DriveLock

WHITEPAPER

THE HUMAN-AGENT ENTERPRISE

5 Questions CEOs Need to Ask in the Age of the Human-Agent Enterprise

by Udo Riedel and Dr. Philipp Müller

INTRODUCTION

Cloud computing changed the operating model of organizations. Artificial intelligence is about to change the operating model of work itself. For the last decade, most enterprise software behaved like a tool. Humans remained firmly in control. Employees used applications. Managers approved workflows. Developers deployed systems. Organizations designed governance structures around the assumption that people were the operational actors.

That assumption is beginning to change. AI systems are increasingly becoming operational participants inside organizations. They write code, invoke tools, retrieve information, interact with applications, coordinate workflows, and increasingly act on behalf of users. In many companies, this shift is already happening faster than leadership teams realize.

INTRODUCTION

This creates enormous opportunities. Organizations may become dramatically faster and more adaptive. But it also introduces a fundamentally new governance challenge.
The central question is no longer only:

“Who has access to systems and data?”

Increasingly, the question becomes:

“What autonomous actions are allowed to happen inside the organization?”

This is where the idea of Shared Responsibility becomes critically important again.

In cloud computing, the Shared Responsibility Model clarified that cloud providers secure the infrastructure while customers remain responsible for identities, applications, configurations, and data. The AI era extends this logic further. Model providers, cloud providers, software vendors, and enterprises all share responsibility for secure AI operations. But enterprises remain responsible for governing how autonomous systems operate inside their own environments.

Based on our work with organizations across Europe, we believe CEOs and executive teams should begin asking five questions now.

CONTENT

1 Which operational decisions are we comfortable delegating to AI systems?

Most organizations still discuss AI primarily in terms of productivity. That framing is too narrow. The deeper shift is the delegation of operational authority. AI systems are increasingly making or influencing decisions that historically required human judgment. This ranges from software deployment and workflow orchestration to customer interaction and information retrieval.

Leadership teams therefore need to define where autonomous execution is acceptable, where human oversight remains mandatory, and where escalation boundaries exist. This is not a theoretical issue. Organizations are already connecting AI agents to developer environments, internal knowledge systems, APIs, collaboration platforms, and operational workflows. The question is no longer whether this will happen. It already is.

The real challenge is whether organizations consciously define the operational boundaries within which these systems may act.

2 Do we know what our AI systems are actually allowed to execute?

Many organizations focus heavily on access management for AI systems. Access matters, but it is not sufficient. An AI agent with legitimate access rights may still invoke dangerous commands, chain actions unexpectedly, expose sensitive information, or interact with systems in unintended ways. In practice, this means governance must increasingly move from access control toward execution control.

One of the most important lessons from endpoint security over the last two decades is that organizations cannot secure highly dynamic environments solely by trying to detect every possible malicious behavior after the fact. They also need clear operational guardrails that define what is allowed to run in the first place. This principle becomes highly relevant again in the age of agentic systems. Organizations need the ability to govern:

- which AI agents may operate
- which tools they may invoke
- which scripts and applications they may execute
- which process chains are approved
- which execution environments are trusted

Technologies such as application allowlisting and behavioral controls, historically used to harden endpoints and reduce attack surfaces, become strategically important again because they create trusted execution boundaries for autonomous systems. In practice, this means organizations can establish operational guardrails around AI-driven execution instead of relying purely on reactive monitoring.

3 Is our organization able to maintain visibility and control over how information moves?

AI systems are increasingly participating in collaboration itself. They summarize meetings, retrieve knowledge, generate documents, coordinate workflows, and interact across teams and organizational boundaries.

This creates enormous efficiency gains. It also creates new governance challenges around:

- confidentiality,
- classification,
- traceability,
- external sharing,
- regulatory compliance,
- sovereignty.

Historically, organizations often focused on securing storage systems or network perimeters. In agentic environments, however, the challenge increasingly becomes governing how information moves dynamically between humans, systems, agents, and external partners.

This requires more than static access control. Organizations increasingly need collaboration environments where information flows remain controlled, traceable, and policy-driven even as AI systems participate operationally.

One important lesson from secure collaboration environments is that governance becomes much more effective when security follows the information itself rather than relying solely on infrastructure boundaries. This includes granular access management, controlled workspaces, auditability, and clear governance over who or what system may access sensitive information under which conditions. In practice, trusted collaboration becomes part of operational governance itself.

4 Are our most sensitive AI workloads operating in trusted environments?

As organizations delegate increasingly sensitive tasks to AI systems, the execution environment itself becomes strategically important. Many AI discussions still focus primarily on models and applications. Increasingly, however, organizations must also ask:

- where sensitive AI workloads execute,
- who controls the infrastructure,
- which jurisdictions apply,
- how confidential data is protected during processing,
- whether operational environments themselves can be trusted.

This becomes particularly relevant in sectors such as public administration, defense, healthcare, finance, and critical infrastructure, where organizations must balance AI adoption with regulatory requirements, operational resilience, and sovereignty concerns.

One important development in this area is the rise of confidential computing and isolated execution environments designed to protect sensitive workloads even during active processing. Rather than relying solely on perimeter defenses, these approaches create protected execution domains that reduce operational trust dependencies and strengthen control over highly sensitive processing activities.

As AI systems gain operational autonomy, trusted execution environments increasingly become part of governance itself.

5 Do we have a governance model for the Human-Agent Enterprise?

The most important question is ultimately organizational, not technical. Most companies still treat AI primarily as a tooling discussion owned by innovation teams, IT departments, or isolated business units. But agentic systems increasingly affect governance, accountability, operational resilience, compliance, collaboration, and risk management across the enterprise.

This means organizations need a broader operational framework for governing autonomous systems.

Based on our work in cybersecurity, endpoint hardening, secure collaboration, and sovereign infrastructure, we believe three principles are becoming increasingly important:

1. **Organizations need governable execution environments with clear operational guardrails around what autonomous systems are allowed to execute.**
2. **Organizations need trusted collaboration architectures where information flows remain visible, auditable, and policy-driven even as humans and AI systems increasingly work together.**
3. **Organizations need trusted execution environments for sensitive workloads where confidentiality, resilience, and sovereignty remain protected during processing itself.**

Taken together, these principles point toward a broader shift in enterprise governance. The organizations that succeed in the coming years will not necessarily be those adopting AI the fastest. They will be those capable of integrating autonomous systems while maintaining operational control, accountability, resilience, and trust.

The central governance challenge of the AI era is not simply protecting systems from AI. It is ensuring organizations remain governable once software itself begins to act.

CONCLUSION

What this means in practice is that governance in the Human-Agent Enterprise cannot remain abstract. Organizations need operational mechanisms that allow them to define, enforce, and continuously adapt trusted boundaries for autonomous systems. This is where we at DriveLock focus our work.

Over the past years, we have learned that resilient organizations increasingly require three capabilities at the same time: trusted execution, trusted collaboration, and trusted processing. First, organizations need the ability to define what autonomous systems, applications, scripts, and processes are actually allowed to execute across endpoints and operational environments. This is why technologies such as application control, behavioral controls, endpoint hardening, and trusted execution boundaries become strategically important again. They create governable environments in which organizations can safely operationalize AI systems without losing control over execution itself.

Second, organizations need collaboration architectures where sensitive information remains visible, traceable, policy-driven, and protected even as humans and AI agents increasingly interact across organizational boundaries. Secure collaboration can no longer rely solely on perimeter assumptions. Governance increasingly needs to follow the information itself. With idgard's zero-knowledge collaboration environments and protected data-sharing capabilities, organizations can establish trusted digital workspaces where confidential information remains controlled even in highly distributed and AI-supported workflows.

Third, organizations increasingly need trusted environments for highly sensitive processing itself. As AI systems gain operational autonomy, questions of sovereignty, confidentiality, operator access, and trusted infrastructure move from technical considerations to governance questions. This is particularly relevant in sectors such as public administration, defense, healthcare, critical infrastructure, and regulated industries more broadly. Our work around confidential computing, hardened architectures, and sovereign processing environments is therefore ultimately about enabling organizations to operationalize AI adoption while remaining resilient, governable, and in control.

Annex: DriveLock's Governance and Security Controls for the Human-Agent Enterprise

APPLICATION CONTROL

DriveLock Application Control governs which applications, scripts, runtimes, and execution paths are allowed to operate inside organizational environments.
Based on trusted execution principles and allowlisting approaches, it establishes explicit operational guardrails around autonomous execution. In hybrid organizations, Application Control becomes a governance mechanism for machine-driven operational activity.

DRIVELOCK ADVANCED BEHAVIORAL CONTROL

DriveLock Advanced Behavioral Control governs runtime behavior inside operational environments.
It monitors process chains, execution patterns, privilege escalation attempts, suspicious automation behavior, and dangerous combinations of legitimate tools.
The objective is to ensure autonomous systems remain observable, controllable, and aligned with organizational intent during execution itself.

DRIVELOCK IDGARD

idgard is a secure collaboration and information governance platform designed to maintain control over information flows across internal and external operational ecosystems.
It combines encrypted communication, controlled collaboration spaces, traceability, secure sharing, auditability, and granular governance mechanisms to ensure organizations retain visibility and operational control even as humans and autonomous systems increasingly collaborate across organizational boundaries.

DRIVELOCK SEALED CLOUD

Sealed Cloud is a sovereign and confidential processing platform built around dedicated hardware, open-source infrastructure, and isolated execution environments.
It enables organizations to operate sensitive workloads, including AI and LLM environments, inside governable runtime architectures with reduced trust dependencies.
Sealed Cloud supports cloud, on-premise, and edge deployments while strengthening sovereignty, jurisdictional control, and confidential processing capabilities for highly sensitive operational workloads.

DriveLock

CONTACT US!

DriveLock SE
Landsberger Straße 396
81241 Munich
Phone +49 (0) 89 5463649-0
eMail info@drivelock.com

www.drivelock.com