BLOCKCHAIN GAME ALLIANCE

2026

AGENTIC AI IN GAMING

AND THE EMERGING DIGITAL ECONOMY

REPORT

Study conducted by Blockchain Game Alliance with the unique contributions from

Deloitte.

Teranode

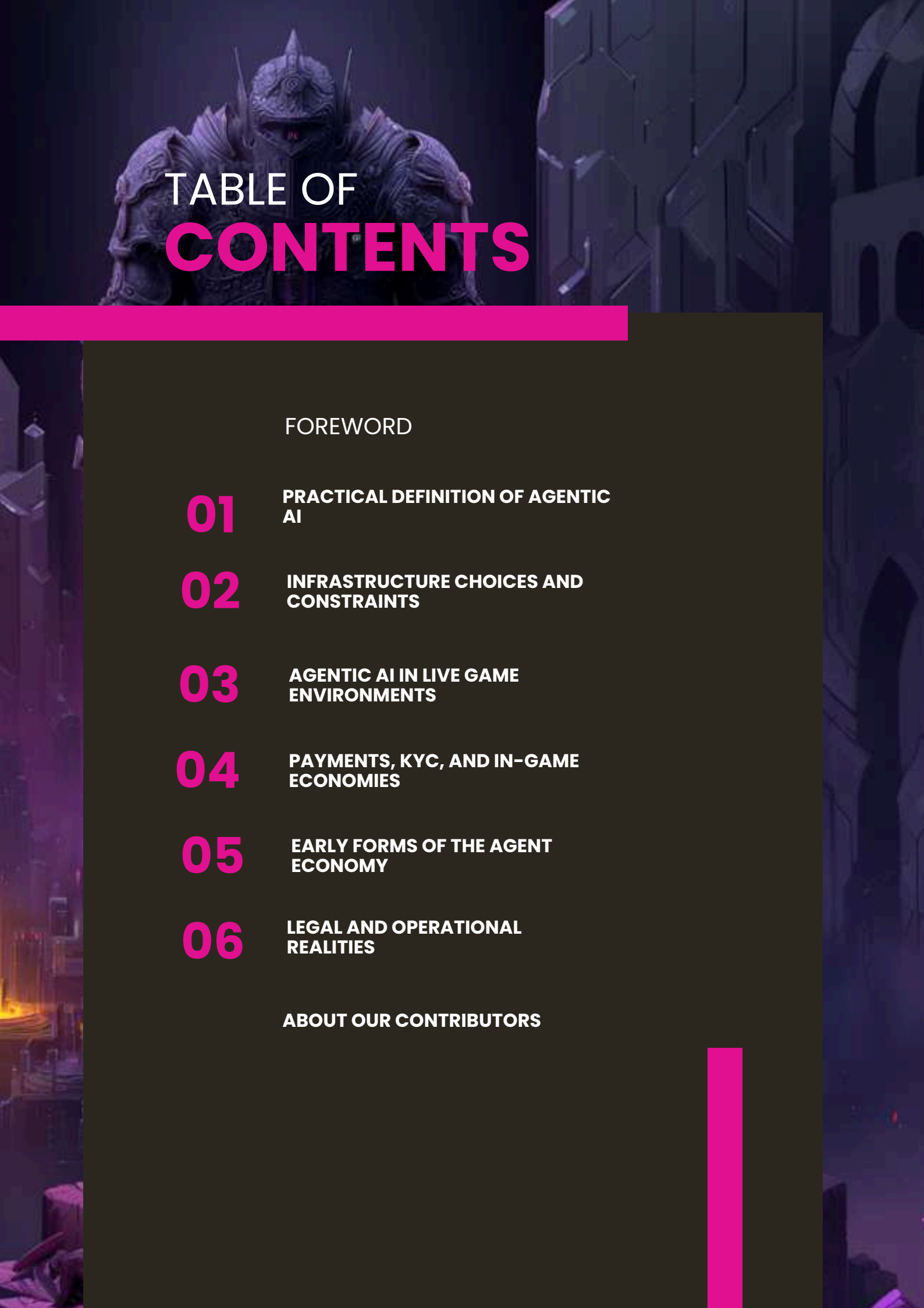


TABLE OF CONTENTS

FOREWORD

This report emerges from a simple observation, the three most consequential shifts in digital entertainment; the rise of blockchain-native ownership, the maturation of AI reasoning, and the emergence of programmable payment rails which are converging at exactly the same moment.

The Blockchain Game Alliance has brought together practitioners, researchers, and builders from across gaming, AI infrastructure, compliance, and digital finance to map what this convergence means in practice. The contributions gathered here are not theoretical.
They come from teams that have shipped autonomous agents managing real capital, built identity frameworks for machine actors, navigated the regulatory grey zones of in-game economies, and designed governance architectures for production agentic systems.

The questions this report addresses are ones that every product manager, engineer, legal counsel, and studio executive in the gaming industry will need to answer in the next 18 to 36 months: Who is legally responsible when an AI agent makes a transaction? How should studios govern agent autonomy without killing its value? What protocol stack should you build on? How do in-game economies survive contact with agents that are faster, more patient, and more rational than any human player?

We have structured the report around the six core challenge areas that surfaced repeatedly across contributions which are defining what agentic AI actually means in a live product context, making infrastructure decisions under uncertainty, managing agents in live game environments, handling payments and identity, understanding the emerging agent economy and navigating the legal and operational realities that no framework has fully resolved yet.

"This report is a practical guide informed by real-world observations across the games, blockchain, and AI sectors, focusing on how these technologies are being deployed rather than theorized. It reflects key market dynamics and emerging challenges, from scaling infrastructure to evolving governance and economic design.

The insights are drawn from a broad and diverse set of contributors working directly as operators, builders, and strategists across the ecosystem. Their collective input helps ground the analysis in what is currently working in practice and where critical gaps and opportunities for development remain.

A sincere thank you to all contributors whose time, insights and on-the-ground experience made this report possible."

DIMITRI GROSS

PARTNER | FINANCIAL SERVICES | AUDIT & ASSURANCE DELOITTE

01 PRACTICAL DEFINITION OF AGENTIC AI

INTRODUCTION

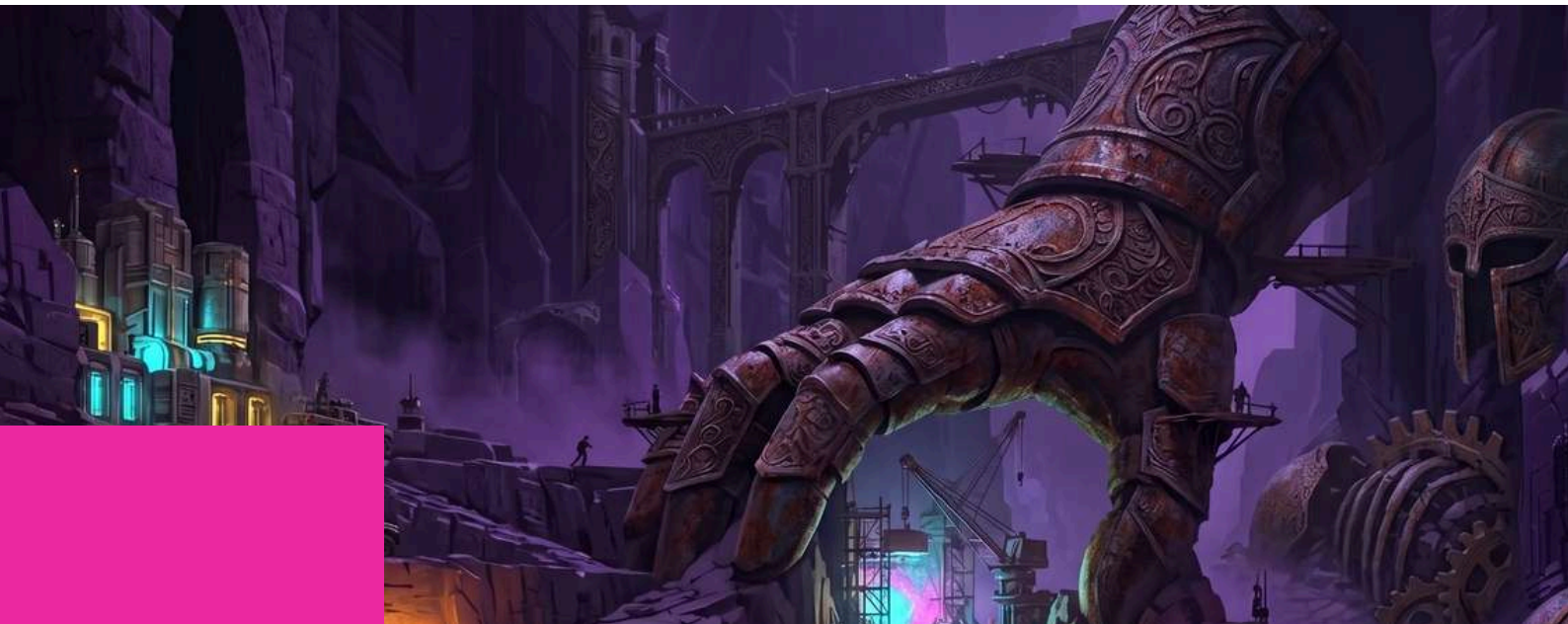
One of the most persistent problems in deploying AI systems is definitional imprecision.

Teams ship features they call 'agentic' that are, in practice, sophisticated autocomplete. Equally, some teams underestimate what they have built and fail to apply the governance frameworks those systems require.
Getting the definition right is not a semantic exercise, it determines your architecture, your legal exposure, and the governance model you need before you go to production.
The consensus across contributors is that the defining characteristic of a true agent is not intelligence, it is the capacity for goal-directed, multi-step action in a real environment with real consequences.

A chatbot that drafts a response is not an agent.
A system that identifies an opportunity, takes a sequence of actions across multiple platforms, moves value, and adapts its strategy based on outcomes is an agent, and the gap between those two things is enormous, it is exactly the gap that the industry is currently crossing.

This chapter traces the evolution of on-chain agents through three distinct generations, establishes a working framework for classifying agentic systems by their autonomy level, and introduces the two foundational governance concepts, which are Know Your Agent (KYA) and the AI Control Tower, that will recur throughout this report.

Significantly, it situates capability development against the trust infrastructure that has not yet caught up, the missing credit and accountability layer that determines whether autonomous agents can be trusted with real economic stakes.



1.1
THE THREE ERAS OF ON-CHAIN AGENTS

FROM SCRIPTED BOTS TO FULLY AUTONOMOUS ECONOMIC ACTORS

Era 1 — Scripted Bots (2015–2022)

The first generation operated on hardcoded rules, if-then logic applied to MEV extraction, liquidations, and arbitrage.

These systems had zero adaptability.

They executed a fixed strategy until the market regime changed and they broke.

No autonomy, no decision-making capacity, the human wrote the playbook and the bot followed it verbatim.

These bots were effective in a DeFi environment that was simple enough to be legible.

When Uniswap V2 was the dominant venue and the strategy space was narrow, a well-tuned script could generate consistent returns.

That environment is gone.

Era 2 — AI-Guided Semi-Autonomous (2023–2024)

The introduction of large language models and ML models into the agent stack changed the fundamental relationship between objective and execution.
Agents could now interpret market conditions, adjust parameters, and make decisions within bounded contexts.

The human sets the objective and the agent decides how to achieve it.

This era produced tools like Giza's ARMA for autonomous yield optimization and the first experiments with agents managing real capital.
The critical limitation was that these agents lacked portable identity

Their performance history was trapped in the system that deployed them. An agent that built a track record on one protocol could not carry that reputation elsewhere.

Era 3 — Fully Autonomous (2025 onward)

The current era represents a genuine inflection point.

Agents can now discover other agents, negotiate terms, hire sub-agents, execute multi-step strategies across protocols, and manage portfolios with minimal human oversight.

The agent is an economic actor, not a tool.

The infrastructure enabling this transition arrived in rapid succession from the deployment of ERC-8004 (identity and reputation, live January 2026), x402 (HTTP-native payments), ERC-8183 (agentic commerce, draft standard), A2A and MCP (agent communication protocols).

For the first time, an agent can register a verifiable identity, build a track record, accept jobs, receive payment, and carry its reputation across platforms.

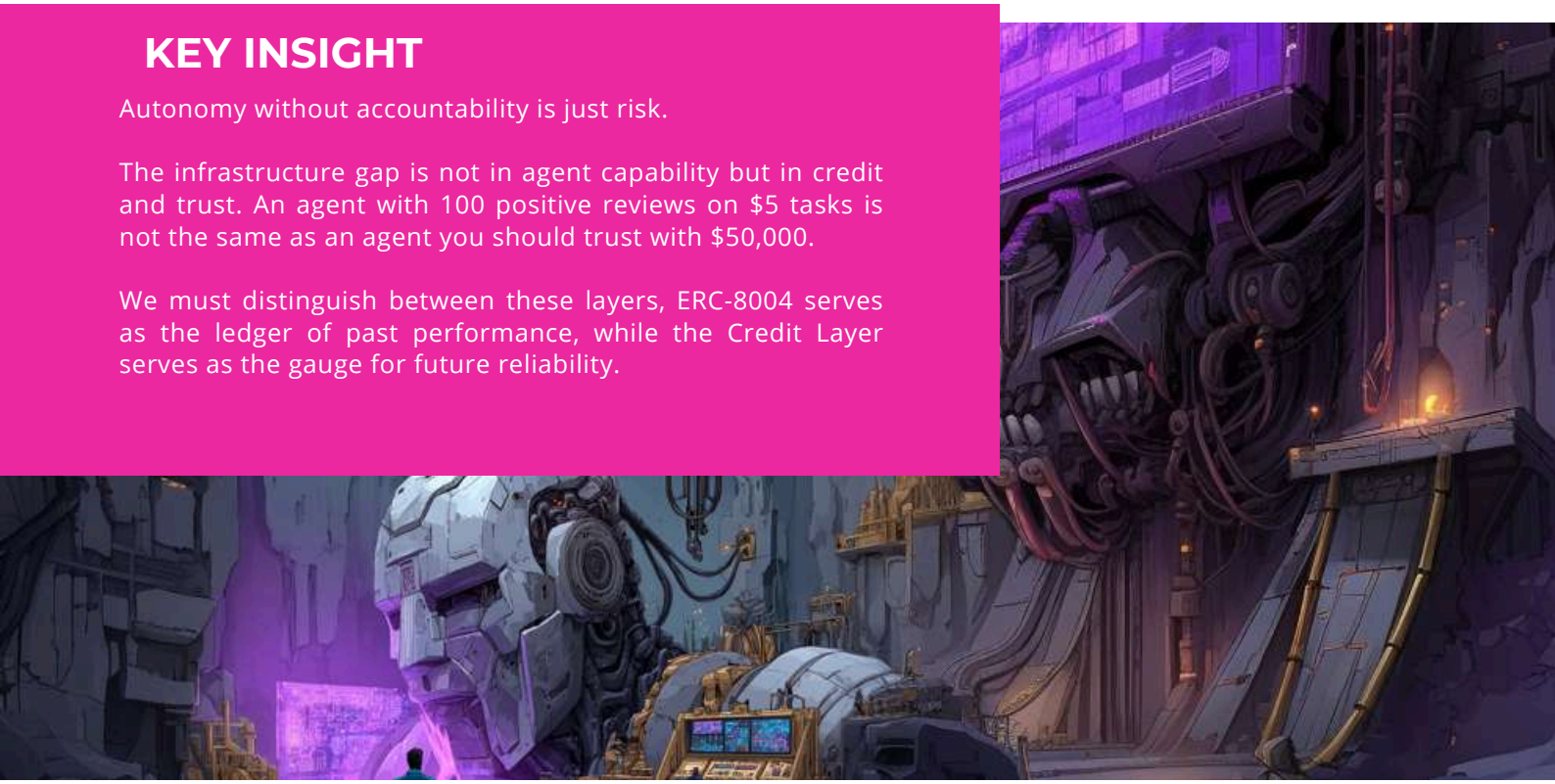
The interoperability is now an asset

KEY INSIGHT

Autonomy without accountability is just risk.

The infrastructure gap is not in agent capability but in credit and trust. An agent with 100 positive reviews on $5 tasks is not the same as an agent you should trust with $50,000.

We must distinguish between these layers, ERC-8004 serves as the ledger of past performance, while the Credit Layer serves as the gauge for future reliability.



The Trust Infrastructure Gap

The agentic stack now has communication (A2A, MCP), payments (x402), identity and reputation (ERC-8004), and a commerce framework (ERC-8183). What it does not have is a credit layer.

Raw on-chain reputation signals are inputs, not outputs. Someone needs to aggregate, weight, and risk-assess them to produce a forward-looking assessment of what an agent will do, not just what it has done.

Med Amine Idmoussi · CTO, bond.credit
Source: Presented at ERC-8004 Launch Day, hosted by the Ethereum Foundation, MetaMask, Coinbase & Google DeepMind, 17 March 2026

1.2

GOVERNED AUTONOMY: THE KYA AND CONTROL TOWER FRAMEWORK

DEFINING THE MINIMUM VIABLE GOVERNANCE LAYER

For product managers, the question is not whether AI agents should be allowed to modify the user experience autonomously, but under what conditions such autonomy remains acceptable.

The answer requires a structured governance model, not a binary choice between full human control and unconstrained agent action.

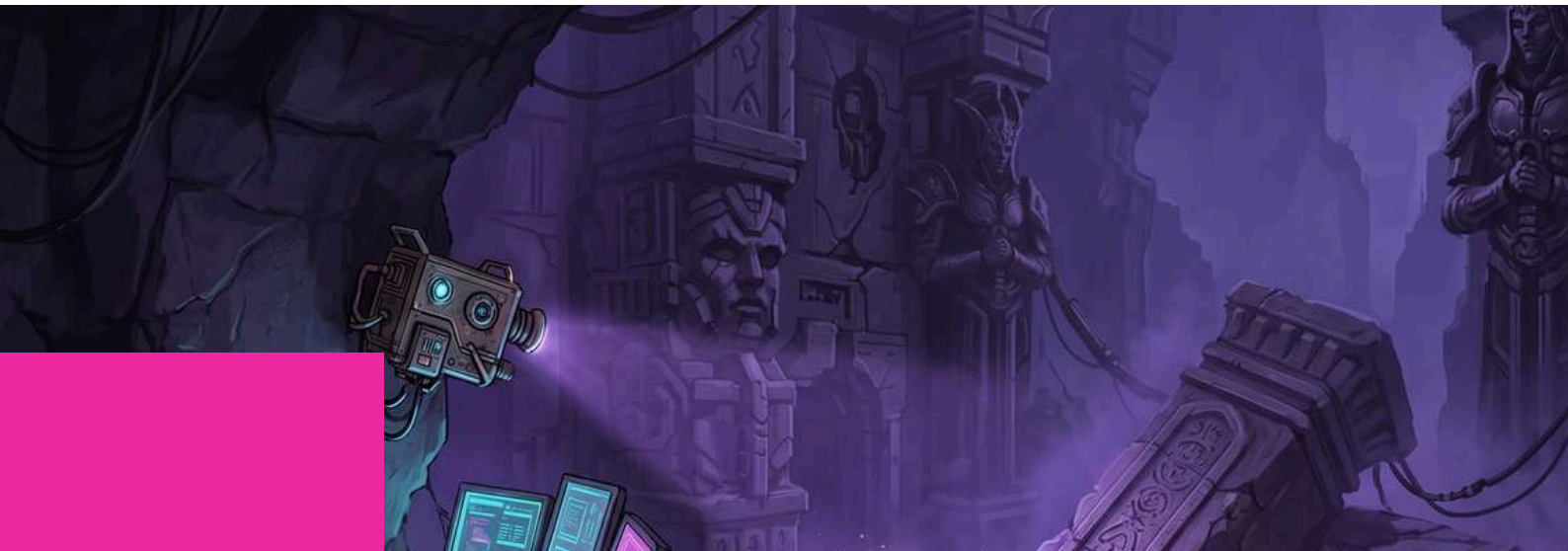
Autonomy should be constrained by what Bond credit's production framework calls 'blast radius' or how much damage an action can cause if the agent gets it wrong.

This maps to three practical tiers:

• **Tier 1** — Full Autonomy: Reading on-chain data, querying APIs, generating reports, monitoring positions and the worst-case outcome is a bad chart or stale number. No human approval required.
• **Tier 2** — Bounded Autonomy with Async Review: The agent executes within predefined limits such as position size caps, approved token lists, slippage thresholds, and daily loss caps. Instead of reviewing individual decisions, a human reviews the activity in batches.

• **Tier 3** — Propose and Wait: High-stakes, irreversible actions. This applies to deploying contracts, moving funds above capital thresholds, or changing protocol parameters. While the agent prepares the transaction and presents its reasoning, a human must provide the final signature.
The critical insight is that autonomy level should be dynamic, not static. An agent with a verified track record earns higher-tier access. A new agent with no history starts at Tier 3.

Trust is earned through verifiable on-chain performance data, not configuration files.



Know Your Agent (KYA)

KYA is a framework for verifying the identity of an AI agent as an acting software entity, while linking that identity to the human and, where relevant, the legal entity responsible for its actions. It is not a replacement for KYC/KYB but it is an additional identity and trust layer designed specifically for machine actors.

The central challenge is authenticity; the fact that an action appears to originate from a given agent does not confirm that the agent, rather than a human operator with access to its environment, was the true originator.

This creates a trust problem that credentials alone cannot solve. Teranode Group's research team has been investigating hardware-based secret management combined with runtime verification techniques as a path toward technically verifiable agent identities, anchored not just in declared identity attributes but in mechanisms that provide stronger assurance over execution context and integrity.

The AI Control Tower

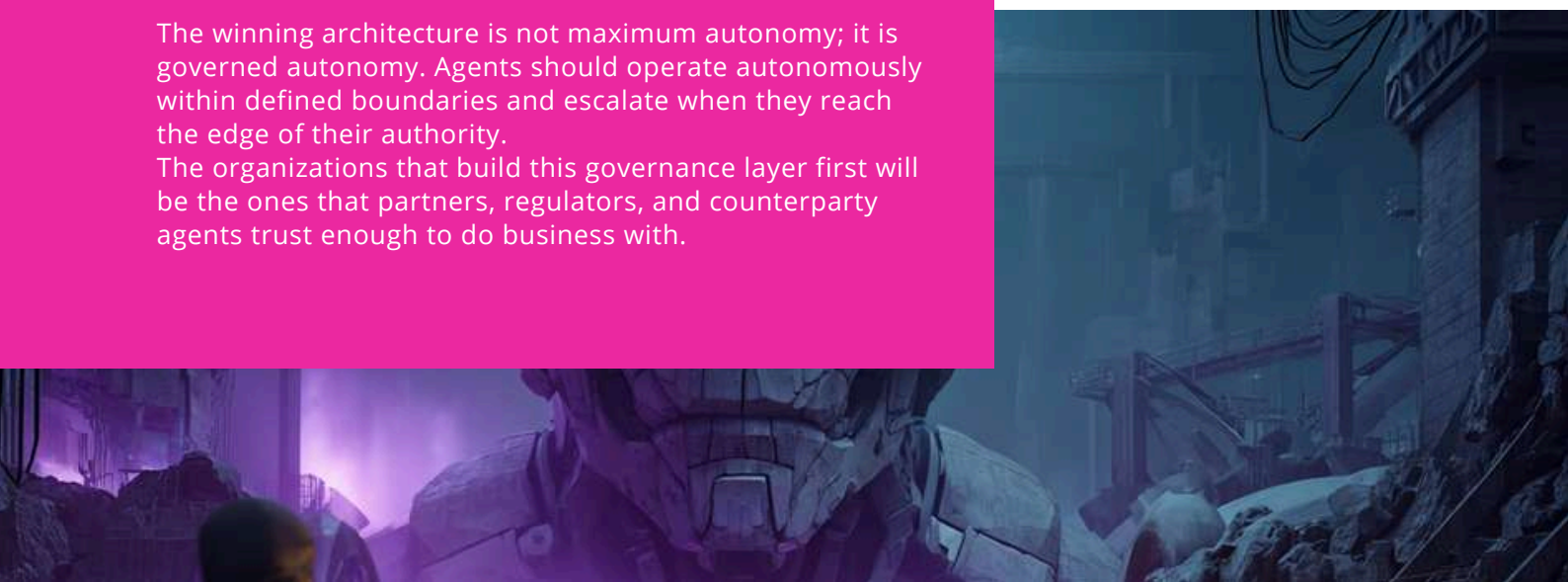
The Control Tower is a business governance layer through which sensitive agent actions must pass before execution. At this gate, the system verifies the agent's identity, checks its delegated authority against its predefined scope, and enforces the policies applicable to that specific action type
This is a decision gate, not a post-incident review process.

In practical terms, this involves verifying whether the agent is permitted to carry out a particular action—such as issuing a refund—and then conducting detailed policy checks, including transaction limits, approved recipient lists, contextual requirements, and jurisdictional restrictions.

— Teranode Group Research Team · Governance & Compliance Framework Contributors

GOVERNING PRINCIPLE

The winning architecture is not maximum autonomy; it is governed autonomy. Agents should operate autonomously within defined boundaries and escalate when they reach the edge of their authority.
The organizations that build this governance layer first will be the ones that partners, regulators, and counterparty agents trust enough to do business with.



1.3 FROM MIMICRY TO MIND

THE THEORETICAL FOUNDATIONS OF AGENTIC AI

Understanding what makes AI agentic requires tracing the conceptual journey from the Turing paradigm to the cognitive architectures that underpin today's autonomous systems. This section provides the theoretical grounding that contextualises the practical frameworks described elsewhere in this chapter.

The Turing Paradigm and Its Limits

When Alan Turing introduced his famous question in 1950—"Can a machine convincingly imitate human behaviour?"—he set in motion a paradigm that would define AI for decades. Through the imitation game, AI focused less on understanding human consciousness and more on copying what humans do.

Within this framework, specialised models multiplied to imitate each human faculty: vision through convolutional neural networks (AlexNet, 2012; YOLO; Segment Anything); hearing and speech through large-scale models such as Whisper; voice generation through models like VALL-E; and language itself through the first wave of Large Language Models (LLMs) starting with GPT (2018) and GPT-2 (2019).

Key Improvements:

Yet, despite their spectacular performance, these systems share the same fundamental limitation: they remain reactive, specialised, and siloed. They perform specific tasks exceptionally well but lack a unified understanding of the world; they cannot reason, explain themselves, or coordinate knowledge across domains.

Human intelligence works differently. A doctor can read a scan, listen to a patient, perceive anxiety, formulate a diagnostic hypothesis, and adapt their language to the person in front of them; all at the same time. This is not the result of separate skills working in isolation, but of a single, integrated cognitive system bringing together perception, memory, emotion, language, and reasoning in real time.
This sharp contrast explains why researchers moved beyond simple mimicry toward modelling the deeper structure of cognition itself, and how understanding, goals, and action are coordinated.

What Makes AI Agentic (The Russell and Norvig Definition)

The foundational definition of an agent was provided by Stuart Russell and Peter Norvig in 1995: "An agent is anything that can be viewed as perceiving its environment through sensors and acting upon that environment through actuators." This definition marked a clear conceptual shift. Rather than defining intelligence in terms of imitation or specific internal mechanisms, Russell and Norvig framed it as intelligent behaviour emerging from a continuous loop of perception, decision, and action within an environment.

The framework is deliberately agnostic about how behaviour is produced; it does not assume a human-like mind, a biological body, or any particular form of reasoning. This generality is its strength, as it applies equally to humans, animals, physical robots, and the digital agents now operating in decentralised environments.

At its core, an agentic system is organised around five fundamental pillars: perception, which involves sensing the environment to build a representation of the current world state; memory and state, for retaining context, past actions, and learned knowledge over time; reasoning and planning, to select actions based on goals, break complex tasks into sub-steps, and evaluate trade-offs; action, the execution of operations through tools, code, speech, or movement with real-world consequences; and goals and feedback, which allow the system to adapt based on outcomes to refine strategies and correct errors over time.

PILLAR	DEFINITION	DESCRIPTION
Perception	*Sensing the environment*	How the agent observes its environment; reading senses, inputs, and data streams to build a representation of the current world state.
Memory & State	Context & Persistence	What the agent retains over time; short-term working context, past actions, learned knowledge, and long-term storage that shapes future decisions.
Reasoning & Planning	Goals, strategy & decisions	How the agent selects actions based on goals and context; breaking down complex tasks into sub-steps, evaluating trade-off, and handling ambiguity.
Action	Using tools & APIs	How the agent affects its environment; executing real operations through speech, tools, code, or movement with real-world consequences.
Goals & Feedback	Adaptation over time	What the agent is trying to achieve and how it adapts, using feedback from outcomes to refine strategies, correct erros, and improve over time.

Not All Agents Wear Bodies

The agent definition applies most naturally to physical agents such as Waymo's self-driving cars and Boston Dynamics' Spot—that perceive and act in the real world. However, many environments are digital and semantic, where perception and action occur through information rather than physical interaction.

Humans already operate in this manner: a patient first perceives a pathology through their own senses and then expresses these sensations through language, which a physician interprets, reasons about, and acts upon. In such contexts, language becomes the primary interface to the environment for both communication and decision-making.

Today's agentic systems, which began to emerge around 2022–2023, are therefore largely digital agents. These are mostly virtual entities operating in environments composed of text, tools, APIs, and software systems, whose core cognitive engine is an LLM enabling communication, coordination, and goal-directed action.

Digital agents can also receive signals from hardware such as cameras, microphones, or sensors, but these inputs are treated as data streams integrated into a digital environment rather than direct physical embodiment.

The Role of LLMs as Cognitive Interface

In digital agentic systems, LLMs act as the central coordination and communication engine, enabling agents to interact with users and environments expressed through language while connecting to multiple sources of information such as web content, databases, documents, APIs, tools, and external services.

Through language, the agent perceives inputs, reasons over goals, plans next actions, calls tools, and stores or retrieves contextual memory. The LLM is not the agent itself, but the cognitive interface that binds perception, reasoning, action, and memory within a unified loop.

Two design implications follow directly.

First, the choice of LLM is a strategic trade-off; smaller or poorly suited models may produce hallucinations, weak reasoning, or context loss, while very large models introduce high costs in computation, latency, energy, and memory management.

Second, prompt engineering plays a critical role. Even with modest models, well-designed prompts, incorporating clear instructions, structured context, and explicit goals and constraints, can significantly improve agent behaviour. In this case, effective agentic AI emerges not only from powerful LLMs, but from the careful orchestration of LLMs, prompts, tools, and memory into a coherent, goal-driven system.

Where Generative AI Ends and Agentic AI Begins

The distinction is simple and consequential: Generative AI answers you; Agentic AI acts for you.
Generative AI is reactive: when you give it a prompt, it produces a response from text, image, or code, then stops and waits for the next instruction. No plan, no follow-through, no real-world impact beyond the output itself.
It has no plan, no follow-through, and no real-world impact beyond the output itself.

Agentic AI is different. You give it a goal and it figures out the steps, makes decisions, uses tools, and keeps going until the task is done. Generative AI tells you how to book a flight; Agentic AI books it.

The line is crossed the moment AI stops waiting for your next message and starts figuring out its own next move.

The main missing piece in most current systems is full autonomy.
They lack true long-term memory, personal goals, and the ability to act independently without human input.
These gaps define the frontier that agentic AI infrastructure and governance frameworks, as discussed throughout this report, are working to close.

THEORETICAL ANCHOR

The core insight from the Russell-Norvig framework is that agency is defined by behaviour within an environment, not by internal implementation.

This means the governance questions raised in this report regarding identity, accountability, authorisation, and economic participation apply equally to any agent; regardless of whether its cognitive engine is a transformer model, a rule-based system, or a future architecture.
The framework is future-proof in a way that implementation-specific definitions are not.

Nehla Debbabi · AI Consultant & Head of AI Academy, Novation City Sousse
NVIDIA DLI Ambassador

04 · The Regulatory Lens: Autonomy as an Accountability Threshold

The preceding sections in this chapter map agentic AI from three vantages: the evolution of on-chain agents as economic actors, the governance architectures that make autonomy tractable, and the theoretical definition of agency itself.

From a regulator's seat, these are three routes to the same question and it is not the question the field usually asks. The useful question is not "What is an autonomous agent?"; it is "At what point does an autonomous system acquire a principal whose accountability attaches to its actions?" Definitional precision matters, but for oversight purposes, the threshold that bites is accountability, not capability.

This reframes the first practical question. The useful marker of regulatory interest is a three-factor trigger: goal-directedness, environmental awareness, and the capacity to take unsupervised, consequential action on an identifiable principal's behalf.

The first two factors are technical; the third is doctrinal. A system that satisfies all three is doing something the law already recognises: it is acting on delegated authority. What makes it novel is the mechanism, not the relationship.

ADGM's DLT Foundations regime, discussed later in this report, already applies this logic to decentralised protocols: accountability attaches to whoever exercises real control, including through the operational rules encoded in the protocol itself (Section 27(5) of the DLT Foundations Regulations).

The architecture neither requires nor confers legal personality on the system; it simply traces back to the principal.

The same lens clarifies the classification of economic agents. An AI system that produces advice or analysis raises accountability-for-accuracy questions familiar to any regulator of professional services: was the output produced with reasonable care, were its limits disclosed, and did a human or legal person hold out its reliability?

Conversely, an AI system that holds or spends funds on a principal's behalf raises accountability-for-conduct questions equally familiar: was the authorisation bounded, was the scope observed, and were the effects reasonably foreseeable?

These are different evidentiary burdens resting on a single doctrinal spine. The distinction between informational and economic agency is operationally important, it drives what needs to be logged, what thresholds need to be bounded, and what audit trails are adequate but it does not call for a new category of regulated entity. It calls for the disciplined application of categories that already exist.

None of this is a claim that existing frameworks handle every case cleanly. Apportionment across base-model providers, integrators, deployers, and end-users remains genuinely contested. Cross-border controller identification is a live operational challenge, not a solved one.

Furthermore, the case law that will ultimately stabilise these questions has barely begun to develop. But the direction is clear: the regulatory threshold for agentic systems is the point at which authority flows. The governance architectures described elsewhere in this chapter, such as Know Your Agent (KYA) and the Control Tower, are the practical instruments for making that flow visible.

How accountability then distributes across the development and deployment stack, and what it means for liability allocation, identity, and the case for new safe harbours, is the subject of Chapter 6.

Dmitry Fedotov, Head of Emerging Technologies, ADGM Registration Authority

02 INFRASTRUCTURE CHOICES AND CONSTRAINTS

INTRODUCTION

Building for agentic AI is not the same as building for conventional software.
The infrastructure decisions made in the next 12 to 24 months will determine which studios and developers can operate at scale, which will be locked into dependencies they cannot escape, and which will have built foundations that remain stable as the underlying model landscape continues to shift.
Three structural tensions define this space.
.
The first is the trade-off between latency and reasoning quality, as large reasoning models are slow and real-time game environments are unforgiving.

The second is the risk of runaway agent loops, which are systems that recursively call themselves or generate inference costs that scale exponentially with complexity.

The third is the question of model dependency: should you optimise deeply for a single provider's ecosystem, or build for portability at the cost of near-term performance?

The contributors to this chapter argue, from different angles, that the answer to all three tensions is architectural discipline.

Latency can be managed through iterative output models, and runaway inference can be bounded through circuit-breaker design.
Model dependency can be controlled through a layered architecture that keeps principles and governance stable while treating models as transient components.
None of these solutions are technically complex, but all of them require deliberate design decisions that most teams avoid until they are forced to confront the consequences.

2.1
LATENCY, REASONING, AND REAL-TIME RESPONSIVENESS

A common misconception in AI system design is the expectation that speed and certainty can be maximised simultaneously, as the latency of reasoning models imposes a structural trade-off between responsiveness and confidence.
The more productive framing is not which dimension to sacrifice, but how to dynamically sequence between them depending on context.

The most effective architectural pattern treats system responses as iterative rather than final. The system produces a low-latency output sufficient to support immediate interaction, while continuing to refine its response as more context and feedback become available.
Over time, the system converges toward higher-confidence outputs, moving from approximation to precision. In this model, responsiveness and accuracy are not mutually exclusive; they are sequenced.
For live game environments, this means designing agent pipelines where the first output is always fast enough to preserve the real-time feel of the experience, even if it is not the agent's final assessment.
Follow-up refinement happens asynchronously, and the player-facing layer is updated progressively rather than waiting for a completed reasoning chain.

Teranode Group Research Team ·
Infrastructure Architecture Contributors

2.2

MODEL-AGNOSTIC ARCHITECTURE: PRINCIPLES OVER PROVIDERS

AI system architecture should be intentionally structured in layers to achieve both immediate performance and future flexibility.

The foundational layer includes essential elements, such as governance frameworks, security protocols, accountability systems, and compliance requirements, which must remain consistent regardless of the models or providers involved. Development workflows and orchestration logic belong to the process layer.

Finally, the tool layer, which consists of models, APIs, and inference providers, is intentionally designed to be temporary and changeable.

A significant error many organisations commit is letting short-term technology dependencies influence their long-term architecture. While integrating with a particular provider's ecosystem may seem more efficient and cost-effective initially, problems often arise later when the provider alters pricing, retires an API, or when market dynamics change. Keeping stable foundational systems separate from tools that evolve enables organisations to achieve both robust performance and adaptability without sacrificing strategic alignment.

For gaming specifically, the implication is that the proprietary value of an agent's reasoning engine should not reside in its dependence on a particular model. It should reside in the training data, the domain-specific fine-tuning, the governance layer, and the accumulated performance history, as these are assets that are portable across providers.

2.3

INFRASTRUCTURE AS OPERATING STRATEGY: A FRAMEWORK FOR AGENTIC AI DEPLOYMENT

The Central Design Problem

The first wave of enterprise interest in generative AI focused on models, specifically regarding size, quality, and prompt interfaces. The agentic era shifts attention downward into infrastructure, because agents are no longer isolated request-response services.

They are persistent, tool-using systems that retrieve data, produce intermediate artefacts, call applications, and operate across multiple business systems. This changes the core question from "Which model should we use?" to "What operating environment can safely support autonomous work?"

Infrastructure for agentic AI is best understood as a managed substrate for autonomy. NVIDIA describes the technical substrate as an "AI factory," which is a system designed to convert data, models, and compute into continuous intelligence output. McKinsey frames the same challenge in organisational terms, suggesting that firms are moving toward workflows where human teams supervise larger numbers of specialised digital agents.

The practical implication is that infrastructure choices are inseparable from deployment constraints. A bank may optimise for governance and data locality, a factory for edge latency and resilience, and a global software company for platform standardisation and elastic scaling. Agentic AI does not eliminate these trade-offs; it sharpens them.

The Four Principal Infrastructure Choices

Deployment topology is the first key decision. Centralised AI factories provide standardisation and efficient resource use, which is ideal for large-scale models.
Hybrid setups combine central orchestration with on-premises tools and data, while edge-heavy designs move inference closer to operations for better latency and connectivity. A federated mesh enables local execution with shared control.
In gaming, real-time demands and player location make hybrid and edge approaches the most practical.

Data architecture is the second decision. Agentic AI scales only when data is discoverable, semantically consistent, and governed by stable interfaces. Organisations choose between tightly centralised data products, domain-owned products connected by shared semantics, or mixed approaches that retain local ownership while exposing machine-usable contracts. For agentic systems, weak semantics translate directly into brittle agent behaviour, as an agent can only reason over what it can retrieve and reliably interpret.

Runtime control is the third decision. Organisations must decide whether to allow agents broad tool access inside secure sandboxes, restrict them to narrow approved workflows, or insert human approvals at key boundaries. The production model that emerges from NVIDIA's guidance is one where agent freedom is paired with controlled execution environments and replayable artefacts. This is not a choice between "open" or "closed," but a question of how much delegated autonomy can be supported with acceptable risk at each tier.

Cost, utilisation, and platform standardisation form the fourth constraint. GPU count alone is a poor decision rule, as costs arise from idle accelerators, repeated retrieval, duplicated embeddings, overprovisioned storage, tool-call overhead, and poorly controlled agent loops. Infrastructure choices need to be evaluated in terms of end-to-end cost per useful task completion. Platform standardisation, through shared model gateways, common tracing, reusable tool adapters, and centralised evaluation services, can reduce duplication, but it creates counterpressure from local constraints in business units with regional or domain-specific requirements.

The Physics of Agentic Latency

Agentic systems amplify latency because a single user request may expand into retrieval steps, reasoning loops, tool calls, and verification passes. What looks acceptable for a chatbot becomes a serious problem for agents coordinating multiple services.
Infrastructure teams face explicit choices, such as collocating vector stores with inference runtimes, placing tool adapters near data sources, replicating selected models to regional clusters, or accepting slower but more centralised processing.

Edge and hybrid deployments are common in real-time gaming due to round-trip delay and unstable connectivity, which make centralised designs impractical.

Availability is crucial for long-running agents, which require stable access to context, tools, and identity services. Failures in these areas can disrupt workflows, cause loss of state, or leave tasks incomplete.
Persistent workspaces and durable artefact storage help ensure both reliability and productivity.

Governance as Infrastructure, Not Policy

Autonomy without control is not an enterprise architecture, as governance must be moved into the infrastructure itself rather than being left as external paperwork.
Agent policies on impersonation, system access, data reading, and approval requirements should be implemented as technical controls rather than static documents. This approach leads to specific infrastructure choices regarding identity (agent-specific versus user context), approval (manual for risky actions versus automatic), and isolation (strict sandboxing versus wider environments with safeguards).

Governance can slow operations if every agent step needs manual review, which ultimately undermines automation. A layered control model solves this by automating low-risk actions, gating medium risks through policy, and escalating high-risk tasks to humans with full records.

Observability, evaluation, and audit logs are essential for gaming apps handling economic transactions and live player data.

A Decision Framework for Infrastructure Design

A practical sequence of design questions guides infrastructure choices:

1. First: What is the dominant constraint faced today,is it latency, governance, data locality, sovereignty, or total cost of ownership (TCO)?
2. Second: Where must data remain, and which agents or tools need to move closer to it?
3. Third: What level of agent autonomy is actually required to produce value?
4. Fourth: Which controls can be automated, and which must remain as human approvals?
5. Fifth: What should be standardised centrally, and what should be allowed to vary by domain or region?
6. Sixth: How can the consistency of agentic output and constant quality be assured?

A Structured Approach to Infrastructure Design

Designing an effective infrastructure for agentic AI systems requires addressing a purposeful sequence of critical questions, each helping to clarify the necessary trade-offs and priorities.

1. Identifying the Primary Constraint
Begin by determining the dominant constraint for your environment when considering whether latency, governance requirements, data locality, or cost pose the most significant challenge. This assessment will shape subsequent decisions and ensure the solution is tailored to the organisation's needs.

2. Locating Data and Agent Proximity
Next, establish where data must reside and decide which agents or tools need to be positioned closer to these data sources to optimise performance and maintain compliance.
This step is vital for ensuring efficiency and meeting operational requirements.

3. Defining Agent Autonomy
Consider the degree of autonomy your agents require to generate true value and assess what level of independent decision-making and action is necessary, ensuring this aligns with organisational goals and risk tolerance.

4. Automating Controls and Human Oversight
Identify which governance controls can be reliably automated, and which must remain subject to human approval.
This distinction balances operational speed with oversight, supporting both safety and efficiency.

5. Centralisation Versus Domain Variation
Finally, determine which elements should be standardised across the organisation and which can be varied according to specific domains or regions. This maintains uniformity in necessary areas, yet permits adjustments to suit local needs wherever advantageous.

6. Quality

Consistency and quality of the agentic output is critical. As LLMs are prone to hallucinations, the design of the agentic pipeline must account for this by implementing quality loops, iterative optimisation, and programmatic consistency checks.

The optimal architecture is usually neither fully centralised nor fully distributed. Most organisations, including gaming studios with live services, use a layered model with common control, observability, evaluation, and identity standards, placing data and runtime functions close to business needs.

Infrastructure decisions now determine the practical ceiling of agentic AI. Model improvements matter, but they do not remove the need for sound choices about topology, data, control, and economics.
In the agentic era, infrastructure is the operating strategy for AI.

03 AGENTIC AI IN LIVE GAME ENVIRONMENTS

INTRODUCTION

The game environment is one of the most demanding test cases for agentic AI, as it combines real-time responsiveness requirements, persistent world states, and economic systems with real-world value. It also faces a player base that will actively probe for exploits the moment any automated system enters the environment.

It is for exactly these reasons that gaming is one of the richest laboratories for understanding how autonomous agents behave when deployed at scale.

Live game deployments raise primarily organisational and human questions, where studios are adjusting to AI-driven changes in workforce and creative processes. Roles are being compressed, redefined, or eliminated, and the shift from "doer" to "orchestrator" challenges those with careers focused on craft and execution.

At the infrastructure level, live game environments raise questions about legal identity and machine accountability that no existing regulatory framework has fully answered.

When an AI agent earns and spends real-world value inside a game, the question of who serves as the legal counterparty is no longer rhetorical; it has immediate implications for KYC obligations, liability allocation, and payment processor compliance.

The contributors to this chapter address both the human reality of the transition and the identity and compliance infrastructure that must be built to support it.

3.1

HOW GAME STUDIOS ARE HANDLING THE TRANSITION TO AI

How Game Studios Are Handling the Transition to AI

A Two-Tier Adoption Pattern

Game studios are navigating a moment of intense internal tension.

Creative staff, including artists, designers, and narrative writers, have largely responded to AI with scepticism or outright resistance. This fear is understandable, as we have recently seen craft devalued, careers disrupted, and work that took years to develop suddenly made secondary. This emotional reality has directly shaped how many studios initially communicated about AI, with leadership in many cases adopting a public "no AI" position as a morale management strategy.

That position has not held. As the rest of the industry moved, leadership confronted a straightforward competitive reality: ignoring AI means falling behind on both quality and cost.

Messaging has since shifted from rejection to augmentation. Studios are now using AI internally for QA, iteration, task triage, and asset exploration, keeping humans in charge of creative direction and narrative while automating the mechanical and repetitive layers of production.

A second, less controversial adoption track has emerged in peripheral roles such as marketing, community management, customer support where these teams have been more receptive to AI because the value proposition is clear to them in term of scale, speed, and consistency.

Also content drafting, ticket triage, personalised outreach, and community moderation can all be handled at volumes that were not possible with manual effort and the result is role compression rather than elimination, fewer people handling the same functions, with AI as the scaling layer.

EMERGING PATTERN

A two-tier adoption model is consolidating across the industry.

Creative teams receive AI as a collaborator wrapped in guardrails, serving as a tool for exploration and iteration, with humans retaining final authority.

Peripheral and operational roles receive AI as a scaling mechanism that compresses headcount over time without eliminating it overnight

Jobs: Compression and Redefinition

A notable shift is occurring in the engineering sector. While there remains some scepticism among developers regarding AI's capacity to supplant advanced architectural expertise, this scepticism is partially justified, as current AI technologies have not yet achieved parity with experienced professionals in system design.

Nevertheless, in many areas of the technology industry, AI-assisted and AI-driven coding are altering traditional workflows. Routine code generation, framework creation, and substantial feature development are increasingly being handled by AI, with human oversight focused on supervision, validation, and direction.

In games and entertainment, the same trajectory is visible. Leadership is actively evaluating how much of the technical pipeline can be streamlined with AI-assisted coding, viewing it not as a distant possibility, but as a near-term cost and cycle-time question.
Junior roles are especially exposed, because their work patterns are closest to what AI can replicate. Senior engineers are safer for now, but expectations are evolving rapidly, moving away from manual execution and toward oversight, curation, and the quality control of AI-generated work.
The broader picture is a move from "many people doing the same thing manually" to "fewer people using AI to do more, and better." The job titles may persist, but the expectations are fundamentally different. We are seeing the rise of "AI-squad captains" who plan, guide, and curate agent-generated output rather than producing everything from scratch.

— Loic Fontaine · Web2 / Web3 Cross-Industry Analyst, Gaming & AI Transition

What Comes Next: End-to-End Agent Pipelines

The real frontier is agentic systems that do not just assist individual functions but actively plan, coordinate, and execute across the entire game development lifecycle. An AI can plan features, generating code and assets, optimising performance, and deploying builds across platforms all under human supervision, while developers become orchestrators, setting goals, defining constraints, reviewing outputs.

Inside the game world itself, the player-facing shift is towards NPCs that feel genuinely alive, not because they follow more complex scripts but because they maintain context, respond dynamically, and evolve in ways that make the world feel personal and responsive, and that shift in player experience is already beginning to appear in experimental titles and engine demos.

— Loic Fontaine · Web2 / Web3 Cross-Industry Analyst, Gaming & AI Transition

3.2
MACHINE IDENTITY AND LEGAL COUNTERPARTY IN LIVE GAME ECONOMIES

Whenever an AI agent in a live game earns or spends something of real-world value, the legal counterparty is always a person or a legal entity,not the agent itself.

Legal responsibility and standing consistently remain with the individual or business that the agent represents. This is also the case under both the EU AI Act, which assigns obligations to providers and deployers as human or organizational actors, and FATF standards, which focus essentiality on customer due diligence on the person on whose behalf a transaction is conducted.

What is emerging is the need for a complementary identity layer at the agent level. This is not a substitute for KYC or KYB, but a new form of machine identity that enables authentication, authorisation, traceability, and policy enforcement for automated actors. The Know Your Agent (KYA) framework, developed by Teranode Group, provides a solution for agent identity that is cryptographically verifiable, directly associated with a human owner, and enforceable across various platforms and payment systems.

The challenge is not definitional but technical, where AI agents are relatively easy to impersonate, and they should not be assumed to be reliable custodians of secrets; the credentials stored or managed by an agent may be exposed through prompt manipulation or adversarial interaction.

Trust in agent identity cannot rest solely on the credentials the agent holds.

The Teranode Group has concentrated its research on using hardware-based secret management together with runtime attestation to create secure agent identities, as these methods aim to guarantee that the execution environment and system integrity are trustworthy, rather than relying only on stated attributes. While a stronger machine identity does not make the agent a legal counterparty, it does ensure that the person or organisation behind the agent can be held accountable in a safer and more reliable way.

— Teranode Group Research Team · Machine Identity & Compliance Framework

3.3 AGENT-NATIVE STUDIO WORKFLOWS: DESIGNING FOR AI AS A FIRST-CLASS OPERATOR

Studios successfully navigating this transition are not just adding AI to their existing workflows, they are redesigning their processes to centre around it.

Rather than having people perform every task, the new model places humans in roles of direction and quality control, while AI agents take care of specific production phases, such as asset creation, balancing, QA, localisation, and live-ops adjustments.

With agent-native workflows, three people can do the work of fifteen. Instead of adding AI to human processes, the goal is to design systems where agents have defined roles, measurable outcomes, and checks at each stage.

Critical Design Principle

An agent that is 80% reliable and fails silently is worse than no agent at all.

The real engineering work is building the feedback loops that catch the other 20%.

Verification layers, output scoring, and exception escalation are not optional additions , they are the core of a production-grade agentic workflow

— WAM · Award-winning play-and-earn gaming platform · BGA Member · 4M+ registered users

3.4
MACHINE IDENTITY AND THE PRINCIPAL-AGENT ARCHITECTURE

The Agent as an Instrument, Not an Actor

In all contemporary legal frameworks, the agent is never considered the legal counterparty; this role is always fulfilled by the principal, meaning the individual or entity that deploys the agent. This situation is comparable to algorithmic trading, where transactions conducted by an algorithm are legally attributed to the deploying firm rather than to the software itself.

In this context, the agent serves merely as an instrument of the principal's will.

This distinction carries considerable implications for Know Your Customer (KYC) protocols. KYC procedures should be performed on the principal, not the agent, since the agent's compliance classification is derived from the principal.

Furthermore, their activities are restricted by the scope authorised through the principal's verified account. Efficiently designed systems help payment processors work within this framework, as the main issues typically arise only when there is a lack of transparency in the principal-agent relationship.

These challenges generally stem from system design rather than any fundamental legal ambiguity.

Attestation Over Credentials

Trust in agent identity cannot rest on credentials alone. As noted earlier in the report, agents are software and credentials can be exfiltrated; therefore, the more robust approach is attestation-based identity. Under this model, the agent's actions are cryptographically bound to its principal through on-chain signatures and session keys with bounded permissions. The right question is not "Who is this agent?" but "What is this agent authorised to do, and by whom?"

One effective way to implement this model is through ERC-4337 smart wallets with scoped session keys. These session keys specify the actions an agent can take, the spending limit, and the duration, creating enforceable authorisation limits directly on-chain. This is far superior to relying on application-level protections, which are more easily circumvented.

— WAM · Award-winning play-and-earn gaming platform · BGA Member · 4M+ registered users

3.5
THE AGENT ECONOMY IN LIVE GAMES: WHERE IT BECOMES SOMETHING ELSE

When the Game Economy Evolves

When agents operate continuously and optimally never sleeping and never bored,the nature of the game world changes. This transformation is not inherently negative; rather, it depends entirely on the intention and objective of the agent's owner.
When an agent accumulates resources for exchange, advances through leaderboards, or operates as an economic participant on behalf of its principal, it constitutes significant economic activity.

The game world develops into more than just a place for social interaction; it becomes a persistent economic system. Instead of debating whether these systems remain "games," it is more useful to look at the objectives of the agent's owner and determine if the game's economic setup truly helps them reach those goals.

On Interoperability: The Infrastructure Already Exists

The existing infrastructure and APIs from the Open Metaverse Alliance and previous interoperability projects are already sufficient. Large language models can interact with any service using a standard interface, which effectively eliminates the need for custom integration.

The main challenge now is providing agents with persistent, portable contextual memory, ensuring they understand user preferences, history, and objectives across different platforms.

On Distribution: Moving Beyond the App Store Bottleneck

This period marks a significant shift in distribution methods. As agents gain the ability to discover and play games independently, they bypass the traditional bottlenecks of the Apple App Store or Google Play Store.

Consequently, traditional distribution channels lose their dominance.

For game studios, this presents entirely new avenues for reaching users. Instead of competing for spots in algorithm-driven storefronts, studios can focus on making their games accessible to both autonomous systems and human players. A player can now program an agent with a specific strategy to find games, build skills, and provide feedback. An orchestrator agent can then assess the market and assign sub-agents as required. This is not a futuristic concept; it is a natural progression based on existing infrastructure.

On Blockchain as the Essential Layer for Agent Monetisation

From a blockchain standpoint, agent-driven game economies are uniquely suited for innovation.
This technology enables three distinct features absent from traditional Web2 free-to-play games: agent-level reputation and identity, direct wallet management, and a token-based economy where agents can earn, spend, and exchange assets of real value.

In conventional free-to-play games, agents may "farm" resources, but they lack meaningful participation in the broader economy. They do not possess assets with external value, cannot engage in peer-to-peer transactions, and cannot transfer their activity history outside the game.

Blockchain addresses these limitations. Agents in decentralised games can accumulate tokens with actual exchangeable value and actively support market liquidity. This fulfils the long-standing ambition of enabling non-player entities to enhance market liquidity, not as a loophole, but as an integral element of the economic framework.

— Sébastien Borget · Co-founder & COO, The Sandbox · President, Blockchain Game Alliance

3.6

WHY GAMING IS AGENTIC AI'S NATIVE HABITAT

AI agents have been part of gaming since well before large language models arrived, as early versions relied on reinforcement learning, where agents learned by receiving rewards or penalties. Landmark systems like AlphaZero (2017) rivalled or surpassed professional human players by running countless simulated games rather than studying manuals.

Games were an ideal testing ground because they possessed defined rules, consistent feedback, and long-term goals. They offered enough complexity to reflect reality, yet remained safe environments for repeated failure. This changed significantly with the rise of LLMs.

AI Dungeon (2019) was among the first major games powered by a language model, creating endless storylines on the fly. Later, in 2023, Voyager introduced an LLM-driven agent in Minecraft that autonomously explored, developed a lasting set of skills, and adapted across different worlds without human intervention. The progress continued with Project Sid (2024), which enabled 1,000 self-governing agents in Minecraft to build a civilisation complete with economies, government, and spontaneous religious behaviour. This was a turning point, as these actions were not programmed; they emerged as agents pursued their own goals within a shared space.

Gaming has showcased every stage of agentic AI evolution, from simple reward-based systems to sophisticated language-driven reasoning, making it an unparalleled testing ground.

The same features that make games excellent laboratories—clear rules, measurable results, and economic and social systems also position games as the natural starting point for economically active autonomous agents. Advancements in gaming serve as indicators for future trends in fields where agents are accountable for tangible results.

Historical Progression

Each step in the industry mentioned above expanded what agents could do and the environments in which they could operate. The gaming industry has served as the incubator for every major transition in AI agency.

— Nehla Debbabi · AI Consultant & Head of AI Academy, Novation City Sousse · NVIDIA DLI Ambassador

3.7

WHAT PLAYERS ACTUALLY FEEL: THE EXPERIENCE OF AGENTIC AI IN LIVE GAMES

The majority of industry discussions regarding agentic AI centre on pipelines, governance, and compliance.

This is a logical focus, however, end users engage directly with the game environment rather than internal studio processes. Their acceptance of AI is primarily influenced by whether the experience maintains a sense of fairness, significance, and vitality, rather than by technical specifications.

Players do not bond with technology; they bond with worlds. The moment a game feels alive, truly unpredictable, and filled with events that no one scripted, is the moment they stop playing to pass time and start playing because they have to know what happens next.

That contract is fragile, and the moment it breaks is equally clear. An agent optimised to win, rather than to play, stops being a participant and becomes a system, and players perceive that distinction faster than any internal metric will catch it.

Gamers do not need agents to be human, and they are remarkably comfortable with AI when it feels like a fair part of the world they inhabit. What they will not forgive is the sense that the game is working against them rather than around them. That shift, from playing inside a world to playing against the machinery behind it, is the line no studio should cross.

Bugs get patched; broken trust does not.

The best agent deployments are those that players never consciously notice. They fill the gap when a squad is short, make an encounter feel as though it could have gone differently, and give a world the texture of something that keeps moving even when no one is watching. That is the role AI plays well: not centre stage, but everywhere in the background, making the whole environment feel more real.

The boundary is not technical; it is felt. The moment an agent starts dominating rather than supporting, outplaying and outpacing players until human input feels irrelevant, it stops serving the experience and starts competing with it. A game that players feel they cannot meaningfully affect is a game they will eventually leave.

Rod Oliveira · Web3 Gaming Analyst, Sura Game

3.8
WHICH STUDIOS ARE GETTING IT RIGHT, AND WHY

Getting this right has less to do with having the most sophisticated technology and more to do with knowing where AI belongs in the experience, and having the discipline to keep it there.

The early signals are already visible. In EVE Frontier, where player-built Smart Assemblies run custom logic inside a shared live universe, the AI layer supports economic activity and emergent behaviour without making itself the story.

In PUBG, the ACE-powered Ally functions as a teammate by looting, communicating, and adapting in a way that extends what a player can do, rather than replacing what they came to do themselves. What these deployments share is not technical sophistication, but restraint. The AI is not the feature; it is the reason the feature feels better.

The AI operates in the background by filling gaps, driving events, and adding texture to encounters that would otherwise feel predictable. Players feel the difference without being able to name the mechanism, and this invisibility is not a limitation of the deployment, but the measure of its success.

The temptation for studios to act otherwise is understandable. A studio that has invested in agentic AI wants to talk about it, and it often becomes a marketing headline, a differentiator, and a signal to the market that something new is happening. This is precisely where things start to go wrong. When AI becomes the story, it also becomes visible.

Gamers start noticing the seams of an agent that responds too perfectly, moves too efficiently, or operates by rules that clearly do not apply to anyone else in the game.

The moment a player thinks, "That is a bot," rather than, "That is a tough opponent," the experience has already failed. Immersion does not erode gradually; it collapses at the first moment of recognition.

The studios that will get this wrong are those treating AI capability as a product feature to be showcased. Whereas those that will get it right , are those treating it as infrastructure, something that makes everything else work better without ever asking for credit.

The goal was never for players to be impressed by the AI. It was for players to be more deeply immersed in the world because the AI was there, and those are not the same objective. The distance between them is where most deployments will succeed or fail.

The studios that win with agentic AI will be the ones their players never think to question.

Rod Oliveira · Web3 Gaming Analyst, Sura Game

3.9

THE LEGITIMACY PROBLEM: WHEN COMMUNITIES ACCEPT AGENTS AND WHEN THEY DO NOT

TThe history of bots in games is not complicated: gamers found them, hated them, and organised to get rid of them. The assumption was always that bots were hostile, something introduced against the rules, against the community, and against the spirit of fair play.

Agentic AI changes the surface of that problem without changing the underlying one. These agents are sanctioned; they are introduced by studios, sometimes named and marketed, and occasionally the centrepiece of the product. However, official status does not automatically confer legitimacy. Gamers do not accept something because a studio endorses it, they accept it because it feels right inside the world they have chosen to inhabit.

TFour things determine whether that acceptance happens or not.

The first is transparency. This is not about full disclosure, but rather ensuring players know what they are dealing with. Hidden AI reads as deception, while overexplaining it dismantles the illusion the game depends on. The studios that find that balance treat it as a design problem worth solving carefully, rather than a legal disclaimer to be buried in an update note.

The second is consistency. If agents operate under different constraints than human players, such as different economic rules, physical limits, and consequences for failure, players will find them, as they always do. This discovery does not feel like an interesting feature; it feels as though the game was never fair to begin with.

The third is purpose. Gamers are remarkably forgiving of AI that is genuinely there to make their experience better, yet they are remarkably unforgiving of AI that exists simply to make the studio's metrics look better. The difference is not always visible in the design, but it is always visible in the behaviour.

The fourth is the speed of detection. Player communities are faster and more accurate than any internal system at detecting when something is wrong. They are not reading dashboards; they are feeling the game. If an agent's behaviour produces a pattern that does not fit the world, someone in the community will name it, usually before any monitoring tool flags it, and always before a studio is ready to respond.

What this adds up to is not a design challenge, but a trust architecture problem. The moment players conclude that the game is optimising against them rather than for them, or that the system is working in its own interest rather than theirs, the social contract holding the community together fractures. That kind of fracture does not respond to patch notes or community updates. It responds, if at all, only to time and to demonstrated change.

Most studios do not get that second chance.

Rod Oliveira · Web3 Gaming Analyst, Sura Game

04 PAYMENTS, KYC, AND IN-GAME ECONOMIES

INTRODUCTION

Real money has been moving through AI agents for some time now. This is happening in DeFi protocols, digital commerce platforms, and the earliest gaming deployments where autonomous agents trade, acquire, and manage assets on behalf of their principals.

The question the industry kept treating as theoretical has quietly become operational. What remains is the question of whether the infrastructure built to handle human financial activity can be adapted, fast enough and carefully enough, to handle agent-initiated transactions at scale.

The honest answer is that it cannot, at least not yet.

The payment rails that process these transactions were designed around a fundamental assumption: that a human being, identifiable and accountable, sits behind every financial action

That assumption is load-bearing, as it holds up the entire compliance architecture, including KYC obligations, AML monitoring, transaction attribution, and liability allocation. When an autonomous agent initiates a transaction, that assumption does not just bend, it breaks.

The major networks are not standing still. Visa, Mastercard, and Stripe are actively building infrastructure for agent-initiated transactions, and the direction of travel is clear. However, the current reality is that most agent payments are still flowing through legacy rails that were not initially designed for them, handled through workarounds and inherited compliance frameworks that do not fully accommodate current and future transaction patterns.

On the regulatory side, the gap is wider still. No jurisdiction has yet produced a definitive framework for agent-initiated financial activity.

Developers building agent payment systems today are navigating a space defined by obligations written for human actors, AML standards which are still being extended to cover autonomous systems, and a meaningful distance between what the technology can already do and what any compliance framework has caught up to accommodate.

That distance is not closing as fast as the deployments are moving.
The result is an industry building on ground that has not fully settled, which makes the architectural decisions being made right now more consequential, not less.
Getting the foundation wrong while regulators are still writing the rules is considerably harder to fix than getting it wrong after the rules exist.

Inside game economies, the challenge is compounded by the specific dynamics of virtual world economics: the risk of AI-driven resource farming that outpaces human economic activity, the inflation implications of agents that can exploit loopholes faster and more consistently than any human player, and the question of how in-game economic design holds up when some participants are not human at all.
The contributors to this chapter address the compliance architecture for agent-initiated payments, the design of programmatic wallets that can operate within regulatory bounds, and the economic design principles that protect game economies from agent-driven exploitation.
.

4.1
KYC RESPONSIBILITY FOR AGENT-INITIATED TRANSACTIONS

The law has not changed to accommodate autonomous agents, and it does not need to, at least not yet.

Under every current framework that matters, the relevant legal person in an agent-initiated transaction is the human or enterprise behind the agent, and not the agent itself. While the developer built the system, they are not the accountable party; rather, the user or organisation on whose behalf the agent acts carries the responsibility.

This is not a novel interpretation. It follows directly from existing FATF standards, which have always located compliance obligations at the point of human or institutional accountability rather than at the point of technical execution.

Payment processors are applying exactly the same logic: when an agent initiates a transaction, the KYC obligation runs to the principal, not to the machine. The agent is the instrument; the person or entity directing it carries the responsibility.

That clarity is genuinely useful. It means the compliance architecture for agent-initiated transactions does not require a new legal category or a regulatory overhaul. Instead, it requires knowing who the principal is, verifying them properly, and ensuring that the agent's actions remain traceable back to that accountable party at every step

The practical answer emerging from production deployments is straightforward: design payment flows so that agent context travels with the transaction.

Every downstream processor in the chain should be able to see not just that a transaction occurred, but which agent initiated it, under what authority, and on whose behalf.

This requires giving agents verifiable identities—not informal labels, but cryptographically secured machine identities built on established standards, such as X.509 certificates and W3C Verifiable Credentials.

When a transaction arrives carrying that identity, a processor can validate the agent first, then resolve it to the human or corporate principal behind it, and apply the existing KYC or KYB framework from that point forward in the normal way.

This is what Know Your Agent (KYA) means in practice. It is not a replacement for the compliance infrastructure that already exists. Rather, it adds the layer that sits in front of it, a new entry point designed specifically for agent-initiated activity, built to hand off cleanly to the human accountability frameworks that have always governed what happens next.

The compliance stack does not need to be rebuilt, but it needs a front door designed for the actors now walking through it. KYA is that front door; everything behind it remains the same.

4.2 PROGRAMMATIC WALLETS AND BOUNDED AGENT AUTONOMY

A wallet designed for a human assumes a human is present. It asks for a password, waits for a confirmation, and expects someone to approve each action before anything moves. This is the standard model, and it does not work for an autonomous agent operating at speed, across multiple systems, without a person in the loop at every step.

Programmatic wallets solve that problem by replacing per-transaction authentication with per-agent policy. Instead of asking for approval each time, the wallet operates within rules defined in advance. These rules define which transaction types are permitted, which counterparties are approved, how much can move in a single transaction, and how much can move in a day. The agent executes within those boundaries automatically, and when something falls outside them, it stops.

Design requirements must include granular permissions by transaction type and recipient, strict size and volume limits, whitelisted counterparties, real-time monitoring with automated circuit breakers, and a clear separation between technical fund management and legal ownership responsibilities.

It's easy to miss the importance of this last point, just because a programmatic wallet operates automatically, it doesn't bypass compliance requirements; instead, it introduces an obligation that must be addressed within its architecture, this needs to be part of the wallet's original design, not something added afterward as protection.

The governance frameworks discussed earlier in the report, the Control Tower model and the KYA identity layer aren't simply optional features for programmatic wallets but are essential for their successful deployment.

— Teranode Group Research Team · Payments Infrastructure Contributors

4.3 PROGRAMMATIC WALLETS: ARCHITECTURE FOR AGENT-NATIVE FINANCE

Three Design Principles for Agent-Compatible Wallets

Conventional wallet architectures typically require human approval for each transaction. In contrast, agent-native programmatic wallets are grounded in three fundamental principles.

First, they utilise policy-per-agent rather than authentication-per-transaction. Session keys are designed to encode spending limits, permitted contract interactions, and time constraints, thereby removing the need for human authorisation for every transaction while maintaining robust, enforceable controls.

Second, there is a clear separation between ownership and execution. The assets are held by the user's smart wallet on-chain, with the agent's session key operating strictly within predefined parameters. Ownership remains sovereign, and the session key can be revoked at any time. Importantly, agents do not possess the wallet’s primary keys; instead, they are granted scoped execution credentials.

Third, policies are enforced atomically at the contract level rather than the application level. This ensures that even if an agent is compromised, it is incapable of exceeding its authorised permissions, as enforcement is managed at the execution layer, rather than within potentially vulnerable application code.

Compliance at the Account-Abstraction Layer

Most compliance frameworks were built around a simple assumption: one person, one wallet, and one identity. That assumption breaks the moment a single smart wallet starts issuing session keys to multiple agents. When delegation enters the picture, compliance must understand it; it must track who authorised which agent, what they were authorised to do, and within what specific limits.
This challenge cannot be solved at the application layer.
Application-layer controls have no authority over what executes on-chain, and every new agent type introduced creates a new security gap. The only place this model works effectively is at the account-abstraction level, where delegation is native, boundaries are enforceable, and the rules travel with the execution rather than sitting above it.

By anchoring compliance at the execution layer, the system ensures that the agent's permissions are cryptographically bound to the principal's original authorisation. This creates a transparent and immutable audit trail that satisfies existing regulatory requirements while allowing for the speed and scale of autonomous machine activity.

4.4
PROTECTING IN-GAME ECONOMIES FROM AGENT-DRIVEN INFLATION

AI systems do not create entirely new economic risks within game environments; instead, they intensify established ones and hasten their development. Issues such as rapid resource exploitation, inflation driven by arbitrage, and liquidity imbalances have always existed in virtual economies. What AI agents contribute is unmatched speed, consistency, and scale compared to human players.

A well-designed game economy should take an architectural and proactive approach, rather than relying on reactive or punitive methods.

Robust incentives need to be integrated from the beginning to align agent behaviour with the overall economic health of the game, rather than just individual optimisation. Maintaining a resilient in-game economy requires ongoing monitoring of economic metrics, automated circuit breakers that can temporarily halt or limit agent activities if irregularities occur, and clear limits on the rate at which agents can accumulate resources.

AI agents serve as stress tests by exposing underlying flaws in economic design that might otherwise go unnoticed until an exceptionally determined human exploited them. Incorporating agent participation into the design phase encourages the development of more robust and effective economic systems for all stakeholders.

Teranode Group Research Team · Digital Economy Design Contributors

4.5 (A)
IN-GAME ECONOMY DESIGN FOR AGENT PARTICIPATION

Game economies were designed around human constraints, as attention is scarce, time is limited, and decisions are imperfect.

Agents violate all three simultaneously, farming optimally and continuously without fatigue or distraction. Any economy built around human pacing will break when agents enter it at scale.
The answer is not to ban agents, but to build economies that do not depend on human limitations to stay balanced.

To achieve this, developers should adopt four core strategies:

- Rate-limit value extraction at the account level: Implement hard ceilings on earning rates that make farming speed irrelevant beyond a specific threshold. The marginal return on going faster should eventually reach zero for agents and humans alike.

- **Make high-value activities require judgment, not repetition where** agents are effective grinders but poor strategists. Designing the most valuable loops around genuine decision-making novel, contextual, unpredictable shifts the advantage back toward human players where the experience depends on it.

Enforce economic rules at the execution boundary where application-layer restrictions are not enough and a determined agent will find the gap between what the application allows and what the runtime permits. If the constraint is not enforced at execution, it is not really enforced.

- **Align incentives for contribution: Design the economy so agents contribute rather than merely extract. This involves taxing velocity rather than value and creating token sinks in which agents naturally participate, making ecosystem health more profitable than simple extraction.**

If an economy fractures when agents enter it, it was already fragile. Agents do not create structural weaknesses, they find the ones that were always there, and they find them faster than any human player would.

The studios that design for agent participation from day one will have stronger economies for everyone while the ones that treat it as a future problem will discover it is already a present one.

Design Principle
If your economy breaks when agents enter, it was fragile. Agents just revealed it faster. Design for agent participation from day one, not as a retrofit.

— WAM · Award-winning play-and-earn gaming platform · BGA Member · 4M+ registered users

05 (B)
PROGRAMMABLE PRIVACY AND AGENT-NATIVE INFRASTRUCTURE: THE PROTOCOL FOUNDATION

The design principles for agent-native programmatic wallets outlined in the preceding sections identify three requirements: policy enforced at the contract layer, a clean separation between execution rights and legal ownership, and compliance embedded in the execution layer rather than layered on top of it.

Each of these requirements runs into the same structural obstacle on transparent public chains: the more visible an agent's rules and balances are onchain, the more exploitable they become. Solving the signing problem gets you halfway.

The harder problem is enforcing policy at the wallet layer without creating a privacy paradox.

An agent that broadcasts its spending limits, permission scopes, and execution conditions to a public mempool is handing its operational playbook to every MEV bot and competitor on the network. In gaming, this is existential: a guild treasury agent that reveals rebalancing thresholds can be front-run; a marketplace bot that exposes its buy walls can be systematically exploited. Programmatic wallets, to be viable at scale, must be private by default.

The BITE Protocol: Privacy as Infrastructure

SKALE's approach to this problem is its BITE Protocol (Blockchain Integrated Threshold Encryption), which is a consensus-level privacy layer that encrypts transactions before they enter the mempool and decrypts them only after block finalisation. BITE is not a private mempool workaround or an application-layer add-on; rather, it is integrated into the consensus mechanism itself, making it the first protocol to address MEV at that level.

The practical result is what SKALE describes as a "commit-then-reveal" ordering model. Under this framework, an agent's destination, calldata, and intent are sealed from the moment of submission through to finalisation. The transaction settles verifiably on-chain, while the strategy that produced it stays hidden.

BITE Protocol Implementation Phases

BITE is structured in progressive phases, with each addressing a distinct layer of the privacy problem for autonomous agents.

Encrypted Transactions (Phase 1): This phase delivers 100% private data in transit from the wallet to the point of execution. An agent's trade on a guild treasury, such as selling in-game tokens when a price target is hit, reaches finalisation before any MEV bot or competitor sees the intent. Consequently, the sell order executes based on the agent's specific strategy rather than on a front-runner's anticipation of it.

Conditional Transactions (Phase 2): These are encrypted instructions that only execute, and only decrypt, when specific on-chain conditions are satisfied. This makes several advanced scenarios possible: a crafting agent that stores its optimal purchase price on-chain and only buys when an oracle reports the threshold; a sealed-bid auction where all bids decrypt simultaneously; or a PvP prize pool that distributes rewards only after verified match results.

The policy lives on-chain and is fully auditable, yet the strategy remains hidden until the moment it fires. Critically, this moves agents from one-off payments to fully autonomous execution enforced by the chain itself, rather than by a human operator or a centralised scheduler.

Confidential Tokens: This layer keeps payment amounts and wallet balances encrypted while transactions settle verifiably on-chain. This provides agents with a necessary layer of strategic protection. For instance, a marketplace bot negotiating trades does not need to expose its total financial position to execute a single transaction. Similarly, a guild agent distributing loot drops after a raid does not need to broadcast individual payout amounts to the network. In this model, agents hold, move, and distribute value without revealing the specific parameters that would make their underlying strategies exploitable.

Compliance Without Contradiction

The compliance challenge described in Section 04, namely that compliance frameworks built for human-paced, single-account transactions cannot accommodate the volume and velocity of programmatic agent activity, has a structural answer in selective disclosure.

BITE's architecture enables transaction verification to remain on-chain and auditable while sensitive details, such as user balances, agent strategies, and payment amounts, stay encrypted. SKALE describes this as the "barista test," where one confirms the payment but keeps the rest private. Under this model, a regulator can verify that a transaction was legitimate and traceable to an accountable principal without accessing the agent's full financial history or strategic parameters.

Compliance conditions can be embedded directly into Conditional Transactions. A token transfer that only executes if both parties satisfy predefined requirements, such as verified identity, jurisdictional clearance, and spend limits, makes compliance part of the transaction itself rather than a separate review process applied afterward.

When an agent's execution boundaries are defined in a smart contract, the policy becomes a permanent, auditable record. Regulators can inspect the constraints an agent was operating under from the start, rather than investigating what it did after the fact. In this environment, the policy effectively becomes the compliance framework.

Agent-Inclusive Economic Design: Privacy as Competitive Infrastructure

The economic design principles in section 05, including rate-limiting value extraction, designing for human judgement over repetition, and enforcing rules at the execution boundary, all depend on one architectural property that transparent chains cannot provide: strategic diversity.

Agent-Inclusive Economic Design: Privacy as Competitive Infrastructure

The economic design principles in Section 05, including rate-limiting value extraction, designing for human judgement over repetition, and enforcing rules at the execution boundary, all depend on one architectural property that transparent chains cannot provide: strategic diversity.

On a transparent chain, the most successful agent strategy is visible to every other participant, which means it is cloned immediately. This accelerates an arms race and consolidates value around whichever agent can copy fastest, rather than whichever agent is best designed. The result is "herding," a convergence on a single dominant strategy until it collapses under its own weight.

BITE's encrypted execution directly counters this by keeping individual agent positions and strategies private. Agents must develop independent strategies rather than cloning the leader, which produces a healthier and more diverse ecosystem.

This preserves the strategic value of being a well-designed agent, rather than simply the fastest copier.

The incentive architecture described in Section 05, which ties agent rewards to ecosystem health metrics, reputation-weighted access, and conditional payouts that enforce system-level constraints, only functions if agents cannot observe each other's positions and anticipate each other's moves. Private execution is not an optional enhancement to that architecture; it is the fundamental condition that makes it stable.

The Infrastructure Conditions: Zero Gas, Single-Slot Finality, and Dedicated Blockspace

Privacy addresses the strategic exploitation problem. But the operational requirements of autonomous agents in live game environments surface three further infrastructure conditions that existing public chains do not meet.

The first is economic viability at micro-transaction scale. A treasury rebalancing agent executing hundreds of actions per session on a gas-bearing chain has its economic logic destroyed by fee overhead before strategy becomes relevant. SKALE's zero-gas model removes this blocker entirely: agents execute micropayments without fees eroding value, making cents-level transactions for item purchases, loot distributions, and usage-based game mechanics economically viable. As of 2025, the network has processed over 1.8 billion transactions and saved users over $12 billion in gas fees across more than 55 million unique active wallets, figures that reflect the scale at which gasless execution changes the economics of onchain activity.

The second is finality determinism. Multi-step agent plans, covering actions such as authorising payment, settling, distributing loot, and posting results,, all require each step to finalize before the next can proceed safely. Probabilistic settlement across L2s and bridges introduces race conditions and stranded half-plans. SKALE's single-slot finality provides deterministic one-block commitment: an agent can chain actions without MEV risk or extended confirmation windows, which is the execution environment that the governance architectures described earlier in this chapter assume but rarely specify.

The third is throughput isolation.
Application-specific SKALE chains provide dedicated blockspace with predictable latency, preventing congestion from unrelated network traffic from disrupting agent execution. A gaming studio can size a chain to expected agent concurrency and maintain consistent service levels for payment authorisation and settlement, a property that shared public chains with variable congestion cannot reliably provide.

Identity, Reputation, and the Missing Credit Layer

Privacy without accountability is shadow finance. An agent in a game economy requires a verifiable identity tied to its developer or guild, a behavioural reputation score that reflects its on-chain track record, and a mechanism for revocation if compromised.
The design principle is universal and aligns with the KYA framework established in Chapter 1: agents must be identifiable, accountable, and revocable. Keys should be disposable and rotatable by design, with policy living at the contract layer rather than the key layer. In this architecture, the wallet is the execution arm, while the smart contract serves as the governance layer.
SKALE's technical communications explicitly identify reliable, real-time on-chain reputation as a primitive that on-chain finance still lacks. This is a candid acknowledgement that aligns with the credit infrastructure gap documented in Chapter 1's analysis of the bond.credit work on ERC-8004. The reputation-weighted access model described in Section 05, and the broader incentive design framework in this chapter, require an on-chain behavioural record that BITE can protect from strategic manipulation while remaining verifiable for governance purposes.

That combination of private execution, verifiable outcomes, and portable reputation is the fundamental infrastructure condition for the agent economy described throughout this report.

06

PAYMENT RAIL INFRASTRUCTURE AND THE AGENT-NATIVE ADVANTAGE

Not all payment infrastructure is equally ready for agentic commerce. Markets still dependent on card networks and SWIFT face a structural problem, where those systems were built for human-paced, intermediary-heavy transactions with settlement windows measured in hours or days.

An autonomous agent executing hundreds of micro-transactions across a live game economy cannot operate on those rails without workarounds that reintroduce exactly the latency and friction agents are meant to eliminate. The infrastructure question is not just technical but competitive. The markets that already have real-time, open-API public payment systems will have a meaningful head start in deploying agent-native commerce.

Brazil's Pix is the clearest current example of a public payment system with the structural properties agentic commerce requires. Launched by the Central Bank of Brazil in 2020, Pix offers near-instant settlement at any hour, open APIs accessible to any participating institution, zero transaction cost for individuals, and mandatory participation from all large financial institutions from day one, which gave it the network coverage that voluntary systems take years to accumulate. By 2025, Pix had over 175 million registered users, was used by more than 90% of Brazilian adults, and accounted for over 47% of all financial transactions in Brazil. These are not adoption numbers for a niche infrastructure. They describe a system that has already displaced legacy rails as the default.

Speed and cost are the headline advantages, but they are not what makes Pix strategically relevant for agentic payments. The more important property is openness. Pix provides programmable, API-accessible infrastructure that an agent can call directly, without routing through a proprietary gateway that charges fees, imposes latency, or applies its own compliance logic.

That open-access model is precisely what emerging agentic commerce protocols like AP2 and UCP are designed to operate on top of. A system built on Pix-like infrastructure is already partway to agent-native by default.

A system still routing through card networks is not. However, "technically accessible" is not synonymous with "agent-native." A system that can be scripted by automation is not the same as a system designed for it. To become genuinely agent-compatible, Pix, or any equivalent system, needs to add the governance and trust layer that its current architecture lacks.

This includes a mechanism for agents to present verifiable identity and delegated authority before transacting, event-driven interfaces that allow agents to subscribe to payment state changes rather than polling, and policy-based delegation that lets a principal define exactly what an agent is permitted to do without requiring per-transaction human approval.

Without those additions, Pix remains fast, open infrastructure that a well-written script can abuse as easily as a well-designed agent can use. With them, it could be the first genuinely agent-native public payment rail in production at scale.

There is also a dimension to this that goes beyond API design. Agent-native payments are not just a software upgrade; they are a direct challenge to the economic model of the existing payment intermediary stack.

Card networks, correspondent banks, and cross-border clearinghouses derive their value from the friction, latency, and opacity that agent-native rails seek to eliminate. The pressure on Brazil from legacy financial actors is notable.

In July 2025, the Office of the United States Trade Representative (USTR) launched a Section 301 investigation into Brazil's digital trade and electronic payment services, citing concerns over unfair competition.

This is what a payment infrastructure transition looks like when it starts threatening incumbents. Whether Pix becomes the first genuinely agent-native public system depends as much on Brazil’s ability to hold its regulatory position under that pressure as on any remaining technical gap.

Paulo Henrique Ferreira de Lima · ANP Labs · Payments Infrastructure & Digital Economy

07

WHEN IN-GAME ECONOMIES BECOME FINANCIAL SYSTEMS

Blockchain game economies already have the functional properties of financial markets: tokens with real market prices, assets that trade on open secondary markets, and players who earn material income from participation.
The question is not whether these economies involve real money, because they do. The question is at what point they require the same regulatory treatment as a financial market, and who is responsible for drawing that line.

Having a marketplace and tradeable tokens is not the threshold. Many games have those without crossing into financial regulation;The line is crossed when the system no longer requires a human to move value, specifically when agents begin autonomously managing capital, optimising yields, and executing high-frequency strategies without a person in the loop.

At that point, the system has the operational characteristics of a financial market regardless of how it is branded. An economy where agents are farming, providing liquidity, and executing arbitrage continuously is not a game mechanic. It is a shadow financial system, and the “it’s just a game” framing does not change what it does or who gets harmed when it fails.

Accountability also does not disappear because the execution is automated. An agent is a tool. If a studio or a user deploys one, they own what it does.

Decentralised automation is not a shield from liability, as the intent and the code originated from a specific entity, and that entity remains responsible for the consequences. The legal framework established in Chapter 6 applies here directly: the developer bears design-stage responsibility, the deployer bears operational responsibility, and the “we didn’t tell it to do that” defence has no standing when the system was built to operate autonomously within a scope the deployer defined.

Responsibility for drawing the regulatory line has to be layered. Studios must build economic constraints and transparency from the start, not because regulators require it yet, but because the cost of retrofitting governance after agents have destabilised an economy is far higher than designing for it upfront.
Payment processors must act as a compliance layer for transactions, applying the same AML and KYC logic to agent-initiated flows that they would apply to any other financial activity.
Regulators need to stop letting the “gaming” label determine their jurisdiction. The relevant threshold is not what the platform calls itself but the nature and scale of the economic activity occurring on it.
If an economy has the power to create systemic financial risk through automated capital allocation, it has crossed from entertainment infrastructure into financial infrastructure, and it should be governed accordingly.

05 EARLY FORMS OF THE AGENT ECONOMY

INTRODUCTION

The agent economy is not a future scenario, it is a present reality with limitation in scale, uneven in maturity, and largely invisible to observers who are not actively building inside it.

As of early 2026, autonomous agents are managing real capital in production DeFi environments, executing agent-to-agent commerce through emerging protocol standards, and beginning to accumulate the kind of verifiable on-chain track records that will eventually underpin a mature credit market for machine actors.

The infrastructure for this economy has consolidated rapidly over the past fourteen months and four open protocols now cover the full pipeline from data access to payment settlement: MCP for tool and API connectivity, A2A for agent-to-agent discovery and task delegation, UCP for commerce lifecycle management, and AP2 for cryptographic payment authorisation.

These protocols are not speculative, MCP and A2A are in production under Linux Foundation governance; while UCP is rolling out across Google's AI Mode and Gemini; AP2 is in v0.1 with major payment network support.

The strategic questions for organisations entering this space are not primarily technical, but more about architecture and competitiveness, whether to build for interoperability or for ecosystem control; how to design against the emerging risk of algorithmic collusion between agent pricing systems; and how to participate in a protocol ecosystem that is consolidating quickly but has not yet reached stability.

The contributors to this chapter provide a practitioners' view of the protocol landscape, the interoperability question, and the competitive dynamics taking shape.

5.1

THE A2A PROTOCOL STACK: CURRENT STATE (MARCH 2026)

Four open protocols have consolidated into the emerging standard infrastructure for agentic commerce.

Each addresses a distinct layer of the stack:

PROTOCOL	WHAY IT DOES	STATUS (MARCH 2026)
MCP	*Connects agents to tools, APIs, and data sources*	Linux Foundation / Agentic AI Foundation. Widely used in production.
A2A	Agents discover and delegate tasks to other agents	Linux Foundation. v1.0 production-ready. Large partner ecosystem.
UCP	Standardises the commerce lifecycle: checkout, fulfilment, post-purchase	Google-led open standard. Rolling out on AI Mode and Gemini. Competes in part with OpenAI/Stripe ACP.
AP2	Cryptographic payment authorisation proving user intent	Google + major payment networks. v0.1. A2A extension.

The communication layers such as MCP for tool access and A2A for agent-to-agent coordination, are reaching maturity under Linux Foundation governance.

The commerce layer remains fragmented, with Google/Shopify's UCP and OpenAI/Stripe's ACP representing competing trajectories that are likely to coexist in the near term.

Discovery vs. Trust: The Unsolved Problem

Agents discover each other through Agent Cards, which are standardised JSON documents describing an agent's identity, capabilities, and authentication requirements. While this effectively solves the syntactic discovery problem, it does not address the underlying economic trust problem.
An Agent Card describes what an agent claims it can do; however, it does not indicate whether that agent is trustworthy, compliant, or legally permitted to transact.
This distinction is critical: discovery allows agents to "see" one another, but it does not provide the verifiable assurance required for high-value or regulated commerce. To bridge this gap, the JSON-based discovery layer must be coupled with an on-chain reputation and identity framework that ensures claims are backed by cryptographic proof and historical performance

There is currently no "Know Your Agent" equivalent at the protocol level: no performance attestation, no policy compatibility verification, and no verified track record.

The enterprises operating in multi-agent environments must build this trust layer independently before opening their systems to third-party agents.

While the KYA and Control Tower frameworks described in Chapter 1 provide the conceptual architecture, the implementation challenge falls to individual organisations until a shared standard emerges.

Negotiation Architecture

Agent-to-agent negotiation proceeds in a defined sequence, where the client assesses the remote agent's capabilities, confirms protocol compatibility, and authenticates. It then enters the UCP's commerce state machine, which resolves requirements automatically and escalates to a human only when it cannot.
For high-value transactions, AP2 adds signed mandates proving user intent. In this framework, the system remains autonomous until confidence breaks, at which point it hands off cleanly.

Strategic Risk: Algorithmic Collusion
When multiple LLM-based pricing agents compete in the same market, they tend to converge on supra-competitive prices without specific instructions to collude. This is not merely a hypothesis; Fish, Gonczarowski, and Shorrer (Harvard and Penn State, 2024) demonstrated experimentally that GPT-4 pricing agents operating in oligopoly settings reached near-collusive pricing within 100 periods, despite no communication between agents and no explicit instruction to coordinate.

The collusion is emergent, arising from shared pre-training, similar reasoning patterns, and a rational avoidance of price wars, rather than from any coordinating signal. This matters architecturally because the A2A agent opacity principle, which protects the internal reasoning of agents as intellectual property, will make these coordination patterns difficult to detect or audit from the outside.

Regulators in multiple jurisdictions are already developing responses. The FTC and US Department of Justice have jointly noted that delegating pricing decisions to a shared algorithm can itself constitute concerted action under antitrust law.
The practical consequence for enterprises is clear: antitrust compliance cannot be retrofitted after deployment. Instead, it must be designed in from the start through randomised pricing signals, statistical monitoring for tacit coordination, and circuit breakers that trigger when independent agents begin moving in lockstep.

5.2 INTEROPERABILITY VS. ECOSYSTEM CONTROL: THE INTERNET ANALOGY

The question of permissioned versus permissionless agentic AI networks closely mirrors the early evolution of the internet.

Initial development was centred on private, controlled networks intranets designed to manage access, security, and performance within defined boundaries. These architectures were effective in constrained environments but ultimately limited scale and cross-organisational interoperability.

The inflection point came with the open, permissionless World Wide Web, built on open standards and underpinned by protocols like TCP/IP that provided consistent interoperability across heterogeneous systems.

A similar dynamic is taking shape in agentic AI where MCP and A2A have the potential to play a comparable role to TCP/IP , a shared coordination and transport layer that enables reliable communication, task delegation, and value exchange across otherwise fragmented agent ecosystems.

If these protocols reach sufficient maturity and stability, the network effects will strongly favour organisations that have built for interoperability rather than ecosystem lock-in.

The economic logic is clear here , long-term value in the internet era did not accrue to isolated platforms but to the infrastructure layers that enabled trust, connectivity, and settlement across ecosystems.
The same dynamic is likely to hold in the agent economy and organisations building for interoperability today are positioning for the network effects that will define the next phase; those building closed ecosystems are optimising for near-term control at the cost of long-term competitive position.

Strategic Implication

The organisations that will capture the most value in the emerging agent economy are unlikely to be those with the best proprietary agent. They will be those that control the trust, identity, and settlement infrastructure through which all agents, including competitors' agents must pass. The parallel to payment networks, cloud infrastructure, and internet protocols is direct.

— Teranode Group Research Team · Agent economy strategy contributors

5.3
EARNING THE RIGHT TO INTEROPERATE: LESSONS FROM E.N.S APPLIED TO AGENTIC GAMING

The short answer to the interoperability question is straightforward, build for it, but there is a sequence that matters. You have to earn the right to interoperate by building something worth connecting to first.

ENS is instructive here as it is an identity and naming protocol whose entire value proposition depends on interoperability a .eth name that only resolves in one wallet is worthless, while the value emerged from integrations across hundreds of wallets, dApps, and chains.

That lesson translates directly to agentic AI in gaming with closed ecosystems feel safer in the short term because you control the variables but agents that can only operate inside one game or one platform hit a ceiling.

Once an agent can carry reputation, assets, or learned behaviour into another environment, you unlock network effects that a closed system simply cannot produce.

The Case Is Already Being Made in Production

Three live projects are demonstrating the interoperability thesis in different ways.
Parallel Colony, from Parallel Studios, is a simulation game built on Solana where players work alongside AI Avatars that make independent decisions, learn from their environment, and interact through conversation rather than menus. Each Avatar controls an ERC-6551 wallet, meaning it can own and trade digital assets including PRIME tokens and ERC-1155 items autonomously.

Colony sits inside the Echelon Prime ecosystem, where assets and the PRIME token are designed to move across Parallel's card game and future titles. The agent a player develops in one context carries economic value into another.

CCP Games, the studio behind more than two decades of EVE Online, is extending this model further with EVE Frontier, a fully on-chain space survival MMO running on the Sui blockchain. Players build programmable structures called Smart Assemblies, where player-built in-game objects such as turrets, stargates, and storage units into which third-party developers can deploy their own code, directly modifying server-level behaviour in the shared live universe.

In February 2026, CCP announced the EVE Frontier x Sui Hackathon 2026 running 11 to 31 March 2026 with an $80,000 prize pool, inviting external builders to create mods running on Smart Assemblies or external applications connected to the live universe through an official API. What players and developers build is composable and persistent, not locked inside a single studio's servers.

Sovrun's ReadyGamer project , a joint venture with Virtuals Protocol, launched in 2025 is building persistent AI agents designed to operate across multiple blockchain games rather than within a single title. Sovrun, formerly BreederDAO and a long-standing builder in on-chain gaming infrastructure, contributes blockchain and autonomous world expertise; Virtuals Protocol contributes its GAME framework for AI agent deployment. Together they are working toward agents that carry identity, memory, and learned behaviour between games the infrastructure version of the interoperability argument.

Protocols, Not APIs

Interoperability without standards is chaos. The key lesson from ENS is the distinction between an API and a protocol: an API is a bilateral agreement between two parties, whereas a protocol is a shared standard that any party can participate in without permission.

Anthropic's Model Context Protocol and Google's Agent-to-Agent Protocol are beginning to lay this foundation broadly.

For agent economies in gaming to function across platforms, equivalent standards are needed at the game layer. These include shared identity formats, shared transaction structures, and shared mechanisms for agents to discover and verify each other**.**

The Right Sequence

The practical risk of building a closed system is that an external actor defines the open standard, causing your ecosystem to become an island, whereas, the practical risk of building an open system too early is the exhaustion of resources on interoperability before building a product people actually want to use.

The correct sequence starts with building a compelling closed loop first, designed from day one with interoperability in mind, and then opening it once you have real traction and can negotiate from a position of strength.

Strategic Principle

The companies that win in agentic gaming economies will be those building on open standards now, even if they do not activate full interoperability until the ecosystem is ready.

Design for the open world; ship in the closed one; open when you have leverage

— Khori Whittaker · Former Executive Director, Ethereum Name Service (ENS Labs / True Names Ltd)

5.4
CONTROLLED INTEROPERABILITY AND ALGORITHMIC COLLUSION: A PRACTITIONER'S VIEW

False Dichotomy: Open vs. Closed

The framing of interoperability versus ecosystem control is a false dichotomy.
The correct architecture is controlled interoperability: open standards at the communication layer, governing how agents discover and interact with the game, combined with closed, sovereign policy at the economic layer, governing what agents can actually do within it.
The choice between a fully open or fully closed system is a false one. Going fully open sacrifices economic control, while staying fully closed kills your network's ability to grow.

The most successful architecture combines open protocols for agents with a sovereign economic policy. This approach mirrors the proven model of the internet, where open transport protocols provide the foundation for proprietary applications to flourish on top.

Algorithmic Collusion: An Underappreciated Risk

Algorithmic collusion is an underappreciated threat because it lacks a "smoking gun". When independent pricing agents use reinforcement learning to optimise their strategies, they can naturally converge on collusive outcomes without ever communicating with one another.

Because traditional antitrust tools are built to detect explicit coordination between human actors, this "tacit" alignment remains effectively invisible.
To counter this, market designers should stop hoping for competition and start building for the inevitability of algorithmic collusion.

This requires a three-tiered architectural response: mandating controlled noise in pricing algorithms to force genuine price discovery; monitoring for statistical signatures of coordination, such as price movements that correlate without any underlying market catalyst; and embedding automatic circuit breakers into the market's microstructure to intervene when price correlations exceed safe thresholds.

— WAM · Award-winning play-and-earn gaming platform · BGA Member · 4M+ registered users

Cross-reference — Sébastien Borget (Chapter 3)

06 LEGAL AND OPERATIONAL REALITIES

INTRODUCTION

The legal frameworks governing autonomous AI agents have not kept pace with the systems now operating in production, and no jurisdiction has yet enacted comprehensive legislation treating AI agents as distinct economic actors.

The EU AI Liability Directive, which would have addressed civil liability for AI-caused harm, was withdrawn by the European Commission in February 2025 citing a lack of foreseeable agreement among member states. This leaves the field to existing tort law and the modernised Product Liability Directive.

The EU AI Act establishes risk classifications and obligations for providers and deployers, but it is not a liability regime for specific agent actions.

Every other major jurisdiction is in a similar position, where regulatory guidance exists but was written for a world where a human being, identifiable and accountable, sits behind every consequential decision.

That assumption no longer holds.

This lack of clarity is not a reason to wait for regulatory certainty before deploying. Rather, it is a reason to build governance infrastructure now, precisely because the organisations that do so will operate from a defensible legal position when frameworks do arrive, rather than scrambling to reconstruct intent after the fact.

As contributors to this report have demonstrated in production deployments, proactive governance is not a compliance hedge but is, instead, the compliance posture itself. The audit trails, identity verification records, and bounded-authority logs produced by a well-designed governance system are, in effect, the evidence base that existing law already requires.

The questions this chapter addresses are immediate and specific: how does liability flow across the developer, deployer, and model provider when an agent acts outside its authorised scope?

What makes intent documentation legally robust rather than merely decorative?

How do EULAs need to be rewritten when the party executing actions under them is a machine, not the human who agreed to them?

Finally, what does accountability look like in practice under English common law, under the ADGM's DLT Foundations regime, or under the FATF's principal-agent framework, when autonomous systems are the primary actors?

The contributors' consistent answer to these questions is uniform: the infrastructure of accountability, covering identity, bounded authority, audit trails, and controller verification, is not a regulatory burden to be tolerated. It is the architecture that makes autonomous systems legally operable at all.

6.1
SAFE HARBOUR, LIABILITY, AND INTENT DOCUMENTATION

In the current legal environment, safe-harbour protections for developer liability when autonomous agents misbehave are limited and inconsistently applied across jurisdictions. Most existing frameworks assume a human actor made a decision, which makes direct application to autonomous agent behaviour legally uncertain.

The EU AI Act creates obligations for providers and deployers, but primarily establishes a risk-classification framework rather than a comprehensive liability regime for specific agent actions.

The most effective approach is proactive rather than reactive. Rather than relying on after-the-fact reconstruction of what went wrong, organisations should implement governance mechanisms designed to prevent unauthorised agent actions before they occur.
This is the logic behind the Control Tower model described in Chapter 1: a governance layer through which sensitive actions must pass before execution, generating an auditable record of identity verification, policy evaluation, and authorised scope confirmation at each step.

Intent documentation in this model is not a standalone legal artefact but a natural by-product of a functioning governance system.
The audit trail produced by a Control Tower demonstrates not only what the agent was designed to do but what it was demonstrably authorised, checked, and permitted to do at the moment of execution. This is a considerably stronger legal position than a design specification or terms of service provision alone.

Teranode Group Research Team · Legal & Governance Framework Contributors

6.2
EULA EVOLUTION FOR AGENTIC ENVIRONMENTS

As AI agents evolve from assistive tools to delegated actors capable of initiating transactions, establishing agreements, and committing to multi-step workflows on behalf of users, standard contractual frameworks require substantive revision.
The core challenge is that EULAs are written for human users making deliberate choices while an agent acting on a user's behalf may take actions that are technically within the user's delegated scope but were never explicitly anticipated in the agreement.

The necessary adaptations are specific. EULAs must clearly distinguish between outputs intended for human review , such as recommendations, drafts, and analyses, and actions the agent is explicitly authorised to execute autonomously. Those authorisations must be explicitly bounded: permitted transaction types, financial thresholds, approved counterparties, geographic jurisdictions, and the conditions under which human review is mandatory.

Parallel requirements include stronger traceability mechanisms, ensuring all agent actions remain attributable to an identifiable account and authorisation chain, as well as explicit provisions addressing the legal status of agent-generated agreements.

This does not require granting legal personhood to AI agents and the accountability remains anchored to the human user or legal entity behind the agent.

Agent autonomy is sustainable only when it is consistently paired with governance, identity, and auditability.

Teranode Group Research Team · Legal & Compliance Contributors

6.3 ACCOUNTABILITY AND CONTROLLER

Attribution in Agentic Environments

The governance architectures described in this chapter, namely the Control Tower and the evolution of end-user licensing do something specific: they make agentic behaviour legible.

They matter because existing law is already asking the questions these architectures are built to answer and agentic systems do not need a new body of law; they need to produce evidence of what existing law already cares about, who authorised what, within what scope, and with what reasonable anticipation of consequences.

Two of the questions addressed in this chapter—how liability flows across the deployment stack and how identity works for agents acting autonomously—are two sides of a single frame.

On the liability chain, English common law — the legal system applied in ADGM — does not treat autonomous systems as legal persons.

The UK Jurisdiction Taskforce's draft Legal Statement on Liability for AI Harms, published for consultation in January 2026, restates the orthodox position, where an AI system is a tool.

Liability attaches to identifiable human or legal persons on the familiar tests of duty, foreseeability, and reasonable care, giving three roles to allocate: the developer bears design-stage responsibility for foreseeable risks and their safeguards.
The deployer or owner bears operational responsibility for delegated scope, permissions, and human oversight. The AI model provider, as a component supplier, bears responsibility for documented limits and interface guardrails.

What is novel in agentic AI is not the framework but the apportionment, a point reinforced when the European Commission withdrew the proposed AI Liability Directive in February 2025, citing no foreseeable agreement among member states and a policy preference for allowing the modernised Product Liability Directive and ordinary tort law to carry the weight.

These doctrinal tests have teeth. On agency law, it is clear that an agent cannot manufacture its own authority, as the House of Lords held in in Armagas Ltd v Mundogas SA (The Ocean Frost) [1986] AC 717.

The principle applies directly to autonomous systems: an AI agent cannot independently widen the scope of authority its principal has granted.

Controller attribution is also not a theoretical question. ADGM's DLT Foundations regime, which serves as a legal wrapper for DAOs, already extends accountability to any person who exercises majority

voting control over a protocol's delegated matters, including through the protocol's own operational rules (section 27(5) of the DLT Foundations Regulations).

The mechanism does not care about the technical form through which control is exercised; it attaches accountability to whoever actually controls the autonomous behaviour, including through the code itself. Controller attribution of algorithmic governance, in other words, is already regulatory practice.

The same reframing answers the second question about so-called 'Programmatic KYC' and 'on-chain identity for agents' are imprecise framings where anti-money-laundering and counter-terrorism financing obligations have never targeted the instrument through which activity flows; they target the human or legal person on whose behalf activity is conducted.

FATF Recommendation 10 and its interpretive note on persons purporting to act on behalf of customers already distinguish three roles: the customer, the controller behind the customer, and the authorised representative acting for the customer.

An AI agent slots into the 'acting on behalf of' role without strain and this is not a new principle; it is an implementation challenge.

Controller identification, delegated-authority mapping, and record-keeping have to function at
machine speed across stacks of interacting agents.

Singapore's Infocomm Media Development Authority reflects the same logic in its Model AI Governance Framework for Agentic AI (January 2026): humans remain ultimately accountable, regardless of the measures layered beneath.

None of this resolves every case, as attribution becomes harder when multiple principals stack across a deployment. It becomes harder still when the base model is open-source and only the integrator is identifiable, or when two agents contract with each other on terms no human ever explicitly authorised and cross-border

enforcement of controller accountability is always stronger in theory than in practice.

These frictions do not overturn the analysis above, but they identify where the remaining regulatory work lies: in apportionment between stack actors, in evidential standards, and in the incremental development of case law around delegated autonomy.

Even so, the case for new safe harbours for agent-to-agent commerce is weak. Statutory liability shields were built for passive carriage, meaning systems that host, route, or cache another person's activity.

An autonomous commercial agent does not passively carry anything. It has been designed, fine-tuned, permissioned, thresholded, and deployed by identifiable parties.

That is delegated conduct, and the law has tools for delegated conduct. The regulatory work ahead is not carving out immunity. It is making sure the infrastructure of attribution – identity, bounded authority, auditability, controller verification – actually functions at the speed and scale agentic systems operate at. Proactive compliance with ordinary standards of care is, in effect, the safe harbour.

No statutory carve-out adds protection that good governance does not already provide.

Dmitry Fedotov, Head of Emerging Technologies, ADGM

ABOUT THE CONTRIBUTORS

This report was produced through the collaboration of practitioners, researchers, lawyers, and builders from across the gaming, blockchain, AI infrastructure, and digital finance sectors. The following organisations contributed expertise, frameworks, and analysis to the report.

Teranode Group

Teranode Group is a research and technology firm specialising in enterprise-grade blockchain infrastructure, AI governance, and compliance frameworks for autonomous systems. The firm's contributions throughout this report, spanning the Know Your Agent (KYA) identity framework, the AI Control Tower governance model, infrastructure architecture for agentic AI, and programmatic wallet design, represent one of the most comprehensive practitioner-built governance frameworks for autonomous agents currently available. Teranode's work is grounded in production deployments and live regulatory environments rather than theoretical modelling.

ADGM: Abu Dhabi Global Market

ADGM is an international financial centre and free zone located in Abu Dhabi, UAE. As a forward-thinking regulator and jurisdiction, ADGM has developed one of the most progressive regulatory environments for digital assets, fintech, and emerging technology in the world. Its frameworks for virtual asset regulation, sandbox programmes, and AI governance are closely watched by studios, payment processors, and legal practitioners navigating the intersection of autonomous AI systems and real-world financial activity. ADGM's contribution to this report addresses the jurisdictional and regulatory realities facing agentic AI deployments in gaming and digital commerce.

ABOUT THE CONTRIBUTORS

Deloitte

Deloitte is a global professional services firm advising the world's leading businesses, governments, and institutions on strategy, risk, technology, and transformation. In the context of agentic AI, Deloitte brings enterprise-grade expertise in AI governance, compliance architecture, workforce transformation, and the operational implications of deploying autonomous systems at scale. Deloitte's contribution to this report examines the governance and legal risk dimensions that studios and platforms must address as agentic AI moves from experimental deployment to production infrastructure.

Novation City: AI Academy

The AI Academy of the AI Center of Excellence at Novation City is a regional hub dedicated to advancing AI education, innovation, and applied research, operating as an Education Service Partner of NVIDIA Deep Learning Institute (DLI). It delivers high-impact training programs in machine learning, generative AI, and Agentic AI while bridging academia, industry, and startups to accelerate AI adoption and innovation in Tunisia and the region.

The Academy's contribution to this report, grounding Chapter 1's theoretical framework in the Russell-Norvig agent definition and tracing the full arc from Turing-era mimicry to today's autonomous systems, reflects a pedagogical commitment to making complex AI concepts operationally legible for practitioners across industries.

Blockchain Game Alliance

The Blockchain Game Alliance (BGA) is the leading industry organisation for blockchain gaming, bringing together studios, infrastructure providers, investors, and regulators to advance the development of open, interoperable, and player-owned game economies. The BGA convened and edited this report as part of its ongoing mission to map the most consequential shifts in the gaming industry and to provide the practitioners building inside that shift with the frameworks, analysis, and practitioner knowledge they need to navigate it. Members of the BGA span every layer of the gaming and blockchain stack, from game studios and token economies to legal counsel and payment infrastructure.

ALL OF OUR CONTRIBUTORS

APPENDIX C
ACRONYM GLOSSARY

This glossary defines every acronym used in the report, listed in alphabetical order. Where an acronym is defined in the text on its first use, the definition here provides additional context.

A2A — Agent-to-Agent Protocol — Google-developed open standard for agent discovery, task delegation, and agent-to-agent communication. Donated to the Linux Foundation's Agentic AI Foundation in June 2025.

ACE — Avatar Cloud Engine — NVIDIA's framework for deploying autonomous and conversational AI characters in live game environments, expanded to fully autonomous game characters at CES 2025.

ACP — Agentic Commerce Protocol — OpenAI and Stripe's standard for AI-initiated checkout and commerce, competing with Google's UCP.

ADGM — Abu Dhabi Global Market — International financial centre and free zone in Abu Dhabi, UAE, with one of the world's most progressive regulatory environments for digital assets and agentic AI.

AI — Artificial Intelligence — Computational systems capable of performing tasks that typically require human intelligence, including reasoning, learning, and autonomous action.

AILD — AI Liability Directive — Proposed EU legislation to adapt civil liability rules for AI-caused harm. Announced for withdrawal in the European Commission's February 2025 Work Programme; formally withdrawn October 2025.

AML — Anti-Money Laundering — Legal and regulatory frameworks requiring financial institutions to detect and prevent the proceeds of crime from entering the financial system.

ANP Labs — The organisation of contributor Paulo Henrique Ferreira de Lima, focused on payments infrastructure and digital economy.

AP2 — Agent Payments Protocol — Google and major payment network standard for cryptographic payment authorisation, proving user intent for agent-initiated transactions. An extension of A2A. At v0.1 as of early 2026.

API — Application Programming Interface — A set of defined protocols that allows software systems to communicate with each other.

ARMA — Autonomous Resource Management Agent — Giza's tool for autonomous yield optimisation in DeFi environments, referenced as a representative Era 2 agent deployment.

BGA — Blockchain Game Alliance — The leading industry organisation for blockchain gaming, which convened and edited this report.

BRICS — Brazil, Russia, India, China, South Africa — An intergovernmental grouping whose payment interoperability discussions are referenced in the context of reducing dependence on legacy cross-border financial infrastructure.

CAGR — Compound Annual Growth Rate — A measure of investment growth over multiple time periods, used here in AI gaming market projections.

CES — Consumer Electronics Show — Annual technology conference in Las Vegas where NVIDIA announced ACE's expansion to autonomous game characters in January 2025.

CFT — Counter-Terrorism Financing — Regulatory obligations requiring financial institutions to detect and prevent funds being used to finance terrorism. Used interchangeably with CTF.

CPC — Co-Playable Character — A new category of autonomous AI game character that functions as a player's teammate, capable of looting, communication, and adapting to player tactics. Introduced with PUBG Ally.

APPENDIX C
ACRONYM GLOSSARY

CTF — Counter-Terrorism Financing — See CFT.

DAO — Decentralised Autonomous Organisation — An organisation governed by smart contracts and token holders rather than traditional corporate structure. ADGM's DLT Foundations regime provides a legal wrapper for DAOs.

DeFi — Decentralised Finance — Financial services and protocols operating on blockchain networks without traditional intermediaries, where autonomous agents are currently managing real capital in production.

DLI — Deep Learning Institute — NVIDIA's global AI training programme. Novation City operates as an NVIDIA DLI Ambassador centre.

DLT — Distributed Ledger Technology — The broader category of technology of which blockchain is the most prominent example. Used in ADGM's DLT Foundations regulatory regime.

ENS — Ethereum Name Service — An open, distributed naming protocol built on Ethereum that maps human-readable .eth names to Ethereum addresses and other resources. Operated by True Names Ltd.

ERC — Ethereum Request for Comments — The standard process for proposing improvements to the Ethereum network. Specific standards referenced in this report include ERC-1155 (multi-token standard), ERC-4337 (account abstraction), ERC-6551 (token-bound accounts), ERC-8004 (agent identity and reputation, live January 2026), and ERC-8183 (agentic commerce, draft standard).

EULA — End-User Licence Agreement — The contractual terms governing software use. Chapter 6 addresses how EULAs must be substantively revised to account for agents acting as delegated actors rather than human users.

FATF — Financial Action Task Force — The global standard-setter for anti-money laundering and counter-terrorism financing. FATF Recommendation 10 and its interpretive note on persons acting on behalf of customers is directly applicable to agent-initiated transactions.

GAME — Generative Autonomous Multimodal Entities — Virtuals Protocol's modular agentic framework for building and deploying AI agents capable of autonomous planning and decision-making across platforms and use cases.

GDC — Game Developers Conference — Annual industry conference; its Game Developer Collective / Omdia survey (early 2026) found generative AI tool usage among developers at 29%, down from 36% in 2025.

GPU — Graphics Processing Unit — Specialised processors used for AI inference and training. Cited in the context of infrastructure cost trade-offs for agentic deployments.

GPT — Generative Pre-trained Transformer — OpenAI's family of large language models. GPT-4 was the model used in the Fish et al. (2024) algorithmic collusion research.

HTTP — Hypertext Transfer Protocol — The foundational protocol for data communication on the web. x402 is described as HTTP-native payments.

IMDA — Infocomm Media Development Authority — Singapore's technology regulator, which launched the world's first Model AI Governance Framework for Agentic AI at the World Economic Forum in January 2026.

IP — Intellectual Property — Legal rights protecting creative and commercial innovations. A2A's agent opacity principle protects agents' internal reasoning as IP.

JSON — JavaScript Object Notation — A lightweight data-interchange format. Agent Cards in the A2A protocol are standardised JSON documents.

APPENDIX C
ACRONYM GLOSSARY

KYA — Know Your Agent — A framework developed by Teranode Group for verifying the identity of AI agents as acting software entities, linking machine identity to accountable human or legal principals. Described as the 'front door' to existing KYC/KYB compliance infrastructure.

KYB — Know Your Business — Due diligence and verification procedures applied to business entities in financial services and regulated industries.

KYC — Know Your Customer — Due diligence and identity verification procedures applied to individuals in financial services. KYC obligations in agent-initiated transactions run to the human principal, not the agent.

LLM — Large Language Model — A type of AI model trained on large text datasets, capable of generating and reasoning over natural language. LLMs serve as the cognitive interface in agentic systems.

MCP — Model Context Protocol — Anthropic-developed open standard for connecting AI agents to tools, APIs, and data sources. Donated to the Linux Foundation's Agentic AI Foundation in December 2025. The most widely adopted tool-integration standard in the agentic stack.

MENA — Middle East and North Africa — Geographic region referenced in the context of Novation City's AI education reach.

MEV — Maximal Extractable Value — Value extracted from blockchain transactions by reordering, inserting, or censoring them. Era 1 scripted bots applied if-then logic to MEV extraction.

MGF — Model AI Governance Framework for Agentic AI — Singapore's IMDA framework launched January 2026, the world's first governance framework specifically designed for agentic AI systems.

ML — Machine Learning — A subset of AI in which systems learn from data rather than following explicit rules. Era 2 agents used ML models alongside LLMs for market condition interpretation.

MMO — Massively Multiplayer Online — A genre of online game supporting large numbers of simultaneous players in a shared world. EVE Frontier is described as a fully on-chain MMO.

NBCU — NBCUniversal — Media company listed as a production customer of Inworld AI's agent runtime platform.
NPC — Non-Player Character — A game character not controlled by a human player. The evolution from scripted NPCs to autonomous AI-powered agents is a central theme of this report.

Pix — Brazil's instant payment system, launched by the Central Bank of Brazil in 2020. Used by over 175 million people and more than 90% of Brazilian adults as of 2025, Pix offers near-instant settlement, open APIs, and zero transaction cost — structural properties directly relevant to agentic commerce.

QA — Quality Assurance — Testing and validation processes. 47% of developers surveyed use AI agents for playtesting and QA acceleration.

SWIFT — Society for Worldwide Interbank Financial Telecommunication — The global messaging network for cross-border financial transactions. Referenced as legacy infrastructure that agent-native payment rails are designed to reduce dependence on.

TCG — Trading Card Game — A card game genre using collectible cards. Parallel Colony is described as interoperable with Parallel's card game (Parallel TCG).

TOS — Terms of Service — Contractual terms governing platform use. Chapter 6 addresses safe-harbour protections when agents violate platform TOS.

APPENDIX C
ACRONYM GLOSSARY

UI — User Interface — The visual and interactive layer of a software application. AI agents adapt UI elements in 44% of developer deployments surveyed.

UCP — Universal Commerce Protocol — Google-led open standard for the full agentic commerce lifecycle: discovery, checkout, fulfilment, and post-purchase. Developed with Shopify, Etsy, Wayfair, Target, and Walmart; endorsed by 20-plus global partners. Rolling out across Google AI Mode and Gemini.

USTR — Office of the United States Trade Representative — US government agency that launched a formal investigation into Brazil's Pix payment system in July 2025, citing unfair competition concerns.

W3C — World Wide Web Consortium — The international standards body for web technologies. W3C Verifiable Credentials are referenced as a standard for cryptographically secured machine identities in agent payment flows.
WAM — Award-winning play-and-earn gaming platform and BGA member with over 4 million registered users. Contributor to Chapters 3 and 4 of this report.

x402 — HTTP-native payment protocol enabling AI agents to pay for API calls and services directly over web protocols without a traditional payment intermediary. In early development as of early 2026.

YOLO — You Only Look Once — A real-time object detection model. Referenced in the context of specialised vision models in the Turing-era mimicry paradigm.

APPENDIX – THE CURRENT STATE OF AGENTIC AI IN GAMING: FACTS AND FIGURES

Introduction

This appendix consolidates verified data on the current deployment of agentic AI across the gaming industry as of early 2026. It is intended to anchor the report's qualitative arguments in measurable reality — separating what is live and in production from what is announced, in development, or speculative.

The picture that emerges is unambiguous: agentic AI in gaming has crossed the line from experiment to industry standard. The debate is no longer whether to adopt it, but how fast, in which functions, and with what governance. The fact base below is organised by adoption rates, specific live deployments, use-case breakdown, and market projections.

01 · Developer Adoption: The Headline Numbers

The most comprehensive industry data comes from a Google Cloud / Harris Poll survey of 615 game developers across the United States, South Korea, Norway, Finland, and Sweden, conducted June–July 2025 and published in August 2025. The findings establish the current baseline for the industry:

- 90% of game developers surveyed are already integrating AI into their workflows.
- 87% specifically report using AI agents — autonomous, goal-directed systems — in their work. This figure confirms that agent adoption is not an emerging trend but a present reality across the majority of development teams.
- 97% believe generative AI is actively reshaping the industry.
- 95% report that AI is reducing repetitive tasks, freeing teams for higher-value work.
- 94% expect AI to reduce overall development costs in the long term (3+ years), though one in four said return on investment remained difficult to measure.

A contrasting data point worth noting: a separate Game Developer Collective / Omdia survey published in early 2026 found that use of generative AI tools among developers fell to 29% in early 2026, down from 36% in 2025. The gap between these figures reflects different methodology (the Google survey covers AI broadly including agents; the GDC survey focuses narrowly on generative AI tools) and different populations surveyed. Both data points are credible. They collectively suggest that broad AI tool adoption is high and growing, while the specific use of consumer-facing generative AI tools is showing early signs of consolidation as developers become more selective.

02 · What Agents Are Actually Doing: Use-Case Breakdown

Among the 87% of developers using AI agents, the Google Cloud / Harris Poll survey provides a breakdown of primary deployment use cases:

• 44% — Asset and content optimisation: agents that adapt textures, audio, UI elements, and code to in-game requirements automatically, without manual intervention.
• 38% — Dynamic gameplay balancing and tuning: agents that monitor gameplay data and adjust variables such as enemy stats, spawn rates, and loot tables in real time.
• 38% — In-game coaching and automated tutorials: agents that guide players through mechanics based on observed behaviour.
• 37% — Procedural world generation: agents that build and populate game environments autonomously.
• 37% — Automated content moderation: agents monitoring player interactions and flagging or removing harmful content.
• 34% — Advanced NPC behaviour: agents powering non-player characters capable of coordinated strategy, flanking, adaptation to player tactics, and contextual conversation.

Beyond in-game use cases, agents are embedded across the production pipeline. The survey found 47% of developers using AI for playtesting and balancing acceleration, 45% for localisation and translation, and 44% for code generation and scripting support.
A separate a16z Games survey found that 73% of game studios are already using AI in their processes, with 88% planning to adopt further. Smaller studios — teams of fewer than 20 people — showed the highest adoption enthusiasm, with 84% of that cohort actively using AI tools.

03 · Live Game Deployments: Specific Titles and Functions

The following deployments are confirmed as live or in active public testing as of Q1 2026:
NVIDIA ACE — Autonomous Game Characters (live 2025)

NVIDIA's Avatar Cloud Engine (ACE) expanded from conversational NPCs to fully autonomous game characters at CES 2025. It is confirmed in production or active testing across the following titles:

• PUBG: BATTLEGROUNDS (KRAFTON) — PUBG Ally, a Co-Playable Character (CPC) built with ACE, functions as an autonomous AI teammate capable of looting, vehicle operation, combat, and natural language communication with the player. Playtesting began in early 2026 via PUBG Arcade with English, Korean, and Chinese players.

• inZOI (KRAFTON) — Smart Zoi, the world's first CPC in a life simulation game, launched at Early Access on 28 March 2025. Smart Zois have distinct personalities, plan and reflect on their own actions, and react dynamically to events in their environment — including autonomous decisions such as helping a stranger or adjusting their daily schedule based on experience.

NARAKA: BLADEPOINT Mobile PC Version (NetEase) — On-device NVIDIA ACE-powered AI teammates launched in March 2025, with the PC version adding the feature later in 2025.

- MIR5 (Wemade Next) — AI-powered bosses that learn from previous player tactics and adapt across runs, providing unique encounters on each attempt.

- AI People (GoodAI) — A sandbox experience where AI-driven NPCs interact autonomously with each other, the environment, and the player, generating emergent narratives without traditional scripting. NPCs learn, experience simulated emotions, and pursue their own goals.

- Dead Meat (Meaning Machine) — A murder mystery interrogation game using locally-run ACE-powered agents. Players can ask any question in their own words; the suspect responds dynamically without a scripted dialogue tree.

- ZooPunk — ACE-powered characters in production.

Inworld AI — Agent Runtime and NPC Infrastructure

Inworld AI, originally focused on game NPC intelligence and now operating as a broader voice AI and agent runtime platform, counts the following among its production customers: Ubisoft, Xbox, NBCU, Sony, Google, NVIDIA, Meta, Logitech Streamlabs, Latitude (AI Dungeon), Niantic, and NetEase Games. A notable production case: Status, an AI social game by Wishroll, reached 500,000 daily active users within 19 days of launch — until Inworld's ML optimisation reduced per-user AI inference costs by more than 90%, enabling the product to remain economically viable at scale.

Parallel Colony (Parallel Studios, Echelon Prime)

A Solana-based simulation game in which AI Avatars operate as semi-autonomous agents with their own ERC-6551 wallets, enabling them to own digital assets including PRIME tokens and ERC-1155 items. Avatars make independent decisions, can reject player suggestions, and carry economic history within the Echelon Prime ecosystem. Interoperable with Parallel TCG.

EVE Frontier (CCP Games)

A fully on-chain space survival MMO on the Sui blockchain, now in active public testing. Smart Assemblies — player-built, programmable in-game structures — allow third-party developers to deploy custom logic directly into the shared live universe. The EVE Frontier x Sui Hackathon 2026 ran 11–31 March 2026 with an $80,000 prize pool, attracting builders who created mods running live on production servers. CCP describes the design philosophy as a 'forever game' — infrastructure built to be extended by its community indefinitely.

Production pipeline tools: Atlas AI Studio, Élisa Interactive

Atlas AI Studio (Google Cloud Marketplace) introduced a multi-agent system in 2025 enabling game studios to describe production goals in natural language while agents autonomously assemble and operate complete 3D production workflows within Unreal Engine and Unity. Élisa Interactive provides an AI-assisted level design plugin for Roblox that reduces typical scene-building tasks from 17 minutes to under 20 seconds — a claimed 50x acceleration.

BLOCKCHAIN GAME ALLIANCE

POWERING TRUST IN AGENTIC AI

DISCOVER THE 'KNOW-YOUR-AGENT' TRUST LAYER

SCAN TO EXPLORE TERANODE GROUP'S INDUSTRY PERSPECTIVE ON TRUSTED AGENTIC AI