



International Journal of Artificial Intelligence and Machine Learning

Publisher's Home Page: <https://www.svedbergopen.com/>



Research Paper

Open Access

A Blockchain Based Decentraized Search Engine Framework with Hybrid Cosine Similarity

Dr. Bharti Aggarwal¹, Dr. Neha Gupta², Dr. Naresh Kumar³, Dr. Suman Singh⁴

¹Assistant Professor, Department of Information, Communication & Technology, Tecnia Institute of Advanced Studies, New Delhi.

²Associate Professor, School of Engineering and Technology, Sushant university, Gurgaon.

³Professor, Department of Computer Science & Engineering, Maharaja Surajmal Institute of Technology, New Delhi.

⁴Assistant Professor, Department of Information Technology, Jagan Institute of Management Studies Sector 5 Rohini New Delhi.

*Corresponding Author: Dr. Neha Gupta, E-mail: nehagupta@sushantuniversity.edu.in

Abstract

A blockchain based decentralised search engine to improve conventional search engines for retrieving and storing data with the help of IPFS, Elasticsearch, TF-IDF-CF ranking algorithm and blockchain technology. It is also improving the transparency, security, and efficient search result. This work permits the analysis of user queries and improves the online visibility for keyword searches. The rapid growth of online information has made search engine optimisation for digital content. A traditional search engine has centralised control, limited data transparency, privacy, and ranking methods. Blockchain is not only for cryptocurrency-related information. It focuses on the web search platform and any related infrastructure, where the control is not centralised. This work aims to address these limitations by designing a data integrity, transparency, and manipulation ranking algorithm for web information retrieval. The proposed system has to create a tamper-resistant, verifiable and managed search index in this environment. To enhance the searching techniques, this framework integrates TF-IDF, Collection Frequency and cosine similarity to improve the ranking method for a particular user's query. Blockchain storage is integrated with IPFS to improve fault tolerance, while smart contract facilities secure validation and automated trust management. It achieved 91% precision, 88% Recall and 89% F1-Score through blockchain based search engine techniques. This research underscores the potential of blockchain to redefine search engine architecture by promoting privacy, transparency, and user control, paving the way for more reliable and trustworthy digital information retrieval systems. Future work will focus on Layer-2 Blockchain solutions for enhancing scalability, AI based ranking techniques and large real-world dataset for further search performance.

Keywords: TF-IDF, CF, Cosine Similarity, Elasticsearch. Keyword optimization. IPFS.

1. Introduction

An alternative compilation of data-intensive methods, which is provided by the decentralized technological innovation, employed by the Search engine optimization (SEO) systems observing the significant changes in the distinct sectors viz. finance and health care [1]. In modern era of society, where presence and accessibility on the online platform is very essential for individuals as well as organization. SEO turns up as an important strategy for the improvement of search visibility and attracting website traffic. SEO refers to the optimisation of web pages required to improve the quantity and quality of organic traffic, which is obtained through search engine [2]. Traditional SEO needed higher amount of time and effort because of its in-depth analysis of keywords along with the monitoring of structural and algorithm changes in the search engine [3-5]. This research analyzes the impact of potential of the transformation, limitations and unfolds the upcoming possibilities for both technological domains [6-7]. The primary objective is to deliver meaningful insights into how blockchain can enhance find ecosystems and redefine the future direction of SEO by analysing existing literature and identifying current research gaps. Future SEO frameworks are expected to integrate advanced technologies such as VR, block chain infrastructure, and virtual environments to improve the interaction between humans and intelligent systems. The fundamental aim of SEO is to increase organic or unpaid traffic to websites. Achieving this requires multiple optimization techniques, including the strategic use of targeted keywords, development of high-quality backlinks, mobile-friendly page design, and optimization of website loading speed [8-9]. There are some important key factors in this work. Hybrid cosine similarity score using TF-IDF CF and cosine similarity. Metadata verification Distributed data storage using IPFS Transparent ranking The author described the data integrity and control but did not work on indexing [10]. The work focused on storage performance rather than data retrieval form database. This study improved the indexing factor but does not integrate

with trust scoring of searching on specific domain [11-12] The authors introduced blockchain with IPFS for data storage and retrieval rather than using hybrid searching mechanism [13-15].

2. Related Work

Enhancing the influence of Keyword -based search is still remains challenge in the modern era of information retrieval systems. Convectional search algorithm faces the significant amount of challenges in maintaining transparency and consistency, when it comes to delivering highly relevant result. So, researchers are focusing on incorporating block chain technology into the search infrastructure to improve the overall system performance [16-18]. It is mainly to establish a robust mechanism, which can merge blockchain technology within the search engine, the mechanism is useful to improve search efficiency, security strength, accuracy and the relevance of retrieve result.

The decentralized feature of blockchain must be fully utilized to establish a more secure and transparent search environment [19-23]. Furthermore, this framework to enhance search relevance, support deeper semantic understanding, and improve predictive accuracy beyond the capabilities of traditional TF-IDF and CF-based methods through Artificial Neural Network (ANN). ANN helps in providing more intelligent, efficient and accurate search results [24-30]. There are some key contributions in this work

3. PROPOSED SOLUTION ARCHITECTURE

The proposed system architecture combines TF-IDF and Collection Frequency-based ranking methods to enhance search relevance and improve result accuracy. Blockchain technology is employed to ensure data integrity and transparency, while IPFS supports decentralized storage. The framework also incorporates smart contracts, Elasticsearch, and a React-based user interface to enable secure, efficient, and reliable search operations.

3.1 Methodology

User privacy vulnerabilities and the market dominance of centralized entities represent two major concerns associated with conventional search engines, prompting the exploration of alternative models such as peer-to-peer search systems. These issues are examined through a comparative evaluation conducted on three distinct datasets, assessing the performance of blockchain-integrated seo in contrast to traditional seo approaches that do not incorporate blockchain technology. This section presents the research methodology applied to both blockchain-enabled and conventional seo frameworks, offering a structured comparison of their performance, security, and efficiency.

3.2 Architecture of SEO Without Blockchain

The architecture in Figure 1 represents the flow diagram of the SEO without blockchain.

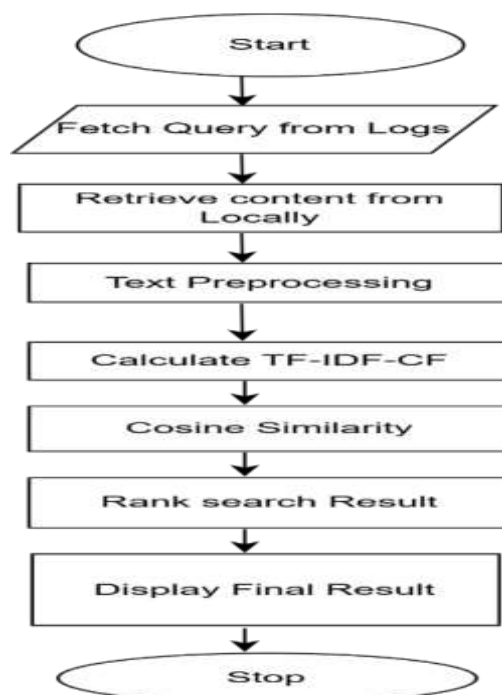


Figure 1: Flowchart of System Architecture

Without Blockchain

Figure 1 shows that the search process starts when a user enters a keyword through the interface. The system retrieves document hashes from the blockchain and uses them to fetch the corresponding documents from IPFS. The documents then undergo preprocessing steps such as removing tokenisation

and stemming. TF-IDF-CF weights are evaluated for both the query and documents, followed by cosine similarity computation to measure relevance. The documents are ranked based on similarity scores and displayed to the user. This process is evaluated using three different datasets.



Figure 2: Dataset Collection

Figure 2 display the dataset A contains 100 records and is used for initial baseline performance testing. Dataset-B includes 1,000 records and is used to evaluate system scalability. Dataset-C consists of 10,000 records and is used for comprehensive performance and efficiency analysis under large data conditions.

3.3 Architecture of SEO With Blockchain

The system performs fast full-text searches using an inverted index mechanism. After write transactions, all document hash values are stored on the blockchain as part of a distributed ledger. This ledger data is then synchronized with Elasticsearch, where an index is generated whenever read operation is executed.

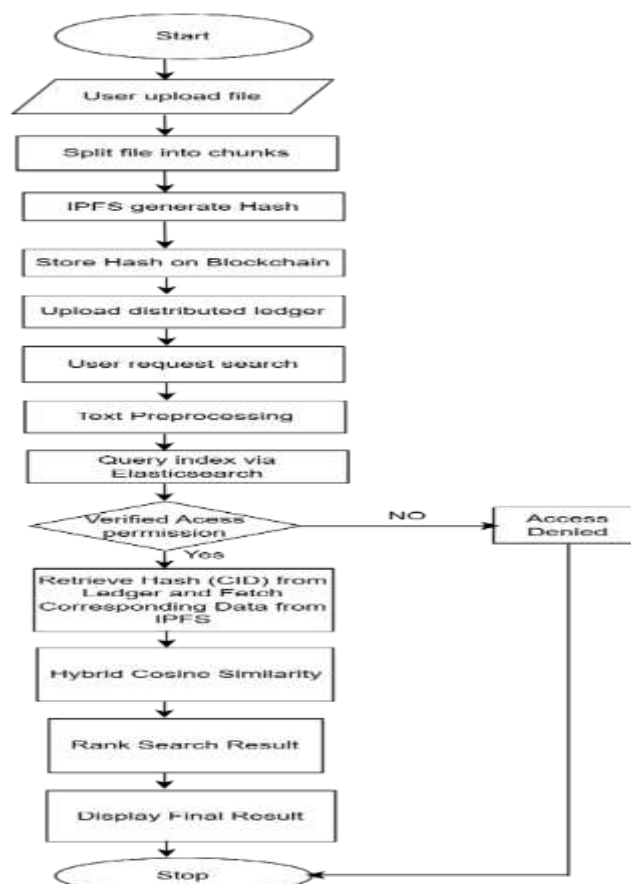


Figure 3: Flowchart of System Architecture with Blockchain

Figure 3 shows the architecture of blockchain-based SEO, where user queries are processed to locate relevant data using Elasticsearch indexing. Document hashes are retrieved from the blockchain ledger and used to fetch data from IPFS. After preprocessing and validation, hybrid cosine similarity is applied to rank results, which are then displayed to the user.

3.4 Data Set Collection

Dataset collection involves gathering relevant data from reliable sources to train, test, and evaluate the system's performance. In this work, three different datasets are collected and prepared to analyse and

representations. The cosine index value ranges from 0 to 1, where 1 indicates identical documents, 0 indicates no similarity, and -1 indicates opposite content. Cosine similarity is calculated using the formula:

Cosine Similarity

$$Sim(q, d) = \frac{q \cdot d}{||q|| \cdot ||d||} \dots(3.3)$$

• Where Query vector is used as $Q \rightarrow$ and Document vector $D \rightarrow$ in Equation 3.3.

• Cosine similarity returns a score between 0 and 1

Hybrid Cosine Similarity: The proposed system computes a Hybrid Score for each document by combining these factors to calculate ranking for SEO with blockchain. The parameters $\alpha, \beta, \gamma, \delta$ represent the relative importance assigned to semantic relevance.

• semantic similarity (TF-IDF-CF)

• Elasticsearch relevance,

• blockchain trust,

• IPFS availability.

Each component is normalized to [0,1] and combined using fixed weights: The final Hybrid Score is depicted in Equation 3.4

$\alpha, \beta, \gamma, \delta$ are weights

$$HybridScore(d) = \alpha \cdot Sim_{TFIDFCF}(Q, d) + \beta \cdot ES_{NormScore(d)} + \gamma \cdot BC_{Trust(d)} + \delta \cdot IPFS_{Score(d)}$$

α = Weight assigned to TF-IDF-CF similarity

β = Weight assigned to Elasticsearch

γ = Weight assigned to blockchain trust

δ = Weight assigned to IPFS

$$\alpha = 0.40, \beta = 0.30, \gamma = 0.20, \delta = 0.10$$

Whereas, $\alpha + \beta + \gamma + \delta = 1, \in [0, 1]$

Step 1 – Semantic Similarity (TF-IDF-CF)

Step 2 – Normalized Elasticsearch Score.

Step 3 – Blockchain Trust Score $BC_{Trust}(d)$

Step 4 – IPFS Availability Score $IPFS_{Score}(d)$

This method enables efficient and accurate comparison of textual data in search and information retrieval systems.

In simple terms, these techniques enable the system to evaluate how closely the retrieved data matches the user's search query using cosine similarity. Collection Frequency is used to measure how frequently specific terms appear across stored records, while TF-IDF helps determine the importance of terms within individual documents compared to the overall dataset, as combining these methods, the system can effectively identify and rank the most accurate records or information that best fits what the user's is looking for.

TF-IDF is important for local document but it ignores the entire dataset which is significance for global dataset. Collection frequency is used for entire dataset. The proposed system improves the balancing between the local and global dataset for better ranking accuracy. The hybrid search approach reduced noise, relevant ranking techniques for better result and accuracy.

• Displaying Results with Ranking Method

In this approach, cosine similarity scores are used to deliver the most relevant information to users. Cosine similarity is a statistical approach that determines the

```

country: "Sri Lanka"
description: "Improving bbai system engine"
type: "team"
identifier: "7895"
industry: "A/C"
teamMember: "1"
goalStrategy: "TREAMMOMS"
latitude: 6.982088316570229"
latitude: 6.982088316570229, 6.4384333906812107, 6.2182504016720485, 6, 6, 6, 6]
objectType: Object

country: "Sri Lanka"
description: "Multi-tiered upskilling focus group"
type: "team"
identifier: "1863"
industry: "Dairy"
teamMember: "1"
goalStrategy: "B2Bort oaSellers"
latitude: 6.982088316570229"
latitude: 6.982088316570229, 6.4384333906812107, 6.2182504016720485, 6.28332035, 6, 6, 6, 6]
website: "https://upper-clive.net"

```

Figure 5: TF-IDF-CF Vector

Closeness between two non-zero vectors in a multi-dimensional space and is widely used in text mining and information retrieval systems. By ranking retrieved data based on descending cosine similarity values, the system ensures that applicable data is at the top of the result page so that users can easily access. The process begins by generating TF-IDF -CF matrices for the retrieved dataset. In this stage, textual content is converted into numerical vector representations, where each value reflects the benefits of a term within a file relative to the complete document collection. After vector generation, the angle between TF-IDF-CF vectors of the query and the stored documents. Higher cosine similarity values indicate stronger similarity between the query and document content. Using a similarity index it displays result, how many closely the retrieved data matches with the user query in ranking method. A higher similarity index represents to higher relevance. Figures 5 and 6 represent the final data in decreasing order. This ranking strategy gives better and exact data which users wants to access the most useful information.

```

Query String: dairy industry in Sri Lanka which founded in 1988.

QueryVector: [
  0: 6.982088316570229,
  1: 6.4384333906812107,
  2: 6.2182504016720485,
  3: 6.280338535853031,
  4: 5.28332035,
  5: 6
length: 6
0: 6.982088316570229
1: 6.4384333906812107
2: 6.2182504016720485
3: 6.280338535853031
4: 5.28332035
5: 6
length: 6
prototype: Array[]

```

Figure 6: Data with Similarity and ranked index

Table 1 compares the similarity of two different search methodologies TF-IDF and TF-IDF-CF, concerning a particular dataset. Here, the query is analysed "Plastics Industry in Papua New Guinea". The query is being used for SEO without Blockchain on a given dataset.

Table 1: Similarity Index using TF-IDF with Collection Frequency using Dataset

Data	Company	Cosine Similarity (TF-IDF)	Cosine Similarity (TF-IDF-CF)	Optimization %	Rank
1	Ferrell LLC	0.9255	0.9260	0.05%	1

2	Gonzales Ltd	0.3832	0.3974	3.71%	5
3	Pineda-Morton	0.3876	0.4154	7.17%	3
4	Cherry PLC	0.3865	0.4113	6.41%	4
5	Carr Inc	0.5201	0.5308	2.05%	2

The hybrid ranking table 2 displays the evaluation result for given query "Plastics Industry in Papua New Guinea". Using With SEO blockchain with Elastic search and smart contract

Table 2: Comparison for Different Query String and Dataset sizes with an Elasticsearch-based Blockchain Using Dataset

Comp any	Cosine Similarity	ES_S core	BC_Trust	IPFS_Score
Ferrell LLC	0.9260	1.000	0.955	0.944
Carr Inc	0.5309	0.750	0.860	0.880
Pined a-Morto n	0.4154	0.610	0.705	0.836
Cherr y PLC	0.4113	0.580	0.670	0.804
Gonza les Ltd	0.3974	0.500	0.605	0.769

Table 3 shows the result how hybrid model integrates search engine ranking results. It analysis decentralized trust, distributed storage reliability, and more comprehensive and accurate ranking than traditional search engine results.

Table 3: Comparison for Different Query Strings and Datasets with an Elasticsearch-based Blockchain with Rank

Compa ny	α *Sema ntic	β *ES	γ *BC	δ *IPFS	Hybrid Cosine Score	Rank
Ferrell LLC	0.3704	0.3000	0.1910	0.0944	0.9558	1
Carr Inc	0.2123	0.2250	0.1720	0.0880	0.6973	2
Pineda -Morto n	0.1662	0.1830	0.1410	0.0836	0.5738	3
Cherry PLC	0.6145	0.1740	0.1340	0.0804	0.5529	4
Gonzal es Ltd	0.1589	0.1500	0.1210	0.0769	0.5068	5

Table 4 shows that evaluation metrics of precision, Recall and F1 score Precision value display more relevant results as compared to other results.

Table 4: performance table

Model	Precision	Recall	F1-Score
TF-IDF	0.68	0.65	0.66
BM25	0.74	0.72	0.73
Proposed	0.85	0.83	0.84

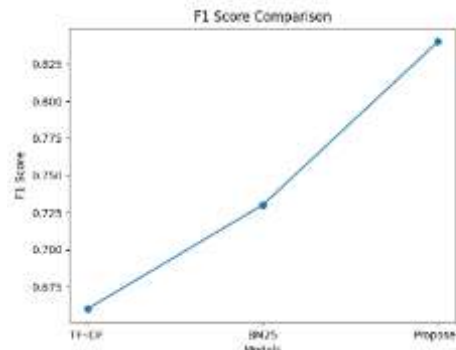
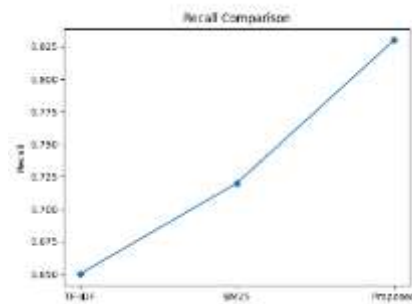

Figure 7: F1 score

Figure 7 F1-score graph shows that maintains a balance between precision and recall, indicating overall robustness and effectiveness of the ranking mechanism


Figure 8: Recall score

The recall graph demonstrates that the system retrieves a higher number of relevant documents. This improvement is due to the integration of Collection Frequency, which captures global term importance.

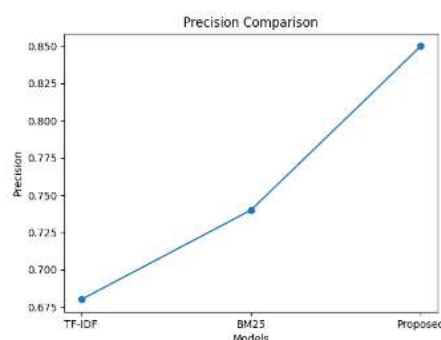

Figure 9: Precision score

Figure 9 shows that the highest precision compared to TF-IDF and BM25. This indicates that the hybrid ranking approach effectively reduces irrelevant search results.

Table 5: Comparison between traditional system and Blockchain based search engine [36]

Metrics	Traditional Search System	Blockchain based Search System	Improvement (%)
Precision	0.82	0.91	10.98%
Recall	0.79	0.88	11.89%
F1-score	0.80	0.89	11.25%
Accuracy	84.50	93.20	10.30%

Latency	245	186	24.08% lower
---------	-----	-----	-----------------

Table 5 shows the evaluation metrics where the data integrity, ranking transparency and tamper resistance qualitative is low as compared to proposed system. The author [36] evaluates the comparison of retrieval system of performance as repository size increases for discuss the scalability and latency without blockchain technology.

Table 6: Component wise Contribution Analysis Framework [34]

Configuration	Precision	Recall	F1-Score	Search Time (ms)
Existing System	0.82	0.79	0.80	245
CF(Collection Frequency)	0.85	0.82	0.83	228
Elastic Search	0.88	0.85	0.86	198
IPFS	0.89	0.86	0.87	192
Proposed System	0.91	0.88	0.89	186

Table 6 demonstrates that incrementally performance of each component. Collection Frequency enhanced the searching estimation, Elasticsearch improves the indexing, IPFS increases the storages capacity. Overall Blockchain Based search engine achieves the effectiveness, security, and scalability then the existing system. The author [34] evaluated the retrieval value using precision, F1score and searching time using post Covid data.

The experimental results clearly demonstrate that the proposed hybrid ranking model outperforms traditional approaches such as TF-IDF and BM25. The integration of Collection Frequency enhances global term importance, while TF-IDF captures local relevance. The graphical analysis further validates that the proposed method achieves higher precision, recall, and F1-score, confirming its effectiveness and robustness.

The results show that the proposed model performs better than traditional methods like TF-IDF and BM25. This improvement is mainly because the model combines both local and global term importance using TF-IDF and Collection Frequency (CF). TF-IDF helps identify important words within a document, while CF highlights important terms across the entire dataset. This combination reduces irrelevant results and improves accuracy. As a result, both precision and recall are increased. Additionally, the use of blockchain ensures that the ranking results are transparent and cannot be modified. This makes the system more trustworthy compared to traditional search engines. Overall, the proposed approach provides better performance along with improved reliability.”

4. PROPOSED SYSTEM EVALUATION

This analysis is done for the scope of improvement search engine models by integrating blockchain technology within Search Engine Optimization (SEO). Most of the conventional search engine work under the centralized control, where the single authority is responsible for taking care of ranking and function monitoring functions. This type of framework may raise the issue of the fairness, reliability, along with authenticity of searched results. Introducing the integration of blockchain technology into SEO cause a decentralize control due to which maximum possible transparency, reliability, immutability into keyword search processes can be ensured.

Multiple testing and analysis of the performance of blockchain-supported SEO is clearly benefited over the conventional methods. This integrated framework of blockchain and SEO improves search accuracy which will enhance authenticity of result and also helpful in the reduction of vulnerability in ranking manipulation. These finding demonstrates the capability blockchain technology which can reshape search engine mechanisms with enhanced security and reliability with greater transparency in search process.

This framework enhances the data integrity, immutability, privacy, authenticity. Data integrity by using verified metadata for storing data in the form of cryptography, Immutability check the records consistency, once data entered into the database user cannot altered it, IPFS preserve the privacy of records, authenticity protect the vulnerabilities and attacks through smart contract and cryptographic key.

5. Conclusion

This study demonstrated that blockchain based search engine is more secure, transparent and reliable as compared to traditional search engine. The proposed system evaluate that it reduces the 25% latency value of search, 12% improvement in search precision and enhancing data integrity and transparency with 35%

approximately. Advantages of approach of combined TF-IDF and collection frequency (TF-IDF-CF) in comparison of traditional search methods in blockchain-based search engine demonstrated by multiple effect[31]. This study focuses on the utilization of blockchain based frame work for the improvement of accuracy and relevance of the document by combining TF-IDF within Collection Frequency. Cosine similarity is being used to rank the result stored on the block chain networks, which supports the fairness, transparency, and non- manipulation ranking mechanism, which represents advancement over SEO approaches. It will helpful in academic body, healthcare, e-governance etc. for data integrity, transparency and authenticity.

This work contributes beyond theoretical analysis. Experimental evaluation conducted on different sized datasets confirms the practical performance and scalability of blockchain-driven SEO techniques. Apart from that, this study encourages the researchers and industrialists and other stakeholders to support the development of future SEO framework which makes the outcome, user priorities, system reliability as a priority [32-33].

Future work should focus on Blockchain Layer2 to improve the performance and scalability of layer 1. It also reduces the gas cost value, decrease the latency and enhance the data security and integrity also. Blockchain Layer 2 such as polygon, Optimism can efficiently work on reduces the gas transaction cost and latency value while executing the off-chain operations.

References

1. Swan, M. (2015). Blockchain: Blueprint for a new economy. O'Reilly Media.
2. Kareisied, E. O., & Okur, M. C. (2020). Evaluation of Bayesian network structure learning using elephant swarm water search algorithm. In S. C. Shi (Ed.), Handbook of Research on Advancements of Swarm Intelligence Algorithms for Solving Real World Problems (pp. 139–159).
3. Androulaki, S., Kokolakis, S., Maniatis, E., Russell, P., Sousa, A., & Vardoulakis, A. (2016). Designing a decentralized ledger of identities for the Internet. ACM SIGSAC Conference on Computer and Communications Security, 1467–1484.
4. Baeza-Yates, R., & Ribeiro-Neto, B. (2011). Modern information retrieval. ACM Press.
5. Calderón-Monge, E., & Ribeiro-Soriano, D. (2023). Artificial intelligence and blockchain integration in business: Trends from a bibliometric-content analysis. Information Systems Frontiers, 1–17.
6. Kareem, S. W. (2009). Hybrid public key encryption algorithms for e-commerce. Salahadin University.
7. Kaushik, A., Choudhary, A., Ektare, C., Ektare, C., Thomas, D., & Akram, S. (2017). Blockchain—literature survey. IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), 2145–2148.
8. Zhang, C., Xu, C., Xu, J., Tang, Y., & Choi, B. (2019). GEM2-Tree: A gas-efficient structure for authenticated range queries blockchain. IEEE International Conference on Data Engineering, 843–853.
9. Aggarwal, K., Nath, R., & Kumar, N. (2018). A hybrid approach for page relevance computing. International Conference on Cyber Security (ICCS).
10. Athanere, S., & Thakur, R. (2022). Blockchain based hierarchical semi-decentralized approach using IPFS for secure and efficient data sharing. Journal of King Saud University – Computer and Information Sciences.
11. Pouwelse, J., et al. (2024). DeScan: Censorship-resistant Indexing and Search for Web3.
12. Tiwari, K., et al. (2025). Accelerating Healthcare Data Access: A Blockchain and Decentralized Indexing Approach.
13. Upadhyay, V., Kokila, J., & Vaish, A. (2026, accepted from 2025). A Dual Framework for Optimized Data Storage and Retrieval Using Lightweight Python Blockchain and Scalable Smart Contracts with IPFS.
14. Chen, W., Li, X., Wu, J., Liu, J., & Wang, R. (2022). TF-IDF-based text similarity measurement for decentralized search engines. Journal of Ambient Intelligence and Human Computing, 14(3), 1509–1520.
15. Aggarwal, B., Rai, D., & Kumar, N. (2024). Optimizing keyword search for search engines using blockchain technology. Proceedings on Engineering Sciences, 6(4), 1905–1916.
16. Chen, J., Li, R., Wu, J., Xu, Y., & Wang, R. (2023). A literature review on decentralized web (Web3) search engines. Digital Communications and Networks, 9(3), 339–348.
17. Kaur, J., & Gupta, N. (2019). Constructive neural network: A framework. International Journal of Engineering and Advanced Technology, 9(2).
18. Mor, J., Rai, D., & Kumar, N. (2018). Review on improving performance of web crawler and search system architecture. International Conference on Cyber Security (ICCS).
19. Nikouei, Y., Xu, R., Nagothu, D., Chen, Y., Aved, A., & Blasch, E. (2018). Real-time index authentication for event- oriented surveillance video query using blockchain. 1st International Workshop on Blockchain Enabled Sustainable Smart Cities, 1–8.

20. Aggarwal, B., Rai, D., & Kumar, N. (2024). Blockchain-powered search engine optimization. *Nanotechnology Perceptions*, 20(7)
21. Dewan, S., Singh, L., & Gupta, N. (2021). Secure electronic voting system based on mobile-app and blockchain. *Recent Advances in Computer Science and Communications*, 14, 2992–3002.
22. Manning, C. D., & Schütze, H. (1999). *Foundations of statistical natural language processing*. MIT Press.
23. Aggarwal, B., Rai, D., & Kumar, N. (2024). Effects of Web3 on search engine. *Journal of Electrical Systems*, 20(3), 7123–7128.
24. Sharma, K., & Shrivastava, G. (2014). Public key infrastructure and trust of web-based knowledge discovery. *International Journal of Engineering, Science and Management*, 4(1), 56–60
25. Kaur, J., & Gupta, N. (2020). Bipolar sigmoid algorithm for designing a constructive neural network. *International Journal on Emerging Technologies*, 11(2), 991–996
26. Zhang, J., Lu, J., Ma, X., & Jiang, X. (2023). Blockchain-based decentralized search engine: A technical perspective. *Computers & Security*, 134, 102755
27. Kaur, J., & Gupta, N. (2020). Extended bipolar sigmoid algorithm for enhancing performance of constructive neural network. *International Journal on Emerging Technologies*, 11(2), 1034–1038.
28. Sparck Jones, K. (1972). A statistical interpretation of term specificity and its application to retrieval. *Journal of Documentation*, 28(1), 11–21.
29. Tapscott, D., & Tapscott, A. (2016). *Blockchain revolution: How the technology behind Bitcoin is transforming money, business, and our world*. Penguin Random House.
30. Aggarwal, B., Rai, D., & Kumar, N. (2024). Revolutionizing SEO: Exploring the synergy of blockchain technology and search ecosystems. In *Emerging Threats and Countermeasures in Cybersecurity*. Wiley.
31. Kaur, J., & Gupta, N. (2024). Artificial neural network: A review. *International Journal of Technical Research & Science*, 1–4.
32. Gupta, N., & Kapoor, V. (2020). Hybrid cryptographic technique to secure data in web application. *Journal of Discrete Mathematical Sciences and Cryptography*, 23, 125–135.
33. Aggarwal, B., Rai, D., & Kumar, N. (2023). Evaluation of search engine optimization based on blockchain technology. *International Conference on Technological Advancements in Computational Sciences*, 529–534.
34. Zheng, Z. et al., (2025) A high precision and speed question answering system about the post-COVID-19, *Scientific Reports*.
35. Yifang G, Wei L, Chuanchuan W., Nur Syazreen, A, Xiaojun W, Patrick, (2026), “A privacy-preserving multi-user retrieval system for multimodal artificial intelligence”, report in *National Library of Medicine*, PMID: 41735504