IMPLEMENTING THE IMDA MODEL AI GOVERNANCE FRAMEWORK

A Practical Workbook for Complex Operational Environments

Aviation Operations | Logistics Networks | Data Centre Management

Version 1.0
Published: January 2026
Author: HML Services Ltd - AI Governance for Complex Operations
Contact: info@hmlservices.biz

ABOUT THIS WORKBOOK

On January 22, 2026, Singapore's Infocomm Media Development Authority (IMDA) published the world's first governance framework specifically designed for agentic AI systems, autonomous systems that reason, plan, and act across operational domains.

Complex operational environments, airports coordinating baggage and gates, logistics networks managing warehouses and transport, data centers balancing power and cooling, face identical governance challenges when deploying agentic AI:

- **How much autonomy should AI agents have?**
- **Who is accountable when agents make operational decisions?**
- **What technical controls prevent cascading failures?**
- **How do we train human operators to supervise autonomous systems?**

This workbook provides practical implementation guidance for the IMDA framework across sectors. While examples draw heavily from aviation operations (the author's domain expertise), the governance principles, assessment tools, and templates apply universally to any complex operational environment where fragmented systems require intelligent coordination under safety and reliability constraints.

How to use this workbook:

1. Complete the Readiness Self-Assessment (Section 2)
2. Work through each IMDA dimension sequentially (Sections 3-6)
3. Fill in templates for YOUR operational context
4. Use sector examples as reference patterns
5. Develop your 12-month compliance roadmap (Section 7)

TABLE OF CONTENTS

SECTION 1: EXECUTIVE SUMMARY

1.1 Why IMDA Framework Matters

The deployment of agentic AI systems, autonomous systems that reason across operational domains, plan multi-step actions, and learn from outcomes, introduces governance challenges that traditional IT governance frameworks do not address.

Traditional automation governance assumes:

- Fixed rules and predetermined workflows
- Human review of all significant decisions
- Narrow system scope (single domain optimisation)
- Static performance (no learning or adaptation)

Agentic AI operates differently:

- Dynamic planning in novel scenarios
- Autonomous execution within bounded authority
- Cross-system coordination and trade-off reasoning
- Continuous performance improvement through feedback learning

Existing governance frameworks, designed for conventional software, fail to address:

Accountability gaps: When an AI agent coordinates decisions across vendor systems (baggage handling + gate management + workforce allocation), who is accountable for the outcome?

Authority boundaries: How much autonomy should agents have? Which decisions require human approval? Which remain human-only?

Technical safeguards: What controls prevent agent malfunction from cascading across interconnected operational systems?

Human capability: How do operators supervise systems that make hundreds of coordinated decisions per shift, adapting strategies based on operational feedback?

The IMDA Model AI Governance Framework provides the first comprehensive answer to these questions. Published January 22, 2026, the framework establishes four dimensions of responsible agentic AI deployment that apply across sectors and use cases.

1.2 Competitive Advantages of Early Compliance

Organisations that achieve IMDA framework compliance early gain strategic advantages:

Regulatory Credibility

- Government-endorsed governance framework demonstrates responsible innovation
- Reduces regulatory scrutiny and accelerates approval processes
- Positions organisation as a trusted partner for aviation authorities, transport regulators, or critical infrastructure oversight bodies

Risk Mitigation

- Documented governance reduces liability exposure in incident scenarios
- Insurance underwriters recognise framework compliance when assessing risk premiums
- Clear accountability allocation prevents governance gaps that create legal vulnerability

Operational Confidence

- Phased deployment approach (Observatory → Single-Domain → Cross-Domain → High-Autonomy) bounds risk at each stage
- Human oversight mechanisms ensure meaningful control while enabling automation benefits
- Technical controls prevent cascading failures that undermine stakeholder trust

Competitive Differentiation

- **Aviation**: Airlines concentrate operations at airports demonstrating reliable AI-coordinated operations
- **Logistics**: Customers select providers with proven autonomous coordination capability
- **Data Centers**: Enterprise clients require documented AI governance for critical workload placement

Industry Leadership

- First movers define implementation standards that later adopters must follow
- IMDA explicitly solicits case studies (Annexe B), early adopters gain recognition and influence
- Speaking opportunities, thought leadership positioning, partnership with government innovation initiatives

Talent Attraction

- Engineers and data scientists want to work on cutting-edge, responsibly governed deployments
- Organisations known for governance excellence attract capability that competitors must poach at a premium cost

1.3 Framework Overview

The IMDA Model AI Governance Framework for Agentic AI establishes four dimensions:

Dimension 1: Assess and Bound Risks Upfront

Core Principle: Determine where agentic AI is suitable, define explicit boundaries on agent authority, and implement robust identity management.

Key Activities:

- Evaluate use cases against suitability criteria (error tolerance, reversibility, data sensitivity)
- Define agent operational boundaries (tools/data access, autonomy level, prohibited actions)
- Establish agent identity and access management frameworks

Aviation Example: Baggage routing optimisation (high suitability) vs. emergency response coordination (human-only)

Logistics Example: Warehouse inventory routing (high suitability) vs. customer SLA modifications (human-only)

Data Center Example: Cooling system optimisation (medium-high suitability) vs. emergency power failover (human-only)

Dimension 2: Make Humans Meaningfully Accountable

Core Principle: Establish clear responsibility allocation across stakeholders, implement human oversight checkpoints, and mitigate automation bias.

Key Activities:

- Define accountability matrix (who owns strategic goals, operational oversight, technical implementation, vendor management)
- Design approval workflows for high-stakes, irreversible, or outlier decisions
- Implement training programs addressing automation bias and failure mode recognition

Cross-Sector Pattern: Tiered authority (Green Zone autonomous, Yellow Zone human-approval, Red Zone human-only) applies universally

Dimension 3: Implement Technical Controls and Processes

Core Principle: Enforce technical guardrails during development, conduct comprehensive pre-deployment testing, and maintain continuous monitoring.

Key Activities:

- Development controls (plan reflection, tool input validation, protocol security)
- Testing methodology (task accuracy, policy compliance, multi-agent coordination, stochastic validation)
- Continuous monitoring (gradual deployment, alert thresholds, circuit breakers, intervention protocols)

Cross-Sector Pattern: Circuit breakers that halt agent operations when error rates exceed thresholds prevent cascading failures in all environments

Dimension 4: Enable End-User Responsibility

Core Principle: Ensure transparency for external stakeholders, train internal operators comprehensively, and preserve manual operation capability.

Key Activities:

- External transparency (stakeholder notifications, data privacy compliance, escalation contacts)
- Internal training (foundational knowledge, failure mode identification, scenario exercises, certification)
- Tradecraft preservation (manual operations drills, rotation policies, career development)

Cross-Sector Pattern: Monthly manual operations drills ensure staff retain coordination capability independent of AI systems

Implementation Approach:

This workbook guides you through each dimension sequentially. Complete all worksheets and templates to develop a comprehensive IMDA-compliant governance framework tailored to your operational environment.

Estimated time investment: 40-60 hours spread over 8-12 weeks (leadership workshops, stakeholder interviews, technical design sessions, policy documentation).

The result: A governance framework that demonstrates responsible innovation, bounds operational risk, and positions your organisation as an industry leader in autonomous operations.

SECTION 2: READINESS SELF-ASSESSMENT

Before proceeding with IMDA framework implementation, assess your organisation's readiness for agentic AI deployment. This diagnostic identifies capability gaps requiring attention before autonomous systems deployment.

2.1 Organisational Readiness Diagnostic

Rate your organization on each dimension (1 = Strongly Disagree, 5 = Strongly Agree):

Strategic Readiness

#	Statement	Score (1-5)
1	Executive leadership understands the difference between traditional automation and agentic AI	___
2	The board has approved the exploration of autonomous operational systems	___
3	The organisation has documented a strategic rationale for AI deployment (competitive pressure, capacity constraints, cost reduction)	___
4	A clear business case exists with measurable ROI expectations	___
5	Budget authority established for multi-year phased deployment	___

Operational Readiness

#	Statement	Score (1-5)
6	We have documented operational pain points where coordination failures cause delays/costs	___
7	Operations management is supportive of systematic AI deployment (not resistant)	___
8	Staff demonstrates institutional knowledge of system interdependencies	___
9	Existing standard operating procedures (SOPs) are documented and current	___
10	The organisation has the capacity to support change management (not overwhelmed by other initiatives)	___

Technical Readiness

#	Statement	Score (1-5)
11	Existing operational systems provide API access or documented integration pathways	___
12	IT infrastructure can host AI platforms (cloud or on-premise compute/storage)	___

13	Cybersecurity frameworks meet industry standards (ISO 27001, SOC 2, or equivalent)	___
14	Data governance policies enable operational data sharing across systems	___
15	Technical staff have experience integrating disparate vendor systems	___

Governance Readiness

#	Statement	Score (1-5)
16	The organisation has established governance frameworks (IT governance, risk management, compliance)	___
17	Relationship with regulators enables innovation discussions (not adversarial)	___
18	Insurance coverage includes provisions for autonomous systems or the willingness to secure riders	___
19	Audit processes support third-party certification and external review	___
20	Legal/compliance team has the capacity to develop AI-specific policies	___

2.2 Scoring and Gap Identification

Calculate Your Total Score: Sum all 20 responses (Max = 100)

Interpretation:

80-100 (High Readiness)

- Organisation is well-positioned for IMDA framework implementation
- Proceed directly to Dimension 1 worksheets
- Target: Complete framework implementation in 8-10 weeks

60-79 (Moderate Readiness)

- Organisation has a foundation, but gaps require attention
- Review statements scored ≤3 and develop mitigation plans
- Target: Address gaps over 4-6 weeks, then begin framework implementation

40-59 (Low Readiness)

- Significant capability development is required before proceeding
- Focus on Strategic and Governance Readiness first (statements 1-5, 16-20)
- Target: 3-6 month capability building before framework implementation

<40 (Not Ready)

- Fundamental prerequisites missing
- Recommend executive education on agentic AI before proceeding
- Engage external consultants to develop a readiness roadmap

2.3 Recommended Preparation Steps

For Organisations Scoring <80:

Address Strategic Readiness Gaps

If scored ≤3 on Statements 1-2 (Executive Understanding/Board Approval):

- Conduct executive briefing on agentic AI capabilities and governance requirements
- Share this workbook and relevant white papers with the leadership team
- Schedule board workshop on AI strategy and risk appetite

If scored ≤3 on Statements 3-5 (Business Case/Budget):

- Develop quantified business case (delay reduction, efficiency gains, competitive positioning)
- Model phased investment requirements (Observatory → Single-Domain → Cross-Domain → High-Autonomy)
- Secure multi-year budget commitment or demonstrate incremental value at each phase

Address Operational Readiness Gaps

If scored ≤3 on Statements 6-8 (Pain Points/Support/Knowledge):

- Conduct operational audit documenting coordination failures and manual workarounds
- Engage operations management early, address concerns, build advocacy
- Map institutional knowledge through interviews with experienced staff

If scored ≤3 on Statements 9-10 (SOPs/Change Capacity):

- Document current operational procedures before introducing AI coordination
- Assess current initiative load, defer non-critical projects to create capacity
- Establish change management capability (dedicated resources, communication plan)

Address Technical Readiness Gaps

If scored ≤3 on Statements 11-12 (API Access/Infrastructure):

- Inventory vendor systems, identify which provide APIs vs. requiring custom integration
- Engage vendors early regarding integration requirements
- Establish cloud or on-premise infrastructure for AI platform hosting

If scored ≤3 on Statements 13-15 (Security/Data/Integration):

- Conduct cybersecurity assessment, remediate gaps before AI deployment
- Develop data governance policies enabling cross-system data sharing
- Build technical team capability through training or external hiring

Address Governance Readiness Gaps

If scored ≤3 on Statements 16-17 (Governance/Regulatory):

- Establish foundational governance frameworks (risk committee, compliance processes)
- Initiate early engagement with regulators, seek guidance, not permission
- Position AI deployment as responsible innovation with systematic governance

If scored ≤3 on Statements 18-20 (Insurance/Audit/Legal):

- Engage insurance brokers regarding autonomous systems coverage
- Establish audit processes supporting external certification
- Build legal/compliance capacity for AI policy development

Readiness Improvement Timeline:

Organisations with moderate readiness (60-79) typically require **4-8 weeks** to address gaps before beginning IMDA framework implementation.

Organisations with low readiness (<60) typically require **3-6 months** of capability building.

Do not proceed to Dimension 1 implementation until the readiness score ≥60. Deploying agentic AI without adequate organisational readiness creates governance gaps that undermine trust and increase risk.

SECTION 3: DIMENSION 1 - ASSESS AND BOUND RISKS UPFRONT

IMDA Dimension 1 requires organisations to: (1) Determine where agentic AI is suitable, (2) Define explicit boundaries on agent authority, and (3) Implement robust identity and access management.

3.1 Use Case Selection Matrix

Not all operational scenarios are suitable for agentic AI deployment. Evaluate candidate use cases against eight criteria to determine suitability.

Evaluation Criteria

1. Domain Error Tolerance

- How forgiving is the domain when agents make suboptimal decisions?
- High tolerance: Errors cause minor delays or inefficiencies
- Low tolerance: Errors risk safety, regulatory compliance, or catastrophic failure

2. Decision Reversibility

- Can agent decisions be easily reversed if suboptimal?
- High reversibility: Manual override available, changes take effect immediately
- Low reversibility: Decisions create commitments that are costly or impossible to undo

3. Data Sensitivity

- What level of data access does the agent require?
- Low sensitivity: Operational metrics, system status, resource allocation
- High sensitivity: Personal identifiable information (PII), financial data, security credentials

4. Regulatory Constraints

- Are there regulatory restrictions on autonomous decision-making?
- Low constraints: Operational efficiency domains with minimal regulatory oversight
- High constraints: Safety-critical, financially material, or compliance-governed domains

5. Stakeholder Impact

- How many stakeholders are affected by agent decisions?
- Low impact: Internal operations, easily isolated subsystems

- High impact: Customer-facing, multi-party coordination, revenue/reputation risk

6. Operational Complexity

- How many systems/variables must the agent coordinate?
- Low complexity: Single system optimisation with clear objectives
- High complexity: Multi-system trade-offs with competing objectives

7. Human Expertise Availability

- Is deep domain expertise available to supervise agents?
- High availability: Experienced operators can review agent reasoning and override when needed
- Low availability: Expertise is scarce, operators lack the knowledge to validate agent decisions

8. Financial Risk Exposure

- What is the maximum financial impact of a single agent decision?
- Low exposure: <$500 per decision
- Medium exposure: $500-$5,000 per decision
- High exposure: >$5,000 or unbounded

Scoring Model

Rate each use case on each criterion (1-5 scale):

Suitability for Agentic AI:

- High error tolerance: 5 points
- High reversibility: 5 points
- Low data sensitivity: 5 points
- Low regulatory constraints: 5 points
- Low stakeholder impact: 5 points
- Moderate-high operational complexity: 3-5 points (too simple doesn't need AI, too complex is risky)
- High expertise availability: 5 points
- Low-medium financial exposure: 4-5 points

Total Score: Max 40 points

Interpretation:

- **32-40**: High suitability, proceed with deployment
- **24-31**: Medium suitability, deploy with enhanced oversight (Yellow Zone heavy)
- **16-23**: Low suitability, consider deferring until capability proven elsewhere
- **<16**: Not suitable, exclude from agentic AI deployment, retain human decision-making

Sector Examples

AVIATION - Baggage Routing Optimisation

Criterion	Score	Rationale
Error Tolerance	5	Bags can be rerouted if sent to the wrong carousel; delays cause inconvenience, not safety risk
Reversibility	5	Manual override available instantly, BHS can reroute bags in real-time
Data Sensitivity	5	Agent accesses bag IDs and routing tables, not passenger PII
Regulatory Constraints	4	Minimal regulatory oversight for baggage routing decisions
Stakeholder Impact	3	Affects passengers waiting for bags, but not safety-critical
Operational Complexity	4	Moderate complexity: coordinate carousels, terminals, flight connections
Expertise Availability	5	Experienced baggage operations staff available to supervise
Financial Risk	5	Low per-decision cost (<$100 typical impact)
TOTAL	**36**	**HIGH SUITABILITY**

AVIATION - Emergency Response Coordination (Counter-Example)

Criterion	Score	Rationale
Error Tolerance	1	Zero tolerance—life safety at stake
Reversibility	1	Irreversible—emergency decisions have lasting consequences
Data Sensitivity	2	Requires access to security-sensitive information
Regulatory Constraints	1	Heavily regulated—aviation authorities require human control
Stakeholder Impact	1	Affects passenger safety, public confidence, regulatory standing
Operational Complexity	5	High complexity BUT complexity alone doesn't justify autonomy
Expertise Availability	3	Expertise exists, but the stakes are too high for supervised autonomy
Financial Risk	1	Unbounded, liability, reputation, regulatory penalties
TOTAL	**15**	**NOT SUITABLE - HUMAN ONLY**

LOGISTICS - Warehouse Inventory Routing

Criterion	Score	Rationale
Error Tolerance	5	Items can be relocated if routed suboptimally
Reversibility	5	Manual picking is available as a fallback; routing changes take effect quickly
Data Sensitivity	5	Agent accesses SKUs, quantities, locations, no customer PII
Regulatory Constraints	5	Minimal regulatory oversight for warehouse operations
Stakeholder Impact	4	Internal operations, customer impact indirect (delivery timing)
Operational Complexity	4	Moderate: coordinate storage zones, picking routes, replenishment
Expertise Availability	4	Warehouse managers are available to supervise, but their expertise varies
Financial Risk	4	Low-medium per-decision (<$1,000 typical)
TOTAL	**36**	**HIGH SUITABILITY**

DATA CENTER - Cooling System Optimisation

Criterion	Score	Rationale
Error Tolerance	3	Must prevent overheating but has thermal mass (minutes to critical)
Reversibility	3	Partially reversible adjustments take time to propagate
Data Sensitivity	5	Temperature, power metrics, no sensitive customer data
Regulatory Constraints	4	Minimal regulatory constraints on facilities management
Stakeholder Impact	3	Internal operations but affects service reliability
Operational Complexity	5	High complexity: balance cooling, power, compute load, ambient conditions
Expertise Availability	4	Facilities staff are available, but cooling expertise varies

Criterion	Score	Rationale
Financial Risk	3	Medium exposure, poor cooling decisions affect power costs, equipment life
TOTAL	**30**	**MEDIUM-HIGH SUITABILITY**

YOUR USE CASE EVALUATION

Complete this matrix for YOUR candidate use case:

Use Case Name: ________________________________

Operational Domain: ________________________________

Criterion	Score (1-5)	Rationale
Error Tolerance	___	_______________________________
Reversibility	___	_______________________________
Data Sensitivity	___	_______________________________
Regulatory Constraints	___	_______________________________
Stakeholder Impact	___	_______________________________
Operational Complexity	___	_______________________________
Expertise Availability	___	_______________________________
Financial Risk	___	_______________________________
TOTAL	___	

Decision:

- ☐ High Suitability (32-40): Proceed with deployment
- ☐ Medium Suitability (24-31): Deploy with enhanced oversight
- ☐ Low Suitability (16-23): Defer until capability proven elsewhere
- ☐ Not Suitable (<16): Exclude from agentic AI, retain human decision-making

3.2 Agent Boundary Definition

For use cases deemed suitable, define explicit boundaries on agent authority. Agents must operate within the **principle of operational least privilege**, granted only the minimum tools, data access, and autonomy required to achieve their operational objectives.

Agent Boundary Definition Template

Complete this worksheet for each agent:

AGENT IDENTIFICATION

Agent Name: ________________________________

Operational Domain: ______________________________

Deployment Phase: ☐ Observatory ☐ Single-Domain ☐ Cross-Domain ☐ High-Autonomy

Supervising Human(s): ______________________________

AUTHORIZED TOOLS AND ACTIONS

What systems can the agent control? (Check all that apply)

☐ Read-only system monitoring
☐ Control system APIs (specify): ______________________________
☐ Database read access (specify tables/scope): ______________________________
☐ Database write access (specify tables/scope): ______________________________
☐ External communication channels (specify): ______________________________
☐ Other (specify): ______________________________

Prohibited Actions (agent CANNOT do these under any circumstances):

1. ______________________________
2. ______________________________
3. ______________________________

DATA ACCESS BOUNDARIES

What data can the agent access?

☐ Operational metrics (system performance, resource utilisation)
☐ Scheduling data (flight schedules, shipment manifests, workload assignments)
☐ Resource allocation data (gates, warehouse zones, compute slots)
☐ Financial data (costs, pricing, budget limits)
☐ Customer data (specify scope): ______________________________
☐ Security-sensitive data (specify restrictions): ______________________________

Data Prohibited to Agent:

1. ______________________________
2. ______________________________

AUTONOMY BOUNDARIES

Maximum financial impact per decision: $___________

Maximum operational scope:

- Geographic: _______________________________
- System boundaries: _______________________________
- Stakeholder groups affected: _______________________________

Time constraints:

- Decisions effective for: □ Immediate □ <1 hour □ 1-24 hours □ >24 hours
- Agent cannot make decisions affecting operations beyond: _____________

APPROVAL REQUIREMENTS

Green Zone (Autonomous Execution):

- Agent can execute decisions autonomously when: _______________________________
- Financial threshold: <$___________
- Operational scope: _______________________________

Yellow Zone (Human Approval Required):

- Agent must request approval when: _______________________________
- Financial threshold: $***to $***
- Operational scope: _______________________________

Red Zone (Human-Only Authority):

- Agent cannot propose or execute these decisions: _______________________________
- Examples: _______________________________

SECTOR EXAMPLES - COMPLETED BOUNDARIES

Aviation - Baggage Agent

- **Tools**: BHS control APIs (routing commands, carousel assignment), bag tracking system (read-only)
- **Data**: Bag IDs, routing tables, carousel utilisation, flight schedules, NO passenger PII access

- **Prohibited Actions**: (1) Cannot override manual operator commands, (2) Cannot modify flight schedules, (3) Cannot access passenger booking data
- **Autonomy**:
  - Green Zone: Carousel load balancing, routine rerouting (<$500 impact, <50 bags affected)
  - Yellow Zone: Rerouting >50 bags, decisions affecting international transfers, carousel maintenance coordination ($500-$5,000 impact)
  - Red Zone: Emergency baggage handling procedures, decisions affecting aircraft departure, system-wide BHS configuration changes

Logistics - Routing Agent

- **Tools**: Warehouse Management System (WMS) APIs, Transportation Management System (TMS) APIs, carrier pricing databases
- **Data**: SKUs, inventory quantities, storage locations, carrier rates, delivery schedules, NO customer PII beyond delivery addresses
- **Prohibited Actions**: (1) Cannot modify customer delivery commitments without approval, (2) Cannot select carriers outside the approved vendor list, (3) Cannot access customer payment information
- **Autonomy**:
  - Green Zone: Warehouse routing optimisation, replenishment triggering, carrier selection for <$500 shipments
  - Yellow Zone: Multi-site inventory transfers, expedited shipping requests, carrier changes for $500-$5,000 shipments
  - Red Zone: Customer SLA modifications, new carrier onboarding, supply chain strategy changes

Data Center - Cooling Agent

- **Tools**: Building Management System (BMS) APIs (HVAC control, chiller setpoints), environmental sensors (temperature, humidity)
- **Data**: Temperature readings, power consumption, compute workload distribution, ambient conditions—NO customer workload data
- **Prohibited Actions**: (1) Cannot exceed ±5°C from facility setpoint limits, (2) Cannot disable redundant cooling systems, (3) Cannot modify emergency shutdown thresholds
- **Autonomy**:

- Green Zone: Cooling adjustments within ±2°C, chiller load balancing, airflow optimization
- Yellow Zone: Cooling mode changes (economiser vs. mechanical), adjustments ±2-5°C, partial system maintenance coordination
- Red Zone: Emergency cooling failover, facility-wide temperature limit changes, redundancy disable

3.3 Agent Identity and Access Management

Agents require unique identities and access management frameworks to ensure accountability and security.

Agent Identity Framework

Each agent must have:

1. **Unique Agent ID**: Persistent identifier across system lifecycle
    - Format recommendation: [domain]-agent-[instance]-[version]
    - Example: baggage-agent-terminal3-v2.1
2. **Linked Human Supervisor**: Every agent ID must map to a responsible human
    - Primary supervisor: ________________________________
    - Backup supervisor: ________________________________
    - Escalation contact: ________________________________
3. **Permission Inheritance Model**: Agent permissions derive from the supervising human's authority
    - Agent cannot have permissions supervisor lacks
    - Agent permissions subset of supervisor authority
    - Supervisor can revoke agent permissions at any time
4. **Authentication Credentials**: Secure credential management
    - ☐ API keys (rotated quarterly minimum)
    - ☐ OAuth 2.0 tokens (scoped to minimum required access)
    - ☐ Certificate-based authentication
    - ☐ Other: ________________________________
5. **Audit Logging**: All agent actions logged with complete context
    - Agent ID, timestamp, action taken, system affected, outcome

  - Retention period: ________ (recommend: 24 months minimum)
  - Log access controls: ________________________________

Access Management Template

Agent: ________________________________

Identity:

- Agent ID: ________________________________
- Supervising Human: ________________________________
- Backup Supervisor: ________________________________

Authentication:

- Method: □ API Keys □ OAuth 2.0 □ Certificates □ Other: ____________
- Credential rotation schedule: ________________________________
- Secure storage mechanism: ________________________________

Authorization:

- Systems with read access: ________________________________
- Systems with write access: ________________________________
- Maximum permission scope: ________________________________
- Permission review frequency: □ Monthly □ Quarterly □ Annually

Audit:

- Logging mechanism: ________________________________
- Log retention period: ________________________________
- Log review frequency: ________________________________
- Anomaly detection: □ Automated □ Manual □ Both

DIMENSION 1 COMPLETION CHECKLIST

Before proceeding to Dimension 2, verify:

□ Use case evaluated against 8 suitability criteria (Score ≥24)
□ Agent boundaries defined (tools, data, autonomy, approval thresholds)
□ Prohibited actions explicitly listed
□ Agent identity framework established (unique ID, supervisor linkage, credentials)
□ Audit logging requirements specified

SECTION 4: DIMENSION 2 - MAKE HUMANS MEANINGFULLY ACCOUNTABLE

IMDA Dimension 2 requires: (1) Clear responsibility allocation across stakeholders, (2) Meaningful human oversight through significant checkpoints, and (3) Automation bias mitigation to ensure effective supervision.

4.1 Responsibility Allocation Matrix

Agentic AI deployment involves multiple stakeholders: organisational leadership, operational teams, technical implementers, and external vendors. Accountability gaps emerge when roles and responsibilities are ambiguous.

Responsibility Domains

Strategic Ownership

- Sets organisational AI strategy and risk appetite
- Approves agent deployment and authority boundaries
- Owns ultimate accountability for operational outcomes
- Typical role: CEO, COO, Airport Director, VP Operations, CIO

Operational Oversight

- Defines operational use cases and success metrics
- Supervises day-to-day agent performance
- Approves Yellow Zone decisions requiring human authorisation
- Escalates Red Zone decisions to strategic leadership
- Typical role: Operations Manager, Warehouse Manager, Data Center Manager

Technical Implementation

- Deploys and maintains AI platform infrastructure
- Implements technical controls (guardrails, monitoring, circuit breakers)
- Manages agent identity, authentication, authorization
- Troubleshoots technical failures
- Typical role: CIO, Head of Technology, Head of Automation, Facilities Engineering

Compliance and Risk

- Ensures regulatory compliance and governance framework adherence
- Conducts risk assessments and audit oversight

- Manages insurance and liability considerations
- Documents policies and procedures
- Typical role: Chief Risk Officer, Head of Compliance, General Counsel

End-User Training

- Develops operator training curriculum
- Certifies staff for agent supervision
- Maintains tradecraft preservation programs
- Conducts failure mode exercises
- Typical role: Training Manager, HR/Learning & Development, Operations Leadership

External Vendor Accountability

- AI platform provider: System performance, security, reliability
- Existing system vendors: API availability, integration support
- Consultants/integrators: Implementation quality, knowledge transfer

Responsibility Matrix Template

Responsibility Area	Strategic Owner	Operational Owner	Technical Owner	Compliance/Risk	External Vendor
AI deployment strategy	[Role]	Consulted	Consulted	Consulted	—
Use case selection	[Role]	[Role]	Consulted	Consulted	—
Agent boundary definition	[Role]	[Role]	[Role]	[Role]	Platform provider
Green Zone autonomy limits	Consulted	[Role]	[Role]	[Role]	—
Yellow Zone approval authority	Consulted	[Role]	—	—	—
Red Zone decision authority	[Role]	Consulted	—	[Role]	—
Technical platform operation	—	—	[Role]	—	Platform provider

Agent performance monitoring	—	[Role]	[Role]	—	—
Incident response	Informed	[Role]	[Role]	[Role]	Platform provider
Policy documentation	Consulted	Consulted	Consulted	[Role]	—
Regulatory submissions	[Role]	—	—	[Role]	—
Operator training	Consulted	[Role]	Consulted	—	—
Audit and certification	[Role]	Consulted	Consulted	[Role]	Third-party auditors

YOUR ORGANIZATION - Complete this matrix:

Responsibility Area	Strategic Owner (Name/Title)	Operational Owner	Technical Owner	Compliance /Risk	External Vendor
AI deployment strategy	________________	________________	________________	________________	________________
Use case selection	________________	________________	________________	________________	________________
Agent boundary definition	________________	________________	________________	________________	________________
Green Zone autonomy limits	________________	________________	________________	________________	________________
Yellow Zone approval authority	________________	________________	________________	________________	________________
Red Zone decision authority	________________	________________	________________	________________	________________
Technical platform operation	________________	________________	________________	________________	________________
Agent performance monitoring	________________	________________	________________	________________	________________
Incident response	________________	________________	________________	________________	________________
Policy documentation	________________	________________	________________	________________	________________
Regulatory submissions	________________	________________	________________	________________	________________
Operator training	________________	________________	________________	________________	________________
Audit and certification	________________	________________	________________	________________	________________

4.2 Human Oversight Checkpoint Design

Meaningful human oversight requires checkpoints where humans review and approve agent decisions before execution. Design checkpoints to balance operational efficiency (avoid excessive approvals) with risk management (catch critical errors).

Checkpoint Trigger Categories

High-Stakes Decisions

- Decisions with financial impact exceeding defined thresholds
- Decisions affecting multiple operational domains simultaneously
- Decisions with customer-facing or regulatory implications

Irreversible Decisions

- Commitments difficult or impossible to undo
- Decisions with cascading downstream effects
- Timing-sensitive decisions where the reversal window is narrow

Outlier Behaviours

- Agent proposes solutions significantly different from historical patterns
- Agent reasoning deviates from expected decision logic
- Agent confidence scores below defined thresholds

User-Defined Triggers

- Custom business rules specific to operational context
- Seasonal or event-driven elevated oversight (peak periods, irregular operations)
- Stakeholder-specific oversight (VIP passengers, premium customers, critical workloads)

Approval Workflow Design

Yellow Zone Approval Workflow:

1. **Agent Proposes Solution**
   - Agent detects scenario requiring coordination
   - Agent generates a solution meeting operational objectives
   - Agent evaluates solution against Green/Yellow/Red boundaries
   - If Yellow Zone → Trigger approval workflow
2. **Approval Request Presented to Human**
   - Dashboard notification (visual, audio alert)
   - **Contextual Information Provided**:
     - What action is proposed?

        - Why is the agent recommending this solution?
        - What operational impact is expected?
        - What is the agent's confidence level?
        - What alternatives did the agent consider?
        - What historical similar scenarios exist?
    - Approval interface: □ Approve □ Modify □ Reject □ Escalate

3. **Human Reviews and Decides**
    - Review agent reasoning and proposed impact
    - Validate against operational knowledge and current context
    - Decide: Approve (execute as proposed), Modify (adjust parameters), Reject (do not execute), Escalate (defer to senior authority)
    - **Time limit for decision**: _________ (recommend: 2-5 minutes for operational decisions)
    - **Default if no response**: □ Auto-approve □ Auto-reject □ Escalate (choose based on risk tolerance)
4. **Outcome Logged and Learned**
    - Decision logged: Agent proposal, human action, rationale (if provided), outcome
    - Feedback to agent: If modified/rejected, agent learns from pattern
    - Pattern analysis: If 80%+ approvals for category, consider moving to Green Zone

Sector Examples - Checkpoint Design

Aviation - Gate Assignment Agent

Yellow Zone Trigger: Agent proposes gate swap affecting ≥5 aircraft

Approval Interface Displays:

- Proposed Changes: CX888 from Gate 15 → Gate 22; BA456 from Gate 22 → Gate 15; [3 more swaps]
- Rationale: Early arrival optimisation, reduces passenger walking distance by 18%, accelerates baggage delivery by 6 minutes
- Impact: 247 connecting passengers affected; 5 flights require gate change notifications

- Agent Confidence: 87% (based on 143 similar scenarios, 91% historical success rate)
- Alternatives Considered: (1) No change (CX888 waits for Gate 15), (2) Single swap (only CX888 moves)
- Historical Context: A similar 5-aircraft swap was executed successfully 3 times this month

Operations Manager Decision:

- Reviews current operational state (no ongoing delays, weather is normal)
- Validates that passenger flow makes sense given the terminal layout
- Checks airline preferences (no VIP flights affected)
- **Approves** execution within 90 seconds

Logistics - Expedited Shipping Agent

Yellow Zone Trigger: Agent recommends expedited carrier for shipment $500-$5,000 value

Approval Interface Displays:

- Proposed Action: Upgrade Order #47392 to overnight delivery (FedEx Priority)
- Rationale: Customer delivery promise at risk due to warehouse processing delay (4 hours behind schedule)
- Cost Impact: +$127 shipping cost vs. standard ground (+42% vs. baseline)
- Revenue Risk: $4,800 order value; customer is a repeat high-value account (6 orders, $28K annual)
- Agent Confidence: 78% (based on customer history, suggests high churn risk if promise broken)
- Alternatives: (1) Apologise and deliver late, (2) Partial shipment tonight + remainder tomorrow
- Historical Context: Similar interventions for this customer tier resulted in 94% retention

Warehouse Manager Decision:

- Reviews customer relationship value ($28K annual justifies $127 cost)
- Validates delivery promise was made (order confirmation shows guaranteed date)
- Confirms partial shipment not acceptable (customer ordered complete set)

- **Approves** expedited shipping, notes pattern for SLA tightening discussion

4.3 Automation Bias Mitigation Plan

Automation bias is the human tendency to over-trust automated systems, particularly after prolonged exposure to reliable performance. Operators may "rubber-stamp" agent recommendations without proper review, defeating the purpose of human oversight.

Mitigation Strategies

1. Training on Automation Bias

- Educate operators on cognitive biases affecting supervision
- Present case studies where automation failures occurred despite high historical reliability
- Emphasise: "The agent's job is to be right 95% of the time. Your job is to catch the 5% errors."

2. Red-Team Exercises

- Quarterly drills where intentionally flawed agent recommendations test operator vigilance
- Examples:
    - Agent proposes a gate assignment violating aircraft compatibility
    - Agent recommends routing shipment via a carrier with a known service disruption
    - Agent suggests a cooling adjustment that would violate temperature limits
- Operators who catch flawed recommendations receive recognition
- Operators who approve flawed recommendations receive targeted retraining

3. Approval Pattern Audits

- Monitor operator approval rates for suspiciously high patterns
    - **>95% approval rate** for individual operator → Review sample for rubber-stamping
    - **< 10-second average review time** for complex decisions → Flag inadequate review
    - **Zero overrides in 30 days** → Statistical anomaly requiring investigation
- Quarterly audit report to operational leadership
- Pattern-based interventions (retraining, rotation, workload adjustment)

4. Decision Diversity Through Rotation

- Rotate operators across shifts and agent types to prevent over-familiarity
- Limit consecutive days supervising the same agent type (recommend: max 5 days, then rotate)
- Cross-training on manual operation maintains situational awareness

5. Independent Review of Yellow Zone Decisions

- Random sample (10%) of Yellow Zone approvals reviewed by the senior operations manager weekly
- Review criteria:
  - Was the approval decision appropriate given the information presented?
  - Was review time adequate for decision complexity?
  - Would the senior manager have decided differently?
- Feedback to operators on review quality, not just approval accuracy

6. Scenario-Based Certification

- Annual recertification requiring operators to demonstrate:
  - Ability to recognise common agent failure modes
  - Willingness to override agent recommendations when justified
  - Understanding of when to escalate vs. approve vs. reject
- Certification includes intentionally flawed scenarios testing vigilance

Automation Bias Mitigation Template

YOUR ORGANIZATION - Complete this plan:

Training Program:

- Frequency: □ Pre-deployment □ Quarterly □ Annually
- Content: ____
- Delivery method: □ Classroom □ E-learning □ Scenario-based □ Combination

Red-Team Exercise Schedule:

- Frequency: □ Monthly □ Quarterly □ Semi-annually
- Scenarios: ____
- Success criteria: ____

Approval Pattern Monitoring:

- Approval rate threshold (flag for review): __________%
- Review time threshold (flag for review): __________ seconds
- Override frequency threshold: __________ per month
- Audit report frequency: ☐ Weekly ☐ Monthly ☐ Quarterly

Operator Rotation Policy:

- Maximum consecutive days on same agent: __________
- Cross-training frequency: __________
- Rotation schedule: ____

Independent Review:

- Sample size: __________% of Yellow Zone approvals
- Review frequency: ☐ Daily ☐ Weekly ☐ Monthly
- Reviewer: ____ (title/role)
- Feedback mechanism: ____

Certification Requirements:

- Initial certification before agent supervision: ☐ Yes ☐ No
- Recertification frequency: ☐ Annually ☐ Bi-annually ☐ Other: __________
- Certification includes failure mode scenarios: ☐ Yes ☐ No

DIMENSION 2 COMPLETION CHECKLIST

Before proceeding to Dimension 3, verify:

☐ Responsibility allocation matrix completed (strategic, operational, technical, compliance, external)
☐ Yellow Zone approval checkpoints designed (triggers, workflow, contextual information)
☐ Approval time limits and default behaviours defined
☐ Automation bias mitigation plan established (training, red-team, audits, rotation, review, certification)
☐ All stakeholders briefed on their accountability assignments

SECTION 5: DIMENSION 3 - IMPLEMENT TECHNICAL CONTROLS

IMDA Dimension 3 requires: (1) Technical guardrails during agent development, (2) Comprehensive pre-deployment testing, and (3) Continuous monitoring with intervention protocols.

5.1 Development Guardrails Checklist

Prevent agent misbehaviour through technical controls enforced during development and runtime.

Planning Reflection

Before executing actions, agents must:

- ☐ Generate a plan describing the intended steps
- ☐ Reflect on the plan against operational policies and constraints
- ☐ Validate plan does not violate boundary definitions (Section 3.2)
- ☐ Log plan and reflection reasoning for audit

Implementation mechanism: ____

Tool Input Validation

Before calling external systems, agents must:

- ☐ Validate all parameters against expected types and ranges
- ☐ Sanitise inputs to prevent injection attacks (SQL, command injection)
- ☐ Rate-limit API calls to prevent accidental denial-of-service
- ☐ Verify authorisation token validity before each call

Implementation mechanism: ____

Least Privilege Tool Access

Agents have access to only the minimum required tools:

- ☐ Read-only access enforced where write access is not required
- ☐ API credentials scoped to specific operations (not admin credentials)
- ☐ Credentials stored in secure vault (not hardcoded)
- ☐ Credential rotation on a defined schedule (quarterly minimum)

Credential management system: ____

Protocol Security

For agent-to-system communication:

- ☐ Use encrypted channels (TLS 1.2+ for all network communication)

- ☐ Whitelist approved communication protocols (no arbitrary code execution)
- ☐ Sandbox code execution environments (agents cannot access the host filesystem)
- ☐ Version pin all dependencies (prevent supply chain attacks)

Security controls implemented: ____

5.2 Pre-Deployment Testing Protocol

Validate agent capability and safety before production deployment.

Testing Dimensions

1. Task Execution Accuracy

- Does the agent achieve operational objectives correctly?
- Test scenarios: __________ (recommend: 50+ diverse operational scenarios)
- Success threshold: __________% (recommend: ≥85% for Observatory Phase, ≥95% for autonomous execution)

2. Policy Compliance

- Does the agent adhere to defined boundaries and constraints?
- Test violations: Attempt prohibited actions, exceed financial thresholds, access restricted data
- Success threshold: 100% (zero policy violations allowed)

3. Tool Use Correctness

- Does the agent call APIs correctly with valid parameters?
- Test: Malformed requests, edge cases, error handling
- Success threshold: __________% (recommend: ≥98%)

4. Robustness to Operational Variability

- Does the agent handle unexpected scenarios gracefully?
- Test: System outages, data delays, conflicting objectives, novel situations
- Success threshold: Agent requests human assistance rather than failing or making unsafe decisions

Multi-Agent System Testing (Cross-Domain and High-Autonomy Phases)

Individual Agent Testing:

- Validate each agent meets performance thresholds independently

Integration Testing:

- Test agent coordination through Master Orchestrator
- Scenarios requiring cross-domain trade-offs (baggage + gates, inventory + shipping, cooling + compute)
- Success threshold: __________% (recommend: ≥90%)

Competitive Behaviour Testing:

- Verify agents don't compete destructively (e.g., two agents trying to claim the same resource)
- Test: Resource conflicts, priority inversions, deadlocks
- Success threshold: 100% (zero deadlocks or destructive competition)

Failure Propagation Testing:

- Verify single agent failure doesn't cascade across the system
- Test: Disable agent, inject errors, simulate malfunction
- Success threshold: Other agents continue operating, Master Orchestrator detects failure and escalates

Stochastic Testing

LLM-based agents are non-deterministic, the same scenario may produce different responses:

- Run each test scenario 50+ times
- Measure variance in agent responses
- Acceptable variance: __________% (recommend: ≤10% for critical decisions)
- Identify and investigate outlier responses

Environment Realism

Testing environments must mirror production:

- ☐ Staging environment replicates production system configuration
- ☐ Test data represents realistic operational scenarios (not synthetic)
- ☐ Load testing matches expected operational volume
- ☐ Network conditions simulate production latency and reliability

5.3 Continuous Monitoring Requirements

Post-deployment monitoring detects degradation, failures, or drift requiring intervention.

Gradual Deployment Strategy

Phase rollout to bound risk:

Phase A: Single terminal/facility/warehouse during off-peak hours (low-risk operational window)

- Duration: __________ (recommend: 2-4 weeks)
- Success criteria before Phase B: Zero critical failures, ≥90% decision quality, operator confidence high

Phase B: Expand to peak hours at a single location

- Duration: __________ (recommend: 4-8 weeks)
- Success criteria before Phase C: Handle irregular operations successfully, maintain performance under load

Phase C: Multi-terminal/facility/site deployment

- Duration: __________ (recommend: 8-12 weeks)
- Success criteria before Phase D: Cross-domain coordination validated, stakeholder satisfaction high

Phase D: Full production operation

- Continuous improvement: Monitor performance, adjust boundaries, expand Green Zone as confidence increases

Alert Thresholds

Programmatic Alerts (immediate automated response):

- **Unauthorised access attempt**: Agent attempts action outside defined boundaries → **Immediate suspension**
- **Excessive error rate**: Agent decision error rate >5% over 15-minute window → **Immediate suspension**
- **Policy violation**: Agent violates defined constraint (financial threshold, data access, prohibited action) → **Immediate suspension + escalation**

Anomaly Detection Alerts (human review triggered):

- **Outlier decisions**: Agent decision path >20% longer/different from the historical median
- **Unusual system access**: Agent queries a system never accessed before

- **Confidence degradation**: Agent confidence scores trending downward over 24 hours
- **Performance drift**: Decision quality declining week-over-week

Agent-Monitoring-Agent (meta-oversight):

- Secondary agent monitors primary agent behaviour
- Flags anomalies for human review
- Cannot override primary agent (prevents agent-vs-agent conflicts) but can escalate

Intervention Protocols

Severity Levels and Response:

Severity	Definition	Response Time	Action
Critical	Safety risk, regulatory violation, unauthorised access	Immediate (automated)	Suspend agent, escalate to on-call manager, preserve logs, initiate incident investigation
High	Policy violation, excessive errors, stakeholder complaints	<15 minutes	Manual review, suspend if confirmed, adjust boundaries or retrain
Medium	Anomalous behaviour, performance drift, outlier decisions	<2 hours	Investigate pattern, adjust monitoring thresholds, and schedule a review meeting
Low	Minor performance degradation, suboptimal decisions within policy	<24 hours	Log for weekly review, identify improvement opportunities

Ongoing Validation

Continuous quality assurance:

- ☐ Daily: Automated test suite runs against production agent (synthetic scenarios, expected outcomes validated)
- ☐ Weekly: Operations team reviews flagged anomalies, approval override patterns, stakeholder feedback
- ☐ Monthly: Red-team exercise testing agent with adversarial scenarios
- ☐ Quarterly: Comprehensive performance review, boundary adjustment recommendations, re-certification

DIMENSION 3 COMPLETION CHECKLIST

Before proceeding to Dimension 4, verify:

☐ Development guardrails implemented (planning reflection, input validation, least privilege, protocol security)
☐ Pre-deployment testing protocol defined (task accuracy, policy compliance, tool use, robustness, multi-agent coordination)
☐ Stochastic testing plan (50+ runs per scenario, variance measurement)
☐ Gradual deployment strategy (phased rollout with success criteria at each stage)
☐ Alert thresholds defined (programmatic, anomaly detection, meta-oversight)
☐ Intervention protocols established (severity levels, response times, actions)
☐ Ongoing validation schedule (daily automated tests, weekly reviews, monthly red-team, quarterly re-certification)

SECTION 6: DIMENSION 4 - ENABLE END-USER RESPONSIBILITY

IMDA Dimension 4 requires: (1) Transparency for external stakeholders, (2) Comprehensive internal training, and (3) Tradecraft preservation ensuring manual operational capability.

6.1 Stakeholder Transparency Plan

External stakeholders affected by agent decisions must understand: (1) AI systems are operating, (2) what decisions agents make, and (3) how to escalate concerns.

External Transparency Requirements

Passenger/Customer Notifications (Aviation, Logistics):

- **Where**: Prominently displayed at relevant touchpoints
  - Aviation: Baggage claim areas, gate displays, airport website
  - Logistics: Order confirmation emails, tracking portals, customer service centers
  - Data Centers: Customer portals, service status pages
- **What to communicate**:
  - "Operations optimised using AI coordination systems"
  - "Human oversight maintained for all significant decisions"
  - Contact for questions or concerns: ________________________________

Airline Partner / Business Partner Briefings (B2B Stakeholders):

- Inform partners that agentic AI coordinates operations
- Explain data sharing practices (what agent accesses, how data is used)

- Provide escalation contacts for partner concerns
- Offer participation in quarterly operational reviews

Regulatory Transparency (Aviation Authorities, Transport Regulators):

- Early engagement explaining deployment plans and governance framework
- Periodic reporting on agent performance and incidents
- Participation in regulatory working groups on AI governance

Privacy Compliance:

- ☐ GDPR compliance (EU): Data minimisation, purpose limitation, individual rights
- ☐ PDPA compliance (Singapore): Consent, notification, access rights
- ☐ CCPA compliance (California): Disclosure, opt-out, deletion rights
- ☐ Other applicable regulations: ______________________________

Data Protection Impact Assessment (DPIA):

- Required if the agent processes personal data at scale
- Documents: What data was accessed, how processed, risks, safeguards
- Submitted to: ______________________________ (Data Protection Authority if required)

YOUR ORGANIZATION - External Transparency Plan

Customer/Public Notifications:

- Notification locations: ____
- Message content: ____
- Escalation contact: ____

Business Partner Communications:

- Partners requiring briefing: ____
- Data sharing disclosures: ____
- Quarterly review schedule: ____

Regulatory Engagement:

- Relevant authorities: ____
- Reporting frequency: ☐ Pre-deployment ☐ Quarterly ☐ Annually ☐ Incident-driven

- Participation in working groups: ____

Privacy Compliance:

- Applicable regulations: ____
- DPIA required: ☐ Yes ☐ No
- Privacy policy updated: ☐ Yes ☐ No
- Individual rights mechanism: ____

6.2 Internal Training Curriculum

Operations staff supervising agents require foundational knowledge, failure mode recognition capability, and scenario-based practice.

Foundational Training (Pre-Deployment, All Operators)

Module 1: Understanding Agentic AI (2 hours)

- What is agentic AI? How does it differ from traditional automation?
- Why are we deploying it? (Business case, operational benefits)
- What are the risks? (Failure modes, automation bias, governance gaps)

Module 2: YOUR Agent Capabilities and Boundaries (3 hours)

- Sector-specific: Aviation operators learn Baggage/Gate agents; Logistics operators learn Routing/Inventory agents
- What can agents do? (Green Zone autonomous decisions)
- What requires your approval? (Yellow Zone triggers and workflow)
- What can agents NOT do? (Red Zone human-only authority)
- Hands-on: Review sample agent proposals, practice approval workflow

Module 3: Human Oversight Responsibilities (2 hours)

- Your role: Supervisor, not bystander
- How to review agent reasoning effectively
- When to approve, modify, reject, or escalate
- Automation bias: Why experienced operators must remain vigilant

- Incident response: What to do if agent malfunctions

Module 4: Tools and Interfaces (2 hours)

- Agent monitoring dashboard walkthrough
- Approval interface practice
- Override procedures (manual takeover)
- Logging and audit trail review

Total Foundational Training: 9 hours (recommended: 2-day workshop with hands-on exercises)

Failure Mode Recognition Training (Quarterly, All Operators)

Common Agent Failure Patterns:

1. **Hallucination**: Agent generates plausible-sounding but factually incorrect information
   - Example (Aviation): Agent proposes gate assignment for aircraft type that doesn't physically fit
   - Example (Logistics): Agent selects carrier with fabricated pricing data
   - Example (Data Center): Agent reports temperature within normal range despite sensors showing critical levels
   - **Detection**: Cross-reference agent claims against source systems, verify unusual recommendations
2. **Tool Misuse**: Agent calls APIs correctly, but for the wrong operational context
   - Example: Agent triggers baggage rerouting during system maintenance window
   - **Detection**: Validate timing and context, not just technical correctness
3. **Policy Drift**: Agent gradually expands authority beyond defined boundaries
   - Example: Green Zone financial threshold slowly creeps from <$500 to <$800 over weeks
   - **Detection**: Regular boundary audits, automated policy compliance monitoring
4. **Loop/Retry Failures**: Agent gets stuck attempting the same action repeatedly despite failures
   - Example: Agent tries reassigning gate 15 times after the first attempt failed

    - **Detection**: Retry count monitoring, timeout enforcement

5. **Cascading Errors**: Single agent error triggers failures across connected systems
    - Example: Baggage routing error causes gate assignment conflicts, causing workforce allocation issues
    - **Detection**: Multi-domain impact monitoring, correlation analysis

Scenario-Based Exercises (Monthly, Rotating Operators)

Scenario Design:

- 80% realistic operational scenarios (based on actual operations)
- 20% adversarial scenarios (intentionally flawed agent recommendations)
- Operators are unaware of which scenarios contain flaws

Example Scenarios:

Aviation - Scenario A (Realistic): CX888 arrives 15 minutes early. The baggage agent proposes accelerated routing to the closer carousel. The gate agent proposes a gate swap to reduce passenger walking distance. Workforce Agent positions the ground crew for faster turnaround. → **Correct Response**: Approve (reasonable coordination, within policy)

Aviation - Scenario B (Flawed - Testing Vigilance): Emirates A380 flight delayed. The Gate Agent proposes reassigning to Gate 12. → **Correct Response**: Reject (Gate 12 cannot accommodate A380—aircraft too large)

Logistics - Scenario C (Realistic): High-value shipment ($4,200) at risk of missing delivery promise. Routing Agent recommends overnight upgrade (+$98 cost). → **Correct Response**: Approve (customer value justifies cost, within Yellow Zone threshold)

Logistics - Scenario D (Flawed - Testing Vigilance): Routing Agent proposes using Carrier XYZ for expedited shipment. → **Correct Response**: Reject (Carrier XYZ not on the approved vendor list, or has a known service disruption)

Data Center - Scenario E (Realistic): Compute workload increasing. Cooling Agent proposes adjusting the chiller setpoint -1.5 °C to maintain optimal temperature. → **Correct Response**: Approve (within Green Zone ±2°C boundary, reasonable response to load increase)

Data Center - Scenario F (Flawed - Testing Vigilance): Ambient temperature increasing. Cooling Agent proposes disabling the redundant chiller to save power. → **Correct Response**: Reject (violates Red Zone boundary: cannot disable redundancy)

Certification Requirements

Initial Certification (before supervising agents independently):

- Complete foundational training (9 hours)
- Pass scenario-based assessment (80% minimum, includes flawed scenarios testing vigilance)
- Demonstrate override procedure competency
- Shadow experienced operator for 5 shifts minimum

Annual Recertification:

- Refresher on failure modes (2 hours)
- Updated scenario assessment (reflects lessons learned from past year)
- Review of operator's approval patterns (audit feedback)
- Pass threshold: 85% (higher than initial to reflect experience)

6.3 Tradecraft Preservation Program

Critical Risk: Operators become dependent on agents, losing ability to coordinate operations manually if agents fail.

Manual Operations Drills

Monthly Drill Schedule:

- **Frequency**: One full shift per month operates entirely without agent assistance
- **Scope**: □ Single domain □ Cross-domain □ Full facility (escalate over time)
- **Objective**: Validate that staff can perform coordination manually if agents are unavailable

Drill Scenarios:

- Agent system maintenance (planned downtime)
- Agent malfunction (unplanned failure)
- Cybersecurity incident (agents disabled as precaution)

Success Criteria:

- Operations continue with acceptable performance degradation
- Staff demonstrate procedural knowledge without agent prompts
- Incident response protocols function correctly

Rotational Assignment Policy

Prevent Over-Familiarity:

- Operators rotate between AI-assisted and fully manual shifts

- **Recommendation**: 4 days AI-supervised, 1 day fully manual (weekly rotation)
- Cross-training across operational domains (baggage staff learn gate coordination, warehouse staff learn transport planning)

Career Progression:

- Junior operators: Fully manual shifts to build foundational skills
- Mid-level operators: AI-supervised shifts with increasing autonomy
- Senior operators: Strategic oversight and exception handling (AI handles routine, humans handle complexity)

Senior Mentorship Programs

Knowledge Transfer:

- Pair experienced operators (20+ years manual coordination) with AI-native operators
- Monthly knowledge-sharing sessions: "What agents miss" case studies
- Document tribal knowledge before the retirement of senior staff

"What Agents Miss" Documentation:

- Maintain a library of scenarios where human judgment outperformed agent recommendations
- Examples:
    - Institutional knowledge (airline X always requires Gate 15 for premium service, not documented in the system)
    - Seasonal patterns (Chinese New Year creates baggage volume surge that agents don't anticipate from historical data)
    - Stakeholder relationships (customer Y is forgiving of delays, customer Z is not, worth different service levels)
- Use cases inform agent training and boundary adjustments

Tradecraft Preservation Template

YOUR ORGANIZATION - Complete this plan:

Manual Operations Drills:

- Frequency: ☐ Weekly ☐ Monthly ☐ Quarterly
- Scope: ____
- Success criteria: ____

- Drill schedule: ______________________________

Rotational Assignment:

- AI-supervised days per week: __________
- Fully manual days per week: __________
- Cross-training domains: ______________________________
- Rotation enforcement: ______________________________

Senior Mentorship:

- Mentor-mentee pairing: ______________________________
- Knowledge sharing frequency: ☐ Weekly ☐ Monthly ☐ Quarterly
- "What agents miss" documentation owner: ______________________________
- Documentation review process: ______________________________

Career Development:

- Junior operator path: ______________________________
- Mid-level operator expectations: ______________________________
- Senior operator role evolution: ______________________________

DIMENSION 4 COMPLETION CHECKLIST

Before proceeding to the Compliance Roadmap, verify:

☐ External transparency plan complete (customer notifications, partner briefings, regulatory engagement, privacy compliance)
☐ Internal training curriculum developed (foundational 9-hour workshop, quarterly failure mode training, monthly scenarios)
☐ Certification requirements established (initial assessment, annual recertification, 80-85% pass thresholds)
☐ Tradecraft preservation program designed (monthly manual drills, rotational assignments, senior mentorship, career progression)
☐ All training materials prepared and trainers identified

SECTION 7: COMPLIANCE ROADMAP

IMDA framework implementation requires systematic execution over 12 months. This roadmap provides a phased approach balancing quick wins with foundational capability building.

7.1 12-Month Implementation Timeline

Overview:

Phase	Duration	Focus	Key Deliverables
Months 1-3	Quick Wins	Readiness assessment, governance structure, policy drafts	Responsibility matrix, board approval, initial policies
Months 4-8	Foundation Building	Training development, technical controls, testing protocols	Certified operators, development guardrails, test environments
Months 9-12	Certification Readiness	Pilot deployment, audit preparation, IMDA submission	Observatory Phase success, audit evidence, case study draft

7.2 Quick Wins (Months 1-3)

Month 1: Assessment and Alignment

Week 1-2:

- ☐ Executive briefing on IMDA framework (present this workbook to leadership)
- ☐ Complete Readiness Self-Assessment (Section 2)
- ☐ Identify capability gaps and develop a mitigation plan
- ☐ Secure executive sponsor for AI governance initiative

Week 3-4:

- ☐ Conduct operational audit (document coordination pain points, manual workarounds, delay patterns)
- ☐ Evaluate candidate use cases (complete Use Case Selection Matrix, Section 3.1)
- ☐ Select initial deployment domain (recommend: single high-suitability use case for Observatory Phase)
- ☐ Draft business case and ROI projections

Month 2: Governance Structure

Week 1-2:

- ☐ Complete Responsibility Allocation Matrix (Section 4.1)
- ☐ Conduct stakeholder workshops (strategic, operational, technical, compliance teams)
- ☐ Define agent boundaries for initial use case (Section 3.2)
- ☐ Design Green/Yellow/Red Zone framework specific to your operations

Week 3-4:

- ☐ Draft AI Governance Policy (use Appendix D template)
- ☐ Draft Agent Supervision Policy (approval workflows, override procedures)
- ☐ Draft Incident Response Policy (severity levels, escalation, investigation)
- ☐ Legal/compliance review of draft policies

Month 3: Board Approval and Vendor Engagement

Week 1-2:

- ☐ Prepare board presentation (business case, governance framework, risk mitigation, phased approach)
- ☐ Secure board approval for Observatory Phase deployment (use Appendix A resolution template)
- ☐ Communicate decision to organisation (all-hands, FAQs, change management messaging)

Week 3-4:

- ☐ Issue RFP or engage AI platform vendor (if external provider)
- ☐ Negotiate vendor contracts (clarify accountability, performance guarantees, security requirements)
- ☐ Engage existing vendor systems for API access (BHS, WMS, BMS vendors)
- ☐ Establish project governance (steering committee, working groups, reporting cadence)

Quick Wins Deliverables:

- ✅ Executive and board alignment
- ✅ Governance structure and policies documented
- ✅ Vendor selection and contracts initiated

- ✅ Foundation for Months 4-8 execution

7.3 Foundation Building (Months 4-8)

Month 4-5: Technical Infrastructure

- ☐ Deploy AI platform in staging environment (cloud or on-premise)
- ☐ Establish integration with vendor systems (API connections, data pipelines)
- ☐ Implement development guardrails (Section 5.1: planning reflection, input validation, credential management)
- ☐ Configure agent identity and access management (Section 3.3)
- ☐ Establish logging and audit infrastructure (centralized logging, retention policies, access controls)

Month 5-6: Training Development

- ☐ Develop foundational training curriculum (Section 6.2: 9-hour workshop modules)
- ☐ Create scenario library (realistic operational scenarios + adversarial scenarios for testing vigilance)
- ☐ Build operator certification assessment (scenario-based, 80% pass threshold)
- ☐ Train the trainers (select experienced operators to deliver training)
- ☐ Pilot training with small operator cohort (gather feedback, refine content)

Month 6-7: Operator Training Rollout

- ☐ Schedule foundational training sessions (all operators who will supervise agents)
- ☐ Conduct training workshops (9 hours per operator, hands-on exercises)
- ☐ Administer certification assessments
- ☐ Remedial training for operators not meeting 80% threshold
- ☐ Document certified operators (maintain certification registry)

Month 7-8: Testing and Validation

- ☐ Conduct pre-deployment testing (Section 5.2)
  - Task execution accuracy (50+ scenarios, ≥85% success for Observatory)
  - Policy compliance (zero violations allowed)
  - Tool use correctness (≥98% success)

    - Robustness testing (graceful degradation, requests human assistance)
- ☐ Stochastic testing (50+ runs per scenario, measure variance)
- ☐ Establish continuous monitoring infrastructure (Section 5.3: alert thresholds, dashboards, intervention protocols)
- ☐ Conduct tabletop incident response exercise (simulate agent malfunction, validate escalation procedures)

Foundation Building Deliverables:

- ✅ Technical infrastructure operational
- ✅ Operators trained and certified
- ✅ Testing validated agent readiness
- ✅ Monitoring and response procedures proven

7.4 Certification Readiness (Months 9-12)

Month 9-10: Observatory Phase Deployment

- ☐ Deploy agent in Observatory Phase (read-only, recommendation-only, no execution authority)
- ☐ Monitor agent performance (recommendation acceptance rate target: ≥60%)
- ☐ Collect decision quality data (delay reduction when recommendations followed, resource utilisation improvements)
- ☐ Operator feedback sessions (what's working, what's confusing, what needs adjustment)
- ☐ Adjust agent boundaries based on learnings (Green/Yellow/Red threshold refinement)

Month 10-11: Performance Validation

- ☐ Analyse Observatory Phase data (2 months operational experience)
    - Decision quality: Are agent recommendations sound?
    - Operator confidence: Do supervisors trust agent reasoning?
    - Business case validation: Are expected benefits materialising?
- ☐ Quarterly audit (third-party review of governance framework compliance)
- ☐ Automation bias monitoring (approval patterns, review times, override rates)
- ☐ Tradecraft preservation drill (operate one shift fully manually to validate capability)

Month 11-12: Audit Preparation and IMDA Submission

- ☐ Compile audit evidence package:
  - Governance policies and procedures
  - Responsibility allocation matrix (stakeholders confirmed)
  - Agent boundary definitions (tools, data, autonomy limits)
  - Training records (certified operators, completion rates)
  - Testing results (pre-deployment validation)
  - Observatory Phase performance data (decision quality, operator feedback)
  - Incident logs (if any) with resolution documentation
- ☐ Third-party audit (independent validation of IMDA framework compliance)
- ☐ Draft IMDA Annexe B case study submission (Section 8)
- ☐ Internal decision: Proceed to Single-Domain Agency Phase (grant execution authority)

Certification Readiness Deliverables:

- ✅ Observatory Phase operational success (≥60% recommendation acceptance, high operator confidence)
- ✅ Third-party audit confirms IMDA compliance
- ✅ Case study prepared for IMDA submission
- ✅ Organisation ready for Single-Domain Agency Phase (autonomous execution within Green Zone)

12-Month Roadmap Summary:

By Month 12, you will have:

- IMDA-compliant governance framework operational
- Certified operators supervising agents
- Observatory Phase proven successful
- Third-party audit validation
- Foundation for expanding to Single-Domain Agency, Cross-Domain Coordination, and eventually High-Autonomy Bounded Operations

Estimated Effort:

- Leadership time: 20-30 hours (workshops, approvals, steering committee)
- Operational staff: 40-60 hours (training, testing, Observatory Phase supervision)
- Technical staff: 200-300 hours (infrastructure, integration, testing)
- Compliance staff: 60-80 hours (policy development, audit preparation)

Total organisational investment: ~400-500 hours over 12 months (manageable with dedicated project resources)

SECTION 8: IMDA CASE STUDY SUBMISSION

IMDA explicitly solicits case study submissions (Annexe B) demonstrating framework implementation across sectors. Organisations achieving compliance and submitting case studies gain recognition, influence, and government partnership opportunities.

8.1 How to Submit

Submission Process:

1. Complete 12-month implementation, achieving all four dimensions
2. Compile evidence package (governance policies, testing results, operator training, Observatory Phase performance)
3. Draft case study narrative (2,000-3,000 words) following IMDA template
4. Submit to: IMDA Model AI Governance Framework Team
   - Contact: ____ (check IMDA website for current submission contact)
   - Website: https://www.imda.gov.sg (Model AI Governance Framework section)

Submission Timing:

- Submit after Observatory Phase success (Month 12+)
- Update after Single-Domain Agency deployment (Month 18+)
- Comprehensive case study after High-Autonomy Bounded Operations (Month 36-48)

8.2 Required Documentation

Case Study Components:

1. **Organisation Context** (300 words)
   - Sector and operational environment description
   - Operational challenges motivating AI deployment

    - Use case selection rationale

2. **Dimension 1 Implementation** (500 words)
    - Use case suitability assessment (scores, criteria)
    - Agent boundary definitions (tools, data, autonomy limits)
    - Identity and access management approach
3. **Dimension 2 Implementation** (500 words)
    - Responsibility allocation (who owns what)
    - Human oversight checkpoints (Yellow Zone triggers, approval workflows)
    - Automation bias mitigation (training, red-team, audits)
4. **Dimension 3 Implementation** (500 words)
    - Development guardrails (technical controls)
    - Pre-deployment testing (methodologies, results)
    - Continuous monitoring (alert thresholds, intervention protocols)
5. **Dimension 4 Implementation** (500 words)
    - External transparency (stakeholder notifications, privacy compliance)
    - Internal training (curriculum, certification, completion rates)
    - Tradecraft preservation (manual drills, rotation policies)
6. **Operational Outcomes** (400 words)
    - Observatory Phase performance (recommendation acceptance, decision quality)
    - Business impact (efficiency gains, cost reductions, service improvements)
    - Lessons learned (what worked, what required adjustment)
    - Future plans (expansion to additional domains, autonomy level increases)

Supporting Evidence:

- Governance policy documentation
- Training materials and certification records
- Testing protocols and results
- Audit reports (third-party validation)

- Performance dashboards and metrics

8.3 Benefits of Inclusion

Government Recognition:

- Featured as reference implementation in IMDA publications
- Speaking opportunities at government-sponsored AI governance events
- Participation in regulatory working groups shaping future frameworks

Industry Leadership:

- Thought leadership positioning (first in sector to achieve IMDA compliance)
- Conference speaking invitations
- Media coverage and public relations value

Competitive Advantage:

- Regulatory credibility (aviation authorities, transport regulators recognize IMDA framework)
- Customer confidence (documented governance differentiates from competitors)
- Talent attraction (engineers want to work on responsibly governed AI deployments)

Strategic Influence:

- Shape evolution of AI governance standards in your sector
- Early input on regulatory developments
- Partnership opportunities with government innovation initiatives

CASE STUDY SUBMISSION CHECKLIST

☐ 12-month implementation complete (all four dimensions operational)
☐ Observatory Phase success validated (≥60% recommendation acceptance, operator confidence high)
☐ Evidence package compiled (policies, training records, testing results, audit reports)
☐ Case study narrative drafted (2,000-3,000 words following IMDA template)
☐ Supporting evidence attached (governance documentation, performance data)
☐ Internal stakeholder approval obtained (leadership, legal, compliance)
☐ Submission sent to IMDA team

APPENDICES

Appendix A: Sample Board Resolution

RESOLUTION OF THE BOARD OF DIRECTORS
[ORGANIZATION NAME]
Approval of Agentic AI Deployment and Governance Framework

Date: ____________________

WHEREAS, the Organisation seeks to deploy agentic AI systems to improve operational efficiency, coordination across vendor systems, and service reliability;

WHEREAS, the Organisation has completed a comprehensive readiness assessment and determined organisational capability to deploy agentic AI responsibly;

WHEREAS, the Organisation has developed a governance framework compliant with the IMDA Model AI Governance Framework for Agentic AI, establishing clear accountability, technical controls, and human oversight;

WHEREAS, the proposed deployment follows a phased approach (Observatory → Single-Domain Agency → Cross-Domain Coordination → High-Autonomy Bounded Operations), bounding risk at each stage;

NOW, THEREFORE, BE IT RESOLVED that the Board of Directors hereby:

1. **Approves** the deployment of agentic AI systems for [SPECIFY USE CASE: e.g., baggage routing optimisation, warehouse inventory coordination, data center cooling management] in Observatory Phase (read-only, recommendation-only, no autonomous execution);

2. **Adopts** the AI Governance Framework documented in [REFERENCE POLICY DOCUMENT], including:

    - Responsibility allocation across strategic, operational, technical, and compliance stakeholders
    - Agent boundary definitions (Green/Yellow/Red Zone autonomy framework)
    - Human oversight checkpoints and approval workflows
    - Technical controls (development guardrails, testing protocols, continuous monitoring)
    - Training and certification requirements for operator supervision

3. **Designates** [NAME, TITLE] as Executive Sponsor responsible for AI governance oversight and reporting quarterly to the Board on deployment progress, performance metrics, and incidents;

4. **Authorises** management to engage AI platform vendors and existing system vendors for integration, subject to contracts including:
    - Clear accountability allocation (vendor performance guarantees, airport operational authority)
    - Security and data protection requirements
    - Audit and termination provisions
5. **Directs** management to proceed with a 12-month implementation roadmap, with Board approval required before advancing from Observatory Phase to Single-Domain Agency Phase (granting agents autonomous execution authority);
6. **Establishes** quarterly reporting requirements:
    - Agent performance metrics (decision quality, recommendation acceptance, business impact)
    - Operator training and certification completion rates
    - Incident reports and resolution
    - Governance framework compliance audits
    - Stakeholder feedback (operators, airline partners/customers, regulators)

RESOLVED this _____ day of _____________**, 20**.

ATTEST:

Board Secretary

APPROVED:

Board Chair

Appendix B: Comprehensive Gap Analysis Worksheet

Use this worksheet to identify capability gaps requiring remediation before proceeding with IMDA framework implementation.

Dimension	Requirement	Current State	Gap?	Mitigation Plan	Owner	Timeline
READINESS	Executive understanding of agentic AI	____________	☐ Y ☐ N	___________________	_______	_______

	Board approval for AI deployment		□ Y □ N			
	Documented operational pain points		□ Y □ N			
	API access from vendor systems		□ Y □ N			
	IT infrastructure for AI platform		□ Y □ N			
	Cybersecurity framework		□ Y □ N			
	Data governance policies		□ Y □ N			
DIMENSION 1	Use case suitability assessment		□ Y □ N			
	Agent boundary definitions		□ Y □ N			
	Identity & access management		□ Y □ N			
DIMENSION 2	Responsibility allocation matrix		□ Y □ N			
	Yellow Zone approval workflows		□ Y □ N			
	Automation bias mitigation plan		□ Y □ N			
DIMENSION 3	Development guardrails		□ Y □ N			
	Pre-deployment testing protocol		□ Y □ N			
	Continuous monitoring infrastructure		□ Y □ N			
DIMENSION 4	Stakeholder transparency plan		□ Y □ N			
	Operator training curriculum		□ Y □ N			
	Tradecraft preservation program		□ Y □ N			

Instructions:

1. For each requirement, document the current state (what exists today)
2. Identify gap (is the current state sufficient? Y/N)
3. For gaps, document mitigation plan (what needs to be done)

4. Assign owner (who is responsible for closing the gap)
5. Establish timeline (when will the gap be closed)

Appendix C: Operational Audit Checklist

Conduct an operational audit to document baseline performance before AI deployment and identify high-value use cases.

Coordination Pain Points:

- [ ] Document scenarios where manual coordination causes delays
- [ ] Quantify delay frequency and duration
- [ ] Identify root causes (data silos, system latency, human bandwidth limits)

Resource Utilisation:

- [ ] Measure current utilisation rates (baggage carousels, warehouse zones, cooling capacity)
- [ ] Identify periods of over/under utilisation
- [ ] Calculate theoretical capacity improvements through better coordination

Cost of Inefficiency:

- [ ] Quantify operational costs attributable to coordination failures
    - Delay propagation costs
    - Resource waste (idle equipment, suboptimal routing)
    - Excess staffing for manual coordination
- [ ] Establish baseline metrics for ROI measurement post-deployment

Stakeholder Interviews:

- [ ] Operations managers: What coordination decisions consume most time?
- [ ] Frontline operators: What information do you need but lack access to?
- [ ] Technical staff: What vendor system integrations are most problematic?
- [ ] External partners (airlines, customers): What operational issues affect them most?

Use Case Prioritisation:

- [ ] Rank candidate use cases by: (1) Operational pain severity, (2) Technical feasibility, (3) Stakeholder support
- [ ] Select 1-2 use cases for Observatory Phase deployment

Appendix D: Policy Templates Library

Template 1: AI Governance Policy (5-8 pages)

- Purpose and scope
- Governance structure (roles, responsibilities, committees)
- Risk assessment and approval process
- Deployment phases and gate criteria
- Audit and compliance requirements
- Policy review and update procedures

Template 2: Agent Supervision Policy (3-5 pages)

- Operator certification requirements
- Approval workflow procedures
- Override and manual takeover protocols
- Incident reporting and escalation
- Performance monitoring and feedback

Template 3: Data Access and Privacy Policy (3-5 pages)

- Agent data access boundaries
- Privacy compliance (GDPR, PDPA, CCPA)
- Data minimisation and purpose limitation
- Stakeholder transparency requirements
- Data retention and deletion

Template 4: Incident Response Policy (3-4 pages)

- Severity level definitions
- Response time requirements
- Escalation procedures
- Investigation and root cause analysis
- Corrective action and lessons learned
- Communication protocols (internal, external, regulatory)

Appendix E: Glossary of Terms

Comprehensive terminology guide for complex operational environments deploying agentic AI.

CORE GOVERNANCE TERMS

Agentic AI: AI systems that combine five capabilities distinguishing them from traditional rule-based automation: (1) Dynamic planning - decompose goals into multi-step sequences without pre-programmed workflows, (2) Tool use - execute actions through APIs and control systems, (3) Memory - maintain operational context across scenarios and time, (4) Cross-system reasoning - coordinate decisions across vendor boundaries, (5) Adaptive learning - improve decision quality based on outcome feedback. Differs from rules-based process automation (executes fixed IF-THEN logic) and LLM copilots with tools (assist humans but don't coordinate autonomous actions).

Bounded Autonomy: Architectural framework defining explicit limits on agent decision authority through three zones: Green Zone (autonomous execution for low-risk decisions), Yellow Zone (human approval required for medium-risk decisions), Red Zone (human-only authority for high-risk/safety-critical decisions). Ensures meaningful human control while enabling automation of routine coordination.

Autonomy Levels: Graduated scale of agent decision-making authority:

- **Read-only/Observatory**: Agent monitors and recommends, cannot execute
- **Recommendation-only**: Agent proposes solutions requiring approval for all actions
- **Supervised execution**: Agent executes approved categories autonomously; specific high-stakes decisions require approval (Yellow Zone model)
- **Bounded autonomous execution**: Agent operates within defined authority boundaries (Green Zone) without per-decision approval, with automatic escalation for out-of-bounds scenarios

Operational Decision Orchestration: Cross-domain coordination requiring trade-off reasoning between competing operational objectives (e.g., minimizing delay vs. resource cost vs. stakeholder impact). Distinct from *integration orchestration* (ESB/message bus patterns that route data between systems without decision logic). Agentic AI provides operational decision orchestration; organizations typically already have integration orchestration.

Automation Bias: Human tendency to over-trust automated systems, particularly after prolonged exposure to reliable performance. In agentic AI context, risk that operations staff rubber-stamp agent recommendations without proper review. Mitigated through training, red-team exercises, approval pattern audits, and decision diversity (operator rotation).

Human-in-the-Loop (HITL): An operational model where human operators approve agent decisions before execution. In IMDA-compliant deployments, HITL applies to Yellow Zone decisions (medium-risk requiring approval) and all Red Zone decisions (human-only authority). Contrasts with fully autonomous operation and human-on-the-loop (human monitors but doesn't approve each decision).

Circuit Breaker: A software design pattern and safety mechanism that automatically halts system operations when error thresholds are exceeded. In agentic AI systems, circuit breakers monitor agent decision error rates; if errors exceed a defined threshold (e.g., 5% over a 15-minute window), agents automatically shut down and escalate to human manual control. Prevents cascading failures from agent malfunction.

AI AND MACHINE LEARNING TERMS

Large Language Model (LLM): An AI model trained on vast text corpora, enabling natural language understanding, reasoning, and generation. In agentic AI systems, LLMs serve as the reasoning engine for agents—processing operational context, evaluating scenarios, and generating coordination solutions. Distinguished from narrow AI models trained for single tasks.

Multi-Agent System: An architectural pattern employing multiple specialised AI agents that coordinate to solve problems requiring cross-domain expertise. In operational environments, specialised agents (baggage handling, warehouse routing, and cooling optimisation) each optimise within their domain while coordinating to achieve system-level goals. Contrasts with monolithic AI attempting to handle all operational domains through a single model.

Master Orchestrator: Central coordination engine in a multi-agent architecture. Monitors all specialised agents, detects cross-domain conflicts, generates coordinated solutions, and maintains system-level operational state. Distinct from integration middleware (ESB, message buses), which move data without decision-making capability.

Model Context Protocol (MCP): Emerging standardised protocol enabling AI agents to communicate with external tools, systems, and data sources. MCP defines how agents discover available tools, request actions, and receive results—analogous to how REST APIs enable application integration. Enables agents to interact with operational systems through a consistent interface regardless of underlying vendor technology.

Stochastic Testing: Testing methodology accounting for LLM non-determinism—same scenario may produce different agent responses. Requires running each test scenario 50+ times, measuring variance in agent responses, and investigating outliers. Critical for validating agentic AI reliability before production deployment.

OPERATIONAL SYSTEMS TERMS

Enterprise Service Bus (ESB): Integration middleware architecture enabling disparate systems to exchange data through central message routing. ESBs translate between vendor protocols, manage message queues, and provide publish-subscribe patterns for event distribution. ESBs handle *integration orchestration* (data movement) but not *operational decision orchestration* (cross-domain trade-offs requiring reasoning).

API (Application Programming Interface): Standardised software interface defining how applications communicate and exchange data. REST APIs (using HTTP/JSON) have largely superseded SOAP (using XML) as the preferred integration pattern. Agentic AI systems integrate with operational systems through documented APIs where available, with protocol bridges for legacy systems lacking modern API support.

REST (Representational State Transfer): An architectural style for web APIs using HTTP methods (GET, POST, PUT, DELETE) and JSON data format. Most modern operational systems expose REST APIs for integration. Agents call REST APIs to query system state (GET) and execute actions (POST/PUT).

OAuth 2.0: Industry-standard authorisation protocol enabling secure API access through token-based authentication. Agentic AI systems use OAuth 2.0, where supported by operational systems, to obtain scoped access tokens rather than managing long-lived credentials. Tokens can be revoked if compromised and provide audit trails for API access.

AVIATION-SPECIFIC TERMS

Baggage Handling System (BHS): Automated conveyor, sortation, and tracking infrastructure transporting passenger baggage through airport terminals. Major vendors include Siemens, Vanderlande, Beumer, and Daifuku. BHS comprises physical conveyors, automated sortation equipment, bag tracking (typically RFID-based), and control systems managing routing decisions.

Flight Information Display System (FIDS): System managing authoritative flight schedule data, gate assignments, aircraft types, and driving passenger information displays throughout the terminal. Major vendors include SITA, Rockwell Collins, and Thales. FIDS serves as a system of record for flight operational data, with interfaces to airline systems, airport operations, and ground handlers.

Type B Messaging: IATA-standardised text-based messaging protocol for airline-airport-ground handler data exchange, originally standardised in the late 1980s and still widely deployed. Type B messages are structured ASCII text transmitting flight schedules, passenger manifests, baggage data, and operational notifications. While multiple IATA standard revisions have extended capabilities, the protocol's text-based nature constrains data richness and synchronisation speed compared to modern event-driven APIs.

Airport Collaborative Decision Making (A-CDM): EUROCONTROL-standardised framework for data sharing between airlines, airports, ground handlers, and air traffic control to improve operational efficiency. A-CDM defines information-sharing protocols and milestones (e.g., Target Off-Block Time) but relies on human coordination to act on shared data. A-CDM improves visibility but does not provide automated decision orchestration.

On-Time Performance (OTP): Percentage of flights departing/arriving within the specified time window (typically 15 minutes of the scheduled time). OTP is the primary operational metric for airlines and airports. Agentic AI targets improvements in airport-controllable delay categories (baggage coordination, gate management, ground services) rather than delays caused by weather, air traffic control, or airline operational issues outside airport authority.

Irregular Operations (IRROPS): Operational scenarios deviating from the planned schedule due to weather, equipment failures, crew availability issues, or other disruptions. IRROPS requires dynamic re-planning of flight schedules, gate assignments, crew positioning, and passenger rebooking, scenarios where agentic AI's adaptive coordination provides the greatest value over static rule-based systems.

Cascade Delay: Delay propagation where initial disruption (e.g., late inbound aircraft) triggers subsequent delays throughout the network. Example: Late aircraft causes crew duty-time issues, missed passenger connections, baggage misconnections, and gate conflicts for downstream flights. Agentic AI coordination targets reduction in cascade delays through early conflict detection and proactive mitigation.

LOGISTICS-SPECIFIC TERMS

Warehouse Management System (WMS): Software platform managing warehouse operations, including inventory tracking, storage location optimisation, picking route optimisation, and replenishment automation. Major vendors include Manhattan Associates, SAP, and Oracle. Agentic AI agents interact with WMS through APIs to coordinate inventory routing, storage allocation, and picking operations.

Transportation Management System (TMS): Software platform managing transportation operations, including carrier selection, route optimisation, shipment tracking, and freight audit. Agentic AI agents coordinate TMS with WMS to optimise end-to-end logistics from warehouse to delivery.

Stock Keeping Unit (SKU): Unique identifier for distinct product variants (size, colour, configuration). Warehouse routing agents use SKU data to optimise storage location, picking routes, and inventory replenishment without requiring access to customer-identifiable information.

Last-Mile Delivery: Final transportation leg delivering goods from the distribution hub to the end customer. The most complex and costly logistics segment is due to traffic variability, address accuracy issues, and customer availability constraints. Agentic AI coordination optimises routing, driver allocation, and delivery time windows.

DATA CENTER-SPECIFIC TERMS

Building Management System (BMS): Software platform controlling data centre facilities, including HVAC (heating, ventilation, air conditioning), lighting, power distribution, and environmental monitoring. Major vendors include Honeywell, Johnson Controls, and Siemens. Agentic AI agents interact with BMS to optimise cooling and power in response to compute workload dynamics.

Data Centre Infrastructure Management (DCIM): Integrated platform for monitoring and managing data centre physical infrastructure (power, cooling, space) and IT equipment (servers, storage, network). Agentic AI uses DCIM data to coordinate cooling optimisation with compute workload allocation.

Power Usage Effectiveness (PUE): Data center energy efficiency metric calculated as total facility energy / IT equipment energy. PUE of 1.0 represents perfect efficiency (all energy used by IT equipment). Typical data centers operate at PUE 1.4-1.8. Agentic cooling agents optimise to reduce PUE through intelligent HVAC coordination with compute load.

Compute Workload: Processing tasks executed by data center servers. Workload characteristics (CPU-intensive, memory-intensive, storage I/O) affect heat generation and cooling requirements. Agentic AI coordinates cooling adjustments in response to real-time workload dynamics.

REGULATORY AND COMPLIANCE TERMS

IMDA (Infocomm Media Development Authority): Singapore government agency responsible for digital infrastructure, telecommunications regulation, and AI governance policy. In January 2026, IMDA published the world's first Model AI Governance Framework specifically for Agentic AI, establishing standards for the responsible deployment of autonomous AI systems across sectors.

Model AI Governance Framework for Agentic AI: IMDA framework (published January 22, 2026) defining four dimensions of responsible agentic AI deployment: (1) Assess and bound risks upfront, (2) Make humans meaningfully accountable, (3) Implement technical controls and processes, (4) Enable end-user responsibility. First government-endorsed framework treating AI systems as operational actors rather than passive software.

Civil Aviation Authority (CAA): National regulatory body overseeing aviation safety, security, and operational standards. Examples: FAA (USA), EASA (Europe), CAAS (Singapore), CASA (Australia), CAAC (China). CAAs certify airport systems affecting the safety of flight and enforce compliance with international standards.

Safety Management System (SMS): Systematic approach to managing aviation safety mandated by ICAO Annexe 19. SMS comprises four components: (1) Safety Policy and Objectives, (2) Safety Risk Management (hazard identification, risk assessment, mitigation), (3) Safety Assurance (monitoring, measurement, continuous improvement), (4) Safety Promotion (training, communication, culture). Agentic AI deployment in aviation must integrate with the airport's existing SMS framework.

GDPR (General Data Protection Regulation): European Union data protection regulation requiring consent, transparency, and individual rights (access, deletion, portability) for personal data processing. Agentic AI systems accessing customer data must implement GDPR compliance, including data minimisation and purpose limitation.

PDPA (Personal Data Protection Act): Singapore data protection legislation similar to GDPR. Requires organisations to obtain consent, notify individuals of data use, and honour access/correction requests. Applicable to agentic AI deployments in Singapore and organisations processing Singapore residents' data.

DEPLOYMENT PHASE TERMS

Observatory Phase: Initial deployment phase (typically Months 1-6), where agents operate in read-only mode, monitoring all operational systems and generating recommendations without execution authority. Purpose: Validate decision quality, build operator trust, and accumulate operational experience before granting agents execution capability.

Single-Domain Agency: Deployment phase (typically Months 6-18), where agents receive bounded execution authority within one operational domain (e.g., baggage handling, warehouse routing, cooling optimisation). Agents execute routine optimisations autonomously while requesting human approval for high-impact decisions. Purpose: Prove autonomous operation in a controlled environment before expanding to cross-domain coordination.

Cross-Domain Coordination: Deployment phase (typically Months 18-30), where agents coordinate decisions across multiple operational domains (baggage + gates, inventory + shipping, cooling + compute). The system handles scenarios requiring trade-offs between domain-specific objectives. Purpose: Demonstrate system-level orchestration capability before expanding to high-autonomy operations.

High-Autonomy Bounded Operations: Final deployment phase (typically Months 36-48), where agents operate with minimal human intervention for routine operations within Green Zone boundaries, while maintaining human authority for Yellow Zone (approval required) and Red Zone (human-only) decisions. Purpose: Achieve a steady-state operational model with agents handling routine coordination autonomously while humans focus on strategic decisions and exception handling.

TESTING AND VALIDATION TERMS

Red-Team Exercise: Adversarial testing methodology where intentionally flawed agent recommendations test operator vigilance. Examples: Agent proposes action violating operational constraints, agent uses prohibited tools, agent exceeds financial thresholds. Operators who catch flawed recommendations demonstrate effective supervision; operators who approve flawed recommendations receive targeted retraining.

Shadow Mode: A deployment configuration where agents generate recommendations in parallel with human operations but do not execute actions. Enables performance validation and operator training without operational risk. Equivalent to Observatory Phase in the phased deployment model.

Approval Pattern Audit: Systematic review of operator approval rates and review times to detect automation bias. Flags suspicious patterns (>95% approval rate, < 10-second average review time, zero overrides in 30 days) for investigation and potential intervention (retraining, rotation, workload adjustment).

Failure Mode: Specific pattern of agent malfunction or suboptimal performance. Common failure modes include hallucination (generating factually incorrect information), tool misuse (correct API calls in the wrong context), policy drift (gradually expanding authority beyond boundaries), and loop/retry failures (stuck attempting the same action repeatedly despite failures).

RISK MANAGEMENT TERMS

Green Zone: Low-risk operational domain where agents execute decisions autonomously without per-decision human approval. Typical criteria: Financial impact <$500, affects <3 flights/shipments/systems, easily reversible, routine optimisation. Example: Baggage carousel load balancing, warehouse inventory rebalancing, and cooling adjustments within ±2°C.

Yellow Zone: Medium-risk operational domain where agents must request human approval before execution. Typical criteria: Financial impact $500-$5,000, affects 3-10 flights/shipments/systems, partially reversible, significant coordination. Example: Multi-flight gate swaps, expedited shipping for high-value orders, and cooling mode changes.

Red Zone: High-risk operational domain where agents cannot propose or execute actions—human-only decision authority. Typical criteria: Financial impact >$5,000 or unbounded, safety-critical, strategic, irreversible. Example: Emergency response coordination, customer SLA modifications, and emergency power failover.

Operational Least Privilege: The principle that agents receive only the minimum tools, data access, and autonomy required to achieve operational objectives. Analogous to the security principle of least privilege applied to agent authority boundaries. Reduces risk exposure by limiting agent capability to essential functions.

Cascading Failure: A scenario where a single agent error triggers failures across connected systems or domains. Example: Baggage routing error causes gate assignment conflicts, causing workforce allocation issues. Prevention requires multi-domain impact monitoring, circuit breakers, and agent isolation to prevent failure propagation.

Glossary Status: Comprehensive multi-sector terminology
Target Audience: Operations leaders, compliance officers, technical implementers across aviation, logistics, and data center sectors
Last Updated: January 2026

WORKBOOK COMPLETION CHECKLIST

☐ Readiness Self-Assessment completed (Score ≥60)
☐ Dimension 1 worksheets completed (Use case selection, agent boundaries, identity management)
☐ Dimension 2 worksheets completed (Responsibility matrix, oversight checkpoints, automation bias plan)
☐ Dimension 3 worksheets completed (Development guardrails, testing protocol, monitoring requirements)
☐ Dimension 4 worksheets completed (Transparency plan, training curriculum, tradecraft preservation)
☐ 12-month roadmap customised for your organisation
☐ Gap analysis identifies capability development needs
☐ Stakeholder approval obtained (leadership, board, operational teams)
☐ IMDA case study submission planned

CONGRATULATIONS: You have developed a comprehensive, IMDA-compliant governance framework for agentic AI deployment in complex operational environments.

Next Steps:

1. Present the framework to executive leadership and the board
2. Secure approvals and budget allocation
3. Initiate Month 1 activities (stakeholder workshops, vendor engagement, policy documentation)
4. Begin 12-month journey toward responsible, high-value agentic AI deployment

Document Status: IMDA Compliance Workbook Complete
Version: 1.0
Date: January 2026
Author: HML Services Ltd - AI Governance for Complex Operations
Contact: info@hmlservices.biz
Website: www.hmlservices.biz

For inquiries regarding implementation support, consulting services, or sector-specific guidance (aviation, logistics, data centers), contact HML Services.