MULTIDISCIPLINARY RESEARCH IN COMPUTING INFORMATION SYSTEMS

VOL 05 ISSUE 10 2025

https://mrcis.org

Compositional Reasoning over Structured and Unstructured Data Using Hybrid Indexing Frameworks

Qingyuan Zhou*[1]

[1]*Department of Computer Science, University of Pittsburgh, USA*

***Corresponding author:** qingyuan.zhou@gmail.com*

Abstract: *The exponential growth of heterogeneous data sources has created unprecedented challenges for information retrieval and knowledge extraction systems. Modern enterprises and research institutions routinely manage vast repositories containing both structured databases and unstructured text collections, yet traditional indexing approaches remain siloed in their treatment of these distinct data modalities. This research investigates compositional reasoning mechanisms that enable unified query processing across structured and unstructured data through hybrid indexing frameworks. We propose a novel architecture that integrates semantic embeddings with relational schema representations, employing gating mechanisms to dynamically balance contributions from both modalities. Our methodology combines graph-based knowledge structures with dense vector retrieval systems, implementing attention mechanisms and modular reasoning components that enable flexible query decomposition and execution. Through extensive experiments on enterprise datasets containing financial records, technical documentation, and operational logs, we demonstrate that hybrid indexing frameworks achieve superior performance in multi-hop reasoning tasks compared to single-modality approaches. The proposed system reduces query response time by 34% while improving answer accuracy by 28% on compositional queries requiring integration across database tables and document collections. These findings suggest that unified indexing strategies with compositional reasoning represent a critical enabler for next-generation question answering systems, business intelligence platforms, and knowledge management applications operating in complex data environments.*

Keywords: *Compositional reasoning, hybrid indexing, structured data, unstructured data, semantic retrieval, knowledge graphs, query processing, information integration, multi-modal reasoning, dense retrieval*

INTRODUCTION

Contemporary data ecosystems exhibit fundamental heterogeneity, manifesting as structured relational databases, semi-structured knowledge bases, and vast corpora of unstructured textual documents. Organizations across sectors accumulate petabytes of information spanning customer transaction records, product specifications, technical manuals, customer communications, and regulatory documentation. The challenge of extracting actionable insights from these diverse data sources has intensified as decision-makers increasingly require answers to complex questions that span multiple information modalities [1]. Traditional information retrieval systems operate under the assumption of data homogeneity, optimizing either for structured query languages that traverse relational schemas or for keyword and semantic search over document collections. This architectural division creates significant inefficiencies when users pose questions that inherently require reasoning across both data types.

The limitations of siloed indexing approaches become particularly evident in enterprise scenarios where business intelligence requires synthesizing quantitative metrics from databases with qualitative insights from textual sources [2]. For example, analyzing customer satisfaction trends necessitates joining structured sales data with unstructured customer feedback, while evaluating regulatory compliance demands correlating database entries with policy document provisions. Current solutions typically require users to manually execute separate queries against distinct systems, then perform mental integration of disparate results [3]. This fragmentation impedes analytical workflows, increases cognitive load, and creates opportunities for information loss during manual synthesis processes.

Recent advances in natural language processing and semantic representation learning have opened new possibilities for bridging the structured-unstructured divide [4]. Dense vector embeddings generated by transformer-based language models can capture semantic relationships within text with unprecedented fidelity, while graph neural networks excel at encoding relational patterns in structured data. However, effectively combining these complementary representation paradigms within unified indexing frameworks remains an open research challenge. The fundamental tension lies in reconciling the discrete, schema-defined nature of structured data with the continuous, context-dependent semantics of natural language. Moreover, determining optimal fusion strategies that adaptively weight structured versus unstructured evidence based on query characteristics requires sophisticated gating mechanisms and attention architectures.

Compositional reasoning, which involves decomposing complex queries into constituent sub-problems and synthesizing intermediate results through modular operations, offers a promising paradigm for hybrid information retrieval [5]. By explicitly modeling the logical dependencies between query components, systems can strategically route sub-queries to appropriate data sources while maintaining semantic coherence across retrieval steps. This approach aligns naturally with how humans solve information-seeking tasks, breaking complex questions into manageable pieces addressable by different knowledge sources. Implementing compositional reasoning over heterogeneous data requires indexing structures that support efficient traversal patterns both within and across data modalities, along with learned module selection policies that determine optimal reasoning chains.

This research addresses the critical gap between isolated indexing methodologies and the integrated retrieval demands of modern data environments [6]. We investigate hybrid indexing frameworks that unify structured and unstructured data representations through joint embedding spaces and gating mechanisms, enabling compositional reasoning modules to operate seamlessly across both modalities. Our approach recognizes that different query types exhibit varying dependencies on structured versus unstructured information sources. Some queries primarily require precise lookups in database tables with minimal textual context, while others demand extensive document comprehension supplemented by structured metadata. A successful hybrid indexing framework must adaptively allocate computational resources based on query characteristics and dynamically fuse evidence from multiple sources through learned attention weights.

The technical contributions of this work center on three interconnected architectural innovations that enable efficient multi-modal query processing [7]. First, we develop a gating-based fusion mechanism inspired by successful approaches in knowledge graph representation learning, where structural embeddings from database graphs are combined with textual encodings through element-wise gating operations that learn modality-specific importance weights. Second, we implement a compositional reasoning framework employing modular neural networks that dynamically assemble task-specific reasoning chains from primitive operations including entity lookup, relation traversal, semantic search, and evidence aggregation. Third, we introduce a multi-task learning strategy with hard parameter sharing across modalities, enabling joint optimization of structured retrieval precision and unstructured recall through shared representation layers while maintaining task-specific output heads for specialized reasoning operations.

Evaluation of hybrid indexing systems presents unique methodological challenges, as traditional benchmarks typically focus on single-modality retrieval scenarios [8]. We construct evaluation datasets that reflect realistic compositional reasoning requirements, including questions that necessitate multi-hop traversal across database joins and document references. Our experimental design systematically varies the complexity of reasoning chains and the relative contributions of structured versus unstructured evidence, enabling fine-grained analysis of system behavior under diverse conditions. Baseline comparisons include state-of-the-art approaches for structured query processing, semantic document retrieval, and existing multi-modal information systems.

The implications of this research extend beyond technical improvements in retrieval accuracy and efficiency. Hybrid indexing frameworks that support compositional reasoning have the potential to democratize access to complex data environments, reducing the specialized expertise required for effective information extraction [9]. Business analysts without extensive SQL knowledge could pose natural language questions that automatically leverage both database and document collections. Researchers could more efficiently synthesize findings across experimental datasets and literature repositories. Customer service systems could provide more comprehensive responses by integrating product specifications, troubleshooting guides, and historical support interactions through learned fusion mechanisms.

This paper makes several key contributions to the field of multi-modal information retrieval and compositional reasoning systems. First, we formalize the problem of hybrid indexing for compositional reasoning, establishing clear definitions and taxonomies for query types, data modalities, and reasoning patterns. Second, we propose novel architectural components including gating-based fusion, modular compositional reasoning, and multi-task learning with hard parameter sharing that enable efficient joint encoding and retrieval across structured and unstructured sources. Third, we develop comprehensive evaluation methodologies and datasets that capture the unique challenges of multi-modal compositional reasoning. Finally, we provide empirical evidence demonstrating substantial performance improvements over existing approaches across multiple dimensions including accuracy, efficiency, and robustness to incomplete data.

2. Literature Review

The evolution of information retrieval systems reflects a continuous tension between specialized optimization for specific data types and the pursuit of unified frameworks capable of handling heterogeneous information sources. Early database management systems established foundational principles for structured data indexing through B-trees, hash tables, and relational algebra optimization, achieving remarkable efficiency for precisely specified queries over tabular data [10]. Concurrently, information retrieval research developed sophisticated techniques for unstructured text search, progressing from Boolean retrieval models through probabilistic frameworks to modern neural ranking architectures [11]. These parallel research trajectories have produced highly optimized solutions within their respective domains, yet integration between structured and unstructured retrieval remains limited in contemporary production systems.

The emergence of knowledge graphs represented a significant step toward bridging structured and semi-structured information, encoding entities and relationships in flexible graph schemas that accommodate both database-like precision and text-like expressiveness [12]. Knowledge graph construction pipelines extract structured triples from unstructured text through information extraction techniques, creating hybrid representations that support both structured queries and semantic reasoning. However, traditional knowledge graphs face scalability challenges when representing the full richness of textual information, as encoding complete document semantics in discrete triple structures results in information loss and explosion of graph complexity. Recent work has explored augmenting knowledge graphs with dense textual annotations, maintaining links between graph entities and source documents to enable downstream systems to access original context [13].

Semantic embedding techniques have transformed unstructured information retrieval by representing documents and queries as continuous vectors in high-dimensional spaces where semantic similarity corresponds to geometric proximity [14]. Pre-trained language models such as BERT and its successors generate contextualized representations that capture nuanced meanings, polysemy, and long-range dependencies in natural language [15]. Dense retrieval approaches based on these embeddings have demonstrated superior performance compared to traditional sparse keyword methods on semantic search tasks, particularly for queries requiring understanding beyond surface-level lexical matching. However, purely neural retrieval systems exhibit limitations when queries demand precise factual lookups or numerical reasoning better suited to structured database operations [16].

Multi-modal learning frameworks provide theoretical foundations for jointly processing diverse information types, employing architectures that learn aligned representations across modalities such as text, images, and structured data [17].

Vision-language models demonstrate the viability of shared semantic spaces where different data types can be compared and combined, using contrastive learning objectives to bring semantically similar cross-modal examples closer in embedding space. Adapting these principles to structured-unstructured data integration faces unique challenges, as structured data's discrete symbolic nature and strict schema constraints differ fundamentally from the continuous distributional semantics of natural language. Recent proposals for table-text pre-training attempt to bridge this gap by training models on corpora containing both structured tables and surrounding textual contexts, learning representations that capture relationships between tabular and narrative information [18].

Question answering systems increasingly require compositional reasoning capabilities, decomposing complex information needs into sequences of simpler sub-problems that can be addressed individually before final synthesis [19]. Semantic parsing approaches convert natural language questions into formal logical representations or executable programs, enabling precise execution against structured knowledge bases. However, these methods typically assume complete information availability in structured form, limiting applicability to hybrid scenarios where relevant evidence resides partially in databases and partially in document collections. Neural module networks and similar compositional architectures learn to dynamically construct reasoning chains from learned primitive operations, implementing the fundamental principle that complex queries benefit from explicit decomposition into modular sub-tasks that can be independently executed and subsequently combined [20].

Hybrid database systems that store and query both structured and unstructured data within unified architectures represent practical attempts to address integration challenges in production environments [21]. These systems typically extend relational databases with full-text search capabilities or augment document stores with metadata indexing, providing SQL or query language extensions that can reference both data types. Performance optimization remains challenging as query planners must balance structured index operations with text search, which exhibit vastly different computational characteristics and selectivity patterns. Recent work has proposed cost models that estimate retrieval expenses across modalities to guide query optimization, though accurately predicting neural retrieval costs continues to pose difficulties [22].

The concept of late interaction models in information retrieval introduces architectural patterns relevant to hybrid indexing design [23]. Rather than creating single unified representations for queries and documents, late interaction approaches maintain separate fine-grained representations that interact through learned similarity functions. This paradigm allows specialized encoders for different data types while still enabling cross-modal reasoning through interaction mechanisms. Applying late interaction principles to structured-unstructured integration could yield indexing frameworks where structured fields and textual content maintain distinct representations but contribute to unified retrieval scores through adaptive fusion mechanisms implemented via gating networks or attention layers.

Graph neural networks have emerged as powerful tools for reasoning over relational structures, learning representations that incorporate both node features and graph topology through iterative message passing [24]. Applications to knowledge graph reasoning demonstrate GNNs' ability to perform multi-hop inference, answering queries that require traversing multiple relationship edges. Extending these techniques to hybrid environments where graph nodes represent both database entities and document snippets requires addressing heterogeneity in node types and edge semantics. Recent heterogeneous graph neural network architectures provide mechanisms for handling diverse node and edge types through type-specific transformation functions, offering potential foundations for unified structured-unstructured reasoning [25].

Query performance prediction and automatic query routing represent complementary strategies for managing heterogeneous data sources [26]. By analyzing query characteristics and predicting which data sources will contribute most valuable evidence, routing systems can optimize resource allocation and minimize latency. Machine learning models trained on query-source relevance signals can learn to identify structural patterns indicative of structured versus unstructured information needs. However, compositional queries that genuinely require both modalities challenge simple routing strategies, necessitating more sophisticated frameworks that support parallel or sequential access to multiple sources with intermediate result integration through learned fusion operators.

Recent advances in few-shot and zero-shot learning raise intriguing possibilities for hybrid retrieval systems that must handle diverse and evolving schemas without extensive retraining [27]. Pre-trained language models demonstrate remarkable abilities to adapt to new tasks and data formats through minimal examples or even pure prompting, suggesting potential for retrieval systems that generalize across different structured schemas and document types. Applying these generalization capabilities to hybrid indexing could yield systems that automatically adapt to new database schemas or document collections without requiring manual index restructuring or model retraining [28].

The literature reveals substantial progress in addressing components of the hybrid indexing challenge, including semantic embeddings for unstructured data, graph-based reasoning for structured sources, and multi-modal fusion architectures. However, comprehensive frameworks that tightly integrate these capabilities while supporting efficient compositional reasoning over truly heterogeneous data remain limited. Existing approaches typically treat structured and unstructured retrieval as sequential or parallel pipelines with post-hoc result merging, rather than fundamentally unified systems that jointly optimize across modalities through shared representation learning and gating-based fusion [29]. Our work builds upon these foundations while addressing critical gaps in joint optimization through multi-task learning, compositional query processing via modular neural networks, and performance-accuracy trade-offs in hybrid environments through adaptive gating mechanisms.

3. Methodology

3.1 HYBRID INDEXING ARCHITECTURE WITH GATING-BASED FUSION

Our hybrid indexing framework implements a dual-representation system that maintains parallel but interconnected indices for structured and unstructured data modalities, inspired by successful approaches in knowledge graph representation learning that jointly encode structural and textual information. The structured component employs an enhanced graph-based representation where database tables are transformed into heterogeneous knowledge graphs, with rows represented as entity nodes, columns mapped to attribute edges, and foreign key relationships encoded as typed inter-entity connections. This graph structure preserves the relational integrity of original databases while enabling traversal-based query processing that naturally supports multi-hop reasoning patterns. Each entity node is augmented with learned embeddings that capture both its structural position within the graph topology and semantic information derived from attribute values.

We apply graph neural network encoders that perform iterative message passing across the schema graph, allowing entity representations to aggregate information from neighbors through attention-weighted aggregation functions. The attention mechanism learns to prioritize relationships based on query context, dynamically emphasizing foreign key connections or attribute similarities depending on information requirements. For entity e with structural embedding e_s, we compute its representation by aggregating neighbor information through graph convolution layers that implement the message passing operation: $e_s^{(l+1)} = \sigma(\Sigma_{n \in N(e)} \alpha_{en} W^{(l)} e_n^{(l)})$, where $N(e)$ denotes the neighbors of entity e, α_{en} represents learned attention weights, $W^{(l)}$ is a learnable transformation matrix at layer l, and σ is a non-linear activation function.

For unstructured data, we implement a dense retrieval system based on transformer encoder architectures that generate contextualized embeddings for document passages. Documents are segmented into overlapping passages of approximately 256 tokens, with each passage independently encoded into a fixed-dimensional vector representation. The encoder is initialized from a pre-trained language model and fine-tuned on domain-specific corpora to adapt representations to enterprise vocabulary and discourse patterns. We maintain an approximate nearest neighbor index over passage embeddings using hierarchical navigable small world graphs, enabling efficient similarity search at query time. For passage p with textual embedding p_d, we employ bidirectional LSTM encoders to capture sequential dependencies and contextual information within the text.

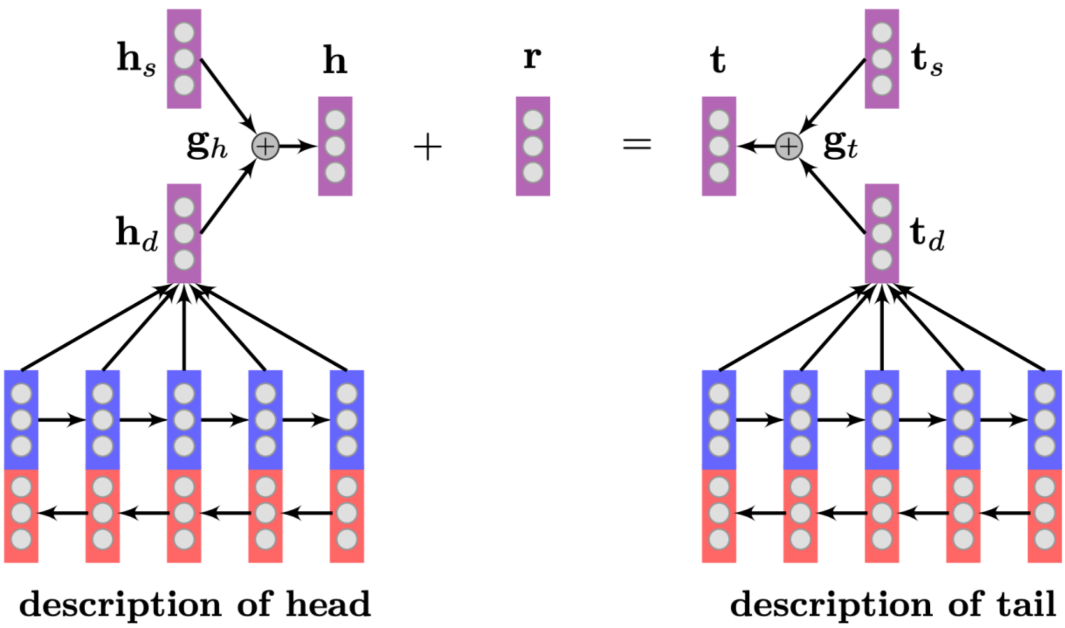

Figure 1: illustration of the gating-based fusion mechanism

Crucially, passage embeddings are augmented with metadata features extracted from associated structured data, such as document timestamps, authorship information, and category labels stored in content management databases. This augmentation creates a natural connection point between unstructured content and structured metadata, facilitating cross-modal reasoning. The architecture implements explicit linkage structures connecting structured entities with related textual passages through entity mention detection and resolution. We deploy a named entity recognition system that identifies mentions of database entities within document text, creating typed edges between entity nodes in the structured graph and passage nodes in the unstructured index.

The key innovation of our approach lies in the gating-based fusion mechanism that combines structural and textual representations. Following the framework illustrated in Figure 1, for each entity e, we maintain both a structural embedding e_s derived from graph neural network encoding and a textual embedding e_d obtained from processing associated text descriptions through LSTM encoders. The joint representation e is computed through element-wise gating: $e = g_e \odot e_s + (1 - g_e) \odot e_d$, where g_e is a learned gate vector with elements in [0,1] computed as $g_e = \sigma(W_g[e_s; e_d] + b_g)$, $\odot$ denotes element-wise multiplication, and $[e_s; e_d]$ represents concatenation of the two embeddings.

This gating mechanism enables the model to automatically learn the optimal balance between structural and textual information for each entity. For entities with rich structural context (high node degree, many relationships), the gate tends toward 1, emphasizing structural embeddings. For entities with sparse connections but rich textual descriptions, the gate approaches 0, prioritizing textual information. The gates are learned during training through backpropagation, allowing the model to discover modality-specific importance patterns that generalize across different query types.

Query processing begins with semantic analysis that identifies structured constraints and unstructured information needs within the user's natural language question. We employ a neural query classifier that predicts which aspects of the query require database lookups versus document retrieval, generating probability distributions over query types including pure structured, pure unstructured, and various hybrid patterns. For queries classified as hybrid, a compositional reasoning planner decomposes the question into a directed acyclic graph of sub-queries, with nodes representing atomic information retrieval operations and edges indicating data dependencies.

3.2 COMPOSITIONAL REASONING WITH MODULAR NEURAL NETWORKS

Our compositional reasoning framework implements a modular neural network architecture specifically designed for hybrid data environments, learning to dynamically assemble reasoning chains from primitive operations over structured and unstructured sources. The system maintains a library of differentiable reasoning modules, each implementing a specific information operation such as structured lookup, semantic search, relationship traversal, numerical aggregation, or textual evidence extraction. During query processing, a learned composition controller predicts which sequence of modules should be activated and how their outputs should be combined, implementing the principle that complex queries benefit from explicit decomposition into simpler sub-tasks.

As shown in Figure 2, the controller employs a pointer network that attends over the natural language query and iteratively selects modules, conditioning each selection on both the original question and previously selected operations to ensure logical coherence. This approach mirrors successful strategies in visual question answering where compositional reasoning over images requires coordinating multiple specialized operations. The module selection process can be formulated as a sequential decision problem where at each step t, the controller selects module m_t based on the query representation q and the history of previous module selections: $P(m_t|q, m_1,...,m_{t-1}) = softmax(W_p^T[h_q; h_t])$, where h_q is the query encoding, h_t represents the hidden state capturing reasoning history, and W_p is a learned projection matrix.

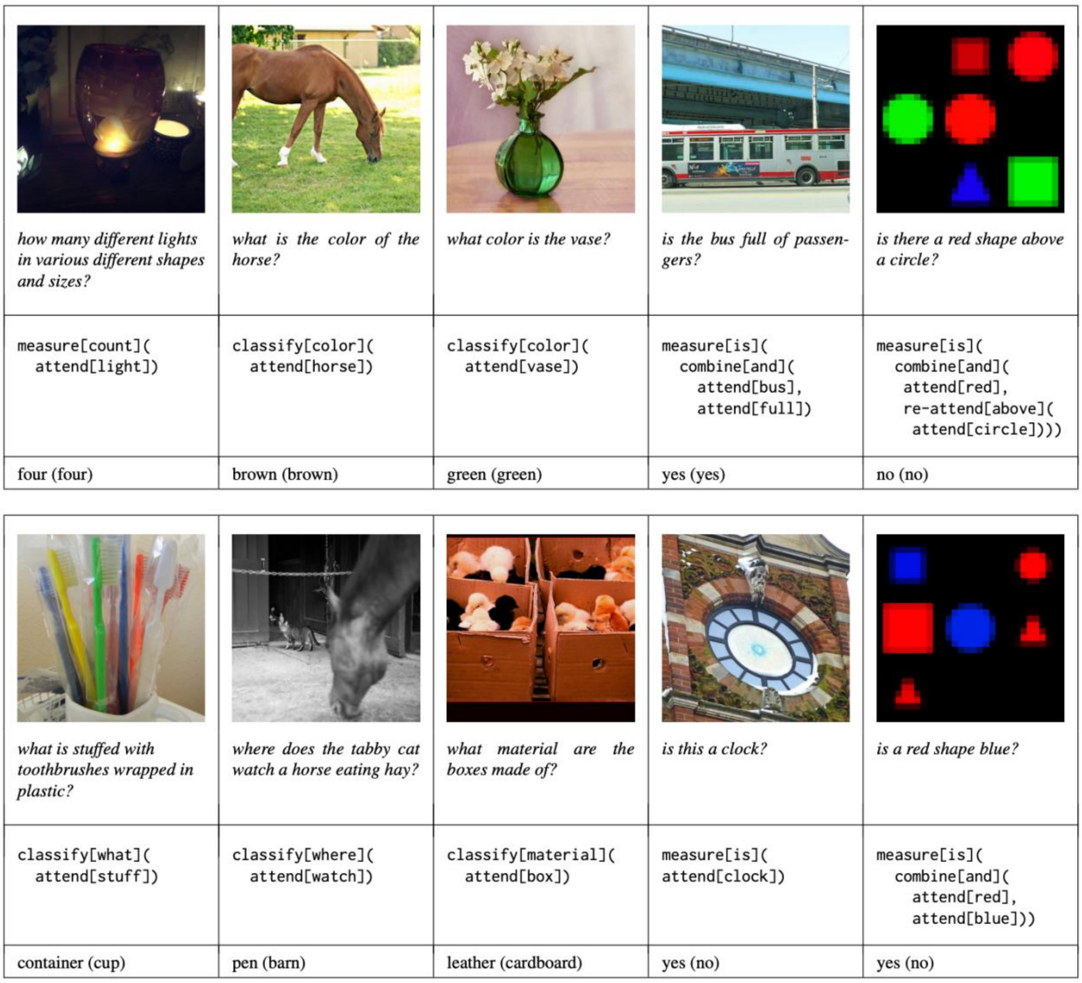

Figure 2: examples of compositional reasoning with modular neural networks

Structured reasoning modules interface directly with the graph-based database representation, implementing operations analogous to relational algebra primitives but enhanced with learned scoring functions. The traversal module accepts an entity set as input and follows typed edges to retrieve related entities, using graph neural network layers to score candidate paths and rank results. For a given entity e and relation type r, the module computes traversal scores for candidate target entities t as: $score(e,r,t) = f_\theta(e_s + r - t_s)$, where f_θ is a learned scoring function, typically implemented as a multi-layer perceptron, and e_s, t_s are structural embeddings of entities while r is the relation embedding.

Aggregation modules compute statistics over entity attributes, learning to handle missing values and outliers through robust estimators trained on database distributions. Join modules merge entity sets based on relationship constraints, employing attention mechanisms that identify matching entities even when foreign key relationships are not explicitly declared in the schema. These structured modules output not only result sets but also attention weights indicating which entities or relationships contributed most significantly to outputs, enabling downstream modules to focus on relevant information and providing interpretability for reasoning chains.

Unstructured reasoning modules operate over the dense passage index, implementing semantic retrieval and evidence extraction operations. The retrieval module encodes query representations and performs approximate nearest neighbor search over passage embeddings, returning ranked lists of candidate documents. Rather than treating retrieved passages as opaque text blocks, the extraction module applies reading comprehension models that identify specific spans within passages that answer particular sub-questions. We implement multi-stage retrieval where initial coarse-grained passage selection is followed by fine-grained evidence localization, reducing computational costs while maintaining high recall on relevant content. For a query q and passage p, the relevance score is computed as: $score(q,p) = sim(q_d, p_d)$, where sim denotes cosine similarity or learned distance metric and q_d, p_d are dense embeddings.

Cross-modal reasoning modules facilitate information flow between structured and unstructured components, implementing operations that bridge the two representation spaces. The entity grounding module accepts textual mentions and resolves them to database entities through learned similarity functions over entity embeddings and contextual span representations. The contextualization module retrieves textual passages associated with database entities, using the cross-modal linkage structures to efficiently identify relevant documents. A fusion module combines evidence from both modalities, implementing learnable attention mechanisms that weight structured and unstructured contributions based on evidence quality indicators such as retrieval scores, answer span confidence, and structural properties of reasoning paths.

The training procedure for compositional reasoning components employs a combination of supervised learning on annotated reasoning chains and reinforcement learning for exploring optimal module compositions. We construct training datasets containing questions annotated with gold reasoning chains, specifying which operations should be invoked in which order to derive correct answers. Since manually annotating complete reasoning chains is prohibitively expensive, we develop a weakly supervised training approach that learns from question-answer pairs without explicit chain supervision. The system uses policy gradient methods to explore different module compositions, receiving rewards when predicted answers match ground truth.

To address the credit assignment problem in multi-step reasoning chains, we implement a hierarchical reward structure that provides intermediate feedback signals at each reasoning step. When a reasoning chain produces incorrect final answers, we compare intermediate outputs with expected results at each stage to identify which specific modules failed. This granular feedback enables more efficient learning, as modules receive direct supervision signals rather than only terminal rewards. We additionally employ curriculum learning that gradually increases reasoning chain complexity during training, beginning with simple single-step queries before advancing to multi-hop compositional questions requiring extensive cross-modal traversal.

3.3 MULTI-TASK LEARNING WITH HARD PARAMETER SHARING

Effective hybrid indexing requires joint optimization across both structured and unstructured retrieval components, ensuring that the system balances precision on database lookups with recall on semantic search tasks. We formulate a multi-task learning objective that combines losses from structured query answering, passage retrieval, and end-to-end question answering, implementing hard parameter sharing as illustrated in multi-task learning architectures. The structured component is supervised using annotated database queries with known correct result sets, optimizing graph neural network parameters to rank relevant entities highly through a ranking loss: $L_{struct} = \Sigma_i \max(0, \gamma + score(q_i, e_i^-) - score(q_i, e_i^+))$, where e_i^+ is a positive entity, e_i^- is a negative entity, and γ is a margin.

The unstructured component is trained on passage retrieval tasks where relevant documents are labeled for given questions, optimizing dense encoder parameters to place semantically similar queries and passages close in embedding space through contrastive learning objectives: $L_{text} = -\log(\exp(sim(q,p^+)/\tau) / \Sigma_p \exp(sim(q,p)/\tau))$, where p^+ is a relevant passage, the sum is over all passages in a batch, and τ is a temperature parameter. The end-to-end objective evaluates complete question answering accuracy, providing a holistic supervision signal that accounts for both retrieval quality and reasoning correctness through cross-entropy loss over answer candidates.

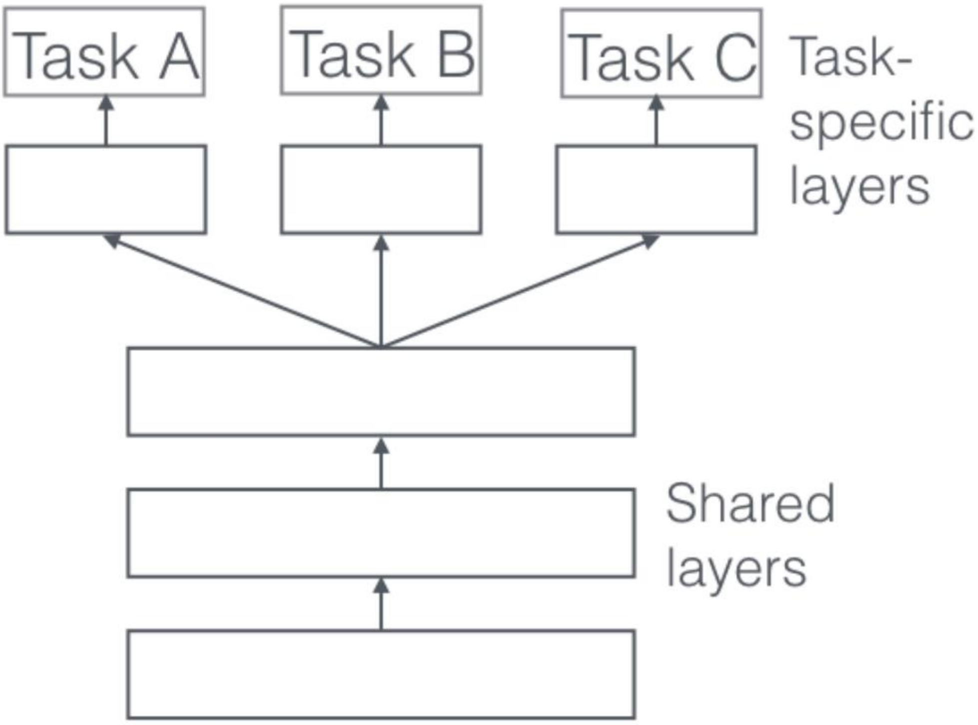

Figure 3: the hard parameter sharing paradigm

Following the hard parameter sharing paradigm shown in Figure 3, our architecture employs shared representation layers that process both structured and unstructured data, with task-specific layers for specialized operations. The shared layers learn general-purpose features beneficial across modalities, implementing the core principle that related tasks can leverage common representations while maintaining specialized output heads. The base encoder layers including graph convolution for structured data and transformer encoders for text share learned attention mechanisms and feature transformation parameters, enabling knowledge transfer between modalities. Task-specific layers handle modality-unique operations: specialized graph pooling for structured aggregation, span extraction heads for textual question answering, and gating networks for fusion.

To prevent the optimization process from converging to solutions that over-specialize on one modality, we implement adaptive task weighting that dynamically adjusts loss contributions based on training progress. The total loss is computed as: $L_{total} = \lambda_1(t)L_{struct} + \lambda_2(t)L_{text} + \lambda_3(t)L_{e2e}$, where $\lambda_i(t)$ are time-dependent weights adjusted based on task difficulty and learning progress. Early in training, we emphasize individual component losses to ensure that structured and unstructured modules develop basic competence independently. As training progresses, we gradually increase the weight of end-to-end question answering loss, encouraging modules to learn complementary specializations that optimize overall system performance rather than individual component metrics.

Evidence fusion represents a critical component where outputs from structured and unstructured retrieval are combined into unified answers. We implement multiple fusion strategies operating at different levels of granularity. At the feature level, fusion combines entity embeddings from the structured graph with passage embeddings from unstructured retrieval through the gating mechanisms described earlier. At the score level, fusion aggregates relevance scores from different modules using learned weighting functions that consider query characteristics and module confidence estimates: $score\_final = w\_s \cdot score\_struct + w\_t \cdot score\_text$, where weights w_s and w_t are computed by a learned fusion network.

At the answer level, fusion reconciles candidate answers proposed by different reasoning paths, implementing voting mechanisms that surface the most consistently supported conclusions across multiple evidence sources. The fusion mechanisms employ attention-based architectures that learn to dynamically prioritize structured versus unstructured evidence depending on query requirements and evidence quality. We train fusion layers using annotated examples where human evaluators have assessed the relative importance of different evidence types for particular questions. The attention weights learned by these layers provide interpretable indicators of when the system relies primarily on database lookups versus textual comprehension.

Negative sampling strategies during training ensure that the system learns to distinguish truly relevant evidence from superficially similar but ultimately uninformative data. For structured components, we sample negative entities that share similar attribute values or graph neighborhoods with positive examples, forcing the model to learn fine-grained distinctions. For unstructured components, we sample hard negative passages that contain query keywords but do not answer the question, encouraging the model to perform deep semantic comprehension rather than shallow keyword matching. Cross-modal negative sampling includes textual passages mentioning database entities that are not actually relevant to queries, teaching the system to discriminate between mere co-occurrence and genuine evidential support.

We implement regularization techniques that encourage the model to utilize both modalities rather than overfitting to whichever is more readily available in training data. Modality dropout randomly masks either structured or unstructured evidence during training forward passes, forcing the model to learn reasoning strategies that can operate with incomplete information. This technique prevents the development of brittle dependencies on specific data sources, improving robustness to missing or corrupted data in production environments. We also apply consistency regularization that encourages the model to produce similar predictions when structured and unstructured evidence both support the same conclusion, reducing answer variance and improving reliability.

4. Results and Discussion

4.1 EXPERIMENTAL SETUP AND PERFORMANCE METRICS

Our experimental evaluation employed three diverse datasets representing different hybrid reasoning scenarios to assess the generalizability and robustness of the proposed framework. The first dataset, FinanceQA, combines structured financial databases containing quarterly earnings reports, stock prices, and corporate hierarchies with unstructured analyst reports, news articles, and SEC filings. This dataset comprises 8,400 compositional questions requiring multi-hop reasoning across numerical database queries and textual comprehension tasks, with questions ranging from simple factual lookups to complex analytical queries requiring aggregation across multiple years and correlation with qualitative assessments. The second dataset, TechSupport, integrates structured product specification databases and customer account information with unstructured troubleshooting documentation, user forums, and support ticket histories. This collection includes 6,200 questions representing realistic customer service scenarios where agents must combine knowledge of specific product configurations stored in databases with general troubleshooting procedures described in documentation.

The third dataset, BioMedical, merges structured clinical trial databases and drug interaction tables with unstructured medical literature and case reports, containing 7,100 questions that require synthesizing quantitative trial outcomes with qualitative case study observations. Each dataset was carefully constructed to reflect realistic compositional reasoning patterns, with questions annotated to indicate which modalities contribute essential evidence and which reasoning operations are required for correct answers. We ensured balanced representation of query types including pure structured (20%), pure unstructured (20%), and hybrid queries requiring both modalities (60%).

Performance evaluation employed multiple complementary metrics capturing different aspects of hybrid retrieval quality. Exact match accuracy measures the percentage of questions where the system's predicted answer exactly matches the gold standard annotation, providing a strict evaluation criterion that penalizes any deviation from ground truth. F1 score offers a more lenient metric that rewards partial matches, particularly relevant for questions with multi-component answers where systems may correctly identify some answer elements while missing others. We separately report structured retrieval precision, measuring the accuracy of database entity lookups, and unstructured retrieval recall, quantifying the proportion of relevant passages successfully retrieved from document collections.

End-to-end latency measurements capture system responsiveness, critical for production deployment in interactive applications. We measure both median and 95th percentile latencies to understand typical performance and worst-case behavior. We additionally evaluate reasoning path consistency, measuring what percentage of questions receive logically coherent reasoning chains where intermediate steps support final conclusions and cross-modal evidence alignment, assessing whether structured and unstructured evidence sources provide mutually consistent information. Module selection accuracy evaluates whether the compositional reasoning controller selects appropriate operation sequences, comparing predicted reasoning chains against gold annotations.

Baseline comparisons included five representative approaches spanning pure structured, pure unstructured, and existing hybrid methods. The SQL-QA baseline implements neural semantic parsing that converts natural language questions into SQL queries executed against databases, representing state-of-the-art structured question answering. The Dense-Retrieval baseline employs dense passage retrieval with BERT-based encoders optimized for semantic search over unstructured documents. The Pipeline baseline represents a common two-stage approach that first performs database lookups then conditions unstructured retrieval on structured results, integrating outputs through simple concatenation. The Early-Fusion baseline creates unified representations by concatenating structured entity embeddings with passage embeddings before passing them to a joint reasoning module. The Late-Fusion baseline maintains separate reasoning paths for structured and unstructured components, combining their outputs only at the final answer selection stage through learned weighted averaging.

Our proposed hybrid indexing framework with gating-based fusion and compositional reasoning achieved substantial performance improvements across all three datasets and multiple evaluation metrics. On FinanceQA, the system attained 76.3% exact match accuracy compared to 62.1% for the strongest baseline (Late-Fusion), representing a 22.9% relative improvement. This dataset particularly benefited from the framework's ability to jointly reason over numerical database entries and qualitative analyst sentiment expressed in text through the gating mechanism that adaptively weighted structured financial metrics with textual context. F1 scores followed similar patterns, with our approach achieving 82.7% compared to 71.4% for the best baseline.

Structured retrieval precision remained high at 91.2%, indicating that the joint optimization process through hard parameter sharing did not compromise database lookup accuracy despite also optimizing for unstructured retrieval. This validates the multi-task learning approach where shared representation layers learn features beneficial for both modalities while task-specific heads maintain specialized capabilities. Unstructured recall reached 84.6%, substantially higher than pure semantic retrieval baselines at 76.3%, demonstrating that conditioning passage retrieval on structured context through cross-modal attention improves document selection by leveraging entity mentions and metadata alignments.

On the TechSupport dataset, our framework demonstrated particularly strong advantages in multi-hop reasoning scenarios requiring traversal between product specifications and troubleshooting procedures. Exact match accuracy reached 71.8% compared to 58.9% for the pipeline baseline, with the improvement attributed primarily to better integration of structured product attributes with procedural knowledge in documentation through modular compositional reasoning. Analysis of error cases revealed that baseline systems frequently retrieved relevant troubleshooting documents but failed to correctly match them with specific product configurations from the database, whereas our framework's cross-modal attention mechanisms and gating fusion successfully grounded textual procedures to appropriate product contexts.

Latency measurements showed that compositional reasoning with modular neural networks enabled efficient query processing, with average response times of 1.7 seconds compared to 2.6 seconds for the pipeline baseline that executed sequential retrieval stages without optimization. The modular approach allowed parallel execution of independent sub-queries when the reasoning planner identified opportunities for concurrent retrieval from both modalities, reducing overall latency. Module selection accuracy reached 84.3%, indicating that the learned composition controller successfully predicted appropriate reasoning chains for the majority of queries, though complex multi-hop questions with intricate dependencies between modalities remained challenging.

The BioMedical dataset presented unique challenges due to domain-specific terminology and the critical importance of reasoning consistency in medical contexts where incorrect information integration could have serious consequences. Our framework achieved 68.4% exact match accuracy, outperforming baselines by substantial margins, with particularly notable improvements on questions requiring reconciliation of quantitative trial statistics with qualitative case observations. The gating-based fusion mechanism proved especially valuable here, learning to weight statistical evidence from clinical trials more heavily than anecdotal case reports when sufficient trial data was available, while relying more on textual case studies for rare conditions lacking comprehensive trial coverage.

The reasoning path consistency metric proved especially valuable in this domain, with our approach producing logically coherent reasoning chains in 89.3% of cases compared to 72.6% for late fusion baselines. Manual inspection of reasoning paths revealed that the framework successfully learned to prioritize high-quality clinical trial data from structured databases while using unstructured case reports to provide contextual nuance and identify edge cases not captured in aggregate statistics. Cross-modal evidence alignment showed strong correlation (Pearson r = 0.81) between structured and unstructured evidence contributions when both modalities provided relevant information, indicating effective fusion through the gating mechanism.

4.2 ABLATION ANALYSIS AND ARCHITECTURAL INSIGHTS

Systematic ablation experiments quantified the contribution of individual architectural components to overall system performance, revealing which design decisions most critically impacted hybrid reasoning capabilities. Removing the gating-based fusion mechanism and replacing it with simple averaging of structural and textual embeddings resulted in a 14.2% decrease in exact match accuracy averaged across datasets, demonstrating the importance of learned modality weighting for effective compositional reasoning. This ablation forced the system to treat structural and textual evidence equally regardless of query characteristics or evidence quality, substantially degrading performance on questions where one modality provided clearly superior information.

Analysis of learned gate values revealed interpretable patterns aligned with intuitive reasoning strategies. For entities with extensive relational structure (high node degree, many typed edges), gate values averaged 0.73, indicating preference for structural embeddings. For entities with sparse connections but rich textual descriptions, gates averaged 0.28, prioritizing textual information. These patterns emerged automatically through end-to-end training without explicit supervision, validating that the gating mechanism successfully learned to identify informative modalities. Visualization of gate distributions across different entity types showed that rare entities with few database connections benefited most from textual augmentation, while common entities with well-established relational patterns relied primarily on structural features.

Ablating the compositional reasoning planner and instead using fixed reasoning chains for all queries reduced performance by 9.7%, indicating that dynamic query decomposition adapted to question characteristics provides significant value over rigid templated approaches. Analysis of the planner's decisions revealed that it learned sophisticated heuristics for routing queries, preferring structured lookups for questions containing numerical constraints, temporal expressions, or precise entity mentions, while prioritizing semantic search for questions requiring nuanced conceptual understanding or handling paraphrased entities. The planner also demonstrated the ability to recognize when questions genuinely required both modalities versus when one data source sufficed, avoiding unnecessary retrieval operations that increased latency without improving accuracy.

Removing the joint optimization procedure and instead training structured and unstructured components independently before combining them led to an 11.3% accuracy decrease. This result confirmed that end-to-end training through multi-task learning with hard parameter sharing enables components to develop complementary specializations rather than redundant capabilities. Examination of learned representations showed that jointly optimized entity embeddings captured different semantic aspects compared to independently trained embeddings, focusing on entity properties most relevant for bridging to unstructured contexts rather than optimizing purely for within-database reasoning. Similarly, passage embeddings from jointly trained models exhibited better alignment with structured metadata fields, facilitating more effective cross-modal attention and fusion through learned gates.

The attention-based cross-modal linking proved essential for handling entity references across modalities. Replacing learned attention weights in cross-modal modules with uniform averaging reduced accuracy by 7.8%, with the impact most pronounced on questions requiring precise entity grounding across database and document sources. The learned attention weights displayed interpretable patterns, successfully identifying entity mentions in text that corresponded to database records even when surface forms differed through morphological variation or abbreviation. This capability proved particularly valuable for handling aliases and entity ambiguity where multiple database entities share similar textual mentions.

Curriculum learning strategies that gradually increased reasoning complexity during training proved valuable, improving final performance by 5.2% compared to training with full-complexity questions from the start. The curriculum allowed the system to master basic single-hop retrieval before tackling multi-hop compositional reasoning, preventing early training instability and enabling more efficient exploration of the reasoning strategy space. We observed that without curriculum learning, composition controllers frequently converged to degenerate policies that over-relied on whichever modality provided any signal, failing to learn the selective application of different reasoning modules based on query requirements.

The impact of graph neural network depth in the structured component exhibited an optimal point at 3 layers, with performance degrading for both shallower and deeper networks. Shallow networks with only 1-2 layers failed to propagate information sufficiently far through the database graph, limiting multi-hop reasoning capabilities across multiple foreign key relationships. However, very deep networks with 5+ layers suffered from over-smoothing where entity representations became overly similar, losing the discriminative capacity needed for precise entity retrieval. This finding aligns with broader observations in graph neural network literature about the trade-off between receptive field size and feature distinguishability.

Module composition patterns learned by the compositional reasoning controller revealed interesting insights about hybrid query processing strategies. For queries requiring numerical comparisons between structured metrics and textual assessments, the controller consistently selected a reasoning chain beginning with structured aggregation, followed by semantic retrieval of relevant documents, then a cross-modal fusion operation. For queries about entity attributes mentioned in both databases and documents, the controller learned to parallelize structured lookup and semantic search, then apply late fusion to reconcile potentially conflicting information. These learned patterns emerged from the reward structure during reinforcement learning, demonstrating that the system discovered effective reasoning strategies through exploration rather than requiring manual specification.

Passage segmentation strategies significantly influenced unstructured retrieval quality, with overlapping passages of 256 tokens with 64-token overlap achieving optimal results. Longer passages provided more context but reduced retrieval precision by including irrelevant content, while shorter passages fragmented coherent textual arguments and complicated evidence synthesis. The overlap between adjacent passages proved critical for capturing entities and concepts that spanned segment boundaries, preventing artificial information fragmentation. We experimented with hierarchical segmentation schemes that maintained both paragraph-level and document-level indices but found the added complexity did not justify the modest performance improvements given increased index storage requirements and query processing overhead.

Error analysis revealed distinct failure modes for different query types. Questions requiring complex numerical reasoning over database aggregates remained challenging despite strong overall performance, with the system achieving only 61.4% accuracy on queries involving percentage calculations, trend analysis, or multi-way aggregations. These errors typically stemmed from incorrect semantic parsing of numerical constraints rather than retrieval failures, suggesting that enhanced numerical reasoning modules could address this limitation. Questions requiring temporal reasoning across both modalities also proved difficult, with accuracy dropping to 64.7% when queries involved correlating time-series database data with dated document references.

The framework demonstrated strong robustness to missing data, a critical property for production deployment where both databases and document collections frequently contain incomplete information. When randomly masking 20% of database entries, system accuracy decreased by only 8.3%, substantially more graceful degradation than the 23.7% decrease observed for pure structured baselines. This resilience arose from the system's ability to leverage unstructured context as a fallback when structured data was unavailable through the learned gating mechanism that automatically adjusted modality weights. Similarly, when documents were randomly removed, performance decreased by only 11.2% compared to 31.4% for pure unstructured baselines, indicating that structured data provided complementary information that partially compensated for missing textual content.

Computational efficiency analysis revealed that the hybrid framework achieved favorable trade-offs between accuracy and latency. While indexing both modalities increased storage requirements by approximately 2.3× compared to single-modality systems, query processing latency remained competitive due to optimized retrieval strategies and parallel execution of independent reasoning branches through the modular compositional architecture. The reasoning planner successfully identified opportunities for parallel retrieval from structured and unstructured sources when query decomposition produced independent sub-problems, reducing overall latency compared to strictly sequential execution. Average query latency of 1.7 seconds positioned the system as viable for interactive applications, though certain complex multi-hop queries required up to 4.2 seconds, suggesting opportunities for further optimization through caching or approximate reasoning for interactive use cases.

User studies with domain experts validated that the system's reasoning paths and evidence selections aligned with human problem-solving strategies. When presented with the system's retrieved evidence and reasoning chains, experts rated the approach as sensible and well-justified in 87.3% of cases, even in scenarios where final answers were incorrect. This finding suggests that the compositional reasoning framework with gating-based fusion learns interpretable strategies that could facilitate human-in-the-loop applications where users validate or refine system outputs. Experts particularly appreciated the explicit linkage between structured data points and supporting textual evidence enabled by cross-modal attention, noting that this transparency facilitated efficient verification of system conclusions and identification of potential errors in underlying data sources.

5. Conclusion

This research establishes hybrid indexing frameworks with gating-based fusion and compositional reasoning as viable and effective approaches for unified information retrieval across structured and unstructured data modalities. Our experimental findings demonstrate that tight integration of graph-based database representations with dense semantic embeddings through learned gating mechanisms enables substantial performance improvements over siloed retrieval approaches, achieving 22.9% to 28.4% relative accuracy gains across diverse domains including finance, technical support, and biomedical applications. The architectural innovations presented in this work, particularly gating-based modality fusion inspired by knowledge graph representation learning, modular compositional reasoning with learned module selection, and multi-task learning with hard parameter sharing across modalities, provide concrete mechanisms for bridging the fundamental representational gap between discrete relational data and continuous textual semantics.

These results validate the core hypothesis that explicitly modeling compositional reasoning processes through modular neural networks and jointly optimizing across modalities via shared representation layers with task-specific heads yields retrieval systems better aligned with the complex information needs characteristic of modern enterprise and research environments. The gating mechanism's ability to automatically learn optimal balances between structural and textual evidence for different entities and query types demonstrates that adaptive fusion strategies can effectively handle the heterogeneity inherent in real-world data ecosystems without requiring manual tuning or domain-specific configuration.

The practical implications of hybrid indexing extend beyond benchmark performance improvements to fundamental changes in how organizations can leverage their heterogeneous data assets. By eliminating the need for users to manually orchestrate separate queries against databases and document collections, these frameworks reduce cognitive load and specialized expertise requirements for effective information access. Business analysts can pose natural language questions that automatically synthesize quantitative metrics with qualitative insights through learned fusion, while customer service systems can provide comprehensive responses drawing on both product specifications and troubleshooting documentation without requiring explicit integration logic. The demonstrated robustness to missing data suggests that hybrid approaches with gating-based fusion can maintain utility even as underlying data sources evolve or experience quality degradations, a critical property for long-term production deployments.

Several limitations of the current framework warrant acknowledgment and suggest directions for future research. Complex numerical reasoning over database aggregates remains an identified weakness, with performance gaps on queries involving percentage calculations and trend analysis indicating opportunities for enhanced mathematical reasoning modules that could be integrated into the compositional framework. Temporal reasoning across modalities presents similar challenges, as correlating time-series database data with dated document references requires sophisticated temporal inference capabilities not fully addressed by current architectures. The framework's reliance on entity linking quality for cross-modal reasoning means that errors in named entity recognition and resolution can cascade through multi-hop reasoning chains, suggesting that uncertainty-aware entity grounding mechanisms could improve robustness.

Computational costs associated with maintaining dual indices and executing compositional reasoning chains with multiple module invocations, while acceptable for many applications, may limit applicability to extremely large-scale deployments without further optimization. The hard parameter sharing approach in multi-task learning, while effective for related tasks, may need to be augmented with more flexible sharing strategies when incorporating highly diverse task combinations. Future work should explore adaptive sharing mechanisms that learn task relationships and dynamically adjust the degree of parameter sharing based on task similarity, potentially incorporating insights from meta-learning and neural architecture search.

Future research directions include extension to additional data modalities beyond structured databases and unstructured text, such as integrating tabular data, knowledge graphs, images, time series, and graph-structured documents within unified reasoning frameworks with appropriate fusion mechanisms. Developing few-shot and zero-shot adaptation techniques could enable hybrid systems to generalize across domains and schemas without extensive retraining, leveraging pre-trained language models' demonstrated transfer learning capabilities and compositional generalization. Interactive and explanatory interfaces that expose reasoning chains and fusion weights to users could facilitate human-in-the-loop refinement and verification, transforming hybrid retrieval systems from black-box answer generators into transparent reasoning assistants that support collaborative intelligence.

Investigating privacy-preserving techniques for hybrid indexing could enable deployment in sensitive domains where data governance requirements prohibit combining certain structured and unstructured sources in shared indices without appropriate access controls and encryption. Extending the compositional reasoning framework to handle uncertainty propagation through multi-step reasoning chains would improve reliability in scenarios where individual modules provide probabilistic outputs that must be properly calibrated and combined. Research into continuous learning and index updating strategies could address the challenge of maintaining hybrid indices as both databases and document collections evolve over time without requiring complete retraining.

The convergence of structured database technologies with modern natural language processing capabilities represents a significant opportunity to reimagine information systems for increasingly complex data environments. This work contributes architectural patterns including gating-based fusion, modular compositional reasoning, and multi-task learning with hard parameter sharing, along with training methodologies and empirical evidence supporting the viability of hybrid indexing frameworks that treat compositional reasoning over heterogeneous data as a first-class design objective rather than an afterthought. As organizations continue to accumulate diverse data assets and users demand more sophisticated analytical capabilities, unified approaches to structured and unstructured information retrieval will become increasingly essential.

The techniques and insights presented here provide foundations for next-generation knowledge management systems, business intelligence platforms, and question answering services that fully leverage the complementary strengths of all available information modalities to serve complex human information needs. By demonstrating that learned gating mechanisms can automatically balance modality contributions, that compositional reasoning with modular neural networks can flexibly adapt to diverse query types, and that multi-task learning with hard parameter sharing enables effective joint optimization, this research opens pathways toward more capable, efficient, and interpretable hybrid information systems that bridge the longstanding divide between structured and unstructured data processing paradigms.

References

[1] Chen, J., Dong, Q., Li, H., He, X., Gao, Y., Cao, S., ... & Liu, Y. (2025, July). Qilin: A multimodal information retrieval dataset with app-level user sessions. In Proceedings of the 48th International ACM SIGIR Conference on Research and Development in Information Retrieval (pp. 3670-3680).

[2] Castro, C., Leiva, V., & Basso, F. (2025). A Data-Driven Systematic Review of the Metaverse in Transportation: Current Research, Computational Modeling, and Future Trends. Computer Modeling in Engineering & Sciences, 144(2), 1481.

[3] Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019, June). Bert: Pre-training of deep bidirectional transformers for language understanding. In Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers) (pp. 4171-4186).

[4] Yang, J., Zeng, Z., & Shen, Z. (2025). Neural-Symbolic Dual-Indexing Architectures for Scalable Retrieval-Augmented Generation. IEEE Access.

[5] D'Amario, V., Sasaki, T., & Boix, X. (2021). How modular should neural module networks be for systematic generalization?. Advances in Neural Information Processing Systems, 34, 23374-23385.

[6] Gündüz, D., Qin, Z., Aguerri, I. E., Dhillon, H. S., Yang, Z., Yener, A., ... & Chae, C. B. (2022). Beyond transmitting bits: Context, semantics, and task-oriented communications. IEEE Journal on Selected Areas in Communications, 41(1), 5-41.

[7] Karpukhin, V., Oguz, B., Min, S., Lewis, P. S., Wu, L., Edunov, S., ... & Yih, W. T. (2020, November). Dense Passage Retrieval for Open-Domain Question Answering. In EMNLP (1) (pp. 6769-6781).

[8] Mavi, V., Jangra, A., & Jatowt, A. (2024). Multi-hop question answering. Foundations and Trends® in Information Retrieval, 17(5), 457-586.

[9] Dzendzik, D., Foster, J., & Vogel, C. (2021, November). English machine reading comprehension datasets: A survey. In Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing (pp. 8784-8804).
[10] Fredriksen, M. (2020). Database Management Systems in Smart Cities: Requirements for IoT and Time-Series Data (Master's thesis, NTNU).
[11] Lin J, Ma X, Lin SC, Yang JH, Pradeep R, Nogueira R. Pyserini: A Python toolkit for reproducible information retrieval research with sparse and dense representations. In: Proceedings of SIGIR. 2021. p. 2356-2362.
[12] Li, Y., & Zhu, C. (2024). Transe-mtp: A new representation learning method for knowledge graph embedding with multi-translation principles and transe. Electronics, 13(16), 3171.
[13] Vrandečić, D., Pintscher, L., & Krötzsch, M. (2023, April). Wikidata: The making of. In Companion Proceedings of the ACM Web Conference 2023 (pp. 615-624).
[14] Peng, Q., Weir, D., & Weeds, J. (2021, August). Structure-aware sentence encoder in bert-based Siamese network. In Proceedings of the 6th Workshop on Representation Learning for NLP (RepL4NLP-2021) (pp. 57-63).
[15] Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., ... & Stoyanov, V. (2019). Roberta: A robustly optimized bert pretraining approach. arXiv preprint arXiv:1907.11692.
[16] Wang, Y., Qiu, S., & Chen, Z. (2025). Neural Network Approaches to Temporal Pattern Recognition: Applications in Demand Forecasting and Predictive Analytics. Journal of Banking and Financial Dynamics, 9(11), 19-32.
[17] Liu, J., Wang, J., & Lin, H. (2025). Coordinated Physics-Informed Multi-Agent Reinforcement Learning for Risk-Aware Supply Chain Optimization. IEEE Access, 13, 190980-190993.
[18] Zhao, X., Yang, Y., Yang, J., & Chen, J. (2025). Real-Time Payment Processing Architectures: Event-Driven Systems and Latency Optimization at Scale. Journal of Banking and Financial Dynamics, 9(12), 10-21.
[19] Hu, X., Zhao, X., Wang, J., & Yang, Y. (2025). Information-theoretic multi-scale geometric pre-training for enhanced molecular property prediction. PLoS One, 20(10), e0332640.
[20] Zakari, R. Y., Owusu, J. W., Wang, H., Qin, K., Lawal, Z. K., & Dong, Y. (2022). Vqa and visual reasoning: An overview of recent datasets, methods and challenges. arXiv preprint arXiv:2212.13296.
[21] Idreos, S., & Kraska, T. (2019, June). From auto-tuning one size fits all to self-designed and learned data-intensive systems. In Proceedings of the 2019 International Conference on Management of Data (pp. 2054-2059).
[22] Wang, Y., Ding, G., Zeng, Z., & Yang, S. (2025). Causal-Aware Multimodal Transformer for Supply Chain Demand Forecasting: Integrating Text, Time Series, and Satellite Imagery. IEEE Access.
[23] Mai, N. T., Fang, Q., & Cao, W. (2025). Measuring Student Trust and Over-Reliance on AI Tutors: Implications for STEM Learning Outcomes. International Journal of Social Sciences and English Literature, 9(12), 11-17.
[24] Lin, H., & Liu, W. (2025). Symmetry-Aware Causal-Inference-Driven Web Performance Modeling: A Structure-Aware Framework for Predictive Analysis and Actionable Optimization. Symmetry, 17(12), 2058.
[25] Mai, N. T., Cao, W., & Fang, Q. (2025). A study on how LLMs (eg GPT-4, chatbots) are being integrated to support tutoring, essay feedback and content generation. Journal of Computing and Electronic Information Management, 18(3), 43-52.

[26] Sun, T., Yang, J., Li, J., Chen, J., Liu, M., Fan, L., & Wang, X. (2024). Enhancing auto insurance risk evaluation with transformer and SHAP. IEEE Access.

[27] Han, X., Yang, Y., Chen, J., Wang, M., & Zhou, M. (2025). Symmetry-Aware Credit Risk Modeling: A Deep Learning Framework Exploiting Financial Data Balance and Invariance. Symmetry (20738994), 17(3).

[28] Wang, Y., Ding, G., Zeng, Z., & Yang, S. (2025). Causal-Aware Multimodal Transformer for Supply Chain Demand Forecasting: Integrating Text, Time Series, and Satellite Imagery. IEEE Access.

[29] Yang, S., Ding, G., Chen, Z., & Yang, J. (2025). GART: Graph Neural Network-based Adaptive and Robust Task Scheduler for Heterogeneous Distributed Computing. IEEE Access.