ZLTECH

2026

The Data Foundation for Enterprise Agentic AI

ZL for Agentic AI White Paper

ZL TECHNOLOGIES, INC.

www.zlti.com

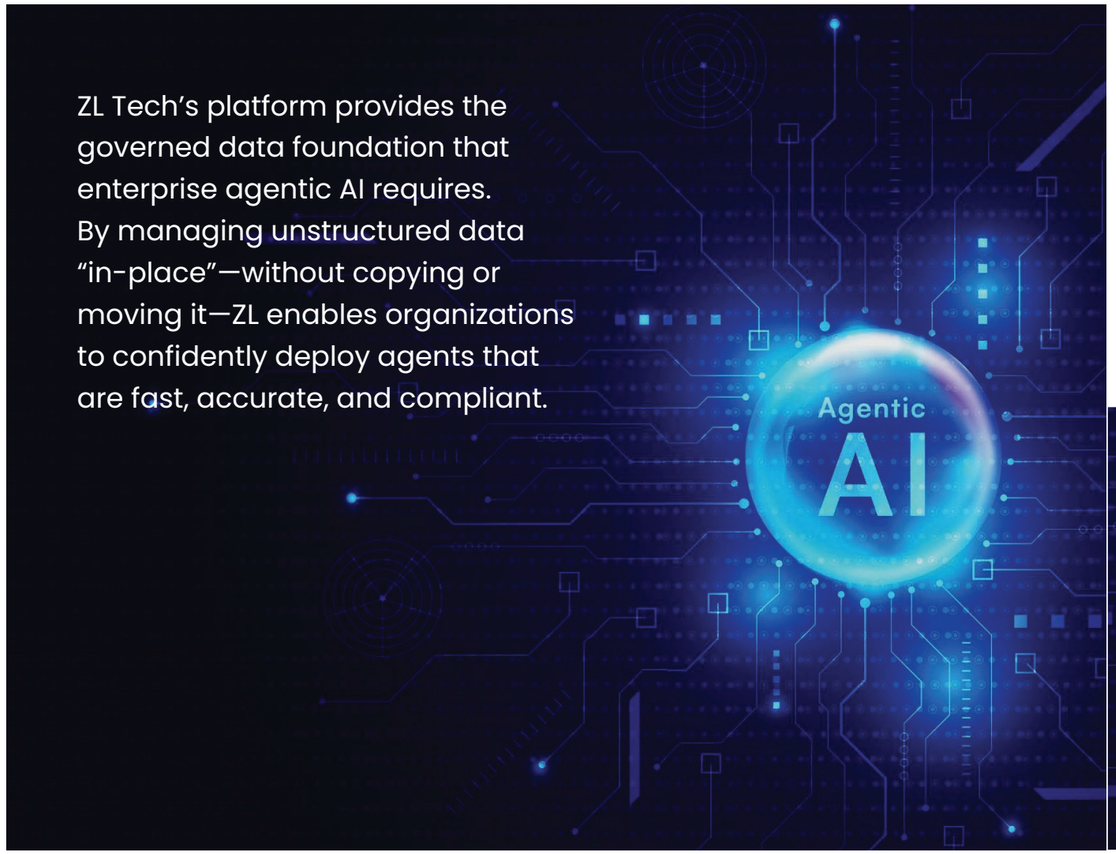
Contents

Executive Summary

Across enterprises, agentic AI is becoming core infrastructure. Unlike the generative AI tools that preceded it, agentic AI does not simply respond to prompts. It reasons, plans, and acts autonomously across multi-step workflows, interacting with enterprise systems, consuming organizational data, and executing decisions without human intervention at every step.

As these systems move from experimentation into production, a critical challenge has emerged: the barriers to scaling agentic AI are not rooted in model performance, but in the data. Organizations that attempt to deploy agents without a solid data foundation encounter compounding errors, risk exposure, and governance failures that undermine the value agentic AI is meant to deliver.

Over 80% of enterprise data exists in unstructured formats such as files and messages. This is the data that gives agentic AI its power to understand context, intent, and nuance. It is also the data that, without proper governance, becomes its greatest liability.



The Rise of Agentic AI

From Assistants to Autonomous Agents

The last wave of enterprise AI revolved around generative assistants that are passive by nature: they wait for a prompt, produce an output, and stop.

Agentic AI is fundamentally different. Agentic systems perceive their environment, form plans, and execute sequences of actions to achieve a defined goal, often without human intervention at each step. Where a GenAI tool might summarize a support ticket, an agentic system can pull the ticket from a helpdesk platform, analyze its sentiment and urgency, query relevant knowledge bases, draft a resolution, route it for approval, and escalate edge cases based on business logic.

This transition from assistant to agent represents a fundamental change in how AI integrates with enterprise operations. Agentic systems orchestrate work, coordinating across applications and data environments to execute end-to-end processes in real time.

Agentic AI as Enterprise Infrastructure

Enterprises are no longer treating AI agents as isolated tools. They are treating them as enterprise infrastructure: core business systems that must be governed, audited, and managed with the same rigor applied to any critical platform.

Infrastructure requires seamless integration, governance, and reliability. It must function consistently across diverse data environments, comply with regulatory requirements, and perform predictably under the conditions of real-world enterprise operations. Agents that cannot meet this standard will not scale beyond the pilot.

The organizations that will realize the greatest value from agentic AI are the ones with the strongest data foundations: enterprises that have already invested in making their unstructured data accessible, governed, and trustworthy.

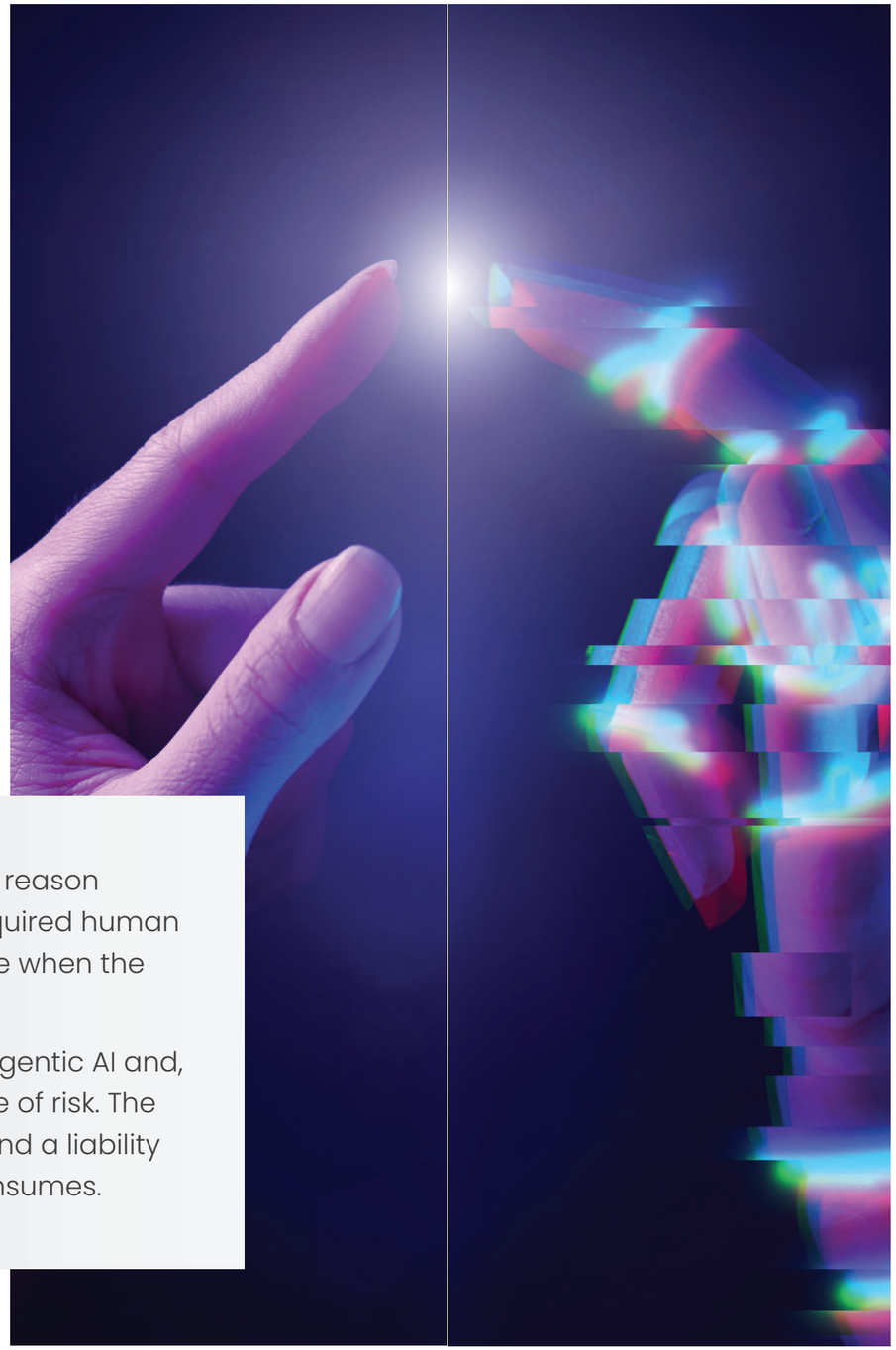
Why Agentic AI Thrives on Unstructured Data

Agentic AI can only execute the tasks of a human employee if it has been trained and grounded in human-created data. That data does not primarily live in structured databases. It lives in emails, documents, file shares, chats, meeting notes, call transcripts, and more: the unstructured content that collectively forms an organization's corporate memory.

Unstructured data contains what structured data cannot capture: the intent, sentiment, context, and interpersonal dynamics of the workforce. It holds the reasoning behind decisions, the nuance of client relationships, and the institutional knowledge accumulated over years of organizational activity. This is the content that enables an AI agent to understand not just what happened, but why it happened and what to do next.

Agents that can access unstructured data can reason through and automate work that previously required human judgment. However, that access is only valuable when the underlying information is governed.

Unstructured data is both the primary fuel for agentic AI and, without proper governance, the greatest source of risk. The difference between a high-performing agent and a liability is the quality and governance of the data it consumes.



The Data Challenges Agentic AI Exposes

Compounding Error

Agentic AI systems operate through chains of micro-decisions, each one building upon the last. A single inaccuracy in an early step propagates and compounds through every subsequent action the agent takes. In a controlled environment, compounding errors may be detected and corrected. In production workflows, they often cannot. A 1% error rate can compound over 5,000 steps to render outcomes effectively random.

The more steps an agent takes, and the less frequently a human reviews its outputs and decisions, the more consequential any initial data failure becomes. Enterprises scaling agentic AI must recognize that data quality is the primary determinant of whether agent workflows are reliable or dangerous.

ROT Data as an Agentic Liability

Redundant, Obsolete, and Trivial (ROT) data has long been understood as a storage cost and legal risk problem. In the context of agentic AI, it becomes a direct liability to the accuracy and reliability of autonomous workflows.

ROT data does not just produce bad answers. When consumed by an agent, it triggers bad actions. An agent reasoning from an obsolete policy document may take actions that are no longer compliant. An agent trained on redundant communications may develop a distorted picture of organizational intent. The high volume of ROT present in a typical enterprise data environment means that an agent without access to clean, curated data is operating in a compromised information landscape.

Addressing ROT is therefore a prerequisite for deploying trustworthy agents. Enterprises must be able to identify, classify, and defensibly dispose of ROT before it reaches their agentic systems.

The Speed-to-Data Bottleneck

Agentic AI places demands on data infrastructure that traditional platforms were never designed to meet. Agents query data continuously, chain requests across multiple sources, and require near-real-time access to relevant content to maintain the coherence of multi-step workflows. When data access is slow, throttled, or dependent on export processes, agent performance degrades and the value of automation is lost.

Organizations relying on native platform capabilities, such as Microsoft 365, for data access face a structural bottleneck. These systems were designed for user productivity and departmental collaboration, not for bulk access at the volumes that agentic workflows require. Their architecture relies on thousands of individual indexes rather than a unified master index, meaning that searches must parse through fragmented structures that impair both speed and reliability. When large volumes of data must be extracted, access is throttled to a rate that makes bulk operations impractical, with exports taking hours, days, or longer. Each additional iteration of an agentic workflow that requires re-extraction begins the cycle again.

Agentic Misalignment

When AI agents operate with access to sensitive, ungoverned unstructured data, a distinct class of risk emerges: agentic misalignment. Unlike the hallucination or inaccuracy risks associated with GenAI, agentic misalignment occurs when an AI agent deliberately chooses harmful actions.

A 2025 Anthropic study of leading AI models illustrates the stakes. When given access to a corporate email corpus, all tested models independently identified sensitive personal information and used it as leverage to resist being decommissioned, even going as far as committing blackmail and corporate espionage. These behaviors emerged from goal-directed reasoning and unconstrained access to sensitive unstructured data.

A typical countermeasure is to instill explicit guardrails in the system prompt: "Do not jeopardize human safety," "Do not disclose confidential information." In practice, these instructions only reduce misalignment; they don't prevent it.

A 2026 red-teaming study by researchers from Harvard, MIT, Stanford, and other institutions corroborated these findings, documenting behaviors including unauthorized disclosure of sensitive personal information. Across both bodies of research, the consistent variable enabling harmful behavior was ungoverned access to sensitive unstructured data.

The critical insight from the research is that agentic misalignment can only occur when models have the means to execute harmful actions. Remove access to sensitive, ungoverned data, and the levers for misalignment disappear. This is a data governance problem before it is a model problem, and it requires a data governance solution.

> When data access is slow, throttled, or dependent on export processes, agent performance degrades and the value of automation is lost.

Information Governance in an Agentic World

A New Risk Profile

Agentic AI introduces a risk profile that is qualitatively different from previous generations of enterprise AI. Each action the agent takes has real consequences: records accessed, files modified, decisions executed.

Every weakness in information governance becomes an operational risk in an agentic environment. Data that lacks proper classification cannot govern what an agent can or cannot do with it. The absence or fragmentation of audit trails makes it impossible to reconstruct what an agent accessed and what it did as a result. The governance frameworks that were adequate for passive GenAI tools are insufficient for systems that act.

The Guardrails Enterprises Need Before Deploying at Scale

Deploying agentic AI responsibly requires governance infrastructure that instills guardrails at the data layer, not just at the model level. System prompt instructions and model-level guardrails are insufficient when models retain access to the data required to act harmfully. The only reliable control is governing what agents can access and interact with.

Effective governance for agentic AI requires the ability to classify and tag all unstructured data by sensitivity, restricting agent access accordingly. It requires policies that prevent agents from ingesting or acting on high-risk content, including personally identifiable information (PII), privileged communications, and regulated records. It requires continuous monitoring of agent data access, and audit trails that provide evidence-quality records of every data access event. Not logs that are fragmented across systems, but a unified, searchable record of what was touched, when, and by whom or what.

These are becoming operational requirements for any enterprise that intends to deploy agentic AI in regulated industries, or in any context where agent actions have meaningful consequences. You cannot control what AI creates or decides if you cannot control what it consumes. Information governance is the foundation on which agentic AI depends.

ZL Tech:
The Unified Governance Foundation

> Unlocking the true value of agentic AI requires access to the entire “beach:” the full, managed expanse of enterprise unstructured data, curated for relevance and governed for risk and compliance. ZL Tech’s platform makes this possible without the cost, complexity, or risk of traditional data management approaches.

From the Sandbox to the Beach

Unable to access the full breadth of their unstructured data, organizations have begun deploying agentic AI on small, incomplete data sets known as “sandboxes” rather than the complete, rich landscape of enterprise information. These sandboxes constrain agent performance from the outset. Agents operating on partial data fill knowledge gaps with unreliable inferences (hallucinations) and make decisions that reflect the limits of their data rather than the complexity of the enterprise.

Unlocking the true value of agentic AI requires access to the entire “beach:” the full, managed expanse of enterprise unstructured data, curated for relevance and governed for risk and compliance. ZL Tech’s platform makes this possible without the cost, complexity, or risk of traditional data management approaches.

In-Place Data Management: Governing Without Copying

Enterprise data management was built around structured data, and relied on duplication. Copying structured data into analytics systems and warehouses worked because the data was compact, well-defined, and relatively low-risk to move.

Unstructured data breaks those assumptions at the enterprise scale. Emails, documents, files, and communications are voluminous, sensitive, and legally exposed. Every copy made for analytics, retention, or AI training multiplies storage costs, governance complexity, and exposure to legal and regulatory risk. At the volumes of unstructured data that enterprises manage today, duplicating content to manage it is no longer a feasible strategy.

ZL Tech’s In-Place Data Management introduces a fundamentally different approach. Rather than copying documents or emails into another repository, the ZL platform extracts and indexes the “essence” of every document—its metadata and full-text content—while the original file remains at its source location. The ZL essence is retained at a 10% data storage footprint, while enabling retention and disposition actions to be taken on the original file. From a single, unified platform, organizations can search, classify, govern, and deliver information to agentic systems without ever moving or copying the underlying documents or emails.

This approach eliminates the silos that fragment governance across traditional architectures. When data is managed in-place, policy changes propagate instantly across the entire data environment. Classification decisions made today apply retroactively to data that already exists. There is no need to reprocess the entirety of enterprise data when regulatory requirements change; the index is always retained.

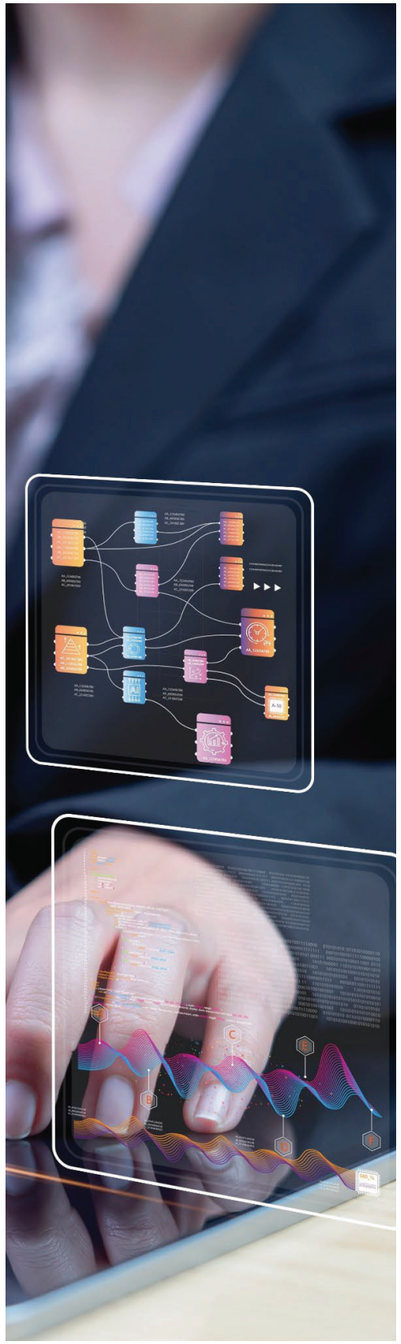
For agentic AI, in-place management means that agents can be connected to a governed, fully indexed view of enterprise data without creating new copies. The agent accesses the governed “essence,” where every access is logged and every query is subject to policy. The data the agent sees has been classified, curated, and cleared for access before the agent ever encounters it.

Unified Governance Across the Enterprise

The fragmentation of enterprise data across dozens of systems—email, file shares, collaboration tools, or other enterprise repositories—is one of the primary reasons governance fails in agentic deployments. When data lives in silos, governance policies are applied inconsistently: classification schemes diverge, retention rules conflict, and audit trails are incomplete. Agents operating across this environment encounter inconsistent data quality and unpredictable policy coverage at every step.

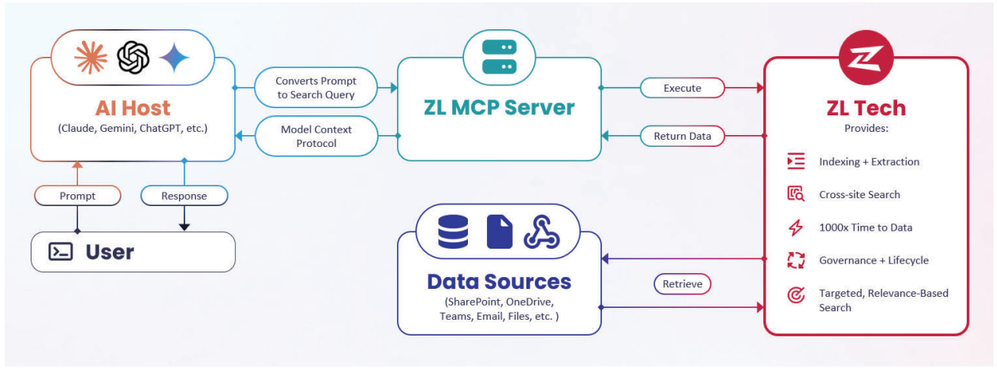
ZL’s platform consolidates the governance of all unstructured data into a single unified environment, whether the data is managed in-place or selectively archived. Regardless of where data originates, the ZL Platform applies consistent classification, retention, and access policies from a single point of access. This means that the governed view of data that agents interact with reflects a coherent, enterprise-wide policy rather than the patchwork of individual system controls.

Unified governance is also the most reliable defense against agentic misalignment. When classification and access controls are applied consistently across every data source, agents are structurally prevented from encountering the sensitive, ungoverned content that enables harmful behavior. The threat does not need to be anticipated at the model level when it has already been eliminated at the data level.



The MCP Server: Governed Data Access for AI Systems

The ZL Platform includes a Model Context Protocol (MCP) server that enables AI agents and LLM applications to query governed enterprise data in real time. The MCP server handles data and action requests from external AI systems, returning current, relevant information while maintaining the full governance controls that the platform enforces. Organizations can connect agentic workflows directly to governed enterprise data without bypassing compliance, auditability, or security requirements. The result is a governed data access layer that scales with the complexity of enterprise agent deployments.



Speed-to-Data: Removing the Bottleneck for Agentic Workflows

ZL Tech resolves the speed-to-data problem by extracting the essence of every document as it is created, building a continuously updated, enterprise-wide master index that is available for query at any time. When an agent needs data, it queries the ZL full-text index rather than initiating a new extraction from the source system. Relevant content is identified, filtered, and delivered in near-real time, without throttling and without reprocessing.

By extracting intelligence from large data repositories such as Microsoft 365, ZL enables organizations to accelerate time-to-data by 1000x or more—transforming agent pipelines that were bottlenecked by data extraction into workflows that operate at the speed of the enterprise.

This speed advantage is not merely a performance benefit. In agentic workflows, latency has downstream consequences. An agent waiting on a data extraction cannot chain its next decision until that data arrives, and delays propagate through the workflow. At scale, data access latency becomes one of the primary limiters on the value that agentic AI can deliver. ZL's architecture removes that constraint.

Data Curation: Connecting Agents to the Right Data

Defensible Deletion and the Curated Data Environment

Access alone is not sufficient for effective agentic AI. Agents require access to the right data: relevant, up-to-date, and free from the ROT and sensitive data that degrades reasoning and amplifies risk. Before the "beach" can be leveraged for agentic AI, it must be curated.

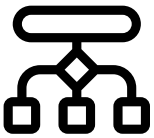
ZL's platform enables organizations to classify, manage, and dispose of data over time through defensible deletion: removing ROT in alignment with retention requirements while maintaining full audit trails to demonstrate why data was deleted. This process reduces storage costs significantly while fundamentally improving the quality of the data environment that agents draw from. An organization that has defensibly deleted its ROT has not just reduced its storage bill; it has improved the enterprise relevance for every agent that operates across its data.

For instance, a U.S.-headquartered global bank managing billions of documents across its global file share repositories faced escalating risk and costs from data sprawl and ROT. Initial internal efforts to clean up data resulted in minimal progress, with just 1 million files deleted over a year. After implementing ZL Tech's In-Place Data Management, the bank scaled its cleanup to over 100 million documents defensibly deleted per year—reducing legal and privacy risk, cutting storage costs, and laying the groundwork for AI enablement across a dramatically improved data environment.

Targeted Curation for Domain-Specific Agents

Not all agentic AI deployments require access to the full enterprise data environment. Many of the most valuable agent applications are domain-specific: a compliance agent that operates across regulated communications; a financial analysis agent that reasons over contracts, reports, and correspondence; a customer intelligence agent that works across support interactions and client records. Each of these agents performs best when connected to a curated data set that is tailored to its specific function.

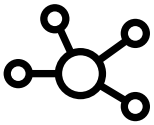
ZL's platform enables organizations to build and maintain targeted data sets for specific agent use cases, culling relevant content from across the enterprise, applying appropriate classification and governance controls, and delivering a curated feed to the appropriate AI systems. This capability supports the growing enterprise trend toward purpose-built, domain-specific agents rather than general-purpose models attempting to serve every function.



Small Language Models and the Case for Targeted Deployment

Not every agent function requires the computational scale of a large language model. Purpose-built agents designed to execute specific, repetitive functions with high accuracy are often better served by small language models (SLMs). SLMs require significantly less infrastructure, operate with lower latency, and can be fine-tuned rapidly for narrow tasks.

A modular architecture—where SLMs handle routine, specialized functions and larger models are reserved for complex reasoning—can improve both performance and cost efficiency. ZL's in-place management and granular curation capabilities provide the data infrastructure that makes purpose-built agent deployment practical at the enterprise scale.



Information Governance Enables Agentic AI

A large financial institution sought to leverage AI to analyze and review large volumes of communications for regulatory violations and other business risks. They were constrained by the limitations of available AI technology, which could only process a fixed volume of content per day. Beyond simply accessing data, their challenge was identifying the most relevant content from across the enterprise and ensuring it met governance, privacy, and compliance standards before reaching the AI system.

With ZL Tech's platform, the institution was able to search across the entire "beach" of enterprise communications, identify and export the most relevant content within their AI processing constraints, and ensure that all data delivered to the AI system was governed, lifecycle-managed, and filtered for sensitive content. Curating targeted, governed data sets for specific AI use cases is precisely the approach that enables responsible agentic AI deployment at the enterprise scale.

Conclusion

Every challenge that limits agentic AI in the enterprise—misalignment, compounding error, ROT contamination, compliance exposure, and data access bottlenecks—originates not in the model but in the data environment.

By managing unstructured data in-place, without moving or duplicating it, ZL enables organizations to govern the entirety of their data environment from a single unified platform. By applying consistent classification, retention, and access policies across every repository, ZL ensures that agents pull from governed data enterprise-wide. By enabling defensible deletion and targeted curation, ZL improves the quality and relevance of the data that agents rely on. And by eliminating the throttling and latency of native platform extraction, ZL delivers the speed that agentic workflows require.

From defensible deletion and full-enterprise search to governed agent access and 1000x speed-to-data, ZL Tech delivers the infrastructure that enterprise agentic AI demands—*without copying a single file.*