both questions, and the dated change made to a threshold, a model, a process or a governance rule, or the recorded decision that no change is warranted.

Table 16.1, The learning loop: the five stages that separate a predictive signal from a governed decision, with the owner and the record for each

Requirement	What it establishes	What done looks like
Comparison with the existing method	That the model is better than what it replaces, not merely better than nothing.	The current forecasting method run on the same data over the same period, with the difference stated in the units the forum already uses.
Project-level holdout testing	That performance is not an artefact of the projects the model was fitted on.	Whole projects held out of training, not random rows. Rows from the same project leak. Projects do not.
Temporal, or out-of-time, validation	That the model works forward in time, which is the only direction it will ever be used in.	Trained on data up to a cut-off and tested on the period after it, with no information from the test period available at training.
Coverage across project types and lifecycle stages	That the model is not silently restricted to the conditions it was built on.	Performance reported separately by project type and by lifecycle stage, not as a blended average that hides a weak segment.

Requirement	What it establishes	What done looks like
Calibrated prediction intervals	That the stated uncertainty means what it says.	An 80% interval that contains the outcome about 80% of the time in test, with the observed coverage reported alongside the nominal.
False-positive and false-negative analysis	That the error the model makes is the error the organisation can afford.	Both rates stated at the operating threshold actually used, with the cost of each named. See Failure mode 10 in Chapter 19.
Sensitivity and scenario testing	Which inputs the output actually depends on, and how it behaves at the edges.	Output movement under plausible input variation, and behaviour tested against missing, late and abnormal data.
Data and concept drift assessment	Whether the relationship the model learned is still the relationship that holds.	Input distributions and model performance monitored over time, with defined limits and a named owner for the alert.
Monitoring for regime change	That an economic, organisational or delivery-regime shift has not invalidated the training conditions.	A named list of the conditions the model assumes (contract form, supply market, delivery structure, reporting regime) and a review triggered when one of them changes.
Documented limitations and prohibited uses	Where the model must not be used, stated by the people who built it.	A written statement of scope, known weaknesses and out-of-bounds uses, published with the model rather than held separately.
Retraining, recalibration and withdrawal triggers	That the model has a defined end, not only a beginning.	Triggers agreed in advance for retraining, for recalibration and for withdrawal from service, with the withdrawal decision owned by a named person.

Table 16.2, Validation requirements to be evidenced before a predictive model is approved for operational use

Template 18.1, Role Profile Delta Table

Traditional PMO role	Decision-Grade PMO role	Capability change required
PMO Lead	Service Portfolio Owner	From process governance to service ownership: managing a catalogue of decision services, each with its own LoA, gate, and assurance commitment.
PMO Manager	Service Owner (per domain)	Accountable for one domain service end-to-end: data quality, gate operation, output defensibility. Reads QA reports as a first-class artefact.

Traditional PMO role	Decision-Grade PMO role	Capability change required
Reporting Analyst	Evidence Spine Custodian	From building decks to maintaining a controlled, queryable evidence base. SQL/data-model literacy. Comfortable refusing to publish.
Cost Engineer	Cost Service Owner	Same craft, plus accountability for the AI-assisted forecast challenge service: knowing when to override, knowing when to escalate.
Planner / Scheduler	Planning Service Owner	Same craft, plus operating the schedule-quality QA service at LoA 1–2: deciding which flagged exceptions are real.
Risk Manager	Risk Service Owner	Same craft, plus operating QRA-quality and risk-pattern detection at LoA 2: signing off on cluster outputs before they reach the gate.
Assurance Lead	Assurance Partner	From periodic review to control-test design: building and operating a sampling regime against published outputs and the gates that produce them.
(new)	Data & Tooling Owner	Accountable for the integration layer, data foundation and AI service health. Sits inside the PMO, not in central IT. Speaks both languages.
(new)	AI Service Reviewer	Human-in-the-loop role for LoA 2–3 outputs. Domain-trained, AI-literate, with documented sign-off authority and a published refusal record.

What is recorded at publication	Why it cannot wait
Model and version identifier	A model updated by a vendor changes the output without anything in the PMO changing. Without the version, a later difference cannot be explained or defended.
Prompt, template or workflow version, where it materially shapes the output	The instruction determines the answer as much as the data does. Governing the data and leaving the instructions uncontrolled is a half-built control.
Source and retrieval record	Which documents and fields were actually retrieved, not which were available. Retrieval is where a defensible answer and a merely plausible one diverge.
Named human reviewer	A role is not accountable. A person is.
Approval timestamp	Establishes what was known at the point of sign-off, which is the only fair basis on which the decision can later be judged.
Record of substantive amendments	Distinguishes a reviewed output from a rubber-stamped one, and is the evidence base for the amendment-rate measure in Table 3.3.

Traditional PMO role	Decision-Grade PMO role	Capability change required
Retention and deletion arrangement	Applies to the output and to everything behind it. An indefinite default is a decision, and usually the wrong one.	

Table 19.1, The provenance record: what is captured when an AI-assisted governance output is published, and why each item cannot be added afterwards

Measure	What it establishes, and what it does not
Recommendation relevance	That what the system offered was worth looking at. It says nothing about what the system did not offer.
Issue-detection recall	The proportion of known issues the system found. It depends entirely on what was counted as a known issue, and it says nothing about how much noise arrived with them.
Precision	The proportion of flags that turned out to be real. High precision with low recall is a system that is right about the few things it chooses to mention.
False-alert rate	The cost the system imposes on reviewers. It is a workload measure, and optimising it alone is the failure described in Failure mode 10 in Chapter 19.
Forecast error	How far the numbers were out. A different question entirely from whether the flags were right, and not answered by any of the measures above it.
Reviewer agreement	That people concurred with the output. Agreement is not correctness, and rising agreement may be automation bias rather than improving quality.
Intervention effectiveness	That the action taken worked. Requires the outcome to have been recorded, and requires the counterfactual to have at least been considered.
Improved project outcomes	That the programme is better off. This is the only one the business case is actually about, it is the hardest to attribute, and it is the one most often claimed on the strength of the first four.

Table 20.1, Eight performance measures that are routinely reported as though they were one, and what each actually establishes

Measure	Baseline	Post (12 months)	Change
Decision cycle time (exception → decision), mean working days	18.0	9.2	−49%

Measure	Baseline	Post (12 months)	Change
Option window protection rate (% decisions made before window closes)	41%	78%	+37 pp
Unplanned executive escalations per quarter	4.1	1.3	−68%
Assurance rework hours per quarterly reporting cycle	380	142	−63%
Cost forecast error at quarter close (quarter-start forecast → quarter-end actual), mean absolute	11.8%	4.4%	−63%
Period-close cycle time (cut-off → steering pack published), working days	11	6	−45%
Publication Gate compliance (% outputs with complete evidence chain)	72%	98%	+26 pp

Table 20.2, Case A measurement set. Baseline reconstructed over the twelve months before Month 0; post-implementation measured over Months 3 to 15.

Measure	Baseline	Post (18 months)	Change
Evidence retrieval SLA, mean hours to controlled pack	72	4	−94%
Reports returned for clarification at Independent Assurance Review	62%	11%	−51 pp
Audit findings per review cycle: minor / major	11 / 1	2 / 0	−9 / −1
ONR interaction turnaround: request → controlled submission, weeks	3.5	0.9	−74%
Period-close cycle, working days	15	8	−47%
Assurance rework per cycle, FTE-months	3.1	0.6	−81%
Evidence lineage traceable to source within one click	48%	96%	+48 pp

Table 20.3, Case B measurement set. Baseline captured over the preceding eighteen months; post-implementation measured at the end of the sixth quarter after go-live.

Model in this book	Themes in the standard that it addresses	Where it is set out
The Nerve Centre model	Responsible and value-led AI adoption; tailoring for organisational and delivery context. The Nerve Centre is the arrangement the principles have to be applied through: services, owners, cadence and gates, rather than tools.	Chapters 2 and 6
Levels of Autonomy 0 to 4	Human accountability; governance and risk management. The ladder converts oversight from a stated intention into a declared, per-output placement carrying named decision rights, override rights, escalation conditions and audit evidence.	Chapters 3 and 4; Tables 3.1 to 3.3
The learning loop	Lifecycle management; responsible and value-led AI adoption. It keeps signal, review, intervention, outcome and learning apart, and feeds evaluated lessons back into thresholds, models, processes and governance arrangements.	Chapter 16; Table 16.1
The predictive-maturity model	Lifecycle management; tailoring for organisational and delivery context. It stages the introduction of predictive capability against demonstrated data and control conditions rather than against ambition, with validation required before operational approval.	Chapter 16; Table 16.2
The Data Product lifecycle	Data quality; lifecycle management. It treats decision-critical data as a product with an owner, a definition, a quality standard, a refresh cadence and a readiness gate.	Chapters 5 and 13; Tables 5.1 and 5.2
Governance and assurance arrangements: the Publication Gate, the Evidence Spine, the Master Failure Mode Taxonomy and the AI Quality Plan	Governance and risk management; ethical and legal considerations; stakeholder expectations; data quality. They supply the release control, the traceable record, the named failure catalogue and the quality artefact that make the principles testable.	Chapters 4, 14, 15 and 19; Appendices C and E

Table B.1, Crosswalk: how the models in this book relate to the themes addressed by The Standard for Artificial Intelligence in Portfolio, Program, and Project Management (PMI, 2026)