AI For U&I

OPERATIONAL AI GOVERNANCE

AI Deployment Governance Lifecycle

ADGL

A Practical Governance Methodology for Deploying AI into Production

DESIGNED FOR ORGANISATIONS IMPLEMENTING

- Microsoft Copilot
- ChatGPT Enterprise
- Azure OpenAI
- Internal AI assistants
- Customer support AI
- Agentic AI
- AI embedded in SaaS products

> This document sets out exactly how AI governance is implemented, from discovery to production.

VERSION 1.0 · JULY 2026

Ramya Amballa · Founder, AI For U&I

www.aiforui.org

THE GAP

Why ADGL Exists

Every AI system eventually reaches production. Governance should arrive before it does.

Most organisations deploying AI already hold:

- ISO 27001
- SOC 2
- GDPR compliance programme

Yet they still lack an operational governance process for deploying AI into production.

What existing certifications do not cover

- The behaviour of a system that generates novel output from customer data
- Who owns the decision to accept an AI risk, and where that is recorded
- Whether personal data in a vector store can be erased on request
- Whether retrieval is scoped so one customer cannot reach another's data
- Whether human review is real, or has become a formality
- Whether an incident can be investigated from the evidence retained

> ADGL fills that gap by providing a practical, implementation-first governance lifecycle that transforms AI governance from policy into operational controls.

WHAT THE FRAMEWORK PROVIDES

Structure

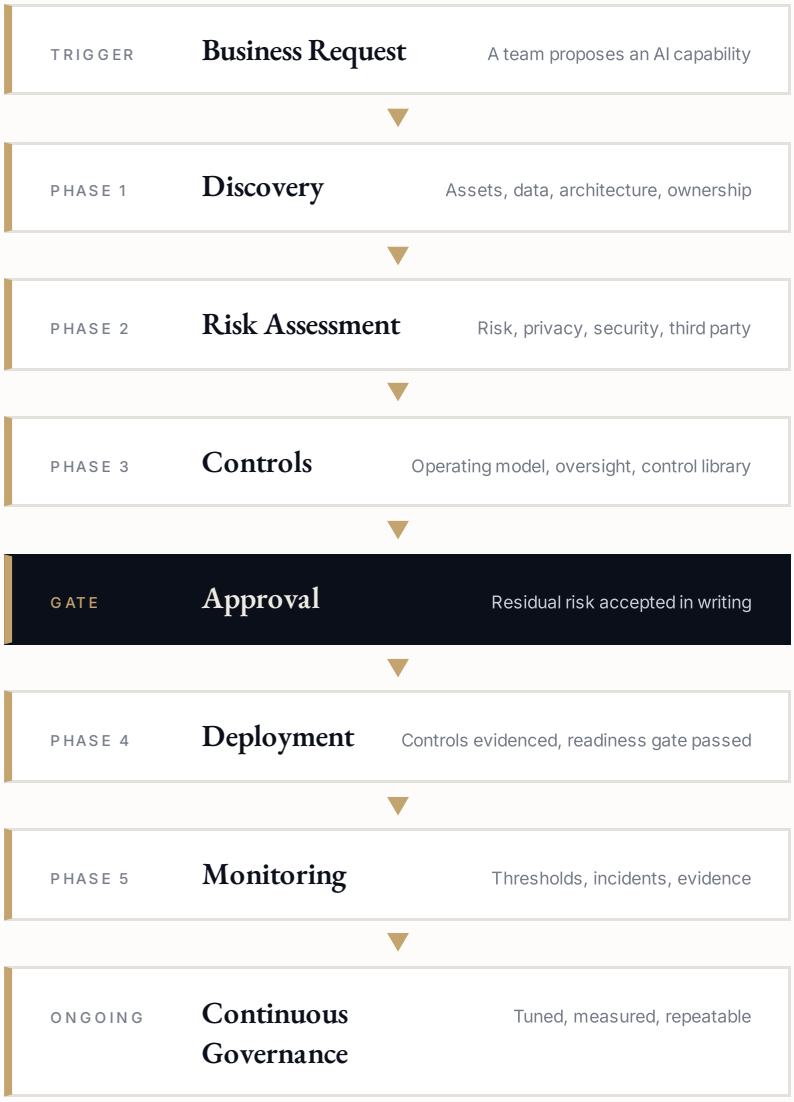
- Five phases with defined activities and exit criteria
- Four decision gates before production
- Twelve-week delivery schedule
- Named owner against every activity

Substance

- Eighteen operational controls with evidence requirements
- Risk scoring and use case classification
- Human oversight design with measurement thresholds
- Monitoring, incident and readiness governance

THE FRAMEWORK

AI Governance Lifecycle



ENGAGEMENT

Delivery Schedule

WK	PHASE	ACTIVITY	OUTPUT
01	Discover	Mobilise. Interviews 1–4. Architecture walkthrough.	Asset register v0
02	Discover	Interviews 5–8. Data flow mapping.	Flow map draft, classification
03	Discover	Validate flow map against codebase. Baselines captured.	**Gate 1**
04	Assess	Risk workshop. Use case classification. Threat modelling.	Risk register v1, threat model
05	Assess	DPIA. Third-party assessment. Regulatory classification.	DPIA, provider assessment
06	Assess	Heat maps. Residual risk accepted in writing.	**Gate 2**
07	Govern	Ownership workshop. Operating model, RACI, approval workflow.	Operating model, RACI
08	Govern	Oversight workshop. Control library. Policy mapping.	Control library, oversight design
09	Govern	Control review for buildability. Logging specification.	**Gate 3**
10	Deploy	Monitoring configured. Logging live. Evidence repository.	Monitoring, evidence repository
11	Deploy	Incident runbook. Training. Adversarial testing in CI.	Runbook, test results
12	Deploy	Readiness evidence pack. Committee review.	**Gate 4**, deployment decision

Sequencing rules

- No control designed before the flow map is validated against the codebase
- Use case classification decided week 4; it sets the control weight
- Logging schema specified week 9, built before deployment
- Evidence collected contemporaneously, never reconstructed

Slippage response

- Stakeholder access, wks 1–2: escalate to sponsor at day 5
- Architecture still moving at wk 3: freeze, or restart Gate 1
- Residual acceptance stalls wk 6: escalate; Govern cannot start
- Controls unbuildable wk 9: revise design, never defer past deployment

PHASE
1 Discover

WEEKS 1–3

PURPOSE

Understand the AI system before governance begins.

OUTPUT

A complete inventory of assets, data, stakeholders and architecture.

DURATION

3 weeks

CLIENT INPUTS REQUIRED

Architecture documentation · access to 8 stakeholders · existing certifications · metrics baseline

ACTIVITIES

- Conduct 8 stakeholder interviews across engineering, security, privacy, operations, legal, product, data platform and sponsor
- Review AI application architecture against the codebase
- Map AI data flow end to end, numbered step by step
- Map prompt construction and retrieval path
- Identify and register every AI asset in build or production
- Classify data permitted in embeddings, prompts and logs
- Review existing certifications for AI coverage
- Record baseline metrics for post-deployment measurement

QUESTIONS ASKED

- *What business problem is this AI solving, and who owns that outcome?*
- *Is the AI making autonomous decisions, or informing a human one?*
- *What customer data is processed, and at which step does it enter?*
- *If it produces wrong output tomorrow, what tells us it happened?*
- *How is retrieval scoped per tenant, and where is that enforced?*
- *Which vendors and models are involved, and is the version pinned?*
- *What has already been committed to customers or to a launch date?*
- *How will success be measured, and against what baseline?*

EXIT CRITERIA

- ✓ Asset inventory completed with named individual owners
- ✓ Data flow mapped and validated against the codebase
- ✓ Data classified for AI processing
- ✓ Business owner identified per asset
- ✓ Architecture frozen for assessment
- ✓ Baseline metrics recorded

GATE 1 — PROCEED

Proceed to Assess

DO NOT PROCEED

Architecture still changing. Freeze for assessment, or restart Gate 1 after the change lands.

DELIVERABLES

AI Asset Inventory | AI Data Flow Mapping | Prompt Flow Map | Architecture Review | Data Classification | Findings Log

PHASE
2 Assess
WEEKS 3–6

PURPOSE
Quantify exposure before any control is designed.

OUTPUT
A scored risk position with named owners and a documented residual.

DURATION
3 weeks

CLIENT INPUTS REQUIRED
DPO availability · security engineering · provider contracts and DPAs · legal input on classification

ACTIVITIES

- Score every risk across likelihood, severity, reach, reversibility and decision autonomy
- Conduct the Data Protection Impact Assessment
- Perform threat modelling with security engineering
- Assess the model provider: training use, residency, retention, versioning, exit
- Assess model behaviour for hallucination, bias and robustness
- Determine regulatory classification and the reclassification trigger
- Build inherent and residual heat maps
- Define the residual acceptance position for committee sign-off

QUESTIONS ASKED

- *What is the worst credible outcome for a customer, and can it be undone?*
- *How many customers or records can a single failure touch?*
- *Can output reach a customer without a human deciding?*
- *Which personal data sits in the vector store, and how is it erased?*
- *What is the blast radius of a hostile input reaching the model?*
- *Does the provider train on our data, and where is it processed?*
- *What does the provider warrant about output correctness?*
- *Which regulation applies, and what would reopen that question?*

EXIT CRITERIA

- [x] Risk assessment completed with owners assigned
- [x] Regulatory classification documented and dated
- [x] DPIA completed and signed
- [x] Third-party assessment completed
- [x] Threat model completed
- [x] Residual risk accepted in writing by a named owner

GATE 2 — PROCEED
Proceed to Govern

DO NOT PROCEED
Residual acceptance stalled. Escalate to sponsor; design cannot begin without it.

DELIVERABLES

AI Risk Assessment | Risk Heat Map | DPIA | Threat Model | Third-Party AI Assessment | Regulatory Assessment

PHASE
3 Govern
WEEKS 6–9

PURPOSE

Design the structure, controls and oversight that reduce risk to the accepted level.

OUTPUT

An operating model and control set the organisation can run itself.

DURATION

3 weeks

CLIENT INPUTS REQUIRED

Committee members for two 90-minute workshops · existing policy hierarchy · engineering review capacity · front-line reviewers

ACTIVITIES

- Design the governance operating model and committee decision rights
- Build the RACI to individual role level
- Design the AI use case approval workflow
- Design human oversight tiers and measurement thresholds
- Build the control library mapped to existing frameworks
- Define the evidence of operation for every control
- Design AI policies and standards into the existing hierarchy
- Specify the interaction logging schema and governance KPIs

QUESTIONS ASKED

- *Who has authority to stop a deployment, and has that been exercised?*
- *Who accepts residual risk when the owner disagrees?*
- *Which decisions must never be automated, under any threshold?*
- *What does a reviewer actually do with generated output?*
- *How would we know if review became rubber-stamping?*
- *Which existing policy does this extend rather than duplicate?*
- *What evidence proves this control operated last month?*
- *Can engineering build this within the remaining timeline?*

EXIT CRITERIA

- ✓ Governance committee established with named decision rights
- ✓ RACI approved to individual level
- ✓ Control library approved and mapped
- ✓ Human oversight designed with thresholds set
- ✓ Policies and standards approved
- ✓ Logging schema accepted by engineering

GATE 3 — PROCEED

Proceed to Deploy

DO NOT PROCEED

Controls unbuildable in the timeline. Revise the design; never defer a control past deployment.

DELIVERABLES

Governance Operating Model | RACI | Approval Workflow | Human Oversight Design | AI Control Library | AI Policies & Standards | Logging Specification | Governance KPIs

PHASE

4 Deploy

WEEKS 9–12

PURPOSE

Prove every control works, then decide.

OUTPUT

An evidenced deployment decision, recorded and owned.

DURATION

3 weeks

CLIENT INPUTS REQUIRED

Engineering capacity for control build · reviewers released for training · committee availability · CI pipeline access

ACTIVITIES

- Configure monitoring signals, thresholds and alert routing to named responders
- Implement interaction logging to the agreed schema
- Build the evidence repository and populate pre-deployment artefacts
- Build the incident runbook and test the escalation path end to end
- Run adversarial and cross-tenant tests; embed them in CI
- Test degraded mode and fallback operation
- Deliver role-based training with scenario assessment
- Assemble the readiness evidence pack and run the gate review

QUESTIONS ASKED

- *Can we show this control operated, or only that it exists?*
- *What suspends the service automatically, and who agreed that?*
- *Who is paged out of hours, and do they know it?*
- *Can blast radius be determined from the logs alone?*
- *Has every reviewer passed the scenario assessment?*
- *Which gate conditions are unevidenced, and who owns each?*
- *What is the fallback if the model provider is unavailable?*
- *Who signs the deployment decision, and is that recorded?*

EXIT CRITERIA

- [x] Controls implemented and evidenced
- [x] Monitoring configured with named responders
- [x] Incident process documented and escalation tested
- [x] Evidence repository created and populated
- [x] Training completed with assessment passed
- [x] Residual risk accepted
- [x] Readiness checklist approved
- [x] Deployment approved by committee

GATE 4 — PROCEED

Authorise production deployment

DO NOT PROCEED

Unevidenced condition. The date moves, or it is recorded as a named exception with an owner and remediation date.

DELIVERABLES

Monitoring Framework · AI Incident Management · Evidence Repository · Test Results Pack · Training Record · Production Readiness Governance

PHASE
5 Operate
ONGOING

PURPOSE

Keep governance working after deployment.

OUTPUT

Tuned thresholds, verified oversight and a repeatable process for the next use case.

DURATION

Quarterly cycle

CLIENT INPUTS REQUIRED

Production monitoring data · QA reviewer time · committee cadence · next use case submitted

ACTIVITIES

- Tune monitoring thresholds against real production volume
- Run independent QA sampling cycles
- Review override rates and review times for oversight degradation
- Route the next AI use case through the approval workflow
- Produce the quarterly executive and board reporting pack
- Test the erasure runbook against a live record
- Re-run adversarial tests after any model version change
- Execute the maturity roadmap

QUESTIONS ASKED

- *Has the override rate fallen below the floor?*
- *Did any model version change without going through the gate?*
- *Can an erasure request still be honoured end to end?*
- *What did the last QA cycle find, and did anything change?*
- *Did the next use case need bespoke governance work?*
- *Has anything changed in the regulatory position?*

EXIT CRITERIA

- [x] Thresholds calibrated against production data
- [x] Independent QA cycle completed
- [x] Executive reporting cadence established
- [x] Next use case governed without bespoke work
- [x] Erasure and version-change controls re-tested

GATE 5 — PROCEED

Governance operates without external support

DO NOT PROCEED

Oversight degrading or controls failing. Return to Govern for the affected control set.

DELIVERABLES

Threshold Tuning Report | QA Sampling Results | Executive Reporting | Governance Review | Maturity Roadmap

ASSESS

Risk Scoring and Approval Workflow

SCORING DIMENSIONS

DIMENSION	QUESTION	SCALE
Likelihood	How often in normal operation?	Rare / Possible / Likely / Frequent
Severity	Worst credible outcome for a customer or the business?	Minor / Moderate / Major / Severe
Reach	How many customers or records can one failure touch?	One / Segment / All tenants
Reversibility	Can the harm be undone once discovered?	Reversible / Costly / Irreversible
Decision autonomy	Can output reach a customer without a human deciding?	Human-approved / Monitored / Autonomous

RISK TIERS AND OBLIGATIONS

TIER	TRIGGER	OBLIGATION
CRITICAL	Irreversible harm, all-tenant reach, or autonomous customer-facing action	Committee approval, external review, deployment blocked until residual reduced
HIGH	Major severity with segment reach, or personal data exposure	Named owner, mandatory control set, monthly review
MEDIUM	Moderate severity, reversible, human-approved	Named owner, control set, quarterly review
LOW	Minor, reversible, contained	Register entry, annual review

AI USE CASE APPROVAL WORKFLOW

QUESTION, ASKED IN ORDER	ANSWER	ROUTE
1. Prohibited practice or listed high-risk area?	**Yes**	Escalate. Committee decision before development. External review
2. Processes personal or customer-confidential data?	**No**	Internal tier. Register entry, standard controls, annual review
3. Can output reach a customer, or act, without a human deciding?	**Yes**	Autonomous tier. Full assessment, committee approval, deployment gate, monthly review
3. Can output reach a customer, or act, without a human deciding?	**No**	Human-approved tier. Full assessment, committee approval, deployment gate, quarterly review

KEY CONSIDERATION

The autonomy decision is taken in week 4, before the interface is built. Moving a use case from autonomous to human-approved typically removes six weeks of control work and most of the regulatory weight. Taken later, the same change is a rebuild.

GOVERN

Governance Operating Model and RACI

COMMITTEE

Chair · AI Governance Lead

Members · CTO, CISO, DPO, head of the operating function, Legal

Cadence · Monthly, 45 minutes, decisions minuted

Quorum · Chair plus three; DPO required for privacy decisions

Escalation · Executive sponsor, then Board Risk Committee

MANDATE: FOUR DECISIONS

- Approve or reject AI use cases at the tier gate
- Accept residual risk in writing for High and above
- Authorise production deployment at the gate
- Direct response to incidents and threshold breaches

RACI

ACTIVITY	AI PLATFORM OWNER	ENGINEERING	COMMITTEE	OPERATING FUNCTION	DPO / LEGAL	RISK & COMPLIANCE
Maintain AI asset register	A	R	C	C	I	I
Classify use case and assign risk tier	A	C	R	C	C	I
Approve new AI use case	C	C	A/R	C	C	I
Accept residual risk, High and above	C	C	A	C	C	R
Design and implement technical controls	A	R	C	C	I	I
Approve model version change	A	R	C	I	I	I
Operate human oversight day to day	I	I	C	A/R	I	I
QA sampling of approved output	I	I	C	A/R	I	C
DPIA and privacy assessment	C	C	C	I	A/R	I
Third-party AI assessment	C	R	C	I	C	A
Monitor thresholds and raise breaches	A	R	C	C	I	I
AI incident response	C	R	A	C	C	I
Evidence retention and audit response	C	C	A	I	C	R
Executive and board reporting	C	I	A/R	I	I	C

A Accountable, one per row · **R** Responsible · **C** Consulted before · **I** Informed after. Residual acceptance never sits with engineering. Daily oversight is accountable to the function that controls workload.

BUILD SPECIFICATIONS

Core Artefact Field Definitions

AI ASSET REGISTER

FIELD	DEFINITION
Asset ID	Immutable identifier, referenced by risk register, logs and evidence
Business purpose	The outcome it produces, one sentence, business language
Lifecycle stage	Proposed / In development / Pre-production / Production / Retired
Model and version	Provider, family, pinned version string, deployment region
Data in scope	Categories, classification, personal data flag, special category flag
Autonomy tier	Output of the approval workflow; drives the control set
Risk tier	Output of the risk assessment; drives review frequency
Accountable owner	A named individual, never a team or function
Regulatory position	Classification, data protection role, residency, reclassification trigger
Evidence location	Where logs, approvals, assessments and reviews are retained

RISK REGISTER ENTRY

- Risk ID and linked asset ID
- Flow map step where it arises
- Cause, stated mechanically
- Impact as consequence, not category
- Inherent rating, five dimensions
- Control response, by control ID
- Residual rating
- Named individual owner
- Acceptance record: who, when, in writing
- Review frequency and next date

INTERACTION LOG SCHEMA

- Interaction ID and timestamp
- Tenant ID and scope check result
- Input, post-redaction
- Retrieved source identifiers, not content
- Model identifier and pinned version
- Generated output
- Reviewer identity and action
- Edit delta where applicable
- Released output, if different
- Immutable, access-controlled, 13-month retention

GOVERN

AI Control Library

ID	CONTROL	PURPOSE	EVIDENCE	OWNER
C-01	**Tenant Isolation**	Prevent cross-tenant exposure	CI test, quarterly probe	Platform Engineering
C-02	**PII Redaction**	Keep personal data out of embeddings	Weekly ingest sample	Platform Engineering
C-03	**Erasure Lineage**	Make deletion requests executable	Quarterly erasure test	DPO
C-04	**Version Pinning**	Prevent silent behaviour change	Change record	AI Platform Owner
C-05	**Retrieval Grounding**	Prevent ungrounded output	Daily grounding rate	Engineering
C-06	**Injection Screening**	Block hostile input	Adversarial suite in CI	Security Engineering
C-07	**Instruction Hierarchy**	Prevent instruction override	Adversarial test	Security Engineering
C-08	**Output Scanning**	Block data exfiltration	Weekly alert review	Security Engineering
C-09	**Human Approval**	Prevent unreviewed output reaching customers	Override rate, review time	Operating Function
C-10	**Senior Review**	Catch high-consequence errors	Monthly routing sample	Operating Function
C-11	**QA Sampling**	Verify oversight is real	Weekly findings log	Risk & Compliance
C-12	**Interaction Logging**	Enable investigation and audit	Daily completeness check	AI Platform Owner
C-13	**Region Pinning**	Meet residency obligations	Quarterly residency pack	Platform Engineering
C-14	**AI Disclosure**	Meet transparency obligations	Reviewed each release	Legal
C-15	**Content Filtering**	Prevent harmful output	Red-team per version change	AI Governance Lead
C-16	**Degraded Mode**	Operate without the model	Semi-annual test	Operating Function
C-17	**Rate Limiting**	Contain abuse and cost	Monthly threshold review	Platform Engineering
C-18	**Log Access Control**	Protect data inside logs	Quarterly access review	CISO

Controls map into the organisation's existing ISO 27001 and SOC 2 estate, entering the current audit cycle rather than creating a parallel programme.

GOVERN

Human Oversight Design

OVERSIGHT TIERS

TIER	APPLIES TO	REQUIRED HUMAN ACTION	PREVENTS
T0 Suggest	Internal results and summaries shown to staff	None	Nothing customer-facing by design
T1 Approve	All customer-facing generated output	Edit or explicit approve. No default send, no bulk approve	Unreviewed output reaching a customer
T2 Senior review	Billing, security, data handling, configuration changes	Second reviewer before release	High-consequence error in dispute-prone areas
T3 Prohibited	Breach notification, legal commitment, contract terms	System disabled, routed to a person	Generated content with legal consequence

MEASUREMENT

SIGNAL	THRESHOLD	FREQUENCY	ACTION ON BREACH	OWNER
Reviewer override / edit rate	**Floor** 15%	Weekly	QA review of that team's approved output	Operating Function
Median review time per item	**Floor** 8 seconds	Weekly	Team-level investigation	Operating Function
Independent QA sample	2% of approved	Weekly	Findings logged, fed to training	Risk & Compliance
T2 escalation rate	Monitored, no target	Monthly	Verify category detection is working	AI Governance Lead

KEY CONSIDERATION

Override rate is measured against a floor, never a target. A falling rate is equally consistent with the model improving and with reviewers no longer reading the output. Measuring it as a floor is what makes oversight evidenceable rather than assumed.

DEPLOY & OPERATE

Monitoring Framework and Executive Reporting

SIGNAL	THRESHOLD	FREQUENCY	ACTION	OWNER
Grounding rate	< 95%	Daily	Investigate within 1 business day; suspend below 90%	Engineering
Cross-tenant retrieval probe	Any failure	Per build	Automatic suspension; incident raised	Platform Eng
Reviewer override rate	< 15% floor	Weekly	QA review of team; retraining	Operating Function
Median review time	< 8 seconds	Weekly	Team investigation with leadership	Operating Function
PII redaction failures	Any	Weekly	Halt ingest; purge batch; notify DPO	DPO
Injection screening alerts	Trend, any success	Daily	Security review; extend test suite	Security Eng
Secret or foreign tenant ID in output	Any	Real time	Block, alert, incident raised	Security Eng
Interaction log completeness	< 100%	Daily	Priority fix; gap documented for audit	AI Platform Owner
Residency assertion	Any mismatch	Quarterly	Suspend region; notify DPO and Legal	Platform Eng
Model version drift	Any unapproved change	Continuous	Roll back to pinned version; regression suite	AI Platform Owner
Cost per tenant anomaly	3× baseline	Daily	Rate limit; investigate for abuse	Platform Eng

Operational pack · monthly

- Threshold status and breaches
- Override rate and review time
- QA findings and trend
- Incidents by class
- Model version changes

Executive and board pack · quarterly

- AI assets in production by risk tier
- Residual risks accepted, and by whom
- Material incidents and remediation
- Control exceptions and trend
- Maturity position against roadmap

DEPLOY & OPERATE

AI Incident Management

INCIDENT	SEVERITY	IMMEDIATE ACTION	ESCALATION	EVIDENCE REQUIRED
Cross-tenant data exposure	AI-1	Suspend the service immediately	DPO within 1 hour; regulator assessment; affected customers	Interaction record, tenant scope result, blast radius query
Personal data disclosed via output	AI-1	Suspend; contain and preserve logs	DPO within 1 hour; Legal; affected customers	Interaction record, output scan alert, affected-record list
Materially incorrect output acted upon	AI-2	Isolate the affected flow; notify the customer	Committee within 4 hours; operating lead	Interaction record, retrieved sources, reviewer action
Systemic grounding failure	AI-2	Suspend if grounding below 90%	Committee within 4 hours; engineering	Grounding trend, sample of ungrounded outputs
Successful prompt injection, no data loss	AI-3	Patch screening; extend test suite	Committee at next meeting; security review	Injection payload, screening log, output scan
Oversight threshold breach	AI-3	QA review of affected team's output	Committee at next meeting; operating lead	Override trend, review times, QA findings
Unapproved model version change	AI-3	Roll back to the pinned version	AI Platform Owner; committee notified	Version change record, regression results
Isolated poor-quality output caught in review	AI-4	Log and trend	QA reporting only	Interaction record, reviewer action

AI incidents route through the organisation's existing security incident process. The classification, immediate action and evidence set are additions specific to AI systems. Every evidence item depends on interaction logging (C-12) implemented to specification before deployment.

DEPLOY

Production Readiness Governance

CONDITION	CONTROL
Cross-tenant isolation verified by automated test and manual probe	C-01
Prototype data purged; production ingest with PII redaction complete	C-02
Erasure runbook tested end to end against a live record	C-03
Model version pinned and recorded in the asset register	C-04
Grounding enforcement active; fail-closed verified	C-05
Injection screening live; adversarial suite passing in CI	C-06, C-07
Output scanning for secrets and foreign identifiers active	C-08
Human approval interface live; no default-send or bulk-approve path	C-09
Senior review routing tested; prohibited categories disabled	C-10
Interaction logging complete and verified against schema	C-12
Region routing verified for every tenant region	C-13
AI disclosure live in interface and outbound content	C-14
Red-team test set executed and passed	C-15
Degraded mode tested	C-16
Monitoring thresholds configured, alerts routed to named responders	All
Incident runbook published; AI-1 escalation path tested	—
Role-based training delivered; scenario assessment passed	—
DPIA signed; sub-processor notice issued per contract	—
Residual risk accepted in writing by the accountable owner	—
Evidence repository populated with pre-deployment artefacts	—

GATE 4 — PROCEED

Production deployment authorised

DO NOT PROCEED

The date moves, or the condition is recorded as a named exception with an owner and remediation date. Decision recorded, approver named, dated.

OPERATE

Governance Maturity Roadmap

0	1	2	3	4
ABSENT	**DOCUMENTED**	**OPERATING**	**MEASURED**	**ASSURED**
No inventory, assessment or owner	Register and policy exist; governance is a document	Controls run and produce evidence; monitoring live	Effectiveness trended; oversight independently verified	External assurance; survives staff change
Start	Week 6	Deployment	Q2–Q3	Q4–Q6

TWELVE-MONTH EXECUTION

QUARTER	FOCUS	OUTCOME
Q1	Tune thresholds against production volume. First independent QA cycle. Second use case through approval workflow. Erasure runbook tested live	Governance proven under real load
Q2	Apply the control set to further AI use cases with no bespoke work. Trend control effectiveness. Oversight independently verified	Repeatability demonstrated
Q3	Internal audit of AI controls. Gap assessment against ISO 42001. Standardise customer security questionnaire responses	Evidence withstands external testing
Q4	Certification decision based on customer demand. Annual policy and risk review. Refresh threat model and red-team set	External assurance path confirmed

GOVERN

AI Control Library

ID	CONTROL	PURPOSE	EVIDENCE	OWNER
C-01	**Tenant Isolation**	Prevent cross-tenant exposure	CI test, quarterly probe	Platform Engineering
C-02	**PII Redaction**	Keep personal data out of embeddings	Weekly ingest sample	Platform Engineering
C-03	**Erasure Lineage**	Make deletion requests executable	Quarterly erasure test	DPO
C-04	**Version Pinning**	Prevent silent behaviour change	Change record	AI Platform Owner
C-05	**Retrieval Grounding**	Prevent ungrounded output	Daily grounding rate	Engineering
C-06	**Injection Screening**	Block hostile input	Adversarial suite in CI	Security Engineering
C-07	**Instruction Hierarchy**	Prevent instruction override	Adversarial test	Security Engineering
C-08	**Output Scanning**	Block data exfiltration	Weekly alert review	Security Engineering
C-09	**Human Approval**	Prevent unreviewed output reaching customers	Override rate, review time	Operating Function
C-10	**Senior Review**	Catch high-consequence errors	Monthly routing sample	Operating Function
C-11	**QA Sampling**	Verify oversight is real	Weekly findings log	Risk & Compliance
C-12	**Interaction Logging**	Enable investigation and audit	Daily completeness check	AI Platform Owner
C-13	**Region Pinning**	Meet residency obligations	Quarterly residency pack	Platform Engineering
C-14	**AI Disclosure**	Meet transparency obligations	Reviewed each release	Legal
C-15	**Content Filtering**	Prevent harmful output	Red-team per version change	AI Governance Lead
C-16	**Degraded Mode**	Operate without the model	Semi-annual test	Operating Function
C-17	**Rate Limiting**	Contain abuse and cost	Monthly threshold review	Platform Engineering
C-18	**Log Access Control**	Protect data inside logs	Quarterly access review	CISO

Controls map into the organisation's existing ISO 27001 and SOC 2 estate, entering the current audit cycle rather than creating a parallel programme.

APPLIED SCOPE

Typical ADGL Engagements

ADGL governs the deployment of:

- ✓ Microsoft Copilot Enterprise
- ✓ ChatGPT Enterprise
- ✓ Customer Support AI
- ✓ Internal Knowledge Assistants
- ✓ RAG Applications
- ✓ Agentic AI
- ✓ AI-enabled SaaS Platforms
- ✓ Industry-specific AI Systems

> The lifecycle is deployment-agnostic. The activities, controls and gates hold regardless of model, vendor or industry.

CONTACT

Ramya Amballa
Founder, AI For U&I

www.aiforui.org
hello@aiforui.org
linkedin.com/in/amballa-ramya

LinkedIn