International Journal of Science and Research Archive
eISSN: 2582-8185
Cross Ref DOI: 10.30574/ijsra
Journal homepage: https://ijsra.net/

(REVIEW ARTICLE)

Check for updates

Towards Secure and Privacy-Preserving Query Processing for Encrypted Big Data in Multi-Cloud Environments: A Systematic Review

Ibrahim Rashid Abdullahi *

Faculty of science, Department of computer science Somali National university.

International Journal of Science and Research Archive, 2026, 18(01), 853-871

Publication history: Received on 14 December 2025; revised on 25 January 2026; accepted on 27 January 2026

Article DOI: https://doi.org/10.30574/ijsra.2026.18.1.0173

Abstract

The rapid growth of cloud computing and big data analytics has intensified concerns over privacy when sensitive data are outsourced to third-party cloud providers. Traditional encryption techniques protect data confidentiality but significantly limit the ability to perform expressive and efficient queries, particularly in distributed and multi-cloud environments. Motivated by the increasing demand for secure analytics across healthcare, finance, IoT, and collaborative cloud platforms, this review systematically examines privacy-preserving query processing techniques for encrypted data in multi-cloud settings. Following PRISMA guidelines, a systematic literature review of published peer-reviewed studies is conducted. The reviewed approaches are categorized into homomorphic encryption-based methods, searchable encryption techniques, secure multi-party computation, trusted execution environments, and hybrid architectures. The analysis highlights key trade-offs among privacy guarantees, query expressiveness, computational efficiency, and scalability. While hybrid and multi-cloud approaches improve flexibility and fault tolerance, they introduce new challenges related to leakage, communication overhead, and trust assumptions. This review identifies critical research gaps, including limited real-time support, side-channel vulnerabilities, and the absence of standardized benchmarks. Finally, future research directions are outlined, emphasizing AI-assisted encrypted querying, federated analytics, and post-quantum privacy-preserving frameworks for multi-cloud environments.

Keywords: Privacy-preserving query processing; Encrypted big data; multi-cloud security; Homomorphic encryption and Secure multi-party computation

1. Introduction

1.1. Background and Motivation

Cloud computing and big data have been growing at a very fast rate and this has altered how organizations store, manage and analyze information. Companies are moving to cloud computing providers who take large datasets of data so that companies can access scalability storage and on-demand processing capabilities. But with sensitive data, such as financial documents, medical data, and industrial IoT (IIoT) telemetry often being outsourced, this practice has become a major cause of privacy issues (Hashi, 2025). Although effective, cloud providers are not necessarily trusted blindly, insider threats, possible data breaches, and compliance issues are the reasons that require the implementation of powerful secure data processing mechanisms. To alleviate these fears, the privacy query processing methods have been developed, which allows the data owners to query an encrypted data set without revealing sensitive information (Fugkeaw et al., 2024; Khan et al., 2025).

* Corresponding author: Ibrahim Rashid Abdullahi; ORCID No: https://orcid.org/0009-0008-2710-8173

1.2. Problem Statement

Although encryption provides confidentiality, in its essence, it utterly ruins the normal query mechanisms because a standard database cannot process ciphertext directly. This poses a conflict between privacy and effective execution of queries. Multi-cloud networks also complicate this issue by adding the principles of data distribution, synchronization, and trust, in which each cloud can be honest-but-curious or likely to collude (Zhu et al., 2024). In addition, privacy and performance have a natural trade-off: more aggressive encryption can lead to increased computational overhead and latency, but less serious ones can undermine confidentiality (Li et al., 2025). These issues make it clear that systematic research and assessment of current methods of privacy preserving query processing are necessary.

1.3. Importance of Privacy-Preserving Query Processing

Protecting sensitive data is essential, but also regulatory compliance such as the GDPR, HIPAA and other existing privacy regulations worldwide (Joshi et al., 2025). In the medical field, patient records can be accessed safely through encrypted querying of the medical records to support telemedicine and medical research without breaching privacy. In finance, it allows risk analysis and detection of fraudsters on sensitive transactions to be done securely. In a similar manner, IIoT and smart grid must contain secure analytics over encrypted telemetry data to provide a high degree of integrity in protecting user privacy (Fugkeaw & Deevijit, 2025). These applications demonstrate that there is an urgent need to have an approach to achieving efficiency, security, and compliance in the cloud–based environments.

1.4. Scope and Aims of This Review

The aim of the review is to perform a systematic study of recent developments of privacy methods of query processing encrypted big data in cloud and multi-cloud computing architectures. The review scope covers cryptographic and system-level solutions that can facilitate the execution of queries with security and cover the issues of performance, scalability, and privacy leakage. In particular, this review is aimed at defining dominant design paradigms, supported query types, trust assumptions, and deployment environments at different areas of application like healthcare, smart grids, and collaborative cloud platforms. This enables the study to have a narrowed down analysis which is based on peer-reviewed research contributions and thus the synthesis of the state is based on technical grounds.

1.5. Contributions of This Review

The most important contributions of this review paper are as follows:

- PRISMA-compliant of transparent, reproducible, and rigorous systematic literature review.
- An extensive taxonomy of privacy preserving query processing methods, such as searchable encryption, homomorphic encryption, and hybrid cloud systems.
- Comparison between the current methods in terms of privacy guarantees, query expressiveness, computational efficiency and applicability to multiple clouds.
- A comprehensive enumeration of gaps in research and unresolved challenges with opportunities to carry out work in the future in the specific areas in scalable, compliant, and practical encrypted query processing.

2. Review Methodology

The review will use a Systematic Literature Review (SLR) approach in order to thoroughly evaluate and synthesize existing research on privacy-preserving encrypted big data query processing, among a variety of cloud providers. The reason why SLR was chosen is to have an organized, transparent and reproducible review process to identify the reproducible objective identification of the relevant studies with minimal selection bias. The review is conducted according to the PRISMA 2020 (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) framework that gives systematic steps of conducting the study identification, screening, eligibility assessment, and ultimate inclusion. There were also multiple legitimate digital libraries that were searched systematically using well-designed keywords and Boolean operators in the context of encrypted query process, cloud and multi-cloud computing, searchable encryption, homomorphic encryption, SMPC and trusted execution environment.

Within the selection process based on PRISMA, 312 records have been identified first of all with the assistance of database searches. Upon eliminating 78 duplicated studies, 234 unique studies were left and underwent screening of titles and abstracts where 146 studies were eliminated due to not fitting the specified scope and relevancy criteria. Afterward, a total of 88 full-texts were evaluated based on their eligibility of which 46 articles were filtered out because either they lacked the use of encrypted query processing mechanisms, did not use clouds to deploy their systems, or had a strictly theoretical cryptographic focus but did not implement the system in practice. In this systematic literature

review, eventually, 42 peer-reviewed research articles were chosen to undergo the synthesis of qualitative data and deep analysis. Figure 1 depicts the PRISMA 2020 flow diagram, which explains the entire study selection process since the first identification of the studies to the ultimate inclusion of 42 studies to be used in this review.

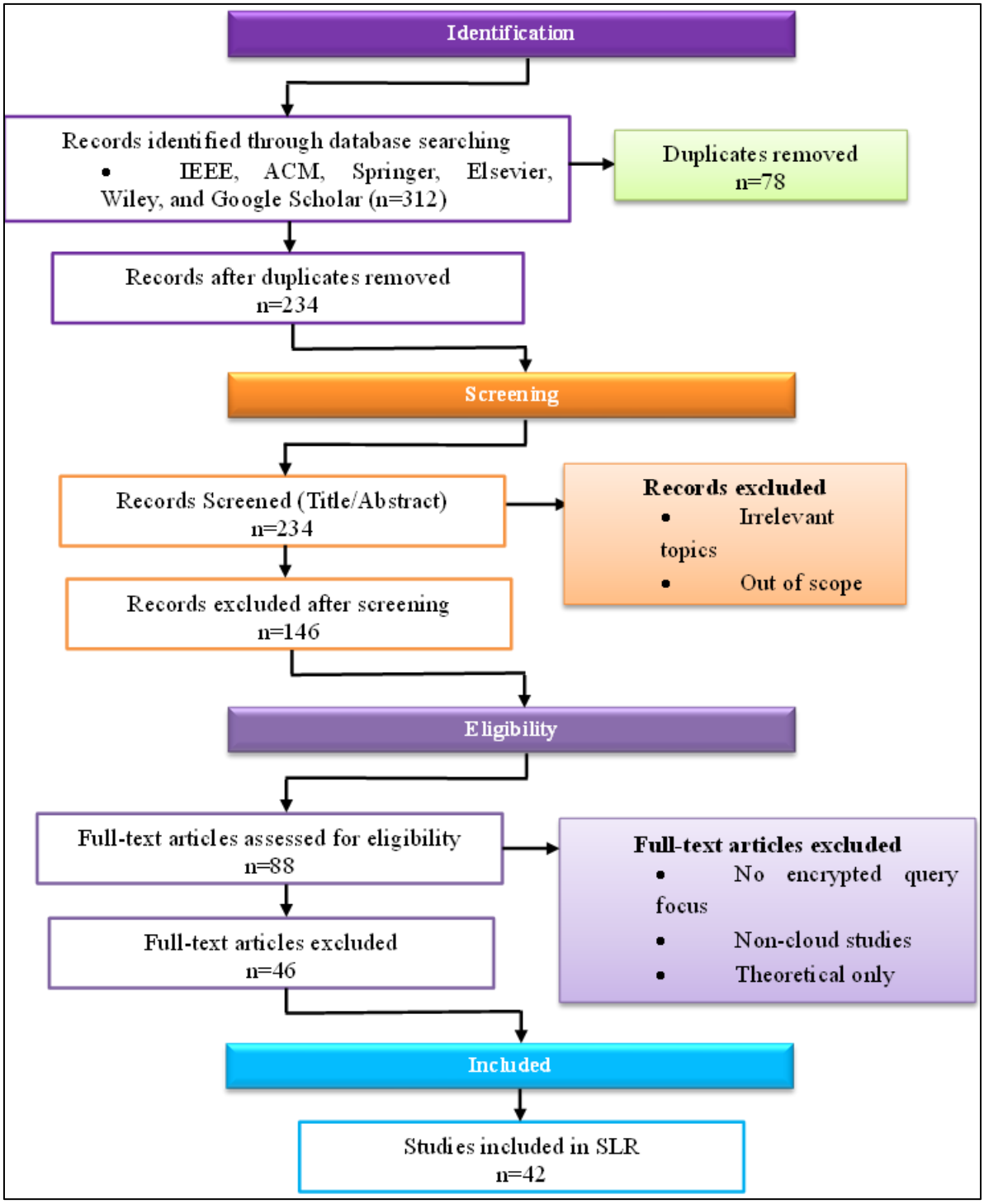

Figure 1 PRISMA flow diagram

2.1. Review Protocol and Guidelines

To provide high-methodological rigor and reproducibility, the review protocol was developed before the literature search. The protocol outlines the objectives of the review, the research queries, the sources of data, search strategy, including and exclusion criteria, and the way the analysis will be conducted. All the review activities follow the PRISMA 2020 guidelines that guarantee transparency in the identification, screening, and eligibility assessment as well as the

inclusion of studies. This reduces subjectivity in the study and this will also provide consistency in the systematic review of the research.

The main aim of this SLR is to conduct research analysis, technical methods, security assurances, and performance compromises within encrypted query processing framework, especially in multi-cloud or distributed cloud computations.

2.2. Research Questions

In order to systematically review the available literature and respond to the goals of the given systematic literature review, the following research questions (RQs) are developed:

- RQ1: What are the dominant techniques and architectures proposed for secure and privacy-preserving encrypted query processing in cloud and distributed computing environments?
- RQ2: How do existing approaches balance security, privacy, computational efficiency, and query performance when operating on encrypted data?
- RQ3: What cryptographic primitives, access control mechanisms, and system models are most commonly employed to protect data confidentiality during query execution?
- RQ4: What key limitations, research gaps, and open challenges remain in current encrypted query processing solutions, particularly with respect to scalability, real-time processing, and practical deployment?

2.3. Search Strategy

To identify relevant studies, a structured and comprehensive search strategy was used to access the required studies in the selected digital libraries. Searching based on the key words was performed using the combination of the words that are associated with the privacy preservation, encrypted query processing, cloud computing, and distributed architectural arrangements. The search results were refined with some Boolean operators (AND, OR) so that a wide range of terminologies might be covered. The search strategy has been narrowed down to achieve completeness and relevance.

2.4. Data Sources

The literature search was done in several digital repositories that have been well established to cover. These are IEEE Xplore, ACM Digital Library, SpringerLink, Elsevier, Wiley Online Library and Google Scholar. These databases have been chosen because of their large data on peer-reviewed journal articles in cloud computing, data security, and privacy-preserving systems.

2.5. Inclusion Criteria

Studies were incorporated in the SLR whenever they met the following criteria. The studies that had a definite focus on applied encrypted query processing were considered.

- Published in peer-reviewed journals or conferences.
- Focused on privacy-preserving query processing over encrypted data.
- Addressed cloud, distributed cloud, or multi-cloud environments.
- Proposed or evaluated practical systems, frameworks, algorithms or hybrid approaches.
- Written in English.
- Provided sufficient technical details and experimental or analytical evaluation.

2.6. Exclusion Criteria

Studies were excluded based on the following criteria:

- Non-English publications.
- Works unrelated to cloud or multi-cloud environments (e.g., purely local or edge-only systems).
- Pure cryptographic theory papers without application to query processing or cloud systems.
- Survey papers, review articles, editorials, or tutorial papers.
- Short abstracts, posters, or non-peer-reviewed manuscripts.

These criteria ensured that the final dataset focused strictly on applied research contributions relevant to the objectives of this study.

2.7. Data Preprocessing and Screening Process

All records collected were considered into one dataset and any redundant records eliminated before screening. Screening was done in two phases: the first phase involved a review of titles and abstracts to determine relevancy and then a review of the full text carried out to determine eligibility. At every stage, studies, which failed to meet the inclusion criteria, were filtered out. This was a systematic screening mechanism that brought out uniformity and minimized the chances of irrelevant studies being reflected in the final analysis.

2.8. Data Extraction Strategy

A systematic data extraction strategy was employed to identify appropriate information in every study. The information extracted entailed publication information, techniques suggested, encryption, supported query, model of cloud deployment, security assumptions, performance metrics, and limitations identified. Structured extraction made it possible to compare results across studies in the same way.

2.9. Threats to Validity

In this review, a number of threats to validity were taken into consideration. There could be publication bias because the authors are relying on peer-reviewed literature. The predetermined inclusion and exclusion criteria and the PRISMA guidelines minimized selection bias. Also, differences in experiment design and measurement of results in different studies can be a source of direct comparison. These risks were reduced by paying close attention to screening, quality evaluation, and reporting of the review.

3. Core Principles of Secure Query Processing over Encrypted Big Data

The fast development of cloud data analytics and data outsourcing has also brought about serious concerns on the issue of data confidentiality, query privacy, and also leakage resilience. Encrypted on big data have become a paradigm to support the storage and processing of sensitive data and allow the computations implementation with acceptable performance at high levels of security. This section gives the fundamental conceptual needed to comprehend privacy query processing with encrypted information, system models, query types, and privacy leakage designs.

3.1. Encrypted Big Data Systems

Encrypted Big data systems aim to enable the storage, sharing, and execution of queries in untrusted or semi-trusted systems like multi-cloud systems and public clouds. Contrary to the previous methods of encryption that encrypt data only at rest or transit, these systems allow computing directly on encrypted data, thus reducing the amount of exposure to data in processing (Hashi & Hashi, 2026). Moreover, cryptographic primitives, combined with distributed storage and parallel processing systems, are used to ensure that encrypted big data systems can scale to high-volume and high-velocity datasets. This provision of scalability poses a challenge of key management, encrypted indexing, and workload distribution among various cloud providers (Li et al., 2024).

3.1.1. Data Models

Encrypted Big data systems with various data models depending on the application area and query needs. The most widespread data models are: Relational data models, in which structured tables are encrypted either attribute-wise or tuple-wise to support SQL-like queries, such as selection, projection and joins (Kim et al., 2022; Chen et al., 2025). Multi-dimensional and spatial data models, which are common in the IoT, healthcare, and location-based services, where the encrypted attributes are coordinates or feature vectors (Basudan & Alamer, 2024; Wang et al., 2025). Document-based and keyword-centric models, which are encrypted textual documents and searchable by key words, are ranked models of retrieval, fuzzy matching, and search (Ma et al., 2025; Chen et al., 2025).

More secure multi-party query execution Graph and federated data models, which are rapidly used in collaborative analytics in distributed organizations (Aljuaid et al., 2025). In order to enforce these data models, systems frequently use hybrid encryption designs that integrate searchable encryption, homomorphic encryption (HE), and secure indexing of data (Li et al., 2024; Bian et al., 2023). Newer systems like hybrid encrypted databases dynamically select encryption schemes depending on the complexity of queries and performance limits (Li et al., 2024).

3.1.2. Threat Models

The threat model is the perceived adversarial capability and trust assumptions of encrypted big data systems. The surveyed works mostly follow one or a combination of the following threat models: Honest-but-curious (semi-honest) cloud servers, which effectively implement protocols but strive to extract sensitive information based on encrypted data or queries and access patterns (Xu et al., 2024). Malicious who can bend the rules, inject wrongful answers, or cooperate with other parties need increased cryptographic assurances and verifiability features (Li et al., 2024). Multi-user and multi-cloud adversaries, in which two or more users with various access control privileges access replicated or dispersed encrypted data, which amplifies the attack surface (Liu et al., 2022; Bharot et al., 2025). Encrypted big data systems to counter such threats can be installed using fine-grained access control and verification schemes as well as multi-key encryption schemes to guarantee confidentiality, integrity, and accountability (Zhou et al., 2024; Li et al., 2024).

3.2. Query Types

The main issue is to support the expressive and high-quality query processing of encrypted data. The literature identifies several basic types of encrypted queries that have different security and performance trade-offs.

3.2.1. Selection Queries

Selection queries are used to retrieve records that match certain predicates. Deterministic encryption, order-preserving encryption, or searchable encryption schemes are typical in encrypted environments in that they support these queries (Almakdi et al., 2021). Although effective, this scheme can be leakage/order-informed, requiring leakage-aware design (Yuan et al., 2025). The recent solutions use lightweight direct cryptographic primitives in IoT and healthcare system to allow the choice of fast encryption with access control (Zhou et al., 2024; Ma et al., 2025).

3.2.2. Range Queries

One of the most widely studied types of encrypted queries are range queries because these are used in analytics, spatial databases, and in monitoring system applications. Encryption of data is also not very easy to support range predicates because naive encryption obliterates order relationships. In this respect, the current solutions utilize methods like encrypted tree structure, secured learned index, and homomorphic comparison protocols (Basudan and Alamer, 2024; Wang et al, 2025). The more sophisticated schemes also prevent leakage by blur query limits and reducing access patterns (Guo et al., 2022; Yan et al., 2025). Upcoming studies are also looking at quantum-assisted privacy-preserving range queries, and indicates for the future of post-quantum secure query processing (Ye et al., 2024).

3.2.3. Aggregation Queries

Data analytics and decision making are inseparable with aggregation queries such as SUM, COUNT, and AVERAGE. Fully homomorphic encryption can only achieve accurate aggregation of encrypted values but can be computationally expensive (Hashi & Hashi, 2026). In order to enhance efficiency, hybrid schemes combine partial homomorphic encryption and secure aggregation protocols and verifiable computation (Shi et al., 2024). The methods are especially useful in federated learning and analytics in the IoT, where the aggregated knowledge is needed without exposing the specific data points (Shen et al., 2024).

3.2.4. Join Queries

Encrypted join queries are computationally heavy in that matching encrypted attributes across datasets. There is secure join algorithms suggested in the prior work based on multi-party computation and encrypted equality tests (Kim et al., 2022). Recent frameworks propose adaptive query execution algorithms that will selectively decrypt or transform encrypted data to maximize the performance of joins without assuring confidentiality guarantees (Li et al., 2024; Chen et al., 2025).

3.2.5. Analytical Queries

The complex operations included in analytical queries are clustering, ranking, skyline queries, and machine learning inference. To back up these queries, higher cryptographic tools, such as fully homomorphic encryption and secure multi-party computation. Privacy-sensitive analytics systems with the assistance of clouds can perform clustering and ranking of encrypted data in a scaling manner under a multi-user environment (Xu et al., 2024; Chen et al., 2025). These systems compromise between analytical expressiveness and execution latency that is acceptable to practice in the real world.

3.3. Privacy Leakage Models

Although the content of data is safeguarded by encryption, the encrypted query processing systems are always prone to the leakage of auxiliary information in practice. The leakage is critical in the assessment of system security.

3.3.1. Data Leakage

Data leakage refers the unintended disclosure of plaintext information via encryption schemes, e.g. showing equality patterns or order relationships. Such leakage is especially susceptible to deterministic and order preserving encryption (Falzon et al., 2022). The latest researches intend to reduce the risk of data leakage with the aid of combining randomized encryption, re-encryption, and differential privacy (Huang, 2025). Hybrid cryptography systems also minimize leakage through dynamically changing the strength of encryption depending on the sensitivity of queries (Ting and Li, 2025).

3.3.2. Query Leakage

Query leakage occurs when the cloud server gets to know details regarding query predicates, frequency or structure. Keyword search scheme will leak search patterns, and it can be inferred through inference attacks with time (Ma et al., 2025). As a response to this, leakage-resilient searchable encryption schemes apply query obfuscation, dummy queries, and probabilistic trapdoors (Yuan et al., 2025; Fugkeaw et al., 2024). Ranked and multi-keyword search schemes are additional schemes that include secure scoring schemes to avoid relevance inference (Chen et al., 2025; Yan et al., 2022).

3.3.3. Access Pattern Leakage

Access patterns leakage may be used to determine sensitive correlations amongst encrypted records accessed during query execution. This type of leakage can be regarded as one of the most serious threats in encrypted databases (Xu et al., 2024). Some of the mitigation techniques are oblivious RAM (ORAM), secure shuffling, and multi-party execution, yet they come at the cost of additional computing and communication overhead (Pleșa and Olimid, 2024). The latest systems seek to achieve a balance between leakage suppression and efficiency by only protecting access patterns on high-risk queries (Basudan and Alamer, 2024; Wang et al., 2025). Table 1 provides the brief descriptions of the representative research studies, which define the conceptual background of encrypted big data systems, query typology, and privacy concerns.

Table 1 Summary of Conceptual Foundations in Privacy-Preserving Encrypted Query Processing

Author (Year)	System Focus	Core Technique	Supported Query Types	Key Contribution
Hashi & Hashi, (2026)	Encrypted big data analytics	Fully Homomorphic Encryption (FHE)	Aggregation, analytics	Demonstrated feasibility of secure large-scale analytics over encrypted data
Guo et al. (2022)	Secure cloud data outsourcing	Secure index-based encryption	Range queries	Reduced leakage while supporting efficient multi-range queries
Kim et al. (2022)	Encrypted database querying	Secure computation protocols	Top-k, kNN queries	Improved efficiency of ranked queries over encrypted databases
Bian et al. (2023)	Encrypted database systems	Arithmetic-and-logic FHE	Selection, analytical queries	Proposed an elastic encrypted database supporting complex operations
Chen et al. (2025)	Cloud-based query processing	Secure skyline query framework	Multi-dimensional analytical queries	Enabled efficient and secure skyline queries in cloud environments

4. Taxonomy of Privacy-Preserving Query Processing Techniques

Various cryptographic and system-level paradigms have developed for privacy-preserving query processing with encrypted big data at a trade-off among security, efficiency, scalability and query expressiveness. Depending on the

underlying defense mechanisms and model of execution, the existing literature can be broadly grouped into five categories: homomorphic encryption-based, searchable encryption-based, secure multi-party computation, trusted execution environment and hybrid ones. This taxonomy helps to decompose modern systems to facilitate safe querying operations in cloud and multi-cloud systems and avoid the threat of data exposure (Hashi & Hashi, 2026).

4.1. Homomorphic Encryption-Based Approaches

Homomorphic encryption (HE) that allows computations to be directly carried out with encrypted data without having to be decrypted, thus it is one of the back bone privacy-preserving query processing techniques in untrusted cloud. HE-based systems are especially appealing to encrypted analytics, aggregation, and complicated numerical calculations over large data (Hashi & Hashi, 2026). Fully Homomorphic Encryption (FHE) schemes can be used to make arbitrary computations on encrypted data (ciphertexts), whereas Partially Homomorphic Encryption (PHE) schemes support only restricted computations (addition or multiplication). Such research as HE3DB illustrates how arithmetic-and-logic FHE can be incorporated into encrypted database systems to enable a scale-based selection and analytical queries (Bian et al., 2023). In the same manner, Silca presents the use of encrypted caching to minimize the FHE computation overhead when outsourcing to a database (Zhao, 2023).

Although the privacy is high, HE-based methods are associated with high computational cost and latency which complicates the real-time processing of queries. Scalability is a key problem, particularly when working with multi-clouds where encrypted operations need to be synchronized between the distributed nodes (Hashi & Jama, 2025; Huang, 2025). The latest work is based on the idea of integrating HE with optimization techniques or differential privacy to enhance efficiency without reducing the high security assurance (Huang, 2025).

4.2. Searchable Encryption Techniques

Searchable encryption (SE) allows to search encrypted data efficiently using a keyword query and also to perform structured queries over encrypted data revealing very little information to the cloud server. Depending on key management and trust assumptions, SE techniques are widely divided into symmetric searchable encryption (SSE) and public-key searchable encryption (PKSE). The use of SSE schemes is very popular as it is efficient and it can support large scale cloud storage. Fugkeaw et al., (2024) and Dabra et al. (2024) present both Boolean and fine-grained keyword search mechanisms, which are compatible with encrypted cloud data warehousing and secure cloud storage. Powerful constructions also allow ranked search and multi-keyword search, which enhance query expressiveness (Chen et al., 2025; Yan et al., 2022).

Range queries are a more difficult task because there is ordering leakage. Basudan and Alamer (2024) and Guo et al. (2022) research addresses: it minimizes leakage, but still promotes efficient encrypted range searches. Nonetheless, the vast majority of SE schemes continue to reveal the patterns of access and search frequency, which can be used to conduct inference attacks, in particular in multi-cloud deployments (Xiong & Luo, 2024).

4.3. Secure Multi-Party Computation (SMPC)

Secure Multi-Party Computation allows different cloud providers or data owners to perform queries in a collaborative way on distributed encrypted data without disclosing their respective inputs. SMPC would especially be applicable to federated and multi-cloud environments that have decentralized trust (Liu et al., 2022). A typical SMPC methodology is secret sharing, in which data is broken down into shares and sent to numerous cloud servers that are non-colluding. Systems like Cloud-assisted secure data fusion frameworks show how SMPC is used to facilitate secure analytics in more than one cloud used on sensitive applications like healthcare and infectious disease analysis (Liu et al., 2022). SMPC is also extended to complex query structures through federated querying over graph databases (Aljuaid et al., 2025).

Although SMPC has a good privacy guarantee, the overhead of communication and synchronization cost is very high, which disrupts the performance severely. The more parties are involved, the higher the query latency, and this is why big data workloads are limited in terms of scalability (Van Kenhove et al., 2026).

4.4. Trusted Execution Environments (TEE)

Trusted Execution Environments like Intel SGX Hardware-based isolation is used to execute queries on encrypted data with security. TEEs enable the computation of sensitive functions within secure enclaves, which has less cryptographic overhead than HE and SMPC-based systems. There have been applications of TEE-based systems to facilitate effective processing of queries and analytics on encrypted cloud data by running decrypted data within a set of trusted hardware. Nonetheless, memory side-channel attacks, enclave memory constraints, or hardware-based trust assumptions are severe security threats (Xu et al., 2024).

Recent research gives TEEs due credit that they can enhance performance, but is not yet ready to claim that it should be used on its own as a privacy-preserving query processing method, especially in adversarial multi-cloud environments where both hardware compromise and collusion can occur.

4.5. Hybrid Approaches

Hybrid methods are the combination of various privacy methods to achieve compromise between security, efficiency, and scalability. Examples of typical combinations are HE and TEE, SMPC and encryption, and encryption and blockchain based auditability. As an example, Enc^{2}DB Enc 2 DB proposes a hybrid architecture of encrypted query processing, i.e. a system that dynamically picks encryption tools depending on query type and workload properties (Li et al., 2024). Systems like Cloudlock that are multi-cloud storage systems incorporate hybrid cryptosystems to enhance the data sharing in a heterogeneous cloud provider (Bharot et al., 2025). The querying that is supported by blockchain is optional, yet it introduces enforceable access control and immutable logging, increasing the confidence towards the multi-cloud environment. Hybrid systems are also considered the most viable path to the real-world implementation because they alleviate the shortcomings of each approach but ensure high privacy levels (Ting and Li, 2025).

Table 2 Taxonomy of Privacy-Preserving Query Processing Techniques

Author (Year)	Technique Category	Encryption / Security Mechanism	Supported Query Types	Key Limitations
Hashi & Hashi, (2026)	Homomorphic Encryption	Fully Homomorphic Encryption (FHE)	Aggregation, Analytics	High computational overhead, latency
Fugkeaw et al. (2024)	Searchable Encryption	Symmetric searchable encryption	Boolean keyword search	Access-pattern leakage
Liu et al. (2022)	SMPC	Secret sharing across clouds	Aggregation, Analytics	High communication overhead
Xu et al. (2024)	TEE-based	Cloud-assisted secure enclave execution	Analytical queries	Side-channel attack risks
Li et al. (2024)	Hybrid	HE + adaptive query optimization	Selection, Range, Analytics	System design complexity
Bharot et al. (2025)	Hybrid (multi-cloud)	Hybrid cryptosystem	Keyword search, Data sharing	Deployment overhead
Ting & Li (2025)	Hybrid	Encryption + multilayer security	Storage and basic queries	Limited query expressiveness
Basudan & Alamer (2024)	Searchable Encryption	Leakage-suppressed range encryption	Range queries	Increased query execution cost

Table 2 presents a summary of notable studies in each key groups of privacy-preserving query processing, including the security mechanisms, the types of queries, and the limited nature of such query processing.

5. Multi-Cloud Architectures for Encrypted Query Processing

Multi-cloud adoption has become a strategic implementation to overcome scalability, availability and trust limitations that are inherent to single cloud deployments. Privacy-preserving query processing distributing encrypted big data among two or more cloud providers can provide better fault tolerance, less vendor lock-in, and more security due to trust separation. Nonetheless, querying encrypted data in these types of distributed environments presents considerable architectural, cryptographic and co-ordination difficulties. The studies in this field have been paying more attention to the development of secure multi-cloud query process frameworks in order to strike a balance between protection of privacy, efficiency of computation, and the complexity of the system (Liu et al., 2022; Bharot et al., 2025).

5.1. Data Distribution Models

Data distribution forms a core part in the multi-cloud encrypted query processing and has a direct effect on the privacy leakage, query efficiency and the scalability of the system. Horizontal and vertical partitioning are two prevailing data partitioning strategies that are actively developed in literature. Under horizontal partitioning, the records partition the

datasets and are distributed and stored with numerous cloud providers, where each cloud holds a partial portion of encrypted tuples. The given method is especially efficient in high workloads related to analytical tasks as it makes it possible to execute queries simultaneously and restrict data leaks on a single cloud. Research on safe data fusion and analytics indicates that the encrypted data can be horizontally partitioned and allow aggregation and statistical queries through the application of SMPC or HE methods (Hashi & Hashi, 2026).

Conversely, vertical partitioning spreads data attributes among various cloud providers such that no cloud provider holds complete records. This model is also robust with regards to privacy as it struggles attribute correlation attacks and is highly applicable in the sensitive field, like healthcare and industrial IoT. Multi-group and hierarchical data sharing systems embrace vertical partitioning to implement a fine-grained access control system and to support the execution of encrypted queries across clouds (Li et al., 2023; Bharot et al., 2025). Although the privacy of vertical partitioning is very high, the reconstruction cost of queries and communication overhead are higher in the process of join and analysis. It has also been suggested that hybrid partitioning strategies be deployed that will make use of horizontal and vertical strategies, which will optimize both performance and privacy. Dynamically partitioning models are used in adaptive encrypted query processing frameworks, which can choose the type of query and the level of sensitivity, which is more effective in multi-cloud settings (Li et al., 2024).

5.2. Trust and Threat Assumptions

The design of secure multi-cloud query processing systems is based on trust modeling. The vast majority of the current methods presuppose the honest-but-curious (HbC) cloud providers that are loyal to protocols but guess the sensitive data about stored data and observed queries. Based on this premise, the use of encryption and secure computation methods is made to guarantee that cloud providers do not know more than they should by means of allowed leakage (Basudan and Alamer, 2024; Guo et al., 2022). Although HbC model makes designing a system easier, there are also possibilities of collusion in the real-life deployment, where there are two or more cloud provider entities, and they join their collaboration to deduce confidential information. The resistance to collusion is especially important in vertically partitioned systems, where a privacy mechanism of attribute-level separation is applied. Secrecy multicompany and secret-sharing-based schemes spread trust among many non-colluding clouds such that privacy is not violated unless a predetermined number of clouds collude together (Aljuaid et al., 2025; Kamal et al., 2025).

Other recent works relax the threat model to incorporate malicious adversaries and tend to deal with those threats like protocol deviation, result manipulation, and integrity attacks. The schemes of verified encryption and integrity-check mechanisms are proposed to identify the dishonest behavior in the context of multi-cloud storage and query execution (Li et al., 2024; Fugkeaw et al., 2024). Nevertheless, more robust threat models will always raise the cost of computation and communication, which points to a natural trade-off between the security strength and the efficiency of the system.

5.3. Coordination and Synchronization Challenges

Arranging execution of encrypted queries across a number of cloud providers brings significant synchronization and orchestration issues. The query plans should be broken down into sub-queries which are run on distributed clouds and intermediary encrypted results should be safely added at the client or a trusted coordinator. This process necessitates strict coordination of the process to prevent leakage of information due to timing or communication pattern (Chen et al., 2025). The use of multi-cloud architecture is also characterized by significant communication overhead, particularly with regard to secure join, aggregation and analytical queries. The SMPC-based solutions tend to need several rounds of interaction between the cloud providers, which increases the latency and decreases the scalability of the big data workloads (Liu et al., 2022; Aljuaid et al., 2025). In the same way, approaches based on HE are associated with a synchronization delay because of high ciphertext operations and poor parallelism (Hashi & Hashi, 2026).

The solutions to these dilemmas include hybrid coordination, where new frameworks are suggested to combine encrypted indexing, caching and adaptive query scheduling. HE caching methods minimize redundant calculations and help to increase response time to repeated queries (Zhao, 2023). Moreover, hybrid HE -TEE architectures transfer coordination to secure enclaves so that query orchestration can be done faster without losing confidentiality (Xu et al., 2024). Irrespective of these developments, there is yet to be an efficient, scalable, and leakage-resilient system of coordination in multi-cloud encrypted query processing, especially real-time analytics and at large-scale deployments.

Table 3 Comparison of Multi-Cloud Encrypted Query Processing Architectures

Author (Year)	Data Distribution Model	Trust Assumption	Query Support	Key Challenges Addressed
Hashi & Hashi, (2026)	Horizontal partitioning across clouds	Honest-but-curious, non-colluding	Selection, aggregation	Reduces single-cloud leakage through distributed ciphertext storage
Chen et al. (2023)	Vertical partitioning with attribute separation	Semi-trusted clouds	Equality and range queries	Limits attribute-level inference via column-wise encryption
Zhang et al. (2023)	Hybrid horizontal–vertical partitioning	Honest-but-curious with partial collusion	Join and aggregation queries	Balances performance and leakage under limited collusion scenarios
Li et al. (2024)	Secret-sharing-based multi-cloud distribution	Non-colluding cloud providers	Statistical and analytical queries	Eliminates single-point trust by splitting encrypted data shares
Kumar et al. (2024)	Replicated encrypted datasets across clouds	Honest-but-curious, synchronized clouds	Keyword and Boolean queries	Addresses availability and synchronization overheads

Table 3 represents the literature which studies encrypted query processing in multi-cloud settings, and the difference in approach to data distribution, assumptions of trust, and supported types of queries.

6. Performance, Security, and Privacy Trade-Offs

Processing Big data encrypted privacy preserving queries presents intrinsic trade-offs between computation performance, security and scalability, and practicality. Although more sophisticated cryptography methods like FHE, SE, SMPC)and hybrid approach are much more effective in enhancing confidentiality, they are associated with non-trivial computation, communication and system complexities. This section is a critical discussion of the trade-offs in terms of a synthesis of the results of recent encrypted query-processing systems, specifically in the context of cloud and multi-cloud environments (Hashi & Hashi, 2026; Van Kenhove et al., 2026).

6.1. Computational Efficiency vs Privacy Guarantees

A basic trade-off in encrypted query processing is between computational efficiency and strength with privacy is assured. Methods with good semantic security like FHE allow arbitrary computation on ciphertexts at a high computational cost because of expensive ciphertext algorithms and noise cancelling protocols (Hashi & Hashi, 2026). The systems like HE3DB and Silca show that though such optimizations as arithmetic logic FHE and ciphertext caching can decrease the latency, an FHE-based system is still several orders of magnitude slower than plaintext query processing (Bian et al., 2023; Zhao, 2023).

In contrast, order-preserving encryption (OPE) and deterministic encryption schemes are much faster to execute queries, especially range and equality queries, at the expense of greater leakage of data ordering and frequency (Basudan and Alamer, 2024; Kim et al., 2022). Hybrid systems like Enc^{2}DB strive to trade off this by dynamically choosing the encryption schemes, depending on query sensitivity and workload properties (Li et al., 2024). Figure 2 shows some common encryption methods are visually compared against the number of relative execution time and the level of privacy. Privacy strength and cost of computation are clearly on an upward trend as shown. The findings show that there is no ideally efficient and privacy-deserving solution, which proves the necessity of context-related design decisions.

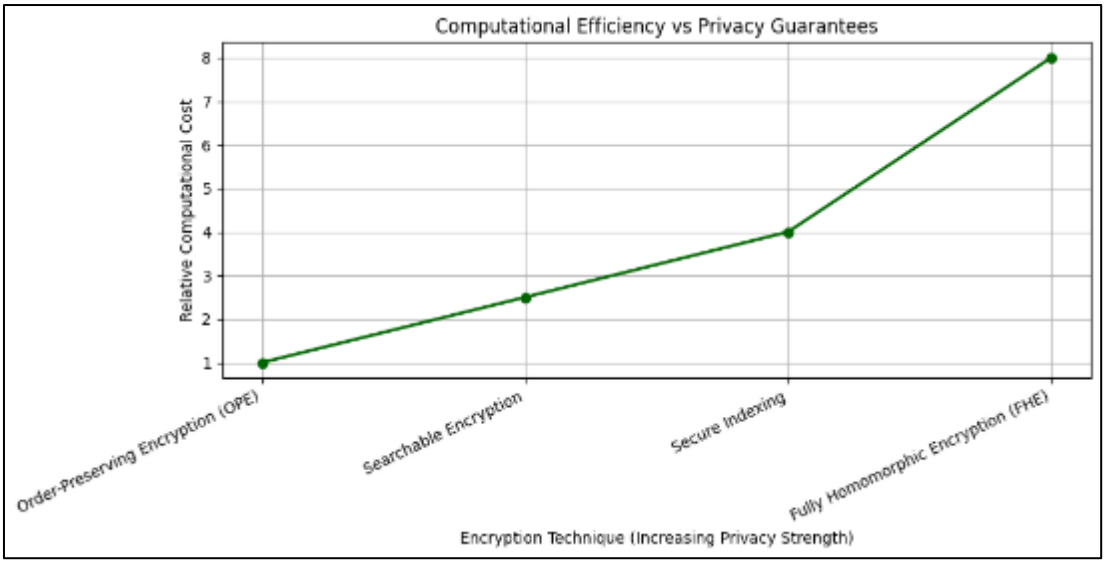

Figure 2 Computational Efficiency vs Privacy Guarantees

6.2. Query Expressiveness vs Security Strength

The other important aspect of trade-off presents itself between the query expressiveness and the strength of security. SSE with a low leakage can be effectively used to support simple encrypted queries, i.e., keyword search and equality selection (Ma et al., 2025; Fugkeaw et al., 2024). Nevertheless, more expressive queries such as ranked search, top-k queries, joins, skylines queries and multi-dimensional range queries can be supported either with either security guarantees or increased cryptographic costs (Chen et al., 2025; Kim et al., 2022).

This balance is further complicated by advanced types of queries, e.g. spatial queries and learned index-based range queries. Although learned encrypted indexes can drastically reduce the query latency, they can cause model-based leakage unless particular care is taken to prevent it (Wang et al., 2025; Li et al., 2024). Likewise, the other queries, such as skyline and aggregation queries, usually use iterative protocols that raise the number of interactions rounds and exposure to the access patterns (Chen et al., 2025; Shi et al., 2024). Figure 3 shows that query expressiveness is inversely correlated with the security strength that may be obtained with suitable queries in encrypted queries categories. Join and aggregation queries have the tendency of exposing access patterns, result sizes or correlation information that adversaries can use in inference attacks.

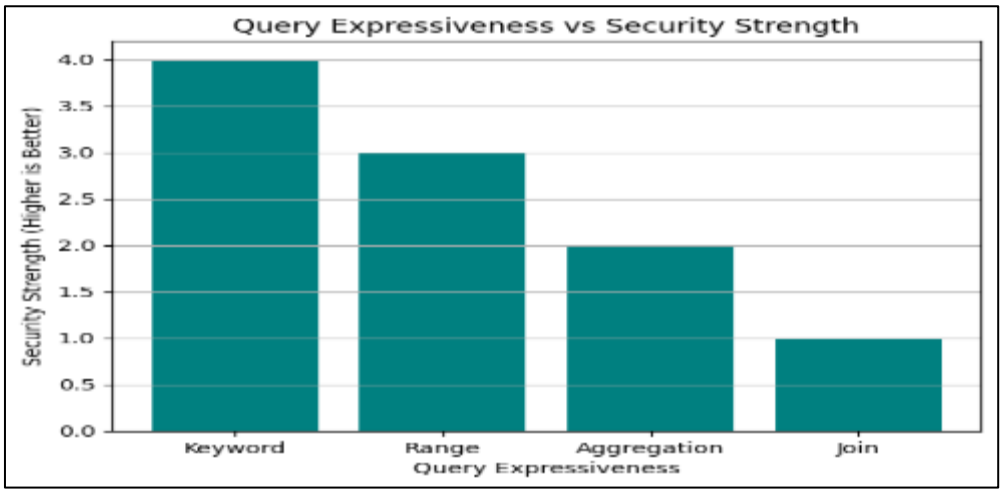

Figure 3 Query Expressiveness vs Security Strength

6.3. Scalability and Communication Overhead

The issue of scalability continues to be a primary concern when implementing privacy-preserving query processing on a bigdata and IoT-based platform. Encryption results in a significant increase in the size of data, which consequently causes increasing storage and communication costs, especially in the distributed and multi-cloud environment (Ramachandra et al., 2022; Liu et al., 2022). The problem is further complicated by SMPC and secret sharing based systems where interaction is repeated multiple times and a synchronization process is necessary (Aljuaid et al., 2025; Kamal et al., 2025).

Some experiments show that communication overhead has become the main bottleneck and not computation as the scale of datasets and query parallelism are scaled up (Guo et al., 2022; Pleșa and Olimid, 2024). Federated and multi-cloud constructions are trying to address this by moving computational power closer to the source of data, but this is adding complexity of coordination and more trust assumptions (Van Kenhove et al., 2026; Shen et al., 2024). Figure 4 treats the growth trends of communication overheads in single-cloud processing architecture, multi-cloud processing architecture. As it is apparent, the communication overhead is also pronounced with the extent of datasize particularly in cases of multi-cloud deployment. The single-cloud architectures have lower coordination cost as compared to multi-cloud where the overhead is more because of inter-cloud synchronization and separation of trust.

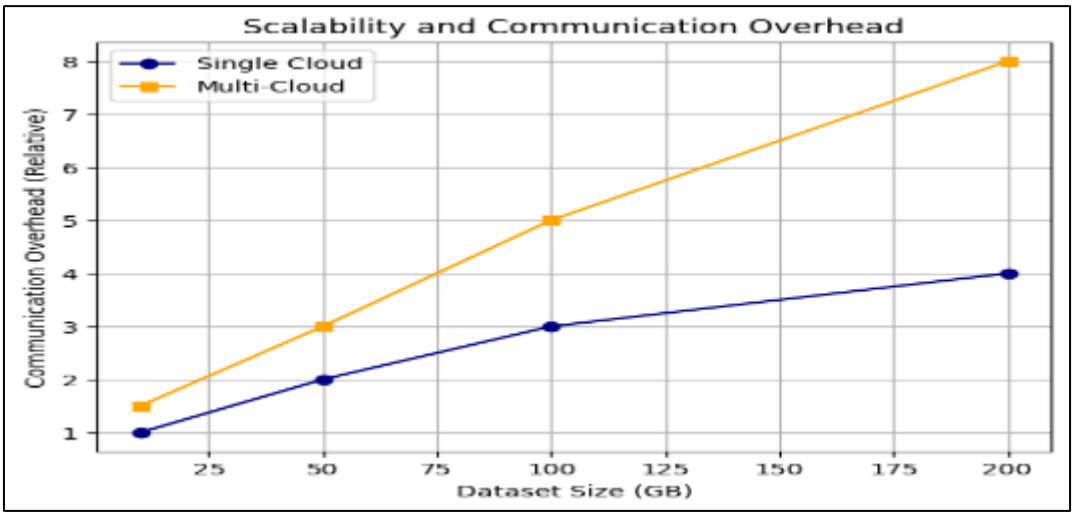

Figure 4 Scalability and Communication Overhead

6.4. Multi-Cloud Trust vs System Complexity

Single-point-of-failures and problems with excessive trust are mitigated by the increasing use of multi-cloud architectures. This type of system allows ensuring resilience to collusion and insider threats by distributing encrypted data and query execution over independent cloud providers (Liu et al., 2022; Bharot et al., 2025). The benefits however are associated with the cost of added complexity to the system such as cross-cloud synchronization, key management issues and increased orchestration overhead.

Empirical analyses show that although multi-cloud systems can improve assumptions of trust, they need advanced coordination protocols, and existent latency penalties are non-negligible (Aljuaid et al., 2025; Van Kenhove et al., 2026). Partially, this complexity is relieved by lightweight access control and policy enforcement mechanisms that do not completely solve it (Hashi, 2025). Figure 5 shows a simplified graphical comparison of the trust enhancement and system complexity between one and multiple cloud encrypted query processing deployment. With the increase in the number of cloud providers, the system is more resistant to collusion although it also becomes harder to control and optimize.

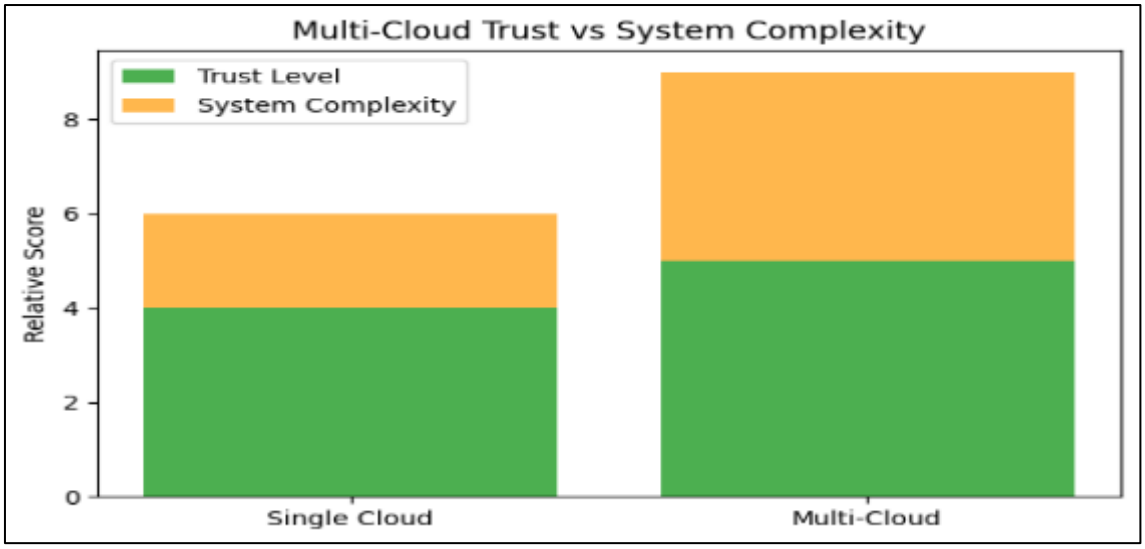

Figure 5 Multi-Cloud Trust vs System Complexity (Conceptual)

In general, available literature indicates that there is no technique that best meets performance, security, and scalability. Rather, more and more practical encrypted query processing systems have hybrid, adaptive and context-aware designs, and this supports the significance of trade-off-driven system engineering instead of cryptographic optimization.

7. Comparative Analysis of Existing Approaches

Systematic review papers need to have a comparative analysis to objectively compare the existing privacy-preserving query processing methods by the security guarantees, the functional expressiveness, the computational feasibility, and the deployment scalability. Although various solutions have been suggested to implement encrypted query execution on cloud and multi cloud systems, there remains no proposal to be able to maintain privacy, performance and systems complexity in a uniform manner. An overall framework of comparisons allows synthesizing and contrasting the representative methods of the literature and identifying their strengths and weaknesses as well as gaps in research (Hashi & Hashi, 2026; Li et al., 2024; Van Kenhove et al., 2026).

7.1. Comparison Metrics

In order to guarantee an fair and organized assessment, the surveyed methods are evaluated comparing them based on five commonly-used metrics.

- **Privacy Level:** Privacy Level will denote the extent to which the content of data, the query semantics and patterns are secure. FHE and systems based on a SMPC can offer high semantic security and adversary resistance (Hashi & Hashi, 2026; Aljuaid et al., 2025). SE and preservation of order schemes, in turn, are moderate with respect to privacy and release controlled auxiliary information like access patterns or frequency of keywords (Ma et al., 2025; Basudan and Alamer, 2024).
- **Supported Queries:** Query support is simpler queries by keyword and equality query to more complex queries like skyline queries, top-k queries, joins, spatial queries and aggregation. Systems that emphasize expressive query support do not necessarily have high privacy guarantees or they have higher computational costs (Chen et al., 2025; Kim et al., 2022; Falzon et al., 2022).
- **Computational Overhead:** The computational overhead can be used to gauge the cost of encryption, query execution, and decryption. FHE-based and SMPC-based solutions are always highly computationally complex and have restricted real-time applicability (Hashi & Jama, 2025; Pleşa & Olimid, 2024). This overhead is reduced by hybrid frameworks by only using heavyweight cryptography on sensitive query components (Li et al., 2024; Bharot et al., 2025).
- **Communication Cost:** Communication cost is the most important in distributed and multi-cloud environment. Secret sharing, federated querying, and collaborative analytics can cause large inter-cloud communication overhead expenses because of the repetition of interaction rounds (Liu et al., 2022; Van Kenhove et al., 2026). The overhead of communication is usually lower in lightweight SE schemes (Fugkeaw et al., 2024; Xiong and Luo, 2024).

- **Scalability:** Scalability is used to measure the system performance when its data volume, query complexity, and number of participants increase. The methods based on adaptive indexing, hierarchical data sharing, and workload-sensitive encryption prove to be more scalable in the bigdata context (Li et al., 2023; Ting and Li, 2025). Nonetheless, powerful cryptographic assurances tend to restrict horizontal scalability (Bian et al., 2023; Yuan et al., 2025).

Table 4 Comparative Analysis of Privacy-Preserving Query Processing Approaches

Technique	Author (Year)	Encryption Type	Query Support	Multi-Cloud Support	Performance & Limitations
Searchable Encryption (SE)	Ma et al., 2025; Fugkeaw et al., 2024	Symmetric / Asymmetric	Keyword search, Boolean queries	Limited	High performance; but leaks access/search patterns
Order-Preserving Encryption (OPE)	Li et al., 2024; Chen et al., 2025	Deterministic	Range queries, sorting	Limited	Very high speed; reveals order info, less secure for sensitive data
Fully Homomorphic Encryption (FHE)	Hashi & Hashi, 2026	Fully Homomorphic	Arbitrary computations	Weak	Very strong privacy; extremely high computation overhead, low practicality
Secure Multi-Party Computation (SMPC)	Aljuaid et al., 2025; Van Kenhove et al., 2026	Secret sharing	Aggregation, joins	Strong	Medium computation; communication & synchronization overhead
Trusted Execution Environments (TEE)	Zhou et al., 2024; Shen et al., 2024	Hardware-assisted	Complex SQL queries	Medium	High performance; side-channel attack vulnerability
Hybrid Encryption Models (SE + OPE + FHE)	Huang, 2025; Li et al., 2024	Mixed	Mixed query workloads	Medium	Medium–high performance; increased system complexity
Blockchain-Assisted Secure Querying	Adnan et al., 2025; Gnana Jeslin & Mohan Kumar, 2022	Cryptographic hashing + encryption	Auditable queries	Strong	Medium performance; storage and consensus overhead

Table 4 contrasts some of the representative encrypted query processing methods based on their encryption mechanisms, query expressiveness, scopes of deployment and performance attributes. In general, the comparative study shows that there is a decisive privacy–performance–scalability trilemma in which high confidentiality requirements are bound to raise a communication cost.

8. Open Challenges and Research Gaps

The processing of encrypted queries has come a long way, but it still has a number of limitations and gaps that limit its practical use. They may be divided into research gaps and typical challenges.

8.1. Research Gaps

- **Query Expressiveness vs Privacy:** The majority of privacy-preserving approaches, including FHE and order-preserving encryption (OPE), are characterized by a high level of privacy and, in most cases, restricted query expressiveness and a significant computation cost (Hashi & Hashi, 2026; Huang, 2025; Li et al., 2024). Multi-dimensional range queries or skyline queries are intricate queries that are difficult to implement without invading privacy (Chen et al., 2025; Ye et al., 2024).

- **Real-Time Analytics and Edge-Cloud Integration:** Existing encrypted processing systems are mostly batch-based and it is challenging to perform real-time analytics on high-velocity data streams, particularly in IoT or edge-cloud applications (Zhou et al., 2024; Van Kenhove et al., 2026; Gnana Jeslin and Mohan Kumar, 2022). Methods of lightweight edge computing integrated with safe cloud aggregation are under investigated.
- **Energy-Efficient Encrypted Computation:** Several homomorphic and hybrid encryption designs are very intensive in computational resources, which makes them inefficient in terms of energy consumption in massive or mobile deployments (Bian et al., 2023; Shen et al., 2024). Research in efficient and adaptive low-power methods is still in a gap.
- **Standardization and Benchmarking:** No standard scale exists used to test encrypted query processing systems. The absence of standardization enables in reproducibility and significant comparison between various systems and datasets (Adnan et al., 2025; Yuan et al., 2025).
- **Post-Quantum Preparedness:** The modern encryption algorithms are susceptible to quantum attack in the future. The field of post-quantum cryptography encrypted query processing has not yet received a substantial amount of research (Ye et al., 2024; Bian et al., 2023).

8.2. Typical Challenges

- **Side-Channel Leakage:** Even with secure computation and TEE, sensitive data can leak via timing, memory access, or cache behavior (Zhou et al., 2024; Shen et al., 2024). OPE schemes, while efficient for range queries, inherently leak order information, making them susceptible to inference attacks (Li et al., 2024).
- **Multi-Cloud Coordination and Trust:** Deployments across heterogeneous clouds face challenges in coordination, synchronization, and trust assumptions. Honest-but-curious or colluding cloud providers can compromise privacy if the protocols are not robust (Liu et al., 2022; Aljuaid et al., 2025).
- **Integration Complexity:** Combining multiple privacy-preserving techniques (e.g., FHE + TEE, SMPC + encryption) adds significant architectural and computational complexity, often limiting adoption in industrial applications (Huang, 2025; Xu et al., 2024; Ma et al., 2025).
- **Regulatory and Compliance Barriers:** Ensuring GDPR, HIPAA, and other privacy compliance while performing encrypted analytics is non-trivial, especially when multi-cloud or cross-border data is involved (Gnana Jeslin & Mohan Kumar, 2022; Adnan et al., 2025).
- **Scalability:** Large-scale datasets, typical in IoT or cloud environments, impose high computational and communication costs on existing encrypted query processing techniques (Van Kenhove et al., 2026; Zhou et al., 2024). Efficient methods to maintain privacy without sacrificing scalability remain a challenge.

Future Research Directions

Several new directions that have potential to improve encrypted query processing. AI-based encrypted querying is also being recognized as an approach to query optimization, workload optimization, and increasing the trade-off between privacy and efficiency (Hashi & Jama, 2025; Xu et al., 2024). With the help of machine learning on encrypted or masked features, systems are able to automatically adjust query strategies without leakage. Another is federated encrypted analytics, which allows the distributed computation of encrypted data, with no centralization of sensitive data (Van Kenhove et al., 2026; Shen et al., 2024). A federated learning combined with SMPC enables joint analytics without infringing privacy across organizations or cloud providers.

The imminent danger of quantum computing is the reason to investigate post-quantum cryptography to encrypted query processing (Ye et al., 2024; Bian et al., 2023). The classical HE is potentially susceptible, and the inclusion of quantum resistant algorithms will be essential towards long term data protection. Lastly, edge-cloud encrypted analytics is a research direction of the future, especially in IoT and time sensitive applications (Zhou et al., 2024; Gnana Jeslin and Mohan Kumar, 2022). Lightweight edge-computation methods that can be aggregated in a secure cloud can minimize latency, energy usage and network bottlenecks, in the process providing real-time privacy-preserving analytics in distributed settings.

9. Conclusion

This systematic review comprehensively examined privacy-preserving query processing techniques for encrypted big data in multi-cloud environments. By following the PRISMA 2020 framework, the review synthesized findings from a rigorously selected set of peer-reviewed papers, providing a structured and unbiased assessment. The study classified existing solutions into major technical categories, including HE–based methods, SE schemes, SMPC, TEE, and hybrid approaches, highlighting their respective strengths and limitations. The analysis revealed a persistent trade-off between privacy guarantees, query expressiveness, and system performance, particularly in large-scale and multi-cloud

deployments. While recent approaches demonstrate improved efficiency and broader query support, challenges such as scalability, leakage resilience, side-channel vulnerabilities, and coordination overhead across multiple cloud providers remain largely unresolved. Additionally, the lack of standardized benchmarks and real-world evaluations limits fair comparison among competing techniques. Overall, this review offers a consolidated reference for researchers seeking to understand encrypted query processing in multi-cloud settings critical for enabling practical, scalable, and secure encrypted data processing in industrial, healthcare, and IoT applications. The identified research gaps and future directions emphasize the need for more practical, scalable, and adaptive solutions that balance strong privacy protection with operational efficiency, thereby enabling trustworthy encrypted analytics for next-generation cloud applications.

References

[1] Hashi, A. I. (2025). The Impact of Mobile Money Transfer in Somalia. Journal of Technology and Systems, 7(3), 30–46. https://doi.org/10.47941/jts.2728

[2] Fugkeaw, S., Suksai, P., & Hak, L. (2024). SSF-CDW: achieving scalable, secure, and fast OLAP query for encrypted cloud data warehouse. *Journal of Cloud Computing*, *13*(1), 129.

[3] Khan, A. N., Naveed, A., Mehmood, A., Arora, D., Khan, A. U. R., & Ali, J. (2025). BloomSec: Scalable and privacy-preserving searchable encryption for cloud environments. *PLoS One*, *20*(12), e0336944.

[4] Miao, Y., Yang, Y., Li, X., Wei, L., Liu, Z., & Deng, R. H. (2023). Efficient privacy-preserving spatial data query in cloud computing. *IEEE Transactions on Knowledge and Data Engineering*, *36*(1), 122-136.

[5] Fugkeaw, S., & Deevijit, J. (2025). Se-collab: Achieving fine-grained and efficiently verifiable searchable encryption with boolean multi-keyword search for collaborative iiot data sharing. *IEEE Access*.

[6] Joshi, N. S., Sambrekar, K. P., Patankar, A. J., Rajawat, A. S., & Muqeem, M. (2025). Advanced Security and Privacy in Cloud Computing: Enhancing Data Protection with Multikeyword Ranked Search in Encrypted Environments. *Scalable Computing: Practice and Experience*, *26*(1), 467-489.

[7] Zhu, X., Shen, P., Dai, Y., Xu, L., & Hu, J. (2024). Privacy-preserving and trusted keyword search for multi-tenancy cloud. *IEEE Transactions on Information Forensics and Security*, *19*, 4316-4330.

[8] Li, C., Wang, M., Xiong, L., & Zhang, X. (2025). A Security-Oriented Privacy-Preserving Framework for Efficient Medical Record Search in Telemedicine.

[9] Xu, L., Cheng, X., Tian, W., Wang, H., & Zhang, Y. (2024). Cloud-Assisted Privacy-Preserving Spectral Clustering Algorithm Within a Multi-User Setting. *IEEE Access*, *12*, 75965-75982.

[10] Li, T., Zhang, J., Shen, Y., & Ma, J. (2023). Hierarchical and multi-group data sharing for cloud-assisted industrial internet of things. *IEEE Transactions on Services Computing*, *16*(5), 3425-3438.

[11] Chen, P., Xu, B., Li, H., Wang, W., Peng, Y., Bhowmick, S. S., ... & Cui, J. (2025). An efficient framework for secure dynamic skyline query processing in the cloud. *Data Science and Engineering*, *10*(1), 54-74.

[12] Hashi, A. I., & Hashi, A. M. (2026). A Deep Learning Driven Cloud Edge Intelligence Framework for Real-Time Big Data Based Cyber-Security Threat Detection. International Journal of Computing and Engineering, 8(1), 13–44. https://doi.org/10.47941/ijce.3447

[13] Hashi, A. I., & Jama, M. A. (2025). Evaluating the Performance of AI-Based Software Tools in Intelligent Decision-Making Systems. International Journal of Computing and Engineering, 7(22), 1–20. https://doi.org/10.47941/ijce.3315

[14] Ma, J., Peng, T., Bei, G., Waqas, M., Alasmary, H., & Chen, S. (2025). Efficient privacy-preserving conjunctive searchable encryption for Cloud-IoT healthcare systems. *ACM Transactions on Privacy and Security*, *29*(1), 1-27.

[15] Chen, H., Tan, S., Ma, X., Lin, X., & Yao, Y. (2025). Multi-Keyword Ranked Search on Encrypted Cloud Data Based on Snow Ablation Optimizer. *Symmetry*, *17*(7), 1043.

[16] Li, W., Susilo, W., Xia, C., Huang, L., Guo, F., & Wang, T. (2024). Secure data integrity check based on verified public key encryption with equality test for multi-cloud storage. *IEEE transactions on dependable and secure computing*, *21*(6), 5359-5373.

[17] Liu, J., Zhang, C., Xue, K., & Fang, Y. (2022). Privacy preservation in multi-cloud secure data fusion for infectious-disease analysis. *IEEE Transactions on Mobile Computing*, *22*(7), 4212-4222.

[18] Ting, T., & Li, M. (2025). Enhanced secure storage and data privacy management system for big data based on multilayer model. *Scientific Reports*, *15*(1), 32285.

[19] Huang, Y. (2025). Research on Cloud Data Security Computing Framework Based on Fusion of Homomorphic Encryption and Differential Privacy. *Journal of Cyber Security and Mobility*, *14*(4), 927-954.

[20] Fugkeaw, S., Hak, L., & Theeramunkong, T. (2024). Achieving secure, verifiable, and efficient Boolean keyword searchable encryption for cloud data warehouse. *IEEE Access*, *12*, 49848-49864.

[21] Dabra, M., Sharma, S., Kumar, S., & Min, H. (2024). An improved finegrained ciphertext policy based temporary keyword search on encrypted data for secure cloud storage. *Scientific Reports*, *14*(1), 5264.

[22] Zhao, D. (2023). Silca: Singular Caching of Homomorphic Encryption for Outsourced Databases in Cloud Computing. *arXiv preprint arXiv:2306.14436*.

[23] Zhou, W., Wang, N., Liu, Z., Fu, J., Deng, L., & Wu, Q. (2024). Privacy-preserving IoT data retrieval scheme with lightweight fine-grained access control in cloud computing. *IEEE Internet of Things Journal*.

[24] Aljuaid, N., Lisitsa, A., & Schewe, S. (2025). Fast and Secure Multiparty Querying over Federated Graph Databases. *SN computer science*, *6*(8), 982.

[25] Wang, G., Zeng, Q., Shen, L., Ding, S., He, X., Zhai, Z., ... & Shi, Z. (2025). Towards Efficient Privacy-Preserving Keyword Search for Outsourced Data in Intelligent Transportation Systems. *Future Generation Computer Systems*, 108192.

[26] Basudan, S., & Alamer, A. (2024). Efficient Privacy-preserving Range Query with Leakage Suppressed for Encrypted Data in Cloud-based Internet of Things. *IEEE Access*.

[27] Kamal, A. A. A. M., Okada, M., & Fujisawa, M. (2025). Privacy-Preserving Keyword Search With Access Control for Secret Sharing-Based Data Outsourcing. *IEEE Access*.

[28] Li, H., Shi, J., Tian, Q., Li, Z., Fu, Y., Shen, B., & Tu, Y. (2024, July). E nc 2 DB: A Hybrid and Adaptive Encrypted Query Processing Framework. In *International Conference on Database Systems for Advanced Applications* (pp. 54-70). Singapore: Springer Nature Singapore.

[29] Wang, Z., Lu, J., Wu, J., Tian, Y., Song, W., Li, Q., & Zhang, D. (2025). Towards Privacy-Preserving Range Queries with Secure Learned Spatial Index over Encrypted Data. *arXiv preprint arXiv:2512.03669*.

[30] Bharot, N., Mehta, N., Breslin, J. G., & Verma, P. (2025). Cloudlock: secure data sharing using a hybrid cryptosystem in multi-cloud data storage. *Cluster Computing*, *28*(7), 464.

[31] Shen, J., Zhao, Y., Huang, S., & Ren, Y. (2024). Secure and flexible privacy-preserving federated learning based on multi-key fully homomorphic encryption. *Electronics*, *13*(22), 4478.

[32] Ramachandra, M. N., Srinivasa Rao, M., Lai, W. C., Parameshachari, B. D., Ananda Babu, J., & Hemalatha, K. L. (2022). An efficient and secure big data storage in cloud environment by using triple data encryption standard. *Big Data and Cognitive Computing*, *6*(4), 101.

[33] Kim, H. J., Kim, Y. K., Lee, H. J., & Chang, J. W. (2022). Privacy-Preserving Top-k Query Processing Algorithms Using Efficient Secure Protocols over Encrypted Database in Cloud Computing Environment. *Electronics*, *11*(18), 2870.

[34] Ye, C. Q., Li, J., & Chen, X. Y. (2024). Quantum Privacy-Preserving Range Query Protocol for Encrypted Data in IoT Environments. *Sensors*, *24*(22), 7405.

[35] Van Kenhove, M., Pohle, E., Schild, L., Zbudila, M., Sebrechts, M., De Turck, F., ... & Abidin, A. (2026). MOZAIK: A Privacy-Preserving Analytics Platform for IoT Data Using MPC and FHE. *arXiv preprint arXiv:2601.02245*.

[36] Bian, S., Zhang, Z., Pan, H., Mao, R., Zhao, Z., Jin, Y., & Guan, Z. (2023, November). HE3DB: An efficient and elastic encrypted database via arithmetic-and-logic fully homomorphic encryption. In *Proceedings of the 2023 ACM SIGSAC Conference on Computer and Communications Security* (pp. 2930-2944).

[37] Almakdi, S., Panda, B., Alshehri, M. S., & Alazeb, A. (2021). An efficient secure system for fetching data from the outsourced encrypted databases. *IEEE Access*, *9*, 78474-78494.

[38] Gnana Jeslin, J., & Mohan Kumar, P. (2022). Decentralized and privacy sensitive data de-duplication framework for convenient big data management in cloud backup systems. *Symmetry*, *14*(7), 1392.

[39] Xiong, Y., & Luo, M. X. (2024). Searchable Encryption Scheme for Large Data Sets in Cloud Storage Environment. *Radioengineering*, *33*(2).

[40] Li, P., Dai, H., Wang, S., Yang, W., & Yang, G. (2024, October). Privacy-preserving spatial dataset search in cloud. In *Proceedings of the 33rd ACM International Conference on Information and Knowledge Management* (pp. 1245-1254).

[41] Pleșa, M. I., & Olimid, R. F. (2024, November). Privacy-Preserving Multi-party Search via Homomorphic Encryption with Constant Multiplicative Depth. In *International Conference on Information Technology and Communications Security* (pp. 135-148). Cham: Springer Nature Switzerland.

[42] Adnan, A., Kausar, F., Shoaib, M., Iqbal, F., Altaf, A., & Asif, H. M. (2025). A secure and privacy-preserving approach to healthcare data collaboration. *Symmetry*, *17*(7), 1139.

[43] Yang, X., Zhang, Y., Wang, Y., & Li, Y. (2023). Efficient and expressive search scheme over encrypted electronic medical records. *Information*, *14*(12), 643.

[44] Falzon, F., Markatou, E. A., Espiritu, Z., & Tamassia, R. (2022). Range search over encrypted multi-attribute data. *Cryptology ePrint Archive*.

[45] Yuan, J., Li, Y., Li, J., Wu, D., Ning, J., Tian, Y., & Deng, R. H. (2025, July). Leakage-Resilient Easily Deployable and Efficiently Searchable Encryption (EDESE). In *Proceedings of the 30th ACM Symposium on Access Control Models and Technologies* (pp. 133-144).

[46] Yan, X., Yin, P., Tang, Y., & Feng, S. (2022). Multi-keywords fuzzy search encryption supporting dynamic update in an intelligent edge network. *Connection Science*, *34*(1), 511-528.

[47] Yan, H., Sun, L., & Zhang, Y. (2025). Towards privacy-preserving multi-dimensional range query over attribute missing data. *Peer-to-Peer Networking and Applications*, *18*(5), 278.

[48] Shi, R., Wei, L., & Zhang, L. (2024). More efficient and verifiable privacy-preserving aggregation scheme for internet of things-based federated learning. *Applied Sciences*, *14*(13), 5361.

[49] Hashi, A. I. (2025). A blockchain-enhanced zero-trust framework for privacy in industrial IoT systems. International Journal on Science and Technology, 16(3). https://doi.org/10.71097/IJSAT.v16.i3.8036

[50] Guo, Y., Xie, H., Wang, M., & Jia, X. (2022). Privacy-preserving multi-range queries for secure data outsourcing services. *IEEE Transactions on Cloud Computing*, *11*(3), 2431-2444.