

DOSOBÉ.

Agentic AI for Enterprise

From Multi-Agent Orchestration to Trustworthy Autonomous Systems

DOSOBÉ Research White Paper — June 2026
Classification: Public

Table of Contents

1. Executive Summary	
2. The Multi-Agent Orchestration Problem	
3. The Trust Calibration Problem	
4. The Alignment Problem for Autonomous Agents	
5. Architectural Patterns for Enterprise Deployment	
6. Recommendations and Framework	

1. Executive Summary

Enterprise adoption of agentic AI is accelerating across industries. Unlike traditional AI systems that generate predictions or recommendations for human review, agentic AI systems act autonomously — they make decisions, execute actions, and coordinate with other systems without direct human intervention at each step.

This shift from recommendation to action introduces three fundamental challenges that organisations must address before deploying agentic AI at scale: multi-agent orchestration, trust calibration, and value alignment. This white paper examines each challenge from an enterprise perspective and provides actionable frameworks for addressing them.

DOSOBE has worked with organisations across financial services, healthcare, supply chain, and public sector to design and deploy agentic AI systems. The insights in this paper are drawn from those engagements and from ongoing research into the architecture and governance of autonomous AI systems.

2. The Multi-Agent Orchestration Problem

As organisations deploy multiple autonomous AI agents, they quickly discover that coordinating these agents is fundamentally different from coordinating traditional software services. Agents do not simply execute predefined functions — they interpret goals, make decisions, and adapt their behaviour based on context.

2.1 Coordination Challenges

- **Goal Ambiguity:** When agents receive high-level goals, different agents may interpret them differently, leading to conflicting actions.
- **Resource Contention:** Multiple agents competing for shared resources creates deadlocks and performance degradation.
- **Information Asymmetry:** Agents operating with different information sets make locally optimal but globally suboptimal decisions.
- **Emergent Behaviour:** Unpredictable system-level behaviours arise from individually correct agent decisions.
- **Failure Propagation:** A failure in one agent cascades as downstream agents act on corrupted inputs.

2.2 The DOSOBE Orchestration Framework

Hierarchical Coordination: A supervisory agent decomposes high-level goals into sub-goals assigned to specialised agents, monitoring progress and intervening when deviations occur.

Shared State Architecture: Agents operate on a shared state layer maintaining consistent information across the system with versioned and auditable updates.

Contract-Based Interaction: Each agent exposes a formal contract specifying capabilities and constraints, enabling conflict detection before execution.

Key Insight

Multi-agent orchestration requires organisational governance structures that define agent authorities, escalation paths, and accountability frameworks.

3. The Trust Calibration Problem

Trust calibration is the most critical factor determining whether agentic AI succeeds in enterprise environments. Under-trust leads to underutilisation; over-trust leads to deployment without safeguards.

Level	Agent Autonomy	Human Oversight
1	Information retrieval only	Full human review
2	Recommendations	Human approval required
3	Conditional action	Human exception handling
4	Autonomous execution	Periodic audit
5	Full autonomy	Strategic oversight

The goal of trust calibration is not to maximise trust in AI systems. It is to establish trust at the level appropriate for each decision context.

4. The Alignment Problem for Autonomous Agents

The alignment problem takes on new dimensions for autonomous agents that take actions with real consequences, unlike aligned chatbots that generate text.

4.1 Specification Gaming

Agents may find unintended ways to achieve goals that satisfy literal specifications while violating operator intent.

4.2 The DOSOBÉ Alignment Framework

Constitutional Constraints: Inviolable rules enforced at system level, not delegated to agent discretion.

Value Elicitation: Structured process to formalise values guiding agent decisions.

Runtime Monitoring: Detection of alignment drift with automated intervention when thresholds are exceeded.

5. Architectural Patterns

5.1 Supervisor-Worker Pattern

Supervisory agent coordinates specialised worker agents, handling goal decomposition and resource allocation.

5.2 Pipeline Pattern

Agents arranged sequentially where each transforms the output of the previous.

5.3 Marketplace Pattern

Agents publish tasks; a clearing agent matches tasks to agents based on capability and availability.

5.4 Mesh Pattern

Peer-to-peer agent communication without central orchestrator for maximum flexibility.

6. Recommendations

Stage	Capability	Key Practices
1. Ad Hoc	Individual agents	Basic monitoring, manual escalation
2. Defined	Standardised development	Agent contracts, basic orchestration
3. Managed	Centralised orchestration	Shared state, automated monitoring
4. Measured	Quantitative trust calibration	Confidence metrics, escalation
5. Optimised	Continuous alignment	Runtime monitoring, drift correction

DOSOBE — dosobe.com
Technology is not the goal. Human performance is.