Enhancing Privacy and Trust in Keyword Search Engines through Blockchain Optimization

Bharti Aggarwal , Dr.Dinesh Rai, Dr. Naresh Kumar
Research Scholar, Professor, Professor
Sushant University, Sushant University,
Maharaja Surajmal Institute of Technology

Abstract: Utilizing the decentralized structure of blockchain technology to optimize search engine optimization (SEO) and increase data accuracy is known as blockchain optimization. By focusing on pertinent keywords and search phrases, SEO specialists can use this strategy to raise visibility, which eventually raises website rankings and increases traffic. Furthermore, classic search engines can gain from improved data precision and possibly rethink digital marketing methods for companies in the cryptocurrency and blockchain industries by integrating blockchain technology. It is important to realize that optimising website for the decentralized web is just as important as appearing for crypto-related keywords in blockchain SEO. Due to modifications to algorithms and privacy issues, existing SEO landscape need fresh approaches. This study investigates how blockchain technology can be incorporated into SEO to increase website visibility while protecting user privacy. The author presents a revolutionary strategy that stresses trust, transparency, and safe data control using blockchain-powered search engines. Users gain appropriate search results and better privacy through the detailed monitoring of their communications on the blockchain, addressing modern data protection issues.

Keyword: Search Engine, SEO, Blockchain, Decentralization, Keyword Optimization

1. Introduction

This decentralised wonder provides a strong substitute for the data-hungry methods of traditional SEO and is already revolutionising a number of industries, including finance and healthcare [1]. Search Engine Optimisation (SEO) has become a critical strategy for obtaining higher search ranks and more web traffic in the quickly changing digital environment, where online presence and accessibility are crucial for both individuals and organisations. Optimising a webpage or website for search engines to increase the volume and quality of natural search engine traffic is known as search engine optimisation or SEO [2]. SEO is the practice of assisting in improving your

website's position on Google along with other search engines [3]. In order to guarantee that content is both discoverable and pertinent to the target audience, SEO methods are constantly adjusted to the search engines constantly evolving algorithms [4].

Concurrently, blockchain technology has surfaced as a disruptive force capable of transforming traditional procedures in a number of industries. Blockchain, which is well-known for being visible, unchangeable, and decentralised, provides creative answers to a variety of problems. The next phase of SEO might be ushered in by integrating Blockchain into the foundation of search operations, providing a safer, reliable, and decentralised method of searching, indexing, and presenting information. SEO demands a substantial time and effort investment, including keyword research and constant monitoring of search engine system changes [5].

This study explores the potential for transformation, obstacles, and possibilities that exist in the nexus of both of these formidable domains—blockchain technology and search engine optimization. The purpose of this article is to provide significant insight into the way. Blockchain can improve the search environment and reshape SEO in the future by synthesising current knowledge and finding gaps in the research. Search engine optimisation will leverage virtual reality, blockchain technology, and the virtual world to improve human-machine interaction. Increased organic or free traffic to that website is the goal of SEO. Enhancing the content of a website and architecture through SEO involves a number of strategies, such as concentrating on targeted keywords, building excellent backlinks, keeping the page optimised for mobile devices, and ensuring a quick page load time [6,7]. Three of the features of blockchain technology are immutability, portability, and decentralisation; the other four are decentralisation, privacy, persistence, and auditability[8-10].

Alternatives to blockchain technology have been explored to solve relevant safety and equity concerns, such as hostile or compromised nodes, violent online keyword guessing, appropriate compensation, data sharing, critical leaks, information security, and consistency [11-14].

The author delves into the actual forces driving this shift, going beyond the abstract promise of decentralisation. One such method is the information retrieval workhorse known as the term frequency and inverse document frequency (TFIDF) metre, which measures the significance of terms both within texts and over a larger corpus [18-19]. But how can the effectiveness of this decentralised approach be evaluated? This is when cosine similarity becomes useful. We can measure the "angle" connecting two documents within a space of high dimensions that is defined by its TF thanks to these mathematical miracle.-IDF-CF character sets [20]. Through the computation of the cosine similarity from the

query provided by the user and the blockchain's content, we are able to rank search results according to their actual relevance rather than manipulated or biassed algorithms [21].

The author [22] hopes to rethink SEO's future with this in-depth examination. Through the utilisation of TF-IDF, CF, as well as cosine similarity in a decentralised blockchain framework, it seems that relevance, privacy, and credibility are the most important factors in a search environment. The author [23-25] envisions a time when people, not businesses, will manage their data and direct their internet experiences, moving past the opaque mechanisms of centralised behemoths.

2. PROBLEM FORMATION

Optimising keyword search within search engines is one of the biggest issues in contemporary information retrieval systems. For typical search algorithms, producing highly relevant results while maintaining transparency and avoiding manipulation is sometimes a challenging task. To solve these issues and enhance search engine functionality, there is growing interest in applying blockchain technology. However, in order to guide this effort, a detailed problem description is needed. To improve the efficiency, security, and relevance of keyword searches, dependable procedures for integrating blockchain technology into search engine architecture must be developed. In this formulation, the decentralised nature of blockchain systems must also be considered, with the aim of exploiting this characteristic to construct an impenetrable and open search ecosystem. Thus, the difficulty is to develop a framework that seamlessly integrates blockchain technology.

2.1 Traditional SEO:

The author [26] faces several challenges in the rapidly evolving landscape of online search and content consumption in traditional search engine optimization:

- **Algorithm Changes**: Search engines like Google frequently update their algorithms to improve user experience and combat spam. These updates can significantly impact search rankings, requiring SEO professionals to continually adapt their strategies.
- **Competition**: The proliferation of online businesses has intensified competition for top search rankings. Ranking for competitive keywords requires substantial effort and resources, making it challenging for smaller businesses to compete effectively.
- **Voice Search**: The rise of voice search technology has changed the way people search for information online. Optimizing content for natural language queries and long-tail keywords poses a challenge for traditional keyword-based SEO strategies.
- **Local SEO**: For businesses targeting local markets, optimizing for local search results is essential. However, local SEO comes with its own set of challenges, such as managing

business listings, obtaining positive reviews, and competing with other local businesses.

- **Technical SEO**: Ensuring proper website architecture, indexation, crawlability, and other technical aspects of SEO requires specialized knowledge and ongoing maintenance. Technical SEO challenges include managing site speed, optimizing structured data, and addressing issues identified through tools like Google Search Console.

2.2 Blockchain Technology

The decentralised and secure characteristics of blockchain technology present a potential remedy for the drawbacks of conventional SEO. It makes search engine optimisation more visible, user-centered, and privacy-preserving [27].

- **Decentralization**: Blockchain operates on a decentralized network of nodes, where transactions are validated and recorded by multiple participants rather than a single central authority. This decentralization ensures greater transparency, security, and resilience against single points of failure.
- **Transparency and Immutability**: Since all transactions are recorded on a public ledger and cannot be altered or deleted, blockchain ensures transparency and immutability of data, reducing the risk of fraud or manipulation.
- **Security**: The cryptographic techniques used in blockchain make it highly secure, protecting sensitive information and mitigating the risk of cyber-attacks.
- **Efficiency and Cost Savings**: By eliminating intermediaries and streamlining processes, blockchain can reduce transaction costs, minimize delays, and increase operational efficiency.
- **Decentralization and Trust**: Blockchain's decentralized nature fosters trust among network participants by removing the need for intermediaries and enabling peer-to-peer transactions.

2.3 Blockchain-Powered SEO Techniques

- **Term Frequency-Inverse Document Frequency**: This measure assesses how relevant keywords are both within papers and throughout a larger corpus [18-19]. It goes beyond keyword stuffing by adding context and generalization to relevance assessments by accounting for collection frequency (CF).
- **Cosine Similarity**: The TF-IDF-CF profiles of two documents define the "angle" that these mathematical tools measure between them in a high-dimensional space [20]. It enables us to prioritize search results over keyword matches by taking into account thematic coherence [21].

The purpose of this study is to evaluate how well blockchain technology works when combined with cosine similarity and TF-IDF for engine optimization (SEO). The evaluation's goal is to do a comprehensive analysis on three different datasets—100, 1000, and 10,000 objects—of

differing sizes. We plan to examine the ability to scale and generalizability of the suggested blockchain-powered SEO solution using this varied dataset methodology. This will enable us to investigate the effects of the quantity of data on search relevancy and accuracy in addition to illuminating the system's capacity to manage varying data sizes. Additionally, studies aim to provide an accurate contrast between the efficacy of a blockchain-based SEO and conventional centralized techniques, offering a comprehension of the possible benefits and limitations of each strategy.

This research plan aims to provide a thorough grasp of the suggested blockchain-powered SEO methodology by comparing it with established centralized methodologies and assessing its adaptability across databases of different sizes. Through these evaluations, we hope to provide insightful information to the SEO community, taking into account the approach's potential for scalability as well as its practical applications for improving search relevance and accuracy.

3 RELATED WORKS

The amalgamation of blockchain technology and search engine optimization, or SEO, is increasingly garnering interest due to its potential to transform internet visibility and tackle privacy issues. One area of interest is blockchain decentralized and secure structure. The purpose of this survey is to investigate the new area of "Blockchain-powered SEO," with a particular emphasis on employing The TF-ID and collection frequency to improve keyword searches in blockchain-powered search engines. By looking at studies, approaches, and real-world examples, one may determine the possible advantages of incorporating blockchain technology into SEO.

- An extensive review of current blockchain-based search engine solutions is given in a survey conducted [21]. Which also highlights the benefits of decentralized management, openness, and user privacy? Although it acknowledges the early stages of development, it provides a starting point for additional research into particular areas, such as SEO integration [21].
- A study in [28] highlights the increasing convergence of blockchain and AI, with possible uses in search engine optimization. Within this integration, it finds search engine-related thematic clusters and recommends more research into powered by AI semantic understanding for you're a blockchain-based SEO framework.
- In 2023, the writers of [29] explore the technological difficulties and potential paths of Web3 engines of search, paying particular emphasis to data security and decentralization techniques. It highlights the requirement for effective index and ranking algorithms, which encourages more research into how your TF-IDF-CF strategy can handle these difficulties in a blockchain setting.
- The author examines numerous possible use cases and current ideas with an emphasis on

blockchain applications for SEO. It highlights the necessity for studies on scalability, spam protection, and incentive mechanisms, and it encourages you to take these into account in your suggested framework [30].

- The data preservation, indexing, and ranking strategies of many blockchain-based search engine concepts are compared in an analysis. It highlights the parallels and divergences across keyword optimization strategies, leading you to assess your TF-IDF-CF strategy against these current recommendations and highlight its distinct advantages.
- In the framework of blockchain search engines, work published in [29] investigates privacy-preserving keyword search strategies. It describes secure search algorithms and encryption techniques, which should make you think about how your framework can include comparable privacy-enhancing features to protect user data.
- The author discuss [31] a technical viewpoint that delves into the design and implementation issues of blockchain-based search engines. It looks at scalability constraints, consensus processes, and possible performance bottlenecks and urges you to take these things into consideration while designing and implementing your suggested framework.
- The author present [34] a semantic keyword search architecture that utilizes the processing of natural language (NLP) and knowledge graphs for decentralized search engines. It draws attention to the possibility of better relevancy and user experience, which should make you consider combining your TF-IDF-CF strategy with other semantic understanding techniques to boost search accuracy.
- In a hybrid ranking system for blockchain-based search engines, PageRank is used for network analysis and TF-IDF is used for term relevance. Although it doesn't directly address CF, it shows how blockchain architecture and conventional information retrieval techniques can be combined to provide ranking [32].
- [30] Presented the TF-IDF for similarity of texts evaluation in decentralised search engines, examining its usefulness and limits. It emphasizes the applicability of TF-IDF for determining word importance in such circumstances even though it does not concentrate on CF.

• The merging of TF-IDF with topic modeling for enhanced retrieval of data in decentralised search engines is examined by the author. It emphasizes the possibility of fusing semantic comprehension with analysis of keywords for improved search accuracy without specifically discussing CF [31].

- An investigation in [32] evaluates the effectiveness of TF-IDF and offers optimization solutions for retrieving data in decentralised search engines through an empirical analysis. Although it doesn't specifically address CF, it offers insightful information on how TF-IDF is actually used in this situation.

- The authors of [34] suggest utilizing TF-IDF weighting in conjunction with semantic indexing to improve the retrieval of texts in decentralised search engines. It stresses the value of combining semantic knowledge with keyword analysis for better search accuracy; even though it doesn't specifically mention CF. this is in line with the emphasis you place on the two TF-IDF and CF.
- In [16] a technique designed to pinpoint these issues in relation to web crawlers is given. has an emphasis on parallel-oriented and web crawlers.
- By employing a XML-driven crawling method, a SE can effectively reduce the amount of shared resources it uses. The returned URLs are given greater weight because of this focused-based architecture, which is built around a highly focused set of terms that are directly tied to each crawler's distinct domain.
- The writer [4] offered a search engine for for the creation of systems that includes online browsing approaches and operational modeling. Reviewing previous work to improve search system architectures is the major objective.

According to the poll, traditional difficulties can be addressed by incorporating blockchain technology into SEO. Blockchain's emphasis on privacy, security, and decentralization is in line with the demands of digital marketing. In blockchain-based search engines, the utilization of The TF-ID and CF using cosine similarity improves keyword search, leading to better user experience and ranking.

By overcoming these obstacles and fusing blockchain, cosine similarity, TF-IDF, and CF, we can create a secure, clear, and consumer-friendly digital environment wherein SEO and user confidentiality coexist.

4. PROPOSED SOLUTION ARCHITECTURE

4.1 Methodology

Overcoming these challenges and combining blockchain technology, cosine similarity, TF-IDF, and CF will enable us to build a transparent, safe, and user-friendly digital environment where user privacy and SEO coexist.

• **Dataset Collection**: A methodical and careful strategy was taken in gathering and selecting datasets for the project. The dataset included organization-specific data, including ID, Organisation ID, Name and Websites, Nation, Definition, Established Date, Industry, and Employee Count. The methodology employed for data gathering was designed to guarantee a thorough portrayal of heterogeneous companies situated in different sectors and regions. The dataset collection procedure was started by utilizing a number of sources. Reputable web platforms, industry-specific archives, publicly accessible databases, and company directories were some of these sources. Information was accessed and gathered via scraping techniques and APIs, guaranteeing that a wide range of companies were included in the dataset. Three different samples sizes—100, 1000, and 10,000 objects—were methodically gathered and

arranged in order to accomplish this. Comprehensive evaluation and scaling tests inside the Blockchain-Powered SEO program framework were made possible by this tiered approach.

Figure 1 - Sample of Collected Dataset

• **Information Storage Using IPFS**: Files are saved among several nodes under the distributed model that is used by IPFS, a peer-to-peer network. As seen in Figures 2, each file's distinct hash guarantees integrity and authenticity, allowing for safe and impenetrable data storage. With IPFS, multiple nodes which store the same material are used to retrieve content instead than relying on a single server. This decentralized structure lessens dependency on a central authority while improving data availability and robustness against single-point failures.

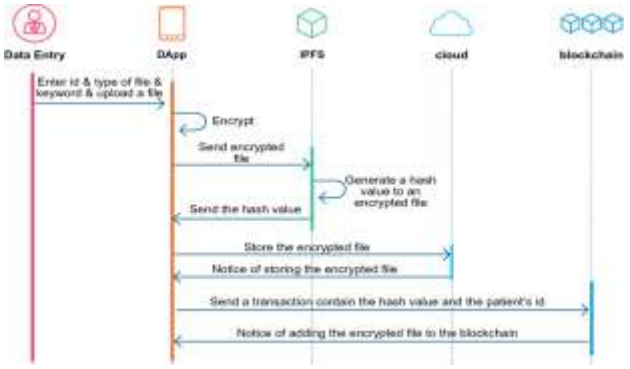

Figure 2 - Sequence Diagram of Data Retrieval from IPFS

The protocols and organizational frameworks of IPFS regulate the storage and retrieval of data. Hashing content provides for efficient replication and retrieval because the same content, regardless of the uploader, always has the same hash. Furthermore, IPFS facilitates data connection via a content-addressed mechanism, making network traversal and referencing simple.

• **Data Retrieval and Pre-processing**: To begin with, a mechanism for retrieving data from IPFS based on user queries must be implemented. Since the Interplanetary File a system, or IPFS, runs on a dispersed network, creating a retrieval mechanism that works well requires employing cryptographic hashes to find and access the precise content that the user has requested. Accurate and secure retrieval of the intended information is guaranteed by this retrieval method. After the information is retrieved, the following stage entails pre-processing and text extraction from the data that was retrieved; this data is displayed in Figure 3. In order to prepare the content for analysis, this step is essential. To

separate textual content and make sure that only pertinent information is taken into consideration for additional processing, text extraction techniques may include parsing HTML or different formats of files. The text data can be cleaned and standardized using pre-processing methods like tokenization, stop word removal, stemming, and lemmatization to prepare it for analysis.

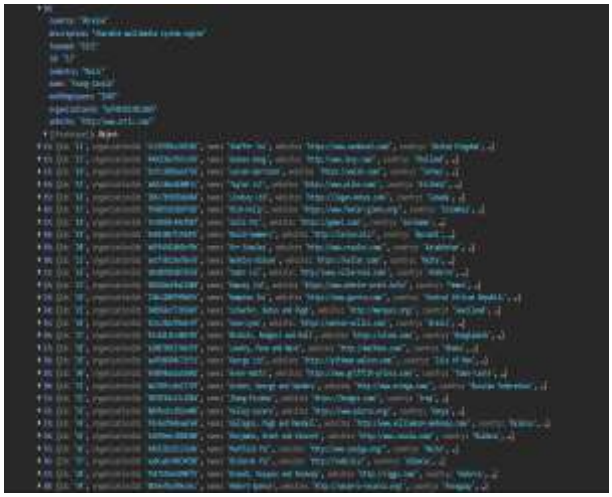

Figure 3. Sample of retrieved Data from IPFS

4.2 TF-IDF, Collection Frequency, and Cosine Similarity Calculation

A key idea in information retrieval and natural language processing is text similarity. Combining cosine similarity with TF-IDF vectorization is a potent technique for assessing text similarity.

The term's frequency of occurrence in a text is measured by term frequency (TF). It is computed by dividing the entire amount of words in the text by the number of times a phrase appears.

$$TF = \frac{Number\ of\ times\ term\ appears\ in\ a\ document}{Total\ number\ of\ terms\ in\ the\ document}$$

The Inverse Document Frequency (IDF) gauges a term's rarity throughout the collection's texts. It is computed by dividing the entire amount of texts by the amount of texts that include the phrase, using the logarithm of that number.

$$\text{IDF} = \log_{10} \frac{Total\ number\ of\ documents}{Number\ of\ documents\ containing\ the\ term}$$

Terms that occur often in a text but seldom in the corpus—the entire collection of text data you are analyzing—are given a greater weight by TF-IDF, which is a combination of TF and IDF.

A metric called cosine similarity is used to calculate the cosine of an angle formed by two vectors. These vectors reflect the TF-IDF descriptions of texts in a given setting of text similarity.

The cosine similarity scale goes from -1 to 1, where a text is considered identical if its cosine similarity is 1, 0 if it is not, and -1 if it is entirely unlike.

To get the cosine similarity between two TF-IDF vectors, divide the product of the dots of the vectors by the combined value of their magnitudes.

$$\cos\theta = \frac{A \cdot B}{|A| \times |B|}$$

- $A \cdot B$ is the dot product of vectors A and B.
- $\|A\|$ and $\|B\|$ are the magnitudes (Euclidean norms) of vectors A and B respectively.

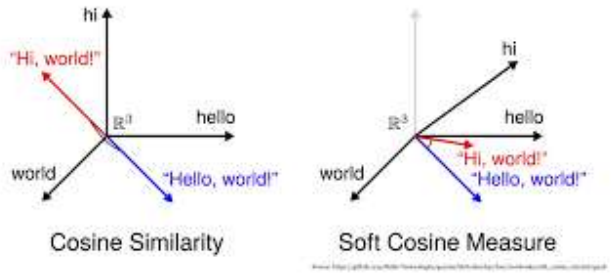

Figure 4 - Example of Cosine Similarity Measures

Essentially, these methods allow the system to determine the degree of similarities among the user's query as well as the information obtained data (cosine similarity), determine the number of terms inside records (collection frequency), and evaluate the significance of sentences in individual documents relative to the entire corpus (TF-IDF), as illustrated in Figure 4. The most pertinent records or information points that closely match the consumer's search goal or query are ranked and identified with the help of this method.

4.3 Ranking and Displaying Results

Cosine similarity ratings are used in this procedure to provide consumers with the most relevant information. Cosine similarity, a statistical metric that assesses how similar two non-zero vectors are in space, is frequently applied in text mining and information retrieval applications. In order to ensure that the most pertinent information is at the front of the list of results for user access, ranking entails sorting the retrieved data in a decreasing sequence based on such cosine similarity scores.

Developing TF-IDF and collection frequency matrices for the acquired data is the first step in the procedure. This entails turning words into numerical vectors that show how significant a phrase is in a document in relation to a set of texts.

Next, the cosine of an angle formed by these TF-IDF-CF vectors is calculated using the cosine similarity method. Greater content similarity is indicated by higher cosine similarity scores. The material is then arranged in order of most similarity and relevance using these ratings.

Results are arranged in order of priority in the data according to their similarity index. The similarity between the query and the obtained data is reflected in this index. The relevance increases with the similarity index. Consequently, the findings are exhibited in descending order such that, as Figures 5, 6 and 7 demonstrate; the data with the greatest similarity index is displayed clearly at the top. By enabling consumers to view the most relevant and closely connected information first, this method improves user experience overall and facilitates efficient consumption. Furthermore, the format used for display—categorized or paginated—complements this ranking and expedites user engagement with the obtained material.

Figure 5 - TF-IDF-CF Vector for Sample User Query

Figure 6 - Ranked Data with Similarity Index Based on Above User's Query

Figure 7 - Ranked Data with Similarity Index Based on Above User's Query

5. PERFORMANCE EVALUATION OF PROPOSED SOLUTION

By examining conventional centralised approaches and using blockchain technology with Search Engine Optimisation (SEO), the study sought to reinvent search engine paradigms. Because they relied on centralised hierarchies, traditional search engines were criticised for their exclusive authority-based monitoring and ranking, which led to questions regarding the veracity of search results. As an illustration of its enormous potential, the use of blockchain technology with SEO techniques brought decentralisation, immutability, and transparency to keyword search strategies. The goal of this solution was to guarantee data integrity inside search engine operations while resolving concerns with trust and dependability.

The comparison between blockchain-powered SEO and traditional SEO revealed significant improvements after thorough testing and analysis. Compared to its centralised predecessor, the blockchain-infused search engine optimisation model performed better, showing gains in search accuracy, legitimacy of results, and resilience to manipulation. These findings demonstrated how blockchain technology is revolutionising search engine algorithms and how it may improve security, dependability, and authenticity in search features.

5.1 Comparison with Result

- **Comparison of TF-IDF vs TF-IDF with Collection Frequency for Different Query String and Data sizes.**

The efficiency of two search techniques on a dataset of 100 documents: TF-IDF as well as TF-IDF with Collection Frequency. "Plastics Industry in Papua New Guinea" is the query string that was examined. The table presents the differences in performance between various methods for finding and prioritising pertinent data in the dataset, demonstrating how differently they may handle the same query in this particular data size context.

Table 1 - Comparison 1 - Similarity Index using TF-IDF vs. Similarity Index using TF-IDF with Collection Frequency

Data	Similarity Index using TF-IDF	Similarity Index using TF-IDF-CF	Optimization %
Data 1	0.9255110658	0.9260039108	0.07 %
Data 2	0.5201745634	0.5328520721	2.25 %
Data 3	0.3878936916	0.4164991400	7.32 %
Data 4	0.3866218201	0.4123086896	6.63 %
Data 5	0.3836112820	0.3984465046	4.01 %
Average Optimization			4.06

6. CONCLUSION

Some points that highlight the superiority of the TF-IDF-CF method over other existing methods in a blockchain search engine:

- **Improved Contextual Understanding**: By integrating contextual elements unique to blockchain data, TF-IDF-CF enables a more sophisticated comprehension of the text. This makes it possible for the search engine to understand the nuances and technical jargon of blockchain technology better.
- **Better Relevance Ranking**: TF-IDF-CF can more effectively rank search results according to how relevant they are for the user's query by taking into account more contextual information. This improves the user experience by producing search results that are more relevant and accurate.
- **Decreased Noise and Inappropriate Results**: By removing irrelevant or less significant material, the contextual filtering features of TF-IDF-CF helps lower noise in

search results. Users will be guaranteed to obtain more focused and targeted results as a consequence, increasing efficiency and saving time.

- **Adaptability to Languages Particular to Blockchain Technology**: Blockchain technology frequently has its own distinct lexicon and linguistic norms. Compared to other general approaches, TF-IDF-CF is especially well-suited for scanning blockchain data because of its capacity to comprehend and adjust to its specific vocabulary.
- **Enhanced for Blockchain Application Cases**: Conventional TF-IDF techniques could fall short in capturing the subtleties of blockchain information, resulting in less-than-ideal search outcomes. TF-IDF-CF is optimised for blockchain applications and performs better than other approaches since it considers the special needs and properties of blockchain engines for searching.
- **Scalability and Performance**: TF-IDF-CF retains sustainability and computational efficiency in spite of its additional complexity, guaranteeing that it can efficiently manage massive amounts of blockchain data. This makes it a sensible and trustworthy option for blockchain search apps that are used in the real world.

This research illustrates a paradigm shift towards more precise, context-aware evaluations of document relevance through a thorough investigation of mechanisms such as Term Frequency-Inverse Document Frequency (TF-IDF) combined with collection frequency (CF) inside blockchain frameworks. The work shows a way towards objective, manipulation-free result rankings by using cosine similarity as a criterion to rank search results inside blockchain-stored data, radically changing the SEO environment.

Most importantly, this research goes beyond theoretical boundaries. Thorough examinations conducted on a range of dataset sizes highlight the effectiveness and scalability of blockchain-based SEO tactics. Beyond scholarly circles, this study serves as an appeal to action, enticing scholars, practitioners, and stakeholders to collaborate in creating an SEO environment that puts user privacy, trust, and actual relevance first.

7. REFERENCE

[1] Swan. M, *Blockchain: Blueprint for a new economy*, O'Reilly Media, Inc, 2015.

[2] S.W. Kareem, “Hybrid Public Key Encryption Algorithms For E-Commerce”, Salahadin University, 2009.

[3] E.O. Kareisied, S., M.C. Okur. “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, 2020.

[4] J. Mor, D.Rai, N.Kumar ,”Review on improving performance of Web Crawler and Search System Architecture” , International Conference on Cyber Security (ICCS), 2018.

[5] K. Sharma, G. Shrivastava. "Public key infrastructure and trust of web based knowledge

discovery." Int. J. Eng., Sci. Manage, 4.1, pp 56-60, 2014.

[6] C. Zhang, C. Xu, J. Xu, Y. Tang, B. Choi,, "GEM2-Tree: A gas-efficient structure for authenticated range queries in blockchain", IEEE International Conference on Data Engineering, pp. 843–853, 2019.

[7] K. Aggarwal, R. Nath, N. kumar," A Hybrid Approach for Page Relevance Computing", International Conference on Cyber Security (ICCS), 2018.

[8] A. Kaushik, A. Choudhary, , C.Ektare, D.Thomas, S.Akram, ." Blockchain—literature survey". IEEE International Conference on Recent Trends in Electronics, Information & Communication Technology (RTEICT), pp. 2145-2148, 2017

[9] B. Aggarwal, D. Rai, N. Kumar, "Evaluation of Search Engine Optimization Based on Blockchain Technology "International Conference on Technological Advancements in Computational Sciences, IEEE Explore, pp 529-534, 2023

[10] Y. Zhang, R.H. Deng, J. Shu, K. Yang, D. Zheng, TKSE, "Trustworthy keyword search over encrypted data with two-side verifiability via blockchain", IEEE Access 6 31077–31087, 2018.

[11] Y. Yang, H. Lin, X. Liu, W. Guo, X. Zheng, Z. Liu, "Blockchain based verifiable multi-keyword ranked search on encrypted cloud with fair payment", IEEE Access 7 140818–140832, 2019.

[12] Chen, W.K. Lee, C.-H. Chang, K.-K. Choo, N. Zhang "Blockchain based searchable encryption for electronic health record sharing", Future Gener. Comput. Syst. 95 420–429, 2019.

[13] J. Niu, X. Li, J. Gao, Y. Han, "Blockchain-based anti-key-leakage key aggregation searchable encryption for IoT", IEEE Internet Things J. pp. 1502–1518, 2020.

[14] Meng Shen, Yawen Deng, N. Guizani, "Privacy-preserving image retrieval for medical IoT systems": A blockchain-based approach, IEEE Netw, pp. 27–33, 2019.

[15] A.-S. Kleinaki, P. Mytis-Gkometh, G. Drosatos, P.S. Efraimidis, E. Kaldoudi, "A blockchain-based notarization service for biomedical knowledge retrieval", Comput. Struct. Biotechnol. J. 16 , pp. 288–297, 2018.

[16] A. Rahman, G. Loukas, S. Abdullah, A. Abdu, S. Rahman, E. Hassanain, Y. Arafa, "Blockchain and IoT-based secure multimedia retrieval system for a massive crowd: Sharing economy perspective", ACM International Conference on Multimedia Retrieval, pp. 404–407, 2019.

[17] S.Y. Nikouei, R. Xu, D. Nagothu, Y. Chen, A. Aved, E. Blasch, "Real-time index authentication for event-oriented surveillance video query using blockchain",The 1st International Workshop on BLockchain Enabled Sustainable Smart Cities, pp. 1–8, 2018.

[18] Salton. G & McGill. M, (1986) *Introduction to modern information retrieval*, McGraw-Hill.

[19] Sparck Jones. K, (1972), A statistical interpretation of term specificity and its application to retrieval, *Journal of Documentation*, vol. 28, 1, pp. 11-21.

[20] Manning. P.D, Schütze. H, (1999), *Foundations of statistical natural language processing* , MIT Press.

[21] Baeza-Yates. R & Ribeiro-Neto. B, (2011), Modern information retrieval, ACM Press.

[22] Androulaki . S, Kokolakis. E, Maniatis. P, Russell. A, Sousa. A, & Vardoulakis. I, (2016), Designing a decentralized ledger of identities for the Internet, *ACM SIGSAC Conference on Computer and Communications Security*, pp. 1467-1484

[23] Buchinger. S, Schranz .M , & Müller. C ,(2023), The Future of SEO: A User-Driven Content Model for the Decentralized Web, *ACM Internet Measurement Conference (IMC)*, pp. 501-512

[24] W, Song, P. Mittal, X. Jiang, and B. Yavuz, (2021), Differentially Private Search: A Survey, *ACM Computing Surveys (CSUR), vo*l. 54, no. 5, pp. 1-50.

[25] Park. J, Seo. M, S. Choi, D. Seo, & J. Kim, (2022), Towards a Human-Centered Search Engine Design: A Review of Human-Computer Interaction Studies, *Journal of Information Science Theory and Practice (JIS)*, *vol*. 13, no. 4, pp. 1-20.

[26] Anderson. M (2010), *Search Engine Optimization: The Beginner's Guide*, O'Reilly Media.

[27] Tapscott. D & Tapscott. A, (2016), *Blockchain Revolution: How the Technology Behind Bitcoin is Transforming Money*, *Business*, *and Our World*, Penguin Random House.

[28] Calderón-Monge E., & Ribeiro-Soriano D, (2023), Artificial intelligence and blockchain integration in business: Trends from a bibliometric-content analysis. *Information Systems Frontiers*, pp. 1-17.

[29] 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), pp. 339-348.

[30] 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), pp. 1509-1520.

[31] Zhang, J., Lu, J., Ma, X., & Jiang, X, (2023), Blockchain-based decentralized search engine: A technical perspective, *Computers & Security*, pp. 134, 102755.

[32] Zhang, Y., Lu, X., & Zhang, J, (2023), A study on the use of TF-IDF in decentralized search engines for information retrieval , *Journal of Network and Computer Applications*, pp. 219, 103820.

[33] Zhang. Y, Lin. M, & He. Q, (2019), Blockchain-based decentralized search engine: Towards a secure, transparent, and trustable web search, *IEEE Access, vol*. 7, pp. 120578-120594.

[34] Nguyen. T, N, Tran, V. A., Doan, A., & Dat, T. T, (2023), Enhancing Text Retrieval in Blockchain-Based Decentralized Search Engines with Semantic Indexing and TF-IDF Weighting, , *International Journal of Semantic Web and Information Systems*, 43(1), pp. 3-18.

[35] Singh, N., Srivastava, D., & Chatterjee, S, (2022), Demystifying decentralized search: A comparative analysis of blockchain-based search engine proposals , *International Conference on Intelligent Systems and Information Management (ICISIM*), pp.