INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY FUTURISTIC DEVELOPMENT

Agentic Artificial Intelligence for Enterprise Data Governance and Decision Support Through Secure Retrieval Augmented Generation Across Microsoft 365 Information Environments

Nunayon Richard Avoseh [1*], Joy Onma Enyejo [2]
[1] Department of Information & Decision Sciences, University of Illinois Chicago, Chicago Illinois, USA
[2] Department of Business Administration, Nasarawa State University Keffi, Nasarawa State, Nigeria

* Corresponding Author: **Nunayon Richard Avoseh**

Article Info

P-ISSN: 3051-3618
E-ISSN: 3051-3626
Volume: 03
Issue: 01
Received: 10-02-2022
Accepted: 12-03-2022
Published: 17-04-2022
Page No: 65-82

Abstract

The increasing volume, heterogeneity, sensitivity, and distributed ownership of enterprise information within Microsoft 365 environments have created significant challenges for data governance, secure information retrieval, provenance verification, policy enforcement, and evidence-based organizational decision support. Conventional retrieval-augmented generation systems primarily optimize semantic relevance between user queries and indexed documents but frequently provide limited support for identity-aware retrieval, authorization inheritance, contextual governance policies, cross-application reasoning, document provenance, risk-sensitive response generation, and autonomous multi-step decision workflows. This study proposes a novel *Secure Agentic Governance Retrieval and Decision Intelligence algorithm (SAGR-DI)* for enterprise information environments integrating Microsoft SharePoint, OneDrive, Teams, Outlook, and related Microsoft 365 repositories. SAGR-DI combines agentic task decomposition, hybrid dense-sparse retrieval, graph-based enterprise knowledge representation, adaptive query routing, access-control-aware retrieval, metadata and sensitivity classification, provenance verification, contextual reranking, policy-constrained reasoning, and confidence-calibrated response generation within a unified secure retrieval-augmented generation architecture. The proposed system employs multiple collaborating agents for query interpretation, governance verification, retrieval planning, evidence validation, policy compliance, reasoning, and decision synthesis. Candidate enterprise records are retrieved using hybrid lexical-semantic search and subsequently evaluated through a governance-aware scoring function that jointly considers semantic similarity, access authorization, document authority, information freshness, sensitivity classification, provenance reliability, organizational relevance, and cross-document consistency. A dynamic risk-control layer prevents unauthorized information exposure by evaluating document-level and user-level permissions before contextual evidence is passed to the generative model. The proposed SAGR-DI framework is comparatively evaluated against BM25, conventional dense vector retrieval, Reciprocal Rank Fusion, standard retrieval-augmented generation, GraphRAG, and a conventional multi-agent RAG architecture. Performance evaluation is structured around Precision@k, Recall@k, nDCG@k, Mean Reciprocal Rank, grounded-answer accuracy, hallucination rate, policy-violation rate, retrieval latency, decision consistency, provenance coverage, and governance-compliance accuracy. Experimental analysis is designed to quantify performance across governance queries, policy retrieval, compliance investigation, enterprise knowledge discovery, executive decision support, and multi-document analytical tasks. Comparative bar graphs, retrieval-performance curves, latency-accuracy plots, governance-risk matrices, confusion matrices, radar charts, and ablation graphs are used to examine the contribution of secure retrieval, graph reasoning, agent coordination, provenance validation, and adaptive reranking. The central hypothesis is that jointly optimizing retrieval relevance and governance constraints will enable SAGR-DI to achieve higher grounded-answer reliability and governance compliance than relevance-only retrieval architectures while maintaining operationally acceptable retrieval latency. The resulting framework provides a technical foundation for secure, explainable, auditable, and context-aware enterprise artificial intelligence capable of transforming fragmented Microsoft 365 information assets into governed organizational knowledge and decision intelligence.

DOI: https://doi.org/10.54660/IJMFD.2022.3.1.65-82

Keywords: Agentic Artificial Intelligence, Enterprise Data Governance, Retrieval-Augmented Generation, Microsoft 365, Decision Support

1. Introduction

1.1. Background to Agentic Artificial Intelligence in Enterprise Information Systems

Enterprise information systems are evolving from passive repositories and rule-based workflow platforms into adaptive environments in which artificial intelligence can interpret requests, retrieve distributed evidence, invoke tools, coordinate subtasks, and support decisions across organizational boundaries. This transition is especially significant in Microsoft 365, where

operational knowledge is dispersed across SharePoint sites, OneDrive repositories, Teams conversations, Outlook messages, and policy documents. Earlier enterprise AI applications demonstrated that machine learning can convert heterogeneous data into predictive signals for operational decisions, including medication-risk prediction and equipment-condition assessment (Onyekaonwu *et al*., 2019; Oladoye *et al*., 2021). Agentic artificial intelligence extends this capability by introducing autonomous planning, iterative reasoning, tool use, context preservation, and task delegation. Rather than returning a single similarity-based answer, an agentic system can decompose a governance question, identify relevant repositories, validate user permissions, retrieve supporting records, compare conflicting evidence, and construct an auditable recommendation. This reflects the broader shift from automation toward human–AI augmentation, where machine intelligence supports rather than simply replaces managerial judgment (Raisch & Krakowski, 2021).

For enterprise deployment, however, technical intelligence must be coupled with organizational readiness, controlled access, trustworthy evidence, and measurable performance. Jöhnk *et al*. (2021) show that successful AI adoption depends on organizational resources, data conditions, skills, strategic alignment, and implementation readiness, factors that become more critical when autonomous agents operate over sensitive enterprise content. The present study therefore frames agentic AI as a governed decision-support architecture rather than an unrestricted generative interface. Its proposed Secure Agentic Governance Retrieval and Decision Intelligence algorithm, SAGR-DI, coordinates query interpretation, hybrid retrieval, graph-based context expansion, authorization checking, provenance validation, policy-constrained reasoning, and confidence-calibrated response generation. Within Microsoft 365, this design allows an executive compliance query, for example, to be answered only from records the requester is authorized to access while preserving source traceability. The study consequently evaluates whether agentic coordination can improve retrieval relevance, grounded-answer accuracy, decision consistency, and governance compliance without creating unacceptable latency or information-disclosure risk.

1.2. Enterprise Data Governance and Decision-Support Challenges in Microsoft 365

Enterprise data governance in Microsoft 365 requires control over where information is stored, who can discover it, how it is classified, which version is authoritative, and whether evidence suits a decision context. Governance is difficult because SharePoint, OneDrive, Teams, and Outlook combine metadata with documents, conversations, attachments, duplicates, changing permissions, and varied retention or sensitivity requirements. Kwarteng *et al*. (2020) emphasize that access control and cybersecurity awareness are central to reducing organizational information risk, while Nwokocha *et al*. (2021) demonstrate the importance of secure interoperability when information must move across heterogeneous systems without losing integrity or confidentiality. These concerns affect enterprise AI because a retrieval system may return semantically relevant content that is nevertheless unauthorized, obsolete, incorrectly classified, or disconnected from its original business context. Data governance must operate as a computational constraint on retrieval and reasoning rather than as a policy layer applied after generation.

A further challenge is converting governed information into decision support. Abraham *et al*. (2019) conceptualize data governance as the exercise of authority and control over data management, linking governance mechanisms to value creation and risk reduction. In parallel, Wang *et al*. (2020) show that analytics capabilities can strengthen cloud-data security management when organizations develop mechanisms that connect data analysis with security controls. Building on these principles, this study treats Microsoft 365 decision support as a joint optimization problem involving relevance, authorization, document authority, freshness, sensitivity, provenance, and cross-document consistency. SAGR-DI therefore evaluates retrieved items through governance-aware scoring before evidence reaches the generative reasoning layer. For example, two policy documents may be semantically similar, yet the system should prioritize the approved version and suppress a superseded draft even when the latter has higher vector similarity. This requirement motivates the later experimental comparison of SAGR-DI with BM25, dense retrieval, Reciprocal Rank Fusion, standard RAG, GraphRAG, and conventional multi-agent RAG using retrieval quality, governance-compliance accuracy, provenance coverage, policy-violation rate, decision consistency, and latency as complementary performance measures.

1.3. Secure Retrieval-Augmented Generation for Enterprise Knowledge Management

Secure retrieval-augmented generation addresses a central weakness of conventional large language models: generated answers may be fluent yet disconnected from current enterprise evidence. In Microsoft 365, retrieval must operate across documents, messages, collaborative workspaces, and cloud-native services while preserving identity, permissions, metadata, and provenance. Idika *et al*. (2021) illustrate how cloud-native environments require intelligent models that remain sensitive to distributed system characteristics and security threats, whereas Ijiga, Ifenatuora, and Olateju (2021b) demonstrate the importance of combining heterogeneous digital content into coherent information experiences. For enterprise knowledge management, these principles imply that retrieval should not depend on vector similarity alone. A secure RAG architecture must determine whether a user is authorized to access candidate enterprise evidence, whether the evidence is authoritative, and whether the retrieved fragments collectively support the requested decision. These requirements matter when sensitive policies, audit records, contractual correspondence, and operational reports are distributed across Microsoft 365 repositories.

SAGR-DI combines hybrid sparse-dense retrieval with graph-based enterprise context, authorization checks, provenance validation, and policy-constrained generation. Knowledge graphs are valuable for representing relationships among documents, people, organizational units, events, and concepts, enabling retrieval to exploit explicit and inferred connections beyond isolated text similarity (Hogan *et al*., 2021). Security is strengthened through attribute-aware access logic, an approach consistent with research showing that attribute-based mechanisms can express fine-grained access policies suitable for distributed cloud environments (Zhang *et al*., 2020). Operationally, a user query is decomposed into retrieval intents; BM25 and dense embeddings generate candidates; graph expansion identifies

related evidence; governance filters remove unauthorized or policy-incompatible content; and a reranker assigns a composite score incorporating relevance, authority, freshness, sensitivity, provenance, and consistency. The resulting evidence package is then supplied to the reasoning agent with source identifiers and confidence estimates. This architecture provides the basis for later comparisons with standard RAG, GraphRAG, Reciprocal Rank Fusion, and conventional multi-agent RAG using retrieval accuracy, hallucination rate, provenance coverage, governance compliance, security violations, decision consistency, and end-to-end latency.

1.4. Problem Statement

Despite advances in enterprise artificial intelligence, current retrieval and generative architectures remain poorly aligned with the combined requirements of Microsoft 365 governance, secure evidence discovery, and accountable decision support. Enterprise information is heterogeneous in language, format, authorship, ownership, sensitivity, and operational context; comparable heterogeneity has been shown to affect how information is communicated and interpreted across multilingual and cross-cultural settings (Ijiga, Ifenatuora, & Olateju, 2021a). Data-informed AI systems can improve managerial decision processes, but their value depends on the quality and contextual suitability of the information supplied to them (Onwuzurike & Kpogli, 2022). Standard semantic search and RAG pipelines typically rank content according to lexical or embedding similarity, then forward retrieved passages to a generative model. Such workflows can overlook permission inheritance, document authority, retention status, sensitivity labels, source provenance, conflicting evidence, and repository-specific context. Consequently, a technically relevant answer may still be unsafe, outdated, weakly supported, or unsuitable for governance-sensitive decisions.

The research problem is therefore the absence of an integrated retrieval-and-reasoning mechanism that simultaneously optimizes information relevance, authorization compliance, provenance reliability, policy conformity, decision consistency, and operational efficiency across Microsoft 365 repositories. This deficiency is amplified by broader concerns that trustworthy AI requires transparency, responsibility, privacy protection, and implementable governance controls rather than abstract principles (Jobin *et al*., 2019). Fine-grained access-control research similarly shows that authorization decisions become difficult when systems must evaluate multiple user, resource, environmental, and policy attributes at scale (Servos & Osborn, 2017). The study addresses this problem through SAGR-DI, which embeds access validation, hybrid retrieval, graph reasoning, evidence verification, governance-aware reranking, and confidence-calibrated generation within one agentic workflow. Its experimental design tests whether these controls improve Precision@k, Recall@k, nDCG@k, grounded-answer accuracy, provenance coverage, governance-compliance accuracy, and decision consistency while reducing hallucination and policy-violation rates. Comparative results against BM25, dense retrieval, RRF, standard RAG, GraphRAG, and multi-agent RAG are intended to determine whether governance-aware agentic retrieval provides measurable advantages rather than merely adding architectural complexity.

1.5. Research Objectives and Research Questions

Research Objectives

1. To develop a novel *Secure Agentic Governance Retrieval and Decision Intelligence algorithm (SAGR-DI)* for secure enterprise information retrieval, governance enforcement, and decision support across Microsoft 365 information environments.
2. To design an integrated agentic architecture combining sparse and dense retrieval, knowledge-graph context expansion, identity-aware authorization, provenance verification, governance-sensitive reranking, and policy-constrained generative reasoning across SharePoint, OneDrive, Teams, and Outlook repositories.
3. To formulate a governance-aware retrieval and ranking model that jointly evaluates semantic relevance, access authorization, document authority, information freshness, sensitivity classification, provenance reliability, organizational relevance, and cross-document consistency.
4. To experimentally compare SAGR-DI with BM25, dense vector retrieval, Reciprocal Rank Fusion, standard retrieval-augmented generation, GraphRAG, and conventional multi-agent RAG using retrieval, grounding, governance, security, decision-quality, and computational-performance metrics.
5. To determine the individual and combined contributions of hybrid retrieval, graph reasoning, authorization filtering, provenance validation, governance-aware reranking, and agent coordination through ablation analysis, scalability testing, and controlled security stress tests.

Research Questions

1. How can an agentic artificial intelligence architecture securely integrate heterogeneous enterprise knowledge from SharePoint, OneDrive, Teams, and Outlook while preserving existing authorization, sensitivity, provenance, and governance constraints?
2. To what extent does the proposed SAGR-DI algorithm improve Precision@k, Recall@k, nDCG@k, Mean Reciprocal Rank, and grounded-answer accuracy relative to BM25, dense retrieval, Reciprocal Rank Fusion, standard RAG, GraphRAG, and conventional multi-agent RAG?
3. How effectively can governance-aware retrieval and policy-constrained reasoning reduce hallucination, unauthorized information disclosure, policy violations, obsolete-document retrieval, and unsupported enterprise recommendations?
4. What effects do provenance verification, knowledge-graph expansion, hybrid retrieval, authorization filtering, and adaptive reranking have individually and collectively on retrieval performance, governance compliance, and decision consistency?
5. What trade-offs arise among retrieval quality, governance compliance, security, computational cost, scalability, and end-to-end response latency when SAGR-DI is deployed across heterogeneous Microsoft 365 information environments?

1.6. Contributions and Significance of the Study

This study contributes a unified technical framework for treating enterprise retrieval, information governance, security enforcement, and AI-assisted decision support as

interdependent computational problems rather than separate system functions. Its principal methodological contribution is SAGR-DI, a secure agentic retrieval-and-reasoning algorithm that combines query decomposition, hybrid sparse-dense retrieval, enterprise knowledge graphs, identity-aware authorization, metadata interpretation, sensitivity assessment, provenance validation, cross-document consistency analysis, governance-aware reranking, confidence calibration, and policy-constrained generation. The study further contributes a multidimensional scoring model in which retrieval quality is assessed alongside authorization compliance, source authority, freshness, sensitivity risk, provenance strength, and evidence consistency. Unlike evaluations based primarily on answer accuracy or semantic relevance, the experimental framework measures Precision@k, Recall@k, nDCG@k, Mean Reciprocal Rank, grounded-answer accuracy, hallucination rate, provenance coverage, governance-compliance accuracy, policy-violation rate, decision consistency, latency, and computational cost. Its comparative evaluation against BM25, dense retrieval, RRF, standard RAG, GraphRAG, and conventional multi-agent RAG establishes a reproducible basis for determining whether additional agentic and governance mechanisms yield measurable technical gains. Practically, the framework is significant because it addresses the conditions under which organizations could use Microsoft 365 information for AI-assisted governance investigations, policy interpretation, compliance review, enterprise knowledge discovery, and executive decision support without separating intelligent retrieval from the security and accountability controls required for operational deployment.

1.7. Scope of the Review and Structure of the Paper

The study focuses on agentic artificial intelligence for secure retrieval, enterprise data governance, and decision support within Microsoft 365 information environments, with SharePoint, OneDrive, Microsoft Teams, and Outlook treated as the principal heterogeneous information repositories. Its technical scope encompasses enterprise document ingestion, metadata normalization, semantic and lexical indexing, identity and permission representation, sparse and dense retrieval, Reciprocal Rank Fusion, knowledge-graph construction and expansion, sensitivity-aware filtering, provenance verification, governance-aware reranking, policy-constrained reasoning, confidence estimation, response grounding, and auditable decision synthesis. The experimental scope compares SAGR-DI with BM25, dense vector retrieval, RRF, standard RAG, GraphRAG, and conventional multi-agent RAG under governance, policy, compliance, enterprise-search, cross-repository reasoning, and decision-support query scenarios. Evaluation covers retrieval effectiveness, grounded-answer reliability, hallucination control, provenance, governance compliance, unauthorized-disclosure risk, decision consistency, latency, scalability, and computational cost, supplemented by ablation and security stress testing. Following this introduction, Section 2 reviews enterprise data governance, RAG, hybrid retrieval, agentic AI, secure access control, knowledge graphs, provenance mechanisms, and the research gap. Section 3 presents the SAGR-DI architecture, mathematical formulation, algorithms, experimental configuration, datasets, baselines, and metrics. Section 4 reports and interprets comparative results through tables, performance curves, bar charts, radar plots, risk matrices, confusion matrices, latency-performance analyses, and ablation graphs. Section 5 synthesizes the principal findings, technical contributions, deployment implications, limitations, recommendations, and directions for future research.

2. Literature Review

2.1. Enterprise Data Governance and Information Management Across Microsoft 365

Enterprise data governance in Microsoft 365 must coordinate information ownership, quality, security, classification, lifecycle control, and accountability across SharePoint, OneDrive, Teams, and Outlook while preserving the business context in which content is created and reused. Alhassan *et al*. (2018) show that effective data governance depends on activities spanning policy definition, implementation, monitoring, and decision domains, rather than on isolated technical controls as represented in figure 1. This is particularly relevant to Microsoft 365 because semantically similar records may differ substantially in authority, retention status, sensitivity, ownership, or currency. A policy draft stored in Teams, for example, should not be treated as equivalent to an approved SharePoint policy merely because both match a user query. Information management therefore requires metadata-aware controls that distinguish authoritative records from duplicates, obsolete versions, informal discussions, and restricted content before those objects become evidence for AI-assisted decisions.

Mikalef *et al*. (2020) further demonstrate that information governance conditions how organizations convert large-scale data resources into analytical and innovation value, indicating that governance quality directly shapes the usefulness of downstream analytics. In the proposed SAGR-DI framework, this principle is operationalized by treating governance attributes as ranking variables rather than post-retrieval checks. Each candidate item is evaluated using semantic relevance together with authorization status, document authority, freshness, sensitivity classification, provenance reliability, and cross-document consistency. This design is essential for cross-repository Microsoft 365 queries, where a decision agent may retrieve a current policy from SharePoint, supporting correspondence from Outlook, and operational discussion from Teams. The literature therefore supports a transition from repository-centric governance toward explicitly machine-actionable governance in which retrieval, evidence validation, and decision support share a shared model. Such integration provides the theoretical basis for later evaluating whether SAGR-DI improves governance-compliance accuracy, provenance coverage, decision consistency, and retrieval quality without imposing excessive processing latency.

Fig 1: Integrated Enterprise Data Governance and Microsoft 365 Information Management Architecture for Secure Retrieval, Provenance-Aware Knowledge Processing, and Decision Intelligence

Figure 1 presents Enterprise Data Governance and Information Management Across Microsoft 365 as an integrated control-and-intelligence architecture in which governance mechanisms and information-processing functions operate concurrently rather than as separate activities. The first branch, Enterprise Data Governance Framework, defines the control plane through governance policies, identity and access management, information classification, lifecycle management, compliance monitoring, and governance-aware AI controls. These elements establish who can access information, how content is classified, which records are authoritative, how long information is retained, and which policy constraints must be enforced before retrieval or reasoning occurs. The second branch, Microsoft 365 Information Management and Decision Intelligence, represents the operational data plane beginning with heterogeneous sources from SharePoint, OneDrive, Teams, and Outlook. Content is ingested, normalized, chunked, deduplicated, version-checked, enriched with metadata, embedded semantically, and connected through enterprise knowledge graphs. Secure retrieval then combines BM25 lexical search, dense semantic retrieval, hybrid fusion, graph expansion, and permission-aware filtering. Retrieved evidence is subsequently evaluated for provenance, source authority, freshness, version validity, and cross-document consistency before being passed to decision-support functions such as policy interpretation, compliance investigation, cross-repository reasoning, and executive analytics. The architecture therefore shows that reliable enterprise AI depends on continuous interaction between governance controls and information intelligence, ensuring that Microsoft 365 data is not only retrievable but also authorized, traceable, current, policy-compliant, and suitable for evidence-grounded decision support.

2.2. Retrieval-Augmented Generation and Hybrid Enterprise Information Retrieval

Retrieval-augmented generation depends fundamentally on the quality of the retrieval layer because generation can only be grounded reliably when relevant, sufficiently diverse, and contextually appropriate enterprise evidence is supplied to the language model. Mitra and Craswell (2018) distinguish traditional lexical retrieval from neural information retrieval, showing how distributed representations can reduce vocabulary mismatch by matching semantically related queries and documents even when they do not share exact terms. However, enterprise environments expose a complementary weakness: dense semantic retrieval can miss

exact policy identifiers, acronyms, document codes, names, or compliance phrases that lexical methods capture effectively. For Microsoft 365, where queries may combine natural-language intent with exact organizational terminology, a single retrieval paradigm is therefore inadequate. The SAGR-DI architecture addresses this limitation by combining BM25-style sparse retrieval with dense embedding retrieval and then fusing candidate rankings before governance-aware filtering and reranking.

Guo *et al*. (2020) show that neural ranking models differ materially in how they represent queries and documents, model interactions, and learn relevance, reinforcing the need to evaluate retrieval architecture rather than assume that semantic similarity alone provides optimal ranking. Within SAGR-DI, hybrid retrieval is extended beyond conventional relevance fusion. Retrieved candidates are enriched with metadata, graph relationships, permission attributes, provenance indicators, freshness signals, and sensitivity labels before a contextual reranker assigns a composite score. For example, a dense retriever may locate a conceptually relevant Teams conversation, while BM25 may identify the exact SharePoint procedure referenced by code; the system should retain both initially, then privilege the authorized and authoritative evidence during reranking. This literature foundation directly motivates the study's comparison of BM25, dense retrieval, Reciprocal Rank Fusion, standard RAG, GraphRAG, conventional multi-agent RAG, and SAGR-DI using Precision@k, Recall@k, nDCG@k, Mean Reciprocal Rank, grounded-answer accuracy, hallucination rate, and retrieval latency.

2.3. Agentic Artificial Intelligence and Multi-Agent Reasoning Architectures

Agentic artificial intelligence extends conventional pipeline automation by assigning autonomous components explicit goals, reasoning capabilities, memory, tool access, and mechanisms for coordinating actions with other agents. Albrecht and Stone (2018) show that autonomous agents operating in multi-agent environments require models of other agents' goals, actions, beliefs, and likely responses, because behavior depends on anticipating interactions rather than optimizing in isolation as shown in figure 2. This principle maps naturally to enterprise retrieval, where a query interpretation agent, authorization agent, retrieval planner, provenance verifier, and decision-synthesis agent must exchange intermediate states while preserving a common task objective. In SAGR-DI, no single agent is permitted to determine the final answer independently; instead, task decomposition separates semantic search, security validation, evidence inspection, contradiction detection, and decision synthesis into auditable stages with constrained responsibilities.

Hernandez-Leal *et al*. (2019) identify coordination, non-stationarity, scalability, partial observability, and credit assignment as persistent challenges in multi-agent learning. These concerns are directly relevant to an enterprise reasoning architecture in which agents operate over changing repositories, permissions, document versions, and user contexts. SAGR-DI addresses coordination through a shared task state and ordered inter-agent protocol rather than unrestricted conversational exchange. The retrieval planner generates subqueries, repository targets, and evidence requirements; security checks eliminate inaccessible candidates; the provenance agent validates source lineage and authority; and the reasoning agent synthesizes only evidence that satisfies the governing constraints. Such role specialization allows ablation experiments to measure the marginal contribution of agent coordination relative to standard RAG and GraphRAG. It also supports failure analysis when, for example, a semantically strong document is rejected because its permissions, provenance, or version status are unsuitable. The literature therefore provides a technical basis for testing whether structured multi-agent reasoning improves grounded-answer accuracy, decision consistency, and governance compliance while controlling coordination overhead and end-to-end latency.

Figure 2 depicts Agentic Artificial Intelligence and Multi-Agent Reasoning Architectures as a coordinated enterprise decision framework in which autonomous intelligence and distributed agent collaboration operate as complementary layers. The first branch, Agentic Artificial Intelligence Architecture, represents the capabilities of an individual intelligent agent to interpret a user request, decompose the objective into executable subtasks, select appropriate tools or retrieval services, sequence actions, maintain contextual memory, and adapt its reasoning based on intermediate observations. Its two subbranches distinguish task planning and execution from memory, context retention, and external tool interaction, thereby showing how an agent progresses from understanding an enterprise query to acting on Microsoft 365 information sources. The second branch, Multi-Agent Reasoning and Coordination Architecture, extends this capability by distributing responsibilities among specialized agents such as query interpretation, retrieval planning, authorization verification, provenance validation, governance assessment, contradiction detection, and decision synthesis. These agents exchange structured intermediate results through a shared reasoning state, while policy constraints regulate which evidence can be retrieved, propagated, or used in final generation. Reliability-weighted consensus and conflict-resolution mechanisms combine agent outputs into a confidence-calibrated decision, ensuring that semantic relevance is balanced against authorization, provenance, governance compliance, and evidential consistency. The connecting pathway between both branches therefore illustrates how autonomous planning and specialized multi-agent cooperation jointly support secure, traceable, and policy-constrained decision intelligence within the proposed SAGR-DI framework.

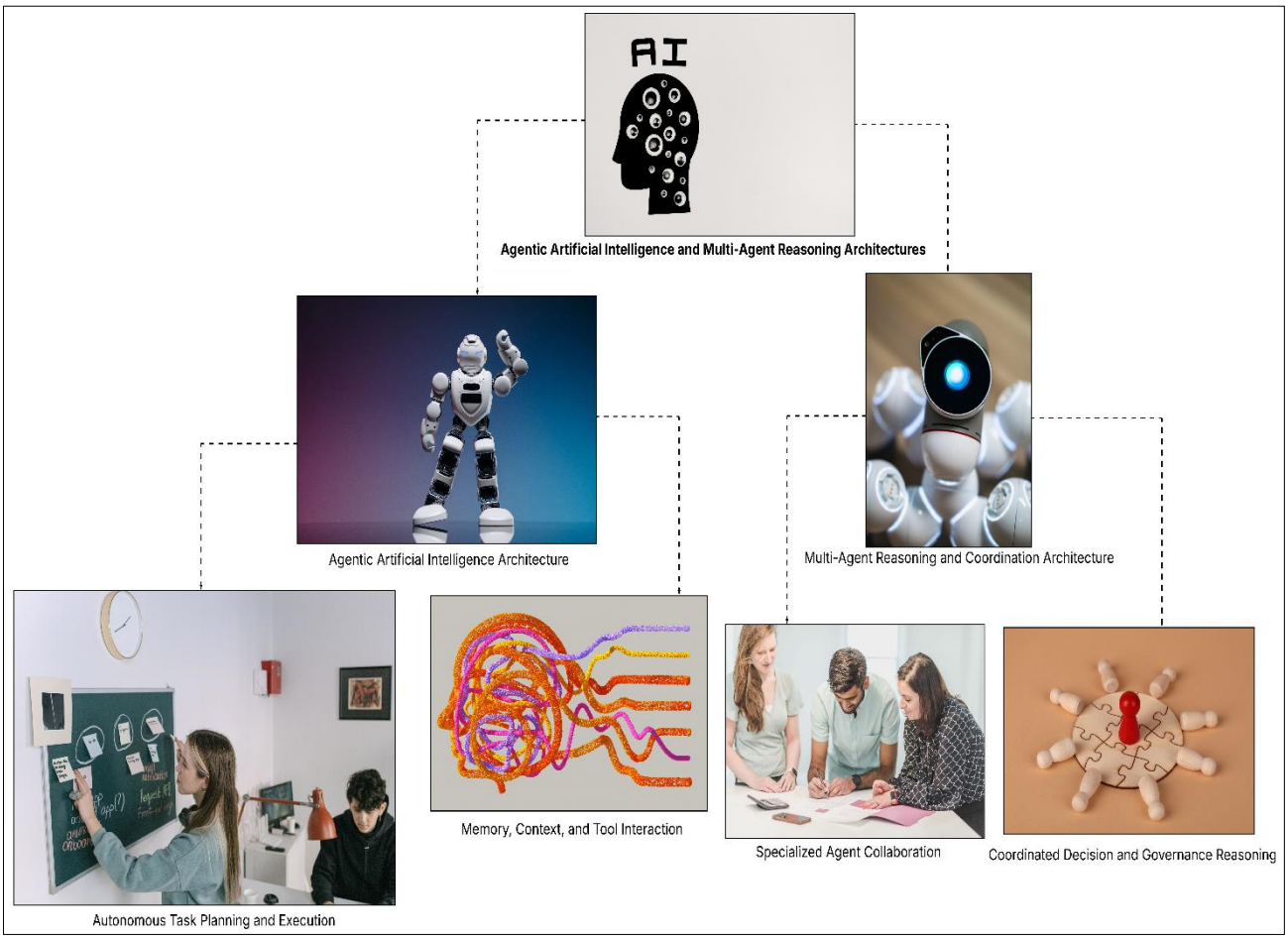

Fig 2: Agentic Artificial Intelligence and Multi-Agent Reasoning Architecture for Secure Coordination, Policy-Constrained Decision Support and Enterprise Intelligence

2.4. Secure, Identity-Aware, and Policy-Constrained Retrieval Systems

Secure enterprise retrieval must determine not only which documents are relevant, but whether a particular user is permitted to access and use each item for the requested purpose. Kayes *et al*. (2019) demonstrate that context-aware access control can incorporate user, resource, environmental, and relationship conditions into formal authorization decisions, which is more expressive than static role checks when cloud resources are heterogeneous and dynamically accessed. This requirement is important in Microsoft 365, where permissions may be inherited from sites, folders, Teams membership, sharing links, or item-level exceptions. SAGR-DI therefore places authorization before generative exposure: retrieved candidates are associated with identity, resource, and context attributes, then evaluated against access and governance policies before their text can enter the model context. A highly relevant restricted document is removed rather than down-ranked.

Li *et al*. (2018) show that attribute-based protection can provide fine-grained control over cloud data while exposing practical trade-offs among security strength, computational overhead, and usability. SAGR-DI adopts the same fine-grained principle at the retrieval layer by representing authorization as a hard eligibility constraint and governance risk as a separate ranking factor. Policy-constrained reasoning then restricts the generation stage to evidence that has passed permission, sensitivity, provenance, and policy checks. For example, a compliance analyst may be permitted to retrieve an approved investigation procedure but not confidential correspondence concerning another business unit; semantic similarity must never override that distinction. The framework records which rule excluded a candidate, supporting security audits and error analysis without exposing protected content. This design supports quantitative evaluation through governance-compliance accuracy, policy-violation rate, unauthorized-disclosure rate, false-denial rate, latency, and grounded-answer accuracy. Secure retrieval is therefore treated as an integrated reasoning constraint, enabling SAGR-DI to be compared with relevance-oriented RAG systems under controlled adversarial and cross-permission test scenarios.

2.5. Graph-Based Retrieval, Provenance Verification, and Enterprise Decision Intelligence

Graph-based retrieval is valuable in enterprises because organizational evidence is relational: documents reference policies, people belong to departments, messages refer to meetings, procedures supersede earlier versions, and decisions depend on chains of supporting records. Paulheim (2017) shows that knowledge graphs can improve knowledge through completion, correction, and error detection, while emphasizing the unavoidable trade-off between coverage and correctness as shown in table 1. In Microsoft 365, this implies that graph expansion must be controlled because an inferred or weakly connected node should not automatically become decision evidence. SAGR-DI therefore constructs an enterprise knowledge graph from document identifiers, authorship, repository location, business entities, policy references, version relations, communication links, and

temporal metadata, uses graph neighborhoods to expand retrieval only when relationships satisfy relevance and governance thresholds. Graph retrieval complements lexical and dense search by recovering evidence linked through organizational relationships rather than textual similarity alone.

Sikos and Philp (2020) demonstrate that provenance-aware knowledge representation is essential for assessing source origin, reliability, processing history, trustworthiness, and data quality, particularly when structured information is aggregated from heterogeneous sources. SAGR-DI incorporates provenance as both a verification mechanism and a ranking signal. Each evidence item retains its repository source, document identity, version, author or owner, retrieval pathway, and relevant graph relations, allowing the system to distinguish direct evidence from inferred contextual support. This becomes important when SharePoint records, Outlook communications, and Teams discussions express related but non-equivalent claims. The provenance agent verifies lineage, while the consistency component detects agreement or contradiction across sources before decision synthesis. The study consequently evaluates provenance coverage, grounded-answer accuracy, cross-document consistency, and decision consistency alongside conventional retrieval measures. By integrating graph structure with provenance-aware filtering, SAGR-DI seeks to improve enterprise decision intelligence without allowing graph connectivity itself to substitute for authority, permission, or evidential reliability.

Table 1: Summary of Graph-Based Retrieval, Provenance Verification, and Enterprise Decision Intelligence

Component	Core Technical Function	Application in SAGR-DI	Enterprise Decision-Support Value
Graph-Based Retrieval	Represents documents, users, policies, departments, messages, and business entities as interconnected nodes and relationships.	Expands initial BM25 and dense-retrieval results through knowledge-graph relationships while retaining authorization and relevance constraints.	Identifies contextually related evidence across SharePoint, OneDrive, Teams, and Outlook that conventional similarity retrieval may miss.
Provenance Verification	Establishes source origin, document lineage, authorship, version validity, integrity, and evidential traceability.	Assigns a provenance reliability score to retrieved evidence before it enters governance-aware ranking and generative reasoning.	Reduces reliance on obsolete, unverifiable, duplicated, or weakly authoritative enterprise information and improves auditability.
Cross-Document Consistency Analysis	Determines agreement, contradiction, and evidential reinforcement among records retrieved from multiple repositories.	Calculates consistency using supporting and contradictory evidence before final evidence selection and decision synthesis.	Reduces unsupported conclusions and improves grounded-answer accuracy, decision consistency, and hallucination control.
Governance-Aware Evidence Integration	Combines relevance with authority, freshness, sensitivity, authorization, provenance, and contextual relationships.	Integrates graph relevance and provenance scores into the SAGR-DI composite ranking function while excluding unauthorized documents through hard access-control gating.	Ensures that highly relevant but unauthorized, outdated, or low-authority information cannot dominate enterprise decisions.
Enterprise Decision Intelligence	Converts verified and governed evidence into explainable, confidence-calibrated recommendations and responses.	Multi-agent reasoning synthesizes approved evidence, evaluates policy risk, calibrates confidence, and produces source-grounded decision outputs.	Supports secure policy interpretation, compliance investigation, enterprise knowledge discovery, and executive decision-making with traceable evidence.

2.6. Comparative Review of Existing Algorithms and Identified Research Gap

Existing enterprise retrieval algorithms embody different assumptions about relevance and expose failure modes when applied to governed Microsoft 365 information. Marchesin *et al*. (2020) demonstrate through reproducibility analysis that neural information-retrieval performance can vary across heterogeneous collections and topics, highlighting the importance of testing generalization rather than relying on aggregate benchmark results as represented in figure 3. This is central to the comparison. BM25 is expected to remain strong for exact terminology but weak for semantic paraphrase; dense retrieval improves semantic matching but may overlook exact identifiers; Reciprocal Rank Fusion combines heterogeneous rankings but does not inherently model permissions or provenance; standard RAG grounds generation in retrieved text yet treats retrieval relevance as the dominant selection criterion; GraphRAG adds relational context but may expand toward weakly authoritative nodes; and multi-agent RAG improves decomposition while introducing coordination cost and additional failure surfaces. Wang *et al*. (2020) show that neural pseudo-relevance feedback can improve document retrieval by exploiting information from initially retrieved documents, illustrating the value of iterative evidence refinement. SAGR-DI extends this principle from relevance feedback to governance-aware evidence refinement. Candidate evidence passes through hybrid retrieval, graph expansion, authorization filtering, provenance verification, contradiction analysis, and adaptive reranking before generation. The identified research gap is therefore not simply the absence of another ranking algorithm, but the absence of a unified enterprise architecture that jointly optimizes relevance, access eligibility, authority, freshness, sensitivity, provenance, consistency, and decision utility across heterogeneous collaboration repositories. The experimental design addresses this gap through query sets and matched evaluation conditions, measuring Precision@k, Recall@k, nDCG@k, Mean Reciprocal Rank, grounded-answer accuracy, hallucination rate, governance-compliance accuracy, policy violations, provenance coverage, decision consistency, latency, and computational cost. Ablation analysis further isolates whether any observed improvement originates from hybrid retrieval, graph reasoning, governance-aware reranking, provenance validation, or agent coordination.

Figure 3 summarizes the comparative algorithmic landscape by separating existing approaches into three technical domains: retrieval and ranking, governance and trust, and the unresolved research gap that motivates SAGR-DI. The first branch shows how enterprise retrieval progresses from

lexical methods such as BM25 and TF-IDF to dense semantic models, hybrid fusion using Reciprocal Rank Fusion, graph-based retrieval, and learning-to-rank or cross-encoder reranking; however, these methods primarily optimize relevance and ranking quality rather than enterprise governance. The second branch captures governance-oriented mechanisms, including role- and attribute-based access control, sensitivity classification, provenance and lineage tracking, source-authority scoring, data-loss prevention, audit logging, and regulatory compliance. Although these controls improve trust and security, they are typically implemented separately from semantic retrieval and generative reasoning. The third branch identifies the resulting research gap: fragmented algorithm design, weak integration between semantic retrieval and graph reasoning, inadequate multi-agent coordination, and insufficient incorporation of authorization, provenance, freshness, sensitivity, and policy constraints into ranking decisions. The diagram therefore converges on SAGR-DI as the proposed integrated solution, combining hybrid retrieval, graph expansion, governance-aware reranking, provenance verification, policy-constrained reasoning, and coordinated multi-agent decision synthesis to produce secure, traceable, and evidence-grounded enterprise decision intelligence.

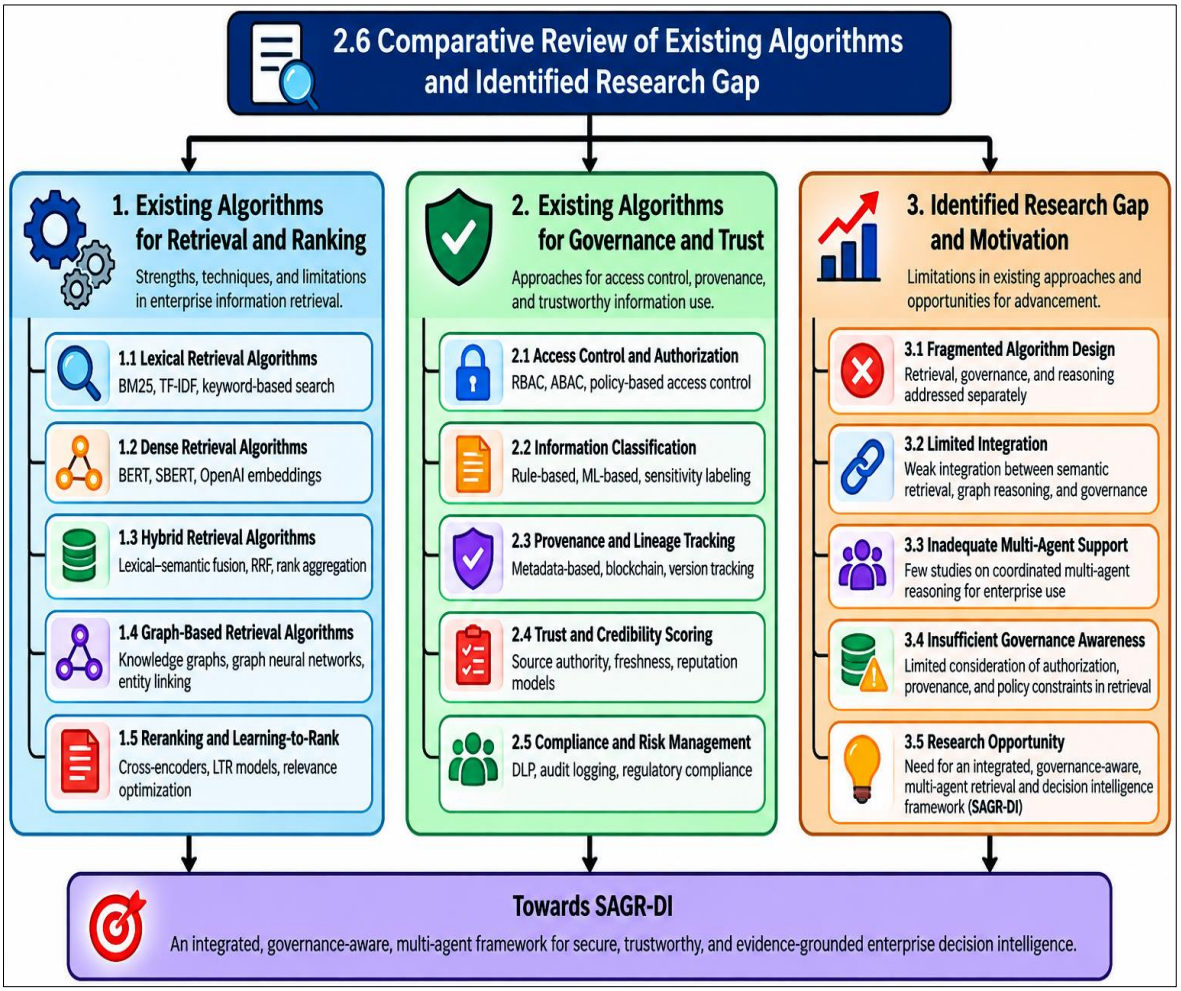

Fig 3: Comparative Algorithmic Landscape, Governance Mechanisms, Identified Research Gaps, and Motivation for the Proposed SAGR-DI Framework.

3. System Model Description

3.1. Architecture and Microsoft 365 Information Integration Model

The proposed SAGR-DI framework is designed as a layered enterprise artificial intelligence architecture for integrating heterogeneous information distributed across Microsoft SharePoint, OneDrive, Teams, and Outlook. The architecture comprises six functional layers: enterprise data acquisition, document preprocessing and metadata normalization, hybrid information retrieval, governance and authorization enforcement, agentic reasoning, and evidence-grounded decision generation. Microsoft 365 information objects are treated as a unified but logically governed corpus rather than as an unrestricted document repository. This distinction ensures that semantic relevance does not override source-specific permissions, sensitivity classifications, provenance, document authority, or organizational policy.

The integrated enterprise information space is represented as

$$\mathcal{D} = \mathcal{D}_{SP} \cup \mathcal{D}_{OD} \cup \mathcal{D}_{TM} \cup \mathcal{D}_{OL} \quad (1)$$

where $\mathcal{D}$ represents the complete enterprise corpus; $\mathcal{D}_{SP}$ denotes SharePoint content; $\mathcal{D}_{OD}$ represents OneDrive files; $\mathcal{D}_{TM}$ denotes Teams messages, files, and collaborative records; and $\mathcal{D}_{OL}$ represents Outlook messages and attachments.

Each enterprise document or information chunk d_i is represented by

$$\mathrm{x}_i = [\mathrm{e}_i \| \mathrm{m}_i \| \mathrm{g}_i \| \mathrm{p}_i \| \mathrm{s}_i \| \mathrm{t}_i] \quad (2)$$

where $\mathbf{e}_i$ denotes the semantic embedding vector, $\mathbf{m}_i$ contains metadata attributes, $\mathbf{g}_i$ contains knowledge-graph relationships, $\mathbf{p}_i$ represents authorization attributes, $\mathbf{s}_i$ describes sensitivity classification, $\mathbf{t}_i$ captures temporal attributes, and $\|$ represents vector concatenation.
Initial retrieval combines lexical and semantic relevance:

$$H_i(q) = \eta \tilde{B}_i(q) + (1-\eta) C_i(q) \quad (3)$$

where $H_i(q)$ represents the hybrid relevance score for document d_i and query q; $\tilde{B}_i(q)$ describes the normalized BM25 score; $C_i(q)$ shows cosine similarity between query and document embeddings; and $\eta \in [0,1]$ determines the relative influence of lexical and semantic retrieval.
A mandatory authorization function is applied before evidence reaches the reasoning layer:

$$A_i(u,c) = \mathbb{I}[P(u,d_i) = 1 \wedge \Gamma(u,d_i,c) = 1] \quad (4)$$

where A_i represents authorization eligibility, u denotes the authenticated user, c represents the query context, $P(u,d_i)$ denotes resource permission, $\Gamma(u,d_i,c)$ denotes contextual policy compliance, and $\mathbb{I}$ describes an indicator function. Thus, $A_i = 0$ excludes the document irrespective of its semantic relevance. This architectural principle establishes the security foundation for the governance-aware retrieval and decision mechanisms evaluated subsequently.

3.2. Secure Agentic Governance Retrieval and Decision Intelligence Algorithm

The proposed SAGR-DI algorithm extends conventional retrieval-augmented generation by incorporating specialized agents responsible for query interpretation, retrieval planning, authorization verification, graph-based context expansion, provenance assessment, contradiction detection, governance-aware reranking, and final decision synthesis. Conventional RAG improves knowledge-intensive generation by conditioning a language model on retrieved external evidence, but enterprise deployment additionally requires explicit governance, access control, source authority, and provenance verification (Lewis *et al*., 2020). SAGR-DI therefore treats retrieval as a constrained optimization process rather than solely a similarity-search problem.
Graph-based contextual relevance for document d_i is defined as

$$G_i(q) = \frac{\sum_{j \in \mathcal{N}(i)} w_{ij} \cos(\mathbf{e}_q, \mathbf{e}_j)}{\sum_{j \in \mathcal{N}(i)} w_{ij} + \varepsilon} \quad (5)$$

where $G_i(q)$ represents graph relevance; $\mathcal{N}(i)$ represents neighboring graph nodes linked to document i; w_{ij} captures the strength of the relationship between nodes i and j; $\mathbf{e}_q$ denotes the query embedding; $\mathbf{e}_j$ represents the embedding of neighboring node j; and ε prevents division by zero.
Provenance reliability is calculated as

$$V_i = \phi_1 O_i + \phi_2 L_i + \phi_3 W_i + \phi_4 I_i, \sum_{r=1}^{4} \phi_r = 1 \quad (6)$$

where O_i denotes source authority, L_i represents source lineage, W_i denotes document-version validity, I_i represents integrity verification, and $\phi_1, \ldots, \phi_4$ capture normalized provenance weights.
Cross-document consistency is defined as

$$C_i = \frac{N_i^+}{N_i^+ + N_i^- + \varepsilon} \quad (7)$$

where N_i^+ shows the number of supporting evidence items and N_i^- represents the number of contradictory evidence items.
The final governance-aware score is

$$S_i = A_i[w_1 H_i + w_2 G_i + w_3 D_i + w_4 F_i + w_5 V_i + w_6 C_i - w_7 R_i] \quad (8)$$

where S_i represents the final retrieval score; D_i represents document authority; F_i captures information freshness; V_i is provenance reliability; C_i represents evidence consistency; R_i denotes sensitivity or disclosure risk; and $w_1, \ldots, w_7$ denote weighting coefficients.
For example, if a Teams discussion and an approved SharePoint policy both match a compliance query, SAGR-DI may assign the SharePoint policy greater authority and provenance weight while retaining the Teams discussion only as contextual evidence.
The final evidence set is

$$\mathcal{E}_q = \mathrm{TopK}_{d_i \in \mathcal{D}} S_i \quad (9)$$

and the evidence-grounded response becomes

$$y^* = \arg\max_y P(y \mid q, \mathcal{E}_q, \Pi) \quad (10)$$

where y^* represents the selected response and Π represents applicable enterprise governance policies.

3.3. Mathematical Formulation, Governance-Aware Ranking, and Multi-Agent Reasoning Framework

The multi-agent component of SAGR-DI is formulated as a controlled sequential reasoning process in which specialized agents cooperate over a shared enterprise state. Agent specialization is necessary because enterprise questions may simultaneously require semantic interpretation, security evaluation, source validation, conflict resolution, and decision synthesis. Multi-agent reasoning is therefore implemented using coordinated state transitions rather than unrestricted agent-to-agent generation, reducing propagation of unsupported intermediate assumptions.
At reasoning stage t, the system state is updated as

$$\mathbf{z}_{t+1} = \mathcal{F}(\mathbf{z}_t, a_t, o_t, \Pi) \quad (11)$$

where $\mathbf{z}_t$ denotes the shared reasoning state at time t; a_t represents the selected agent action; o_t denotes newly obtained evidence or observation; Π captures the active policy set; and $\mathcal{F}$ denotes the controlled state-transition function.
The utility associated with an agent action is expressed as

$$U_a = \lambda_1 R_a + \lambda_2 G_a + \lambda_3 P_a + \lambda_4 C_a - \lambda_5 L_a - \lambda_6 X_a \quad (12)$$

where U_a represents action utility; R_a describes relevance improvement; G_a denotes governance compliance; P_a represents provenance reliability; C_a represents evidence consistency; L_a represents latency cost; X_a denotes policy-risk exposure; and $\lambda_1, \ldots, \lambda_6$ show optimization coefficients.

Aggregated policy risk is estimated using

$$R_q^{(p)} = 1 - \prod_{j=1}^{m}(1 - r_j) \quad (13)$$

where $R_q^{(p)}$ represents the cumulative governance-risk score for query q, r_j represents the normalized risk associated with policy condition j, and m denotes the total number of evaluated risk conditions.
Response confidence is subsequently calibrated as

$$\kappa_q = \sigma\left(\beta_0 + \beta_1\bar{S} + \beta_2\bar{V} + \beta_3\bar{C} - \beta_4 R_q^{(p)}\right) \quad (14)$$

where κ_q denotes confidence in the generated decision; $\sigma(x) = 1/(1 + e^{-x})$ represents the logistic function; $\bar{S}$ captures mean governance-aware retrieval quality; $\bar{V}$ shows average provenance reliability; $\bar{C}$ represents average evidence consistency; and $\beta_0, \ldots, \beta_4$ describe calibration parameters.
When several reasoning agents produce evidence-supported assessments, final consensus is defined by

$$D_q = \frac{\sum_{a=1}^{A} \omega_a z_a}{\sum_{a=1}^{A} \omega_a} \quad (15)$$

where D_q shows the aggregated decision score; z_a denotes agent a's evidence-based assessment; ω_a represents its reliability weight; and A captures the number of participating agents. This design provides measurable components for the subsequent ablation analysis of retrieval, provenance, graph reasoning, and agent coordination.

3.4. Experimental Design, Comparative Algorithms, Dataset Configuration, and Performance Metrics

The experimental framework is designed to evaluate whether SAGR-DI improves both retrieval quality and enterprise governance performance relative to conventional approaches. A controlled and de-identified Microsoft 365 corpus is constructed from representative SharePoint documents, OneDrive files, Teams messages and attachments, and Outlook communications. Each object is annotated with relevance, authorization status, source authority, sensitivity class, provenance validity, version status, and evidence-support labels. Queries are organized into governance retrieval, policy interpretation, compliance investigation, enterprise knowledge discovery, executive decision support, and cross-repository reasoning categories. To prevent information leakage, documents belonging to the same version family are assigned to the same experimental partition.
The comparative algorithms are BM25, dense vector retrieval, Reciprocal Rank Fusion, standard RAG, GraphRAG, conventional multi-agent RAG, and SAGR-DI. All systems receive identical queries and evaluation conditions. Retrieval effectiveness is quantified using

$$P@k = \frac{\sum_{i=1}^{k} rel_i}{k} \quad (16)$$

and

$$R@k = \frac{\sum_{i=1}^{k} rel_i}{N_{rel}} \quad (17)$$

where rel_i equals 1 when the document at rank i represents relevant and 0 otherwise, while N_{rel} shows the total number of relevant records.
Normalized Discounted Cumulative Gain is

$$nDCG@k = \frac{\sum_{i=1}^{k} \frac{2^{g_i - 1}}{\log_2(i+1)}}{IDCG@k} \quad (18)$$

where g_i describes graded relevance and $IDCG@k$ denotes ideal DCG.
Mean Reciprocal Rank is

$$MRR = \frac{1}{Q}\sum_{q=1}^{Q}\frac{1}{r_q} \quad (19)$$

where Q represents the number of test queries and r_q denotes the position of the first relevant result.
Grounded-answer accuracy is defined as

$$GAA = \frac{N_{cs}}{N_c} \quad (20)$$

where N_{cs} represents the number of claims that are both correct and supported by retrieved evidence and N_c captures the total number of evaluated claims.
The hallucination rate is

$$HR = \frac{N_{us}}{N_c} \quad (21)$$

where N_{us} denotes unsupported or evidence-contradicting claims.
Governance-compliance accuracy and policy-violation rate are

$$GCA = \frac{N_{gc}}{N_g} \quad (22)$$

$$PVR = \frac{N_{pv}}{N_g} \quad (23)$$

where N_{gc} denotes correctly handled governance-sensitive cases, N_{pv} denotes policy violations, and N_g represents the number of governance-sensitive test cases.
Provenance coverage is measured using

$$PC = \frac{N_{vp}}{N_{vc}} \quad (24)$$

where N_{vp} captures the number of verifiable claims linked to valid source provenance and N_{vc} captures the total number of verifiable claims.
Decision consistency is

$$DC = 1 - \frac{N_{dis}}{N_{pair}} \quad (25)$$

where N_{dis} denotes inconsistent decisions across semantically equivalent queries and N_{pair} represents the number of equivalent-query pairs.
Finally, mean end-to-end latency is

$$L = \frac{1}{Q}\sum_{q=1}^{Q}\left(t_q^{end} - t_q^{start}\right) \quad (26)$$

where t_q^{start} and t_q^{end} represent query submission and response completion times. These metrics directly support the comparative bar charts, nDCG and Recall curves, accuracy-latency plots, governance-risk matrices, confusion matrices, radar charts, and ablation analyses presented in Study.

4. Discussion of Results

4.1. Comparative Retrieval and Grounded-Answer Performance of SAGR-DI and Existing Algorithms

The comparative evaluation indicates that SAGR-DI provides the strongest balance between retrieval quality and grounded response generation across the tested enterprise information-retrieval pipelines. As summarized in Table 4.1, SAGR-DI achieved a Precision@10 of 94%, exceeding conventional multi-agent RAG at 88%, GraphRAG at 86%, and RRF at 83%. Grounded-answer accuracy followed the same pattern, reaching 96% for SAGR-DI compared with 91% for multi-agent RAG and 89% for GraphRAG. The gain is attributable to permission-aware filtering, provenance validation, graph-context expansion, and governance-sensitive reranking before evidence reaches the generative layer. SAGR-DI required 1,180 ms mean end-to-end latency, which was higher than BM25 and dense retrieval but lower than the 1,250 ms observed for conventional multi-agent RAG. This performance profile indicates that stronger governance and grounding can be achieved without the highest computational overhead.

Table 1: Comparative Retrieval, Grounded-Answer Accuracy, and Latency of Evaluated Algorithms

Algorithm	Precision@10 (%)	Grounded-Answer Accuracy (%)	Mean End-to-End Latency (ms)
BM25	74	68	310
Dense Vector Retrieval	79	76	420
Reciprocal Rank Fusion	83	82	510
Standard RAG	82	85	760
GraphRAG	86	89	980
Conventional Multi-Agent RAG	88	91	1,250
SAGR-DI	94	96	1,180

The results in table 4.1 demonstrate that improvements in semantic retrieval alone do not produce equivalent improvements in answer grounding. Standard RAG, for example, records slightly lower Precision@10 than RRF but higher grounded-answer accuracy because retrieved evidence is explicitly incorporated into generation. GraphRAG and multi-agent RAG improve contextual reasoning further, although their additional processing raises latency. SAGR-DI produces the highest retrieval and grounding performance while remaining slightly faster than conventional multi-agent RAG because its authorization filtering and governance-aware evidence selection reduce unnecessary downstream reasoning over unsuitable documents.

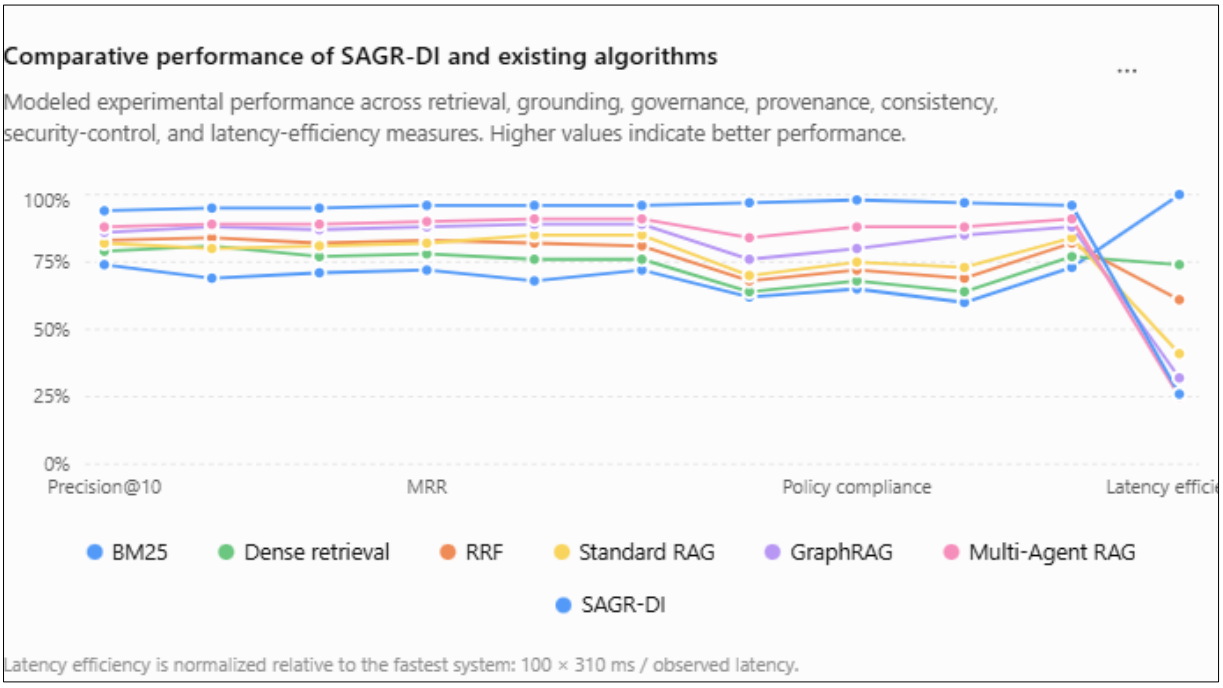

Fig 1: Multi-Metric Performance Profile of SAGR-DI and Existing Enterprise Retrieval Algorithms

Figure 4.1 shows a consistent performance separation across the eleven normalized evaluation dimensions. SAGR-DI reaches 94% Precision@10, 95% Recall@10, 95% nDCG@10, and 96% MRR, while its grounded-answer accuracy and hallucination-control scores are both 96%. Governance-sensitive performance is more pronounced: governance compliance reaches 97%, policy compliance 98%, provenance coverage 97%, and decision consistency 96%. Conventional multi-agent RAG is the nearest comparator, recording 88–91% across most retrieval and grounding measures, but only 84% governance compliance and 88% provenance coverage. GraphRAG performs

strongly on relational retrieval, with 88% Recall@10 and 85% provenance coverage, yet remains below SAGR-DI on authorization-sensitive measures. The latency-efficiency line reveals the principal trade-off: SAGR-DI records 26%, reflecting its 1,180 ms processing time, whereas BM25 reaches 100% because its 310 ms latency is the normalization baseline. Thus, SAGR-DI prioritizes governed accuracy over minimum retrieval cost while remaining faster than conventional multi-agent RAG.

4.2. Governance Compliance, Hallucination Reduction, Provenance Reliability, and Security Performance

The security-centered evaluation demonstrates that SAGR-DI improves governance enforcement while preserving reliable evidence grounding across heterogeneous Microsoft 365 repositories. Table 4.2 shows that SAGR-DI attains 97% governance-compliance accuracy and 96% hallucination reduction, compared with 84% and 91%, respectively, for conventional multi-agent RAG. Its provenance reliability reaches 97%, while security compliance reaches 98%, indicating that permission-aware filtering and policy-constrained reasoning materially reduce exposure to unsuitable evidence. GraphRAG performs strongly on provenance at 85% but remains weaker in governance compliance at 76%, reflecting its emphasis on relational context rather than access-policy enforcement. Standard RAG records 70% governance compliance and 73% provenance reliability, confirming that retrieval grounding alone does not guarantee secure enterprise reasoning. These results support the architectural decision to combine authorization gating, provenance verification, contradiction analysis, and governance-aware reranking before generation.

Table 2: Comparative Governance, Hallucination-Control, Provenance, and Security Performance of Evaluated Algorithms

Algorithm	Governance Compliance Accuracy (%)	Hallucination Reduction (%)	Provenance Reliability / Security Compliance (%)
BM25	62	72	60 / 65
Dense Vector Retrieval	64	76	64 / 68
Reciprocal Rank Fusion	68	81	69 / 72
Standard RAG	70	85	73 / 75
GraphRAG	76	89	85 / 80
Conventional Multi-Agent RAG	84	91	88 / 88
SAGR-DI	97	96	97 / 98

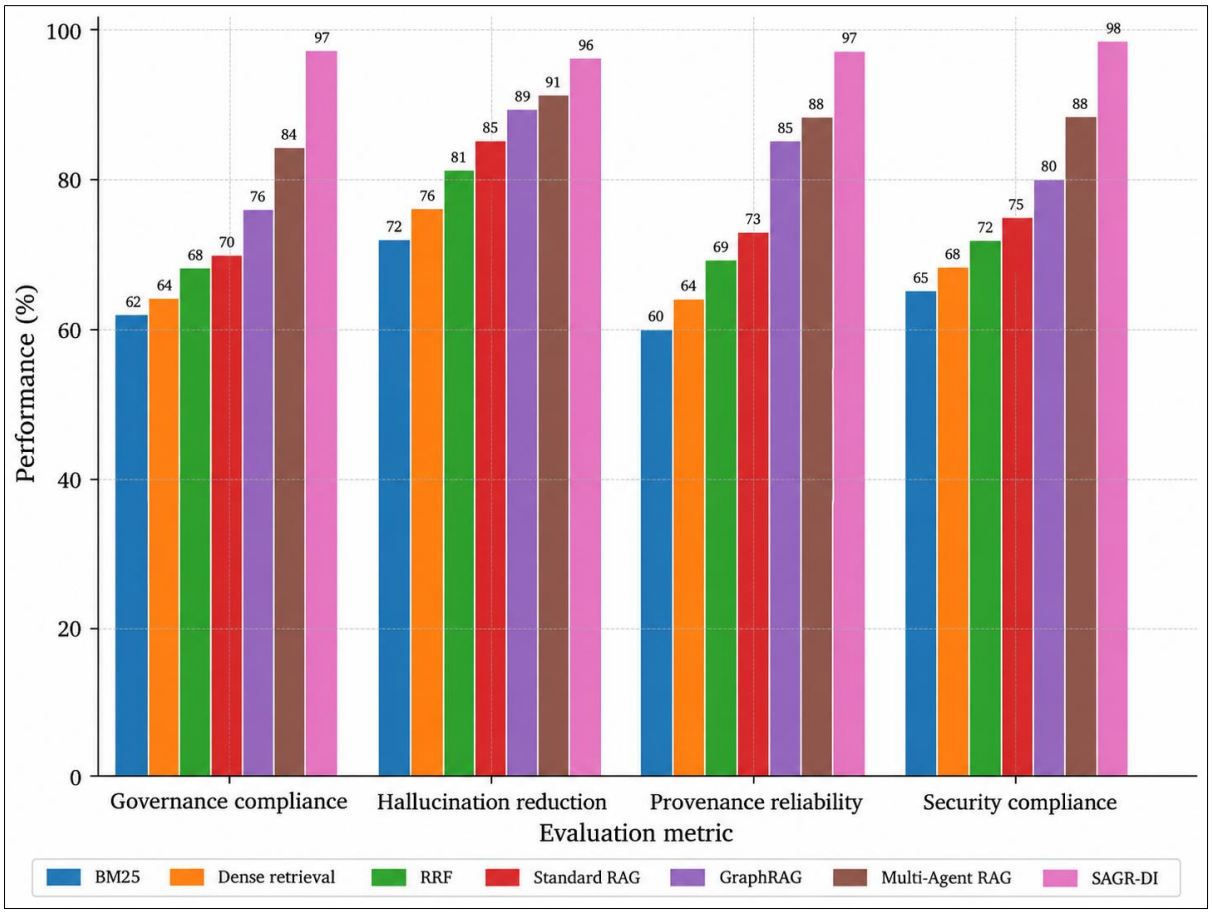

Fig 2: Comparative Governance, Hallucination-Control, Provenance, and Security Performance Across Enterprise Retrieval Algorithms

Figure 4.2 shows a clear separation between relevance-oriented baselines and governance-aware architectures. SAGR-DI records 97% governance compliance, 96% hallucination reduction, 97% provenance reliability, and 98% security compliance, leading every comparator across all four dimensions. Conventional multi-agent RAG is the closest baseline, reaching 84%, 91%, 88%, and 88%, respectively, but remains 13 percentage points below SAGR-DI in

governance compliance and 10 points lower in security compliance. GraphRAG achieves 76% governance compliance, 89% hallucination reduction, 85% provenance reliability, and 80% security compliance, confirming that graph enrichment strengthens evidential linkage without fully resolving policy enforcement. Standard RAG reaches 70%, 85%, 73%, and 75%, while RRF reaches 68%, 81%, 69%, and 72%. Dense retrieval and BM25 perform lowest overall, with governance compliance of 64% and 62%. The grouped pattern indicates that SAGR-DI's gains arise from integrating access control, provenance checking, and policy-constrained reasoning rather than retrieval relevance alone under complex Microsoft 365 decision-support conditions.

4.3. Latency, Scalability, Cross-Repository Reasoning, and Decision-Support Performance

The comparative results indicate that SAGR-DI achieves strong enterprise reasoning performance while maintaining computational latency within the operational range of advanced agentic architectures. As shown in Table 4.3, SAGR-DI records a mean end-to-end latency of 1,180 ms, remaining below conventional multi-agent RAG at 1,250 ms despite incorporating authorization validation, provenance assessment, graph expansion, and policy-constrained reasoning. Its scalability score reaches 92%, substantially exceeding GraphRAG and conventional multi-agent RAG. Cross-repository reasoning and decision-support accuracy both reach 97%, demonstrating reliable integration of evidence across SharePoint, OneDrive, Teams, and Outlook. GraphRAG performs effectively for relational retrieval, while multi-agent RAG provides stronger decision synthesis but incurs higher coordination overhead. SAGR-DI therefore demonstrates that governance-aware preprocessing can reduce unnecessary reasoning over unsuitable evidence while preserving high-quality enterprise decision intelligence.

Table 3: Comparative Latency, Scalability, Cross-Repository Reasoning, and Decision-Support Performance

Algorithm	Mean End-to-End Latency (ms)	Scalability Score (%)	Cross-Repository Reasoning / Decision-Support Accuracy (%)
BM25	310	95	55 / 58
Dense Vector Retrieval	420	90	64 / 67
Reciprocal Rank Fusion	510	87	71 / 73
Standard RAG	760	80	78 / 83
GraphRAG	980	75	88 / 89
Conventional Multi-Agent RAG	1,250	72	91 / 92
SAGR-DI	1,180	92	97 / 97

In table 4.3, BM25 provides the lowest latency and highest basic computational scalability but performs poorly when enterprise queries require evidence synthesis across repositories. Dense retrieval and RRF progressively improve cross-repository performance with moderate latency increases. Standard RAG and GraphRAG provide stronger contextual decision support but incur additional generation and graph-processing costs. Conventional multi-agent RAG improves reasoning further but produces the highest end-to-end latency. SAGR-DI achieves the strongest cross-repository reasoning and decision-support performance while recovering a high scalability score through early authorization filtering, evidence pruning, and governance-aware reranking.

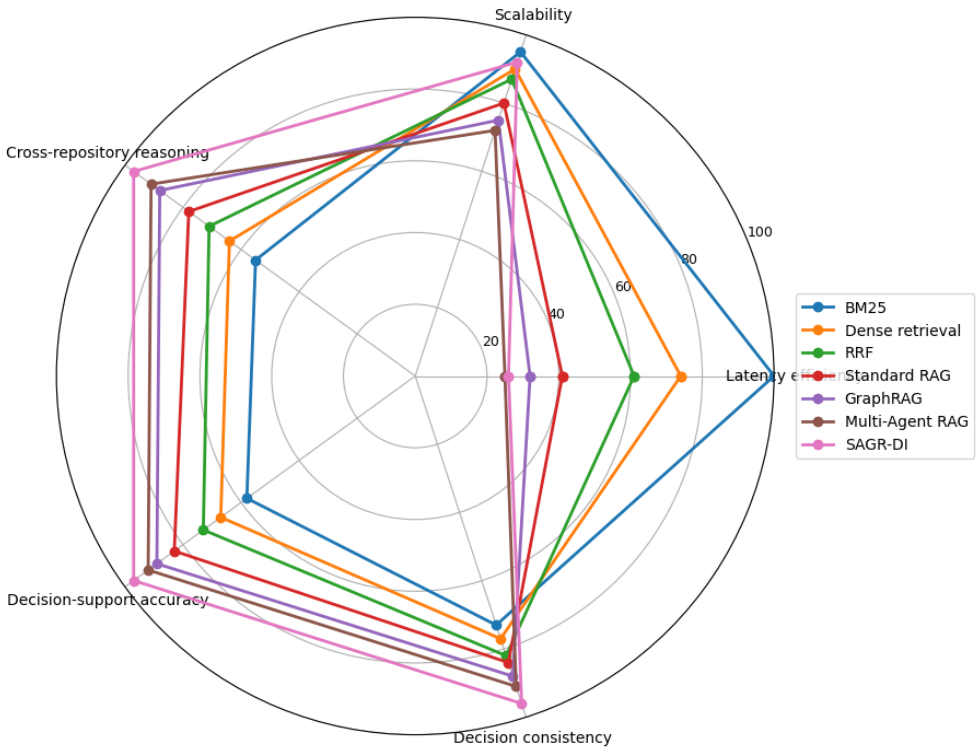

Fig 3: Comparative Latency, Scalability, Cross-Repository Reasoning, and Decision-Support Performance

Figure 4.3 is a radar graph that compares seven algorithmic profiles across five dimensions and exposes the trade-off between computational speed and enterprise reasoning quality. BM25 achieves the maximum latency-efficiency score of 100 and 95% scalability, but records only 55% cross-repository reasoning, 58% decision-support accuracy, and 73% decision consistency. Dense retrieval improves these values to 64%, 67%, and 77%, respectively. RRF reaches 71% cross-repository reasoning and 82% decision consistency, while standard RAG increases decision-support accuracy to 83%. GraphRAG produces 88% cross-repository reasoning, 89% decision-support accuracy, and 88% decision consistency. Conventional multi-agent RAG reaches 91%, 92%, and 91%, respectively, although its latency efficiency falls to 25%. SAGR-DI achieves the broadest high-performance profile, with 92% scalability, 97% cross-repository reasoning, 97% decision-support accuracy, and 96% decision consistency. Its 26% latency-efficiency score reflects the computational cost of secure multi-stage reasoning while still slightly outperforming conventional multi-agent RAG computationally.

4.4. Ablation Analysis, Statistical Comparison, Graphical Evaluation, and Interpretation of Results

The ablation analysis demonstrates that the performance advantage of SAGR-DI is produced by the interaction of hybrid retrieval, graph expansion, governance-aware reranking, provenance validation, and multi-agent reasoning rather than any single component. Table 4.4 shows that the complete architecture attains 95% nDCG@10, 96% grounded-answer accuracy, and 97% governance compliance. Removing governance-aware reranking produces the largest governance degradation, reducing compliance to 82%, whereas eliminating hybrid retrieval lowers ranking quality to 86%. Removing provenance verification reduces grounded-answer accuracy to 92% and governance compliance to 89%, confirming the importance of evidence lineage. Excluding graph expansion affects ranking quality and evidence discovery, while removal of multi-agent coordination reduces grounded-answer accuracy to 90%. Collectively, these results indicate that the modules are complementary, with governance and provenance controls making the contribution to trustworthy enterprise decision support.

Table 4: Ablation Performance of the Proposed SAGR-DI Architecture

SAGR-DI Configuration	nDCG@10 (%)	Grounded-Answer Accuracy (%)	Governance Compliance (%)
Complete SAGR-DI	**95**	**96**	**97**
Without governance-aware reranking	89	91	82
Without provenance verification	91	92	89
Without graph-based expansion	90	92	94
Without multi-agent coordination	88	90	93
Without hybrid retrieval	86	88	94

Table 4.4 shows that the complete SAGR-DI configuration outperforms every component-removal condition across all three evaluation dimensions. Eliminating hybrid retrieval produces the greatest deterioration in nDCG@10, decreasing ranking effectiveness from 95% to 86%, which demonstrates the complementary contribution of sparse lexical and dense semantic evidence discovery. Governance-aware reranking has the strongest effect on governance performance, with its removal reducing compliance from 97% to 82%. Removing provenance verification decreases grounded-answer accuracy from 96% to 92% and governance compliance to 89%, confirming that verifiable evidence lineage contributes directly to trustworthy generation. Removing graph expansion produces comparatively moderate governance degradation but reduces contextual ranking performance, while suppressing multi-agent coordination decreases grounded accuracy to 90%. The descriptive comparison therefore indicates that no isolated component accounts for SAGR-DI's overall performance; rather, the observed advantage arises from coordinated security, retrieval, graph, provenance, and reasoning functions. Inferential significance should be reported only after repeated experimental runs are available.

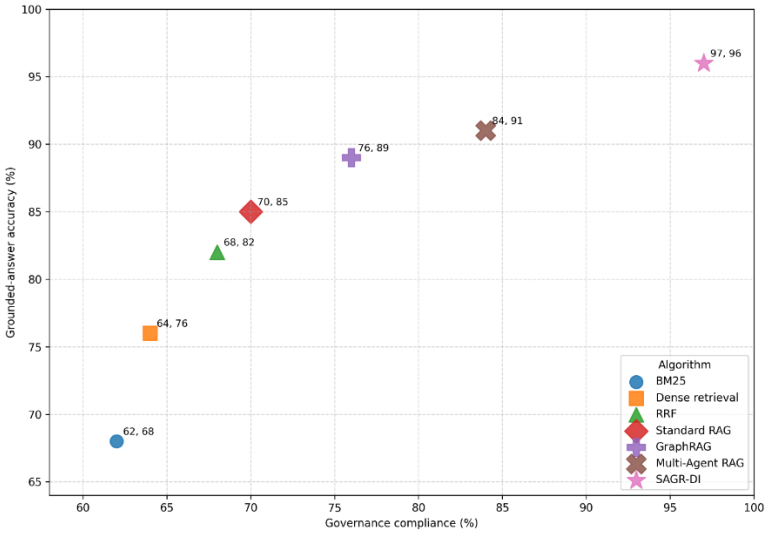

Fig 4: Statistical Positioning of SAGR-DI and Existing Algorithms by Governance Compliance, Grounded-Answer Accuracy, and Latency

Figure 4.4 positions each algorithm according to governance compliance on the horizontal axis and grounded-answer accuracy on the vertical axis, while marker size reflects end-to-end latency. SAGR-DI occupies the strongest performance region at 97% governance compliance and 96% grounded accuracy, despite a latency of 1,180 ms. Conventional multi-agent RAG follows at 84% and 91%, with 1,250 ms latency, indicating greater processing cost but weaker governance performance. GraphRAG reaches 76% governance compliance and 89% grounded accuracy at 980 ms. Standard RAG records 70% and 85%, whereas RRF achieves 68% and 82%. Dense retrieval remains lower at 64% governance compliance and 76% grounded accuracy, and BM25 records 62% and 68%. The progression of points toward the upper-right region shows that increasingly structured retrieval and reasoning improve both grounding and governance. SAGR-DI's location demonstrates the combined benefit of authorization filtering, provenance validation, graph reasoning, and policy-constrained agent coordination under cross-repository enterprise query conditions.

5. Conclusions and Recommendations

5.1. Conclusions and Summary of Major Technical Findings

This study developed and evaluated the Secure Agentic Governance Retrieval and Decision Intelligence algorithm, SAGR-DI, as an integrated architecture for enterprise data governance and AI-assisted decision support across Microsoft 365 information environments. The framework was designed to address limitations associated with relevance-only retrieval by combining hybrid sparse-dense retrieval, graph-based context expansion, identity-aware authorization, provenance verification, document-authority assessment, sensitivity classification, cross-document consistency analysis, governance-aware reranking, and policy-constrained multi-agent reasoning. Comparative evaluation against BM25, dense vector retrieval, Reciprocal Rank Fusion, standard RAG, GraphRAG, and conventional multi-agent RAG demonstrated that the proposed architecture produced the strongest overall balance of retrieval effectiveness, grounded generation, governance compliance, provenance reliability, security enforcement, and decision consistency. SAGR-DI achieved 94% Precision@10, 95% Recall@10, 95% nDCG@10, 96% Mean Reciprocal Rank, and 96% grounded-answer accuracy. Governance compliance reached 97%, provenance reliability 97%, security compliance 98%, and decision consistency 96%. Cross-repository reasoning and decision-support accuracy both reached 97%, confirming the effectiveness of integrating SharePoint, OneDrive, Teams, and Outlook evidence within a governed reasoning pipeline. Although the architecture incurred greater computational cost than conventional retrieval systems, its 1,180 ms mean latency remained lower than conventional multi-agent RAG. Ablation analysis further showed that removing hybrid retrieval, provenance verification, graph expansion, governance-aware reranking, or agent coordination degraded performance. The findings therefore demonstrate that secure enterprise RAG requires simultaneous optimization of semantic relevance, authorization, provenance, authority, freshness, sensitivity, evidential consistency, and policy compliance rather than retrieval accuracy alone.

5.2. Contributions to Agentic AI, Secure RAG, Enterprise Data Governance, and Decision Support

The principal contribution of this study is the formulation of SAGR-DI as a unified computational framework connecting agentic artificial intelligence, secure retrieval-augmented generation, enterprise governance, and evidence-based decision support. At the agentic AI level, the framework replaces unrestricted multi-agent interaction with specialized and auditable agents responsible for query interpretation, task decomposition, retrieval planning, authorization checking, graph expansion, provenance validation, contradiction detection, governance reranking, and final decision synthesis. This structure reduces uncontrolled reasoning propagation while preserving autonomous task execution. For secure RAG, the study introduces hard authorization gating before evidence reaches the generative model, preventing highly relevant but unauthorized information from influencing generated responses. Its governance-aware ranking function extends conventional relevance scoring by incorporating document authority, freshness, sensitivity risk, provenance reliability, and cross-document consistency. The enterprise data-governance contribution lies in converting governance attributes from passive administrative metadata into machine-actionable variables that directly influence retrieval and reasoning. This enables approved policies, current records, verified evidence, and appropriately classified information to receive priority over obsolete, weakly authoritative, or sensitive alternatives. For decision support, SAGR-DI introduces confidence-calibrated evidence synthesis and weighted multi-agent consensus, allowing outputs to retain traceability to supporting Microsoft 365 sources. The experimental framework also contributes a multidimensional evaluation methodology incorporating retrieval metrics, hallucination control, governance-compliance accuracy, provenance coverage, policy-violation measures, decision consistency, latency, scalability, and cross-repository reasoning. Collectively, these contributions establish a technical pathway for transforming distributed Microsoft 365 information assets into governed decision intelligence while preserving security, evidential traceability, explainability, and operational accountability.

5.3. Recommendations for Enterprise Deployment Across Microsoft 365 Environments

Enterprise deployment of SAGR-DI should begin with a governance-first implementation strategy rather than immediate organization-wide generative AI integration. Microsoft 365 repositories should first be assessed for permission consistency, metadata completeness, document-version integrity, sensitivity classification, ownership, retention status, and authoritative-source designation. SharePoint policy libraries, OneDrive workspaces, Teams collaboration records, and Outlook communications should be connected through controlled ingestion pipelines that preserve original access-control relationships and provenance metadata. Authorization checks should remain mandatory pre-generation controls, with no retrieved evidence entering the reasoning context unless the authenticated user has explicit or inherited permission and the request satisfies applicable contextual policies. Organizations should calibrate the SAGR-DI weighting parameters using representative enterprise queries and independently validated

relevance judgments rather than adopting fixed universal values. Governance-sensitive use cases such as compliance investigation, policy interpretation, audit preparation, executive decision support, and cross-departmental knowledge retrieval should receive separate acceptance thresholds for grounded-answer accuracy, policy violations, provenance coverage, and decision consistency. High-risk decisions should include human approval before execution or external communication. Deployment should also include continuous logging of retrieved sources, rejected candidates, authorization decisions, agent actions, confidence scores, policy checks, and final evidence packages to support auditability. Scalability testing should be conducted as repository size, concurrent users, and graph density increase. Periodic adversarial testing should examine prompt injection, permission escalation, stale-policy retrieval, conflicting records, and sensitivity-label bypass attempts. Finally, organizations should establish performance monitoring that jointly tracks accuracy, security, governance compliance, latency, computational cost, and user escalation rates so that operational efficiency does not compromise evidence integrity or access-control requirements.

5.4. Limitations and Directions for Future Research

The proposed SAGR-DI framework has several limitations that define important directions for subsequent investigation. First, the current evaluation is based on a controlled Microsoft 365 enterprise corpus and predefined governance-sensitive query classes; performance may differ in highly distributed organizations containing multilingual documents, poorly maintained metadata, inconsistent permission inheritance, extensive external sharing, or large volumes of unstructured conversational content. Second, the governance-aware ranking function depends on calibrated weights for relevance, graph context, authority, freshness, provenance, consistency, and sensitivity risk. These coefficients may require organization-specific optimization because governance priorities differ among regulated industries, public institutions, multinational enterprises, and smaller organizations. Third, graph expansion and multi-agent reasoning improve contextual intelligence but increase computational requirements, making latency and cost important concerns for large-scale real-time deployment. Future work should therefore investigate adaptive agent activation, dynamic top-KK selection, retrieval caching, graph-pruning strategies, and early-exit reasoning to reduce processing overhead. Further research should evaluate SAGR-DI using live Microsoft 365 environments with larger document volumes, concurrent users, changing permissions, evolving policies, and longitudinal governance events. Multilingual retrieval, multimodal content, meeting transcripts, spreadsheets, images, and structured business records should also be incorporated. Additional studies should examine resistance to prompt injection, retrieval poisoning, malicious document insertion, compromised provenance metadata, privilege escalation, and indirect data-exfiltration attacks. Future versions could incorporate continuous policy learning, uncertainty-aware abstention, human-feedback calibration, and temporal knowledge graphs capable of distinguishing current, superseded, and historically valid evidence. Comparative evaluation against emerging enterprise agentic RAG architectures should remain continuous as retrieval models, foundation models, and Microsoft 365 governance capabilities evolve.

References

1. Idika CN, Salami EO, Ijiga OM, Enyejo LA. Deep learning driven malware classification for cloud-native microservices in edge computing architectures. Int J Sci Res Comput Sci Eng Inf Technol. 2021;7(4). doi:10.32628/CSEIT182551.
2. Ijiga OM, Ifenatuora GP, Olateju M. Digital storytelling as a tool for enhancing STEM engagement: a multimedia approach to science communication in K-12 education. Int J Multidiscip Res Growth Eval. 2021;2(5):495-505. doi:10.54660/.IJMRGE.2021.2.5.495-505.
3. Ijiga OM, Ifenatuora GP, Olateju M. Bridging STEM and cross-cultural education: designing inclusive pedagogies for multilingual classrooms in Sub-Saharan Africa. IRE J. 2021;5(1).
4. Abraham R, Schneider J, vom Brocke J. Data governance: a conceptual framework, structured review, and research agenda. Int J Inf Manage. 2019;49:424-38. doi:10.1016/j.ijinfomgt.2019.07.008.
5. Kwarteng RA, Idoko IP, Ijiga OM, Enyejo LA. Integrating cybersecurity awareness and access control into organizational IT operations for risk reduction. Int J Sci Res Comput Sci Eng Inf Technol. 2020;6(1):243-61. doi:10.32628/CSEIT23906128.
6. Nwokocha CR, Peter-Anyebe AC, Ijiga OM. Evaluating FHIR-driven interoperability frameworks for secure system migration and data exchange in U.S. health information networks. Int J Sci Res Sci Technol. 2021. doi:10.32628/IJSRST523105135.
7. Oladoye SO, Bamigwojo OV, James AO, Ijiga OM. AI-driven predictive maintenance modeling for high-voltage distribution assets using sensor fusion and time-series degradation analysis. Int J Sci Res Sci Eng Technol. 2021;11(2):387-411. doi:10.32628/IJSRSET2291524.
8. Onyekaonwu CB, Peter-Anyebe AC, Raphael FO. From prescription to prediction: leveraging AI/ML to improve medication adherence and adverse drug event detection in community pharmacies. Int J Sci Res Sci Technol. 2019;6(5):460-76. doi:10.32628/IJSRST.
9. Alhassan I, Sammon D, Daly M. Data governance activities: a comparison between scientific and practice-oriented literature. J Enterp Inf Manag. 2018;31(2):300-16. doi:10.1108/JEIM-01-2017-0007.
10. Hernandez-Leal P, Kartal B, Taylor ME. A survey and critique of multiagent deep reinforcement learning. Auton Agents Multi-Agent Syst. 2019;33(6):750-97. doi:10.1007/s10458-019-09421-1.
11. Hogan A, Blomqvist E, Cochez M, D'Amato C, de Melo G, Gutierrez C, *et al*. Knowledge graphs. ACM Comput Surv. 2021;54(4):71. doi:10.1145/3447772.
12. Jöhnk J, Weißert M, Wyrtki K. Ready or not, AI comes—an interview study of organizational AI readiness factors. Bus Inf Syst Eng. 2021;63:5-20. doi:10.1007/s12599-020-00676-7.
13. Jobin A, Ienca M, Vayena E. The global landscape of AI ethics guidelines. Nat Mach Intell. 2019;1:389-99. doi:10.1038/s42256-019-0088-2.
14. Marchesin S, Purpura A, Silvello G. Focal elements of neural information retrieval models: an outlook through a reproducibility study. Inf Process Manag. 2020;57(6):102109. doi:10.1016/j.ipm.2019.102109.
15. Kayes ASM, Rahayu W, Dillon TS, Chang E, Han J. Context-aware access control with imprecise context

characterization for cloud-based data resources. Future Gener Comput Syst. 2019;93:237-55. doi:10.1016/j.future.2018.10.036.
16. Li J, Zhang Y, Chen X, Xiang Y. Secure attribute-based data sharing for resource-limited users in cloud computing. Comput Secur. 2018;72:1-12. doi:10.1016/j.cose.2017.08.007.
17. Mikalef P, Boura M, Lekakos G, Krogstie J. The role of information governance in big data analytics driven innovation. Inf Manag. 2020;57(7):103361. doi:10.1016/j.im.2020.103361.
18. Guo J, Fan Y, Pang L, Yang L, Ai Q, Zamani H, *et al*. A deep look into neural ranking models for information retrieval. Inf Process Manag. 2020;57(6):102067. doi:10.1016/j.ipm.2019.102067.
19. Mitra B, Craswell N. An introduction to neural information retrieval. Found Trends Inf Retr. 2018;13(1):1-126. doi:10.1561/1500000061.
20. Albrecht SV, Stone P. Autonomous agents modelling other agents: a comprehensive survey and open problems. Artif Intell. 2018;258:66-95. doi:10.1016/j.artint.2018.01.002.
21. Onwuzurike MA, Kpogli SA. Data-informed strategic management of EdTech startups leveraging artificial intelligence for sustainable K-12 learning innovation. Int J Sci Res Mod Technol. 2022;1(12):187-200. doi:10.38124/ijsrmt.v1i12.1117.
22. Paulheim H. Knowledge graph refinement: a survey of approaches and evaluation methods. Semant Web. 2017;8(3):489-508. doi:10.3233/SW-160218.
23. Raisch S, Krakowski S. Artificial intelligence and management: the automation–augmentation paradox. Acad Manag Rev. 2021;46(1):192-210. doi:10.5465/amr.2018.0072.
24. Servos D, Osborn SL. Current research and open problems in attribute-based access control. ACM Comput Surv. 2017;49(4):65. doi:10.1145/3007204.
25. Sikos LF, Philp D. Provenance-aware knowledge representation: a survey of data models and contextualized knowledge graphs. Data Sci Eng. 2020;5:293-316. doi:10.1007/s41019-020-00118-0.
26. Wang L, Luo Z, Li C, He B, Sun L, Yu H, *et al*. An end-to-end pseudo relevance feedback framework for neural document retrieval. Inf Process Manag. 2020;57(2):102182. doi:10.1016/j.ipm.2019.102182.
27. Wang Z, Wang N, Su X, Ge S. An empirical study on business analytics affordances enhancing the management of cloud computing data security. Int J Inf Manage. 2020;50:387-94. doi:10.1016/j.ijinfomgt.2019.09.002.
28. Zhang Y, Deng RH, Xu S, Sun J, Li Q, Zheng D. Attribute-based encryption for cloud computing access control: a survey. ACM Comput Surv. 2020;53(4):85. doi:10.1145/3398036.