opentext™

WHITE PAPER

Secure AI content management with OpenText Content Aviator

AI security, data protection, and governance foundations of OpenText Content Aviator

Contents

"... trust issues - relating to data privacy and security, data quality, bias, and accuracy - pose a barrier for both enterprise and consumer gen AI adoption."

Deloitte, 2025 technology industry outlook, February 11 2025

Executive summary

GenAI promises unprecedented productivity gains for knowledge workers, but it also introduces potential cybersecurity and data privacy challenges that must be addressed before organizations can confidently deploy AI at scale. For CISOs and cybersecurity professionals, the question is not whether to adopt AI, but how to do so without compromising enterprise security standards or creating new attack vectors.

OpenText™ Content Aviator™ addresses these challenges through a foundational principle: *Great AI starts with great information management*. Rather than treating AI security as an afterthought, Content Aviator is purpose-built for secure AI content management from the ground up, extending proven content management capabilities with enterprise-grade AI features while inheriting existing governance frameworks. This ensures that AI capabilities never bypass established security boundaries or escalate user privileges.

This paper outlines the platform's approach to data protection throughout the AI lifecycle, from secure data ingestion and vector database management to trusted integration with frontier large language models. Organizations do not need to choose between AI innovation and enterprise security.

Great AI starts with great information management

GenAI is rapidly changing the way users engage with information and get work done. Information has never been so relatable and accessible. Users can now easily search and query for otherwise obscure facts and figures using natural language. AI-powered agents can now simplify complex tasks that once took hours or days into a few minutes. While AI profoundly transforms the productivity of knowledge workers, it can also pose novel risks, vulnerabilities, and threats to cybersecurity if critical precautions are ignored.

AI introduces new security challenges

For CIOs, CISOs, and cybersecurity professionals, GenAI adoption brings with it an entirely new set of issues related to security, governance, and compliance. The very data that GenAI systems leverage to provide value—often an organization's most sensitive intellectual property and customer information—must be rigorously protected. New processes and systems must be put into place to address these challenges.

Proper implementation of a robust content management system mitigates many key cybersecurity concerns:

- **Authentication:** Do you have access to grounding data or the authority to process the information with AI?
- **Versioning:** Is the user working with the latest available version of the content?
- **Managing content sprawl:** Does the use of GenAI tend to contain or exacerbate content sprawl?
- **Data privacy:** Is the GenAI tool properly respecting data privacy issues and compliant uses of GenAI with PII?
- **Data protection:** Are there data leak concerns with the use of GenAI or its output?
- **Jurisdiction:** Will a particular use of AI be compliant with local regulations or industry best practices?
- **Sovereignty:** Does the use of AI break jurisdictional or policy-based data sovereignty requirements?
- **Retention:** Is grounding data past its retention, and would further use of AI result in over-retention of data?

Understanding the issues—as well as planning and implementing new cybersecurity protocols—drains valuable time and can delay AI strategies that organizations are eager to implement.

According to Gartner® "... 57% of organizations estimate their data is not AI-ready." In the same report Gartner® also stated "To scale AI, leaders must evolve data management practices and capabilities to ensure AI-ready data ... can cater to existing and upcoming business demands"

Gartner, The 2025 Hype Cycle for Artificial Intelligence Goes Beyond GenAI, 2025

OpenText Content Aviator is GenAI as an extension of information management

OpenText starts with a simple premise: Great GenAI starts with great information management. Likewise, effective AI governance is an extension of robust information governance. Content management provides a foundation for meeting the various AI governance criteria listed above.

Content Aviator is an extension of a secure, governance-driven content management platform with enterprise-grade data security, privacy, and compliance. OpenText is committed to the highest level of protection for your information so that you can confidently transform user productivity.

Secure generative AI foundation

OpenText Content Aviator is purpose-built for secure AI computing across all use cases, and is deeply embedded within the content management framework, plainly abiding all access controls, privileges, roles, and other governance rules.

Encrypted data protection

Content Aviator inherits and extends OpenText's robust encryption practices, ensuring that customer data is protected throughout its lifecycle. This commitment to encryption excellence is a cornerstone of the platform's security posture.

- **Securing data at rest:** All customer data managed within the Content Aviator ecosystem, including documents, metadata, and the vector database that powers its retrieval augmented generation (RAG) capabilities, is protected by strong encryption (AES 256-bit keys) within the cloud. This ensures that data stored within the platform is rendered unintelligible to unauthorized parties.
- **Active data leak protection:** Integration with Microsoft Purview Information Protection (AIP) provides additional protection of private, proprietary, or sensitive content from data leaks. If a user downloads or emails the document, it remains encrypted, and the user must authenticate every time a document is viewed or opened. OpenText content management ensures that documents are protected in this way through multiple levels of policy management, and AI-based workflows can automate the security review process.
- **Securing data in motion:** All data transmitted between user interfaces, applications, APIs, and backend services—including crucial communications with large language models (LLMs)—is secured using transport layer security (TLS). OpenText's cloud services utilize TLS (HTTPS) enabled servers to encrypt data during transit, protecting it from eavesdropping or interception.

GARTNER is a registered trademark and service mark of Gartner, Inc. and/or its affiliates in the U.S. and internationally and is used herein with permission. All rights reserved.

- **Encryption key management:** To ensure the security of encrypted data, encryption keys must be managed using integrated secure key management applications. This facilitates unique, customer-controlled encryption keys with no direct access by OpenText to your content. A multi-layered approach to key management significantly enhances data segregation and protection, particularly in multi-tenant cloud environments, ensuring that each customer's data remains cryptographically isolated.

Fortified access control and identity management

Content Aviator is designed to integrate seamlessly and securely within an enterprise's existing identity and access management (IAM) framework. This ensures that access to AI capabilities and the underlying data is strictly controlled and adheres to established corporate policies.

Seamless single sign-on (SSO) integration

Content Aviator supports industry-standard SSO protocols, such as Security Assertion Markup Language (SAML) and OpenID Connect, for authentication, as well as OAuth for delegated authorization. This capability can be facilitated by OpenText™ Identity Manager, which provides comprehensive, integrated identity services.

Content Aviator is authenticated simultaneously with its host content management solution, so there is no repeat authentication or risk of authentication confusion for the user or access controls. This not only simplifies user access but also centralizes authentication management under the enterprise's trusted IAM system, ensuring consistent policy enforcement.

Enforcing the principle of least privilege: granular user permissions and entitlements

A core tenet of Content Aviator's security model is the enforcement of the principle of least privilege. This means that both human users and the AI assistant itself are granted only the minimum necessary access rights to perform their tasks. When using Content Aviator, the user's access to information remains the same; the tool will have the same access to privileged information as the user, no more.

All inference tasks are conducted only on content that the specific user is permitted to access according to their established roles and permissions within the enterprise content management system. This prevents the AI from becoming a "super user" or an inadvertent backdoor to sensitive information.

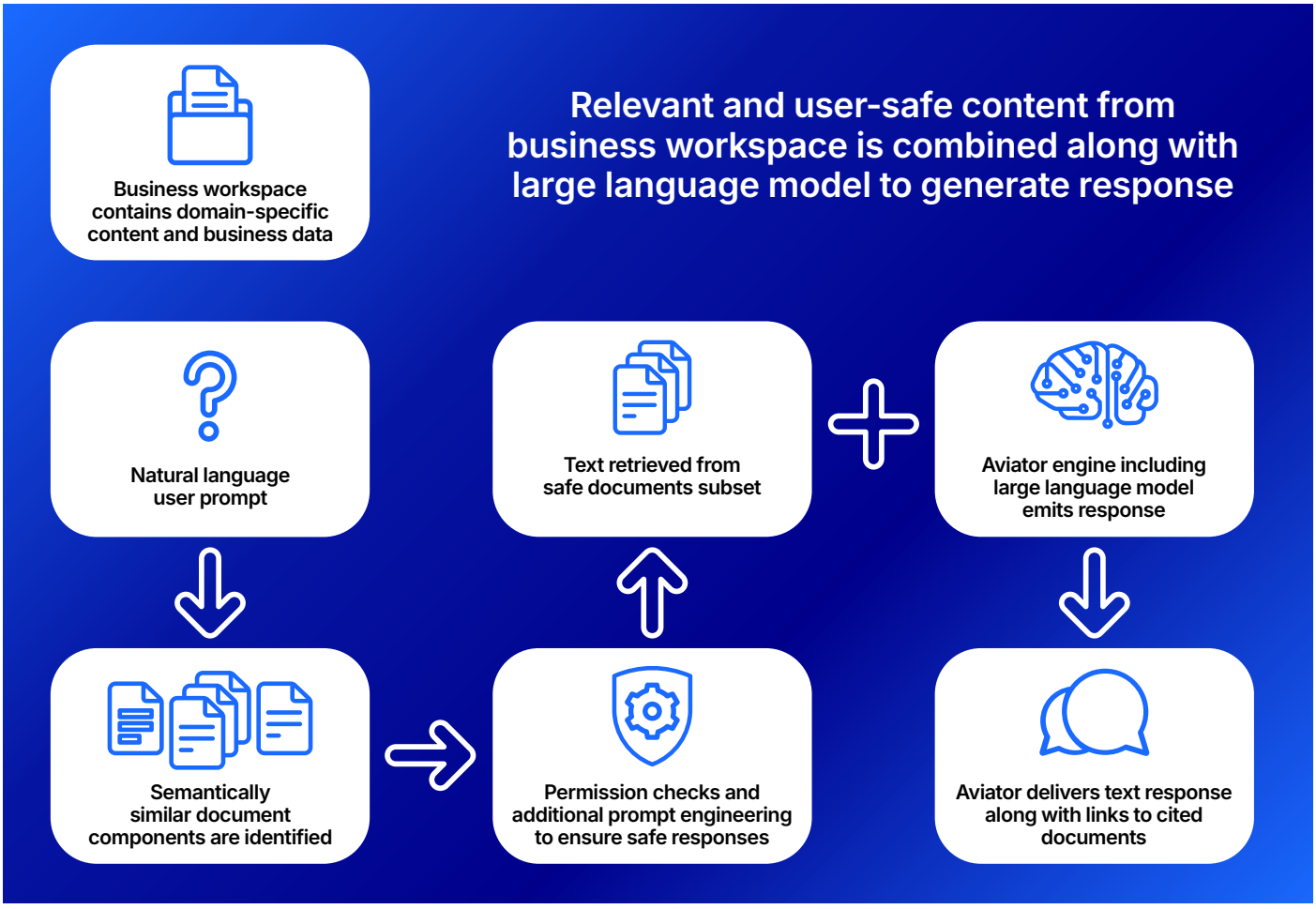
Content Aviator is also part of a larger architecture supporting content management. Host cloud infrastructure and environment are managed according to these same principles, isolating components wherever possible to prevent inadvertent exposure.

Role-based access control within Content Aviator

Content Aviator implements robust role-based access control mechanisms to manage access to its features and the content it processes. Access control filtering ensures that users are only granted access to information for which they have explicit user rights. These controls extend to how Aviator interacts with data, ensuring that searches, summaries, and generated content are all based on the data corpus appropriate for the user's role, location, or department. This granular control ensures that the AI respects and enforces existing enterprise security boundaries, rather than creating new, potentially weaker ones, providing a powerful assurance for CISOs.

Secure GenAI through retrieval augmented generation

OpenText Content Aviator promotes accuracy and prevents data leaks by utilizing RAG. Direct application of permissions to the retrieval process means that grounding is limited to user-permissible content. By dynamically retrieving information from an organization's own authoritative data sources, RAG grounds the LLM's outputs, significantly reducing the likelihood of hallucinations and ensuring that the generated content aligns with enterprise-specific knowledge.



How Content Aviator processes a query

A critical aspect of the OpenText RAG architecture is support for multi-tenancy, ensuring that each user operates in an isolated environment and acts on the user's behalf and with their specific credentials, access controls, and privileges.

Additionally, the RAG architecture depends on a semantic search, backed by a carefully controlled vector database and is protected by the same rules, access rights, and privileges as the original content. No user's AI query can bypass them.

Security of the vector database

Central to the RAG architecture is the vector database, which stores numerical representations (embeddings) of the enterprise data. Securing this vector database is critical:

- **Encryption of embeddings and stored data**
  Vector embeddings, although numerical, are derived from and represent potentially sensitive enterprise information. OpenText ensures these embeddings and the underlying data within the vector database are encrypted both at rest and in transit, consistent with the comprehensive data encryption policies (e.g., AES-256 for data at rest, TLS for data in motion). This renders them unintelligible to unauthorized parties.

- **Access controls and isolation within the vector database**
  Access to the vector database is strictly controlled. Only authorized components of the Content Aviator service are permitted to query or modify it. In multi-tenant deployments, robust data isolation mechanisms are implemented within the vector database to ensure that one tenant's embeddings and data are not accessible by another. Furthermore, the RAG retrieval process itself is designed to be permission-aware. This ensures that the information retrieved to augment the LLM prompt is filtered according to the user's existing access rights.

Ensuring the integrity and confidentiality of grounding data

Grounding content, including that ultimately used for generative AI inference, is subject to strict limits. In addition to the original security policies, data sent to LLM providers for grounding is not stored following inference, nor are responses. Secure processes are implemented for data ingestion, preparation (including document chunking and embedding generation), and storage within the RAG framework.

Mitigating RAG-specific vulnerabilities (OWASP)

OpenText acknowledges the evolving landscape of AI-specific threats, including those identified by the Open Worldwide Application Security Project (OWASP). The OWASP Top 10 for LLM Applications highlights "LLM08: Vector and Embedding Weaknesses," encompassing risks such as embedding inversion attacks, where attackers attempt to reconstruct sensitive input data from its embedding.

OpenText Content Aviator's approach to mitigating these RAG-specific vulnerabilities includes:

- **Encryption of vectors:** Encrypting the vector embeddings themselves is a key defense against unauthorized access and attacks like embedding inversion.
- **Secure API design:** Implementing secure APIs for accessing and querying the vector database, incorporating strong authentication and authorization.
- **Input validation:** Validating and sanitizing inputs to queries against the vector store to prevent injection-style attacks targeting the retrieval mechanism.
- **Permission-aware retrieval:** Ensuring that the RAG process only retrieves data segments that the requesting user is authorized to access, thus limiting the potential exposure even if the embedding itself were somehow compromised.

These measures, combined with the overarching security principles of data minimization and least privilege, demonstrate OpenText's proactive stance in securing the novel components of GenAI architectures. The security of the RAG process is not an afterthought but an integral part of Content Aviator's design, reflecting a deep understanding of the critical role vector databases play and the emerging threats against them.

Secure integration of frontier LLMs: leveraging superalignment for enhanced safety

OpenText Content Aviator offers enterprises flexibility to leverage the power of leading frontier LLMs, including Google Gemini™, OpenAI® on Azure®, and Amazon Nova™. Additionally, OpenText will be releasing its own off-cloud OpenText Model Services to provide Content Aviator in off-cloud configurations, giving customers total control over GenAI infrastructure, model security, and data residency.

The OpenText strategy for integrating these powerful external models is built on a foundation of secure interaction and strict data privacy controls. Architectural measures ensure that all communications with these external LLM APIs are conducted over secure, encrypted channels, adhering to TLS standards.

Ensuring customer data privacy and isolation from LLM providers

A paramount concern for enterprises adopting GenAI is the privacy and control of their data when interacting with third-party LLM providers. OpenText addresses this directly through its core commitment: "Your data is not our product" and "your data is yours alone." This principle is technically enforced in Content Aviator's architecture:

- **No training on customer data:** Customer data, including prompts and the enterprise-specific grounding data sent to an LLM for inference, is explicitly not used to train the foundation models of providers like Google, OpenAI, or Amazon. This prevents proprietary or sensitive information from being absorbed into public or shared model datasets.
- **No data retention by LLM providers:** Data and prompts used for inference by the model are not stored by the model providers after the transaction is complete. This transient handling minimizes the window of exposure and prevents the accumulation of customer data on third-party infrastructure.
- **Architectural data isolation:** The Content Aviator architecture is designed to isolate customer data, ensuring it is only shared with the LLM provider at the precise moment of inference and solely for the purpose of generating a response.

These measures collectively ensure that enterprises retain sovereignty over their data while still benefiting from the advanced capabilities of frontier LLMs.

The superalignment advantage: enhanced LLM safety, ethics, and reliability

The concept of AI "superalignment" refers to the significant and ongoing efforts by frontier model providers (like Google, OpenAI, Microsoft, and Amazon) to ensure their LLMs behave in ways that are safe, ethical, accurate, and aligned with human values and intentions. These providers are investing billions of dollars in research and development dedicated to:

- Reducing harmful or biased outputs.
- Improving factual accuracy and mitigating "hallucinations."
- Enhancing the models' ability to follow complex instructions reliably.
- Building in safeguards against misuse.

Through reliance on with these leading frontier models and additional mitigation offered by OpenText Aviator Model Services, Content Aviator customers will benefit from OpenText's extensive safety and alignment investments. This means Content Aviator leverages LLMs that are continuously being refined for safety, reliability, and ethical behavior by the organizations with the deepest expertise and resources in this domain.

Commitment to industry best practices and AI governance

Content Aviator's security is informed by and aligned with leading industry cybersecurity frameworks and OpenText's own comprehensive principles for trustworthy AI. This commitment ensures a structured, robust, and transparent approach to managing the unique risks associated with generative AI.

Alignment with leading cybersecurity frameworks

Adherence to recognized frameworks provides external validation of OpenText's security practices and offers a common language for discussing risk and compliance with cybersecurity professionals.

NIST AI Risk Management Framework (AI RMF):

The National Institute of Standards and Technology (NIST) AI Risk Management Framework (AI RMF) provides voluntary guidance to better manage risks to individuals, organizations, and society associated with AI. OpenText's development and operational practices for Content Aviator align with the AI RMF's four core functions:

- **Govern:** OpenText establishes clear governance structures, roles, and responsibilities for AI risk management, integrating these into the overall organizational risk culture. This is reflected in policies like the OpenText Bill of AI Obligations.
- **Map:** Context is established, and risks related to AI systems are identified and analyzed throughout their lifecycle. For Content Aviator, this involves identifying potential risks in the RAG architecture, LLM interactions, and data handling.
- **Measure:** Identified risks are assessed, analyzed, tracked, and their impacts evaluated using qualitative and quantitative methods. This includes ongoing monitoring of Content Aviator's performance, security, and user feedback.
- **Manage:** Risks are prioritized and acted upon based on their projected impact. This involves implementing mitigation strategies, such as the security controls detailed throughout this paper, and continuously refining them.

OWASP Top 10 for Large Language Model Applications

OpenText is committed to addressing vulnerabilities outlined in the OWASP Top 10 for Large Language Model Applications, either directly or in partnership with our large language model providers. This means that Content Aviator incorporates measures to counter key risks such as:

- **Prompt injection:** Mitigating unintended or malicious vulnerabilities introduced through GenAI prompts.
- **Insecure output handling:** Outputs from Content Aviator are validated by the LLM or OpenText Model Services before being passed to downstream systems or displayed to users, and reducing the risk of malicious code execution.

- **Training data poisoning:** OpenText partners with trusted frontier large language model providers that are committed to safe, ethical, and reliable model training protocols.
- **Sensitive information disclosure:** Addressed through strict adherence to access controls and privileges as dictated by content management.
- **Vector and embedding vulnerabilities:** Specific measures like encryption of vector embeddings and adherence to access controls for vector data in coordination with content management.

Adherence to Cloud Security Alliance (CSA) GenAI security guidance

The Cloud Security Alliance (CSA) provides valuable guidance on AI governance, security, and risk management for cloud-based AI solutions. OpenText's practices for Content Aviator, as a cloud-delivered GenAI service, align with CSA recommendations concerning the security of GenAI tools and applications, robust third-party/supply chain management for AI components (such as frontier LLMs), and operational considerations for secure AI deployment and use. This includes employee training, secure operationalization, and clear delineation of responsibilities.

OpenText Bill of AI Obligations

The OpenText Bill of AI Obligations serves as the foundational ethical and operational guide for all AI development and deployment at OpenText, including Content Aviator. The key tenets and their manifestation in Content Aviator are:

- **Transparency builds trust:** OpenText strives for transparency in how Content Aviator operates, how it uses data, and the capabilities of its AI models. This white paper itself is an element of that transparency.
- **AI and ethical AI are the same thing:** Ethical considerations are not an add-on but are integrated into the design, development, and deployment lifecycle of Content Aviator.

- **It starts with value-based design:** Content Aviator is designed to provide clear business value while upholding ethical principles and respecting user rights.
- **Your data is not our product:** This is a cornerstone. Customer data processed by Content Aviator is not used to train OpenText's or third-party general-purpose models without explicit consent and is not monetized as a separate data product.
- **Respect intellectual property, images, and likeness:** Content Aviator is designed to operate on customer-provided data within their authorized scope, respecting IP rights associated with that content.
- **Security and privacy remain paramount:** As detailed throughout this document, robust security and privacy controls are fundamental to Content Aviator's architecture and operation.

OpenText Bill of AI Obligations

- Transparency builds trust
- AI and ethical AI are the same thing
- It starts with value-based design
- Your data is not our product
- Respect intellectual property, images, and likeness
- Security and privacy remain paramount
- Dedicated to accurate, verifiable AI results
- Promote the common good

- **Dedicated to accurate, verifiable AI results:** Through RAG and the use of aligned frontier models, Content Aviator aims to provide responses that are accurate, grounded in factual enterprise data, and verifiable.
- **Promote the common good:** OpenText aims for its AI technologies to contribute positively to enterprise productivity and decision-making.

In addition, OpenText maintains a comprehensive AI risk management strategy that encompasses ethical considerations, data governance, and security protocols. This strategy recognizes that strong content management and information governance are prerequisites for trustworthy AI. AI governance frameworks and policies guide the development and deployment of solutions like Content Aviator to ensure responsible and secure practices, addressing potential risks such as data quality issues or model corruption.

Conclusion

OpenText works from a simple premise: Enterprise security and AI innovation are not mutually exclusive. By anchoring AI capabilities within proven content management foundations, Content Aviator combines AI transformation and enterprise-grade data protection. The comprehensive security model—spanning encryption, access controls, RAG architecture, and third-party integration—addresses both current threat landscapes and emerging AI-specific vulnerabilities while enabling organizations to harness frontier LLM capabilities without compromising data sovereignty.

As the AI landscape continues to evolve, OpenText's foundational approach—treating AI governance as an extension of information governance—provides organizations with the flexibility to adapt to new technologies while maintaining consistent security postures.

Content Aviator represents not merely a secure AI tool but a comprehensive platform for responsible AI transformation, enabling enterprises to confidently scale AI capabilities across their workforce while upholding the highest standards of data protection, privacy, and governance.

Copyright © 2025 Open Text • 09.25 | 262-000202-001

opentext™