Multimedia Tools and Applications (2022) 81:36805–36831
https://doi.org/10.1007/s11042-022-13500-z

1219: MULTIMEDIA SECURITY BASED ON QUANTUM CRYPTOGRAPHY AND BLOCKCHAIN

Latent DIRICHLET allocation (LDA) based information modelling on BLOCKCHAIN technology: a review of trends and research patterns used in integration

Chetan Sharma [1] · Shamneesh Sharma [2] · Sakshi [3]

Received: 2 June 2021 / Revised: 16 September 2021 / Accepted: 13 July 2022 /
Published online: 20 August 2022

Abstract
The past decade is known as the era of integrations where multiple technologies had integrated, and new research trends were seen. The security of data and information in the digital world has been a challenge to everyone; Blockchain technology has attracted many researchers in these scenarios. This paper focuses on finding the current trends in Blockchain technology to help the researchers select an area to carry future research. The data related to Blockchain Technologies have been collected from IEEE, Springer, ACM, and other digital databases. Then, the formulated corpus is used for topic modelling, and Latent Dirichlet Allocation is deployed. The outcomes of the Latent Dirichlet Allocation model are then analyzed based on various extracted key terms and key documents found for each topic. All the topic solution has been identified from the bag of words. The extracted topics are thereafter semantically mapped. Thus, based on the analysis of more than 900 papers, the most recent research trends have been discussed in this paper, ultimately focusing on the areas that need more attention from the research community. Also, the meta data analysis has been accomplished, evaluating the year wise and publication source wise research growth. More than 15 research directions are elaborated in this paper, which can direct and guide the researchers to pursuit the research in specific trends and also, find the research gaps in various technologies associated with Blockchain Technology.

✉ Chetan Sharma
chetanshekhu@gmail.com

Shamneesh Sharma
shamneesh.sharma@gmail.com

Sakshi
sakshi@chitkara.edu.in

[1] Chitkara University, Solan, Himachal Pradesh, India

[2] School of Computer Science & Engineering, Poornima University, Jaipur, Rajasthan, India

[3] Chitkara University Institute of Engineering and Technology, Chitkara University, Rajpura, Punjab, India

Keywords Blockchain · Security · Ledger · LDA · Cryptocurrency · Topic modelling

1 Introduction

The blockchain concept is based on databases that record every transaction happening over the network in a decentralized manner [96]. It stores every transaction electronically on a computer system that is part of the network. This database is based on two important aspects, decentralization and transparency [83]. The blockchain system is known as the most accurate system as every transaction proceeding over the blockchain network is thoroughly corroborated and permitted by the innumerable number of nodes that are part of the whole systematic arrangement. Data security increases with the decentralization of the data over the entire network [99]. At any point in time, a new node can be added to the blockchain network, and this information is updated on every node on the network. Spreading this information across the network rather than storing it on a centralized database makes it more arduous to fiddle with [48].

Moreover, most of the blockchain networks are based on public database technology which means the network's transaction information is open to all the nodes. We can simply consider blockchain technology using the concept of transparency. Along with the blockchain, a new term that came into existence is cryptocurrency and is the most sparkling reason behind blockchain technology's popularity. Another reason may be COVID-19 which started the tread of a cashless economy in most countries [26]. Nowadays, some Economic and Governmental agencies are interested in cryptocurrencies, and they have started promoting them. Now when we can see the future as cash-free, the financial transaction needs to be more secure than ever. Cryptocurrencies have come up with a new wave in blockchain technology. The new technologies are evolving at a pace, and researchers are coming up with new integration ideas. Cloud Technology with the Internet of Things [35] is the integrated technology with various segments like supply chain management, healthcare, and autonomous vehicles [91]. Digitalization was also a major concern in the previous decade for most countries, but tremendous growth has been seen in this industry in the last five years. Now the need for data security has also increased, which is again fulfilled by blockchain technology. People have started investing in blockchain technology. These days, blockchain is used in almost every segment of the industry, including baking and finance, currency, healthcare, records of property, smart contracts, supply chain, voting, and the list is endless as shown in Fig. 1. In this study, section 2 comprises the literature review of the article related to blockchain technology. In section 3, challenges in the field of blockchain have been discussed, section 4 provides details about topic modelling, and section 5 comprises the research questions. In section 6, research methodology is detailed, section 7 gives a detailed discussion of results, and section 8 provides the solution to the research question provided in section 5. In the end, the conclusion of the study is provided in section 9.

analysis Some of the major areas where this technology can make an impact are discussed below:

- *Blockchain and Internet of Things:* The common attribute between Internet of Things (IoT) and Blockchain is distributed network management system so that these technologies can be integrated easily. The technology of the Internet of Things (IoT) has been adopted

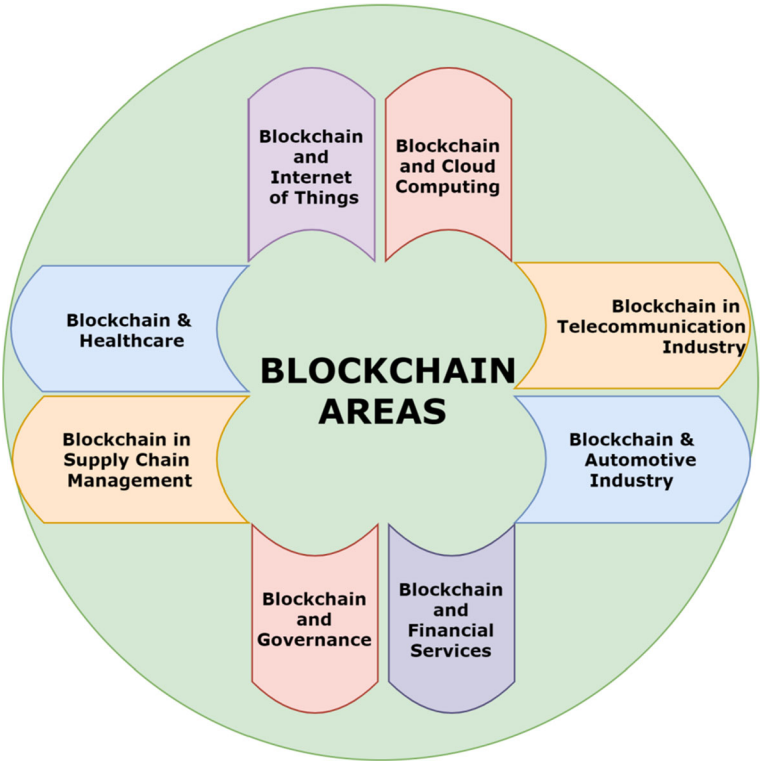

Fig. 1 Blockchain areas

by the major industries; therefore, the concern about data security is increased [110]. The concept of the smart contract can be applied to Internet of Things Ecosystem for the automatic execution of an agreement once certain conditions are encountered.

- *Blockchain and Cloud Computing:* In order to ensure the security of cloud computing, which is based on a multi-structural computing environment, it has been a concern of the technology since its inception [109]. The widespread acceptance of blockchain technology signifies the acknowledgement of a high level of transaction accomplishment in multi-structural computing environments. The combination of these two technologies will open up entirely new vistas in the field of computer science.
- *Blockchain in Telecommunication Industry:* Data Privacy, Operational Efficiency, Customer Experience, and Business Innovations are some of the challenges of the telecommunication industry [102]. The evolution of smartphones and unleashed unprecedented convenience has come to the telecommunication industry, but it has created massive operational complexity. The integration of blockchain in telecommunication will help deal with data privacy, operational efficiency, customer experience, and business innovation.
- *Blockchain and Automotive Industry:* Autonomous vehicles are the future now. Self-driving vehicles are coming into the transportation system, whereas so many legal issues related to security are associated with this technology [73]. The combination of blockchain to automotive vehicles may help the researchers to get these issues solved.

- *Blockchain and Financial Services:* The banking and financial sector [25] may see a major drift in the transaction system using Blockchain Technology. Blockchain technology will bring trust, simplicity, and enhanced customer experience to financial services.
- *Blockchain and Governance:* Smart governance needs the smart handling of data by the administration. By mechanizing the superfluous practices and sharing data among trusted networks in a distributed way, it can bring a new shift to governance and administration. Blockchain technology can decrease the outmoded resistance among organizations and unravel the value elongated ensnared [4].
- *Blockchain and Healthcare:* The major challenge in the healthcare and life science segment is proprietary [75]. After the COVID-19 pandemic, healthcare and life sciences face new challenges like familiarising supply chains to provide shielding to the paraphernalia and promptly emerging treatments with various tests and vaccines. The sanity of these records is very important in case of such a dangerous disease; here, blockchain can help the industry to grow.
- *Blockchain in Supply Chain Management:* The smart tracking system is the main feature that can improve the Supply Chain Management System [101]. Applying blockchain technology to the Supplier Performance Management will provide permanent transparency and validation of transactions among the multiple supply chain partners.

2 Related literature

The concept of blockchain was coined in 1991 by two research scientists who proposed a computationally practical solution to prevent digital documents from tempering [49]. They used a time stamping mechanism so that no one can make them backdated. Their system was based on a cryptographically secured database called block and stored the time-stamped documents. The time stamping was based on the hierarchical hashing mechanism. Later on, in 1992, Markle Trees [5] were introduced in this blockchain design, which allowed the dependent connections between the records stored over the database. This design was patented by Stuart Haber and W. Scott Stornetta in 1992 and got lapsed in 2004. The 1990 decade ignited the researchers to work on this area. In 2004, Hal Finney, a cryptography activist, has developed the concept of Reusable Proof of Work (RPoW) [97] as a prototype for digital cash. The entire journey and evolution of Blockchain technology has been vividly depicted in Fig. 2.

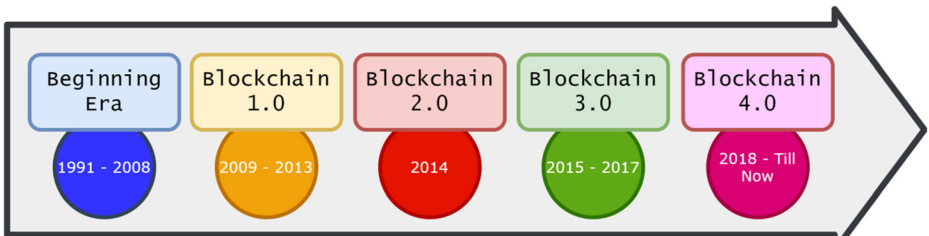

Fig. 2 Evolution of blockchain technology

Further, in 2008, Satoshi Nakamoto abstracted the principle of decentralized ledger through a white paper on the distributed blockchain [86]. Design has been inproved by the author in different manner where they add blocks to first chain and which does not require any sign in by the trusted parties. Currently, Blockchain technology is in the third version. In the year 2009, Blockchain 1.0 [8] was started and was dedicated to cryptocurrencies, whereas Blockchain 2.0 [123] was started in the year 2014 with the concept of smart contracts. After 2015, we have entered into the next generation of blockchain, focusing on the extensions of blockchain applications [70]. Now we have entered into the era of Blockchain 4.0 (W. [135]), where its focus is on developing a perfect ecosystem for this technology. Table 1 is describing the overall development in blockchain technology:

3 Challenges

3.1 Blockchain integration and interoperability

Blockchain technology was introduced in 2008, and currently is not sufficiently mature for enterprise adoption due to limited interoperability. There is still a lot of work to be done before integrating Blockchain with other technologies. Our research has sought to lessen the barriers to broad blockchain awareness by addressing the major blockchain use cases across multiple industries.

3.2 Technical competency space

While blockchain is still relatively new, those who possess these skills are few. Blockchain skills have been highly sought after in the market for some time. Entrepreneurs' further concerns about using blockchain and integrating it with older systems have to do with the expense and complexity of acquiring talent in this field. In order to provide organisations and researchers with the right understanding of the technology, as well as motivate them to pursue the technology, this paper focuses on research topics that are trending right now.

3.3 Indigence of standardization

Despite a number of companies having already adopted the blockchain technology, blockchain adoption in enterprises is still in the early stages. The most recent developments in the industry, like the Internet of Things, healthcare, finance, supply chain management, and energy, have brought hope to the industry, but more needs to be done. There is a necessity for standardisation, greater privacy policies, and better integration. The enterprises will benefit from this paper's analysis of the regulatory and protocol standardisation issues associated with this technology.

4 Topic modelling

Analytics is all about getting the information from the raw data, and in this technological era, data is growing day by day, mostly in an unstructured format. A huge amount of data in

Table 1 Development of blockchain technology

Timeline	Blockchain Version	Year	Development(s)
1991–2008	Beginning Era	1991	Stuart Haber and Scott Stornetta worked on the first Blockchain and presented a scheme for thwarting the digital document from annealing [49]
		1992	Markle Tree-based design was patented by Stuart Haber and W. Scott Stornetta [13]
		2004	The patent of Stuart Haber and W. Scott Stornetta got Lapsed.
			The concept of Reusable Proof of Work (RPoW) by Hal Finney came into existence [97]
		2008	Satoshi Nakamoto released a White Paper on Bitcoin, which was about all the concepts related to Bitcoin currency[85]
2009–2014	Blockchain 1.0	2009	The first bitcoin transaction took place between Hal Finney and Satoshi Nakamoto. [16]
			The Bitcoin Market has started establishing itself after the bitcoin transaction [31]
		2010	The first Bitcoin purchase took place for 10,000 [84]
		2011	Bitcoin exchange value reaches equal with the US dollar [113]
		2013	Bitcoin Marketplace Surpasses $1 Billion
			A white paper on Ethereum was released by Vitalik Buterin [124]
			Bitcoin reached a Market capitalization of 1 Billion dollar
			The first major virtual currency robbery happened from the wallet of Bitcoin Forum Founder
			China's central bank bars financial institutions from handling bitcoin. [121]
	Blockchain 2.0	2014	Bitcoin classified by HMRC as private money and declared new policy describing VAT would not be applicable on mining or exchange of bitcoins.
			Through crowdfunding, the Ethereum Project was propelled as the very first smart contract system [94]
2015–2017	Blockchain 3.0	2015	Ethereum is marked as the second Blockchain
			Linux Foundation revealed the Hyperledger to enhance Blockchain development [7]
			NASDAQ started the trial of a blockchain [68]
			R3 Blockchain Technology was designed and a conglomerate of over 40 inherited financial systems like Arclays, Credit Suisse, Goldman Sachs, JP Morgan, and RBS for implementing Blockchain Technology [133]
		2016	The DAO (Decentralised Autonomous Organisation) sets a record by raising more than 150 million investment [112]
			The DAO lost approximately 50 million dollars in a vulnerability attack
			The industry has started adopting the Blockchain for commercialization in few segments like IoT, AI, Cloud Computing, and many more.
		2017	Seven major European banks announced a DigitalTrade Chain as a collaboration for offering a trade finance policy by using blockchain [27]
			Japan has officially announced the acceptance of virtual currencies in its financial system.
			The was a security breach in the Bitcoin exchange "Bithumb".
			The number of bitcoins in dissemination reached at the number 16.5 million
2018-Till Now	Blockchain 4.0	2018	Switzerland has started receiving its tax disbursements in Bitcoin
			At the beginning of 2018, Bitcoin plummeted to around $10,000 and continued there for some time[115].
			SEC has refused to commend the bitcoin as Exchange-Trade Fund (ETF).
		2019	New cryptocurrency-backed financial vehicles launched [76]
			In August 2019, the Litecoin network practiced its second-ever shared event, and the LTC block reward was abridged to half.
			BSV got the 2GB blocks [82]
			BSV outshines the daily transaction count over BTC and BCH
		2020	Blockchain As A Service (BAAS) was offered by Big Tech Companies. [50]

Table 1 (continued)

Timeline	Blockchain Version	Year	Development(s)
			Amalgamated Blockchain reached out to the Middle Stage Social Networking Problems has got the solution through Blockchain Technology Interoperability in Blockchain Networks became the major concern for the researchers. Economic and Financial services are led by the new Blockchain Applications Integration of blockchain into the Government Agencies of various countries.

unstructured form creates many problems in extracting the relevant and desired information [69]. Technology is enhancing day by day to solve various problems and extract relevant information from the unstructured data topic modeling is one of the most powerful techniques in text mining [18]. It is a machine learning technique used to analyze text data to conclude topics from that automatically. Topic modeling does not require any training type, so it is easy to analyze the data. Topic modeling is a technique used under Natural Language Processing (NLP). Researchers from various fields are using this technique to find the relationship between the topics and documents. This technique is used in fields like medical, engineering, social media analysis, etc. Latent Dirichlet Allocation (LDA) is one of the most popular techniques of topic modeling used by various researchers. In this process, key terms are extracted from the corpus, which further formulates Bag of words (BOW) used to extract features.

BOW creates a vocabulary dictionary from these words, which helps us to analyze the right content rather than traversing the whole document. This technique is used to draw a relationship between the collected corpus and results, which are further represented statistically and graphically. LDA works on probability distribution, which helps each document contain in the dataset, is represented by the probability distribution of topics. Each topic represents the key terms related to that topic. This concludes the connection between the topics and key terms. LDA is used to conclude the patterns or recent trends in any field example, analyzing the data for hot topics or trends posted on Twitter [9]. It is implemented in a various field some recent research is Internet review analysis [55], Agriculture [65], Software Engineering [56], Environment [107], Deep learning (Y. [125]), Medical [137] and many more. In this study, LDA is deployed to the abstraction of trends and emerging topics of blockchain technology.

5 Research questions

The LDA topic modeling has been used identification of research patterns in terms of blockchain technology. For the purpose of the mentioned objectives, the LDA has been deployed on as many as 933 studies published from 2010 to 2021. The core intent for corpus formation and the LDA model is to sublime the core research areas, associated research areas, and research trends. Thus, the research questions are designed to extract and highlight the research areas explored extensively with a keen focus on those research trends that have been less concentrated by the researchers which is represented in Table 2.

Table 2 Research questions and motivation

Research Questions	Motivation
RQ1. What are the extensively explored areas of research in Blockchain technology?	This question corroborates those research areas that have been thoroughly explored and researched by the practitioners and researchers. The recent research era is focused on some research areas; this question aims to highlight those research areas.
RQ2. What have research models been used by the researchers to provide security to different domains using Blockchain?	The intend of this research question is to identify and ferret out the distinct research models providing security that have been deliberately deployed in different applications of Blockchain.
RQ3. Which research areas demand greater attention from researchers?	The rationale behind this research question is to extract and identify those research sections and areas that are suffering from research bias and needs more investigations in the future.

6 Methodology

This section involves all the aspects of experimentation and activities undertaken for performing the analysis of research on Blockchain technology. To predict and depict the research trends in Blockchain Technology, a stepwise procedure has been followed. The detailing of the steps for accomplishing this research has been illustrated in Fig. 3.

6.1 Dataset formation

6.1.1 Sources identification for search

The sources for dataset collection are crucial when we look for research articles according to our required topic. Here the papers related to Blockchain's research have been extensively looked for at online digital databases like IEEE Explore, ACM, Wiley, Springer, Elsevier, Science Direct, and other digital sources. Other than these, the articles have been fetched from various journals and conference portals. To collect the data, various steps are included, which are explained below.

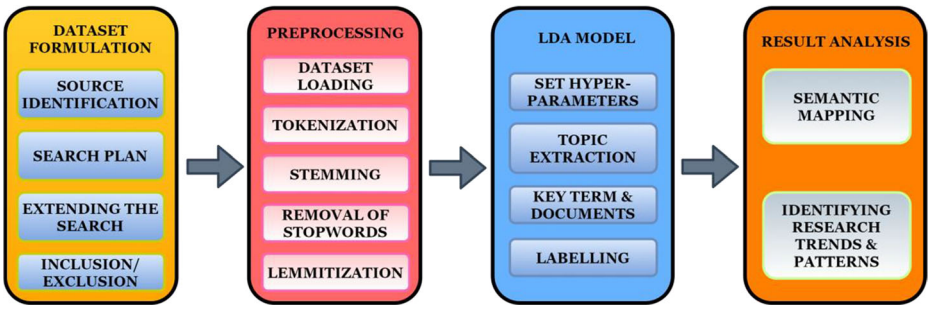

Fig. 3 Research methodology

6.1.2 Search plan

When we want to search the data, some keywords are identified, which is decided by the research question considered in the current study. The key phrases used for searching through various digital libraries are "blockchain" or "blockchain technology". The major search criteria confirmed to relevancy and recency.

6.1.3 Extending the search

Open-source tool Publish and Perish is used to search the studies comes under the keywords mentioned. This tool is used to extract the published information from the main published libraries like Google Scholar, Scopus and Web of Science. *"blockchain"* or *"blockchain technology"* key phrases are used by authors to get maximum results. Some popular databases like IEEE, ACM, Wiley, etc., are also searched with the same keyphrases. The publication databases is searched for Blockchain technology keywords in the publications title for the Year 2000 to Year 2021. A total of 1837 articles has been fetched from the initial search.

6.1.4 Inclusion and exclusion criteria

The author used different inclusion and exclusion criteria in this study in that Firstly, only those articles selected written in the English language. In the database, studies that are relevant to blockchain technology are taken into consideration. One of our inclusion principles is that a study must be published in a reputable, authentic source in order to be considered for consideration. Studies published on and after the year 2000 are included. Studies that are not written in the English language or any regional native language are excluded from the study. Studies published in authenticating journals and conferences are included; the rest are excluded. After this step, 993 studies have been considered for the purpose of the current research and review. The task of the document collection accompanies the process. The publication count over the years is illustrated by Fig. 4, and the dominant publication channels have been represented in Fig. 5, which vividly depicts the contribution of each publication source.

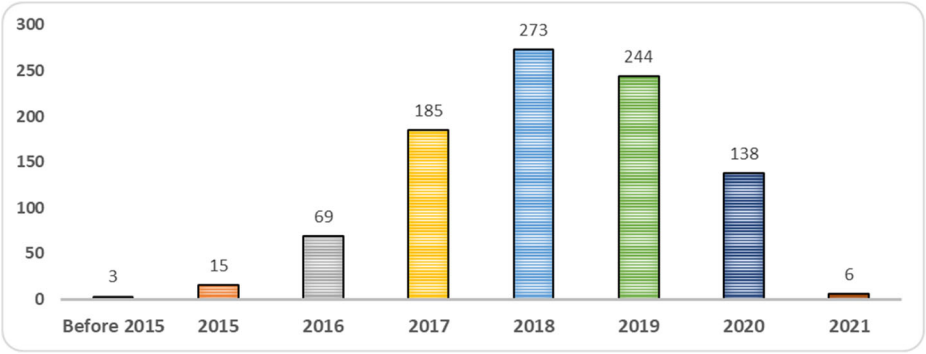

Fig. 4 Year publication analysis

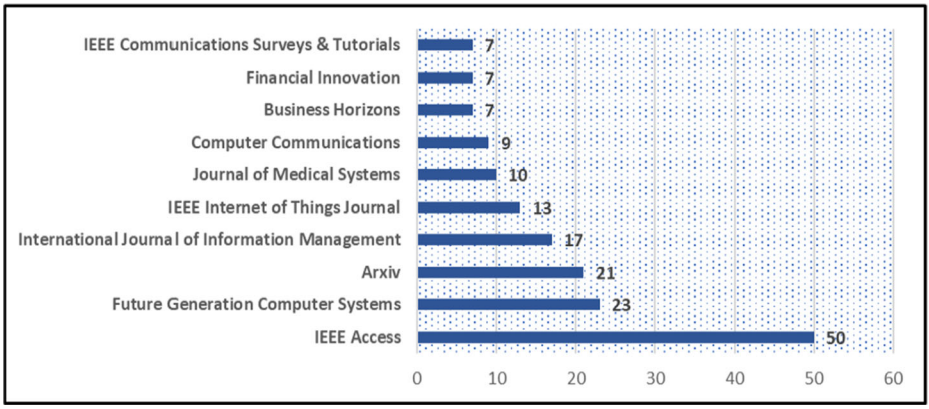

Fig. 5 Dominating journals analysis

6.2 Preprocessing

Datasets are collected from different sources, which are not in the same format, and for that, preprocessing of data is processed on the dataset. Extraneous information is discarded from the collected information. Pre-processing is used to remove the noisy, unwanted, and unrelevant data from any corpus. Here this to remove noisy words, characters different steps are used. This helps us to improve the quality of the dataset. As a result, the profile of further processing becomes more accurate and acceptable. Basis stepts followed to preprocess the corpus is as follows and preprocessing of sample document is shown in Table 3:

Table 3 Preprocessing steps

Sample Document	Blockchain Technology is latest technology which is used by many industry and researchers. It mostly deal in privacy, security and ctyptocurrency. Blockchain technology is used in many real life applications. This technology start evoving itself from the year 2014. Blockchain is used by different researchers in the field of medical, agriculture and engineering.
After Tokenization	"Blockchain", "Technology", "is", "latest", "technology", "which", "is", "used", "by", "many", "industry", "and", "researchers",".", "It", "mostly", "deal", "in", "privacy",",", "security", "and", "ctyptocurrency",".", "Blockchain", "technology", "is", "used", "in", "many", "real", "life", "applications",".", "This", "technology", "start", "evolving", "itself", "from", "the", "year", "2014″,".", "Blockchain", "is", "used", "by", "different", "researchers", "in", "the", "field", "of", "medical", "agriculture", "and", "engineering","."
After Stop Word Removal	"Blockchain", "Technology", "latest", "technology", "industry", "researchers", "deal", "privacy", "security", "ctyptocurrency", "Blockchain", "technology", "real", "life", "applications", "technology", "start", "evolving", "year", "Blockchain", "different", "researchers", "field", "medical", "agriculture", "engineering"
Stemming	"Blockchain", "Techno", "latest", "techno", "industry", "research", "deal", "privacy", "secur", "cryptocurrency", "Blockchain", "techno", "real", "life", "applica", "techno", "start", "evolv", "year", "Blockchain", "different", "research", "field", "medical", "agriculture", "engineering"
Lemitization	"Blockchain", "Technology", "latest", "technology", "industry", "researchers", "deal", "privacy", "security", "ctyptocurrency", "Blockchain", "technology", "real", "life", "applications", "technology", "start", "evolving", "year", "Blockchain", "different", "researchers", "field", "medical", "agriculture", "engineering"

6.2.1 Dataset loading

The raw dataset is collected from the various sources mentioned in the previous section, which further narrows down to a single excel file. After conducting all manual filtration steps like inclusion/ exclusion criteria, repeated data, etc., final data is stored in an excel file, further loaded to the LDA model for topic modeling.

6.2.2 Tokenization

The first step followed in pre-processing is tokenization, in which data contained in the gathered dataset is tokenized into tokens. Further, all the tokens generated were converted to lowercase letters for every data related to created dataset. Now all the punctuation marks, single characters, and other special characters like ";", ",", ".", "/", "\", "brackets", "!" are removed from the corpus. Further, any kind of equation or formula used in the abstract was removed. Numerical values are also discarded from the data[128].

6.2.3 Stemming

Stemming plays an important role in natural language processing and its understanding. This is when identifying the core word from the various words after removing its prefix and postfix. Let's understand from example stem is the main word that can be extracted from the word "stemming", "stemmed", and "stemmer". The stemming process removes the extra part to root out the meaningful word. For the preparation of an effective corpus, words stem from their original form using the Snowball stemmer algorithm [98] and the resulting base keywords stored in the cleansed corpus.

6.2.4 Removal of Stopwords

English is the language that includes some common words like "the", "if", "but", "a" or "an," etc., which make sentences complete, but these words take some space in a corpus, which consumes processing time. So, it must remove the stop words from the corpus, and in this study, Natural Language Toolkit (NLTK) is used to experiment. This toolkit contains the stopwords of sixteen languages [47].

6.2.5 Lemmatization

The words which were previously stemmed needs to be lemmatized. Lemmatization is when the context is considered and stemmed words are converted into more meaningful base words or lemmas. This phase targets removing inflected words and outputs the dictionary form of a word [95].

6.3 Eradicating bigrams and trigrams

Frequently together occurring words are known as bigrams and trigrams. For instance, if two words occur together are mostly known as bigrams. And similar;y the three words that occur together are known as trigrams. Here in this step, the attention is given to eradicating such pairs and triplets which do occur together like "blockchain technology", "Blockchain applications,"

and so on. In this implementation, the gensim library has been used for the removal of such phrases. Gensim's Phrases model can build and identify these bigrams, trigrams, quadgrams, or even n-grams (X. [126]), and thus, we can make removal and make the data cleansing process better.

6.4 Deploying LDA model

After completing the preprocessing part as suggested by [18][74], the LDA model has been deployed. Several steps need to be followed after processing the corpus. Majorly three input parameters are involved in the experimentation process of implementing LDA-based topic modeling. These hyperparameters are

1. Alpha, where alpha represents the distribution of documents per topic. In short, it indicates the number of document-topic density
2. Beta represents the per word weight of the topic of Dirichlet before topic-word distributions. In short, it depicts the topic-word density
3. Number of iterations mentions the required number of iterations needed for converging the model.

The selection of several topics is inspired by the works of [9], and the number of considered iterations is 500. The alpha and beta are considered the smoothening parameters that control the distribution over topics and words. A vigilant initialization is needed to define the values for these parameters, as these hyperparameters contribute to the distribution of high-quality topic results. The bag of words extracted initially is processed in LDA topic modeling where the most frequently and least frequently occurring are removed so that corpus could become absolute. The top frequently occurring words are shown in Fig. 6.

6.4.1 Initializing the Hyperparameters

There is two main hyperparameters used in LDA which is α and β. α value varies or depend upon the number of topic choosen in the study which is 1 / T, where T is number of topics. In this study β value is taken as 0.01 for all topic solutions.

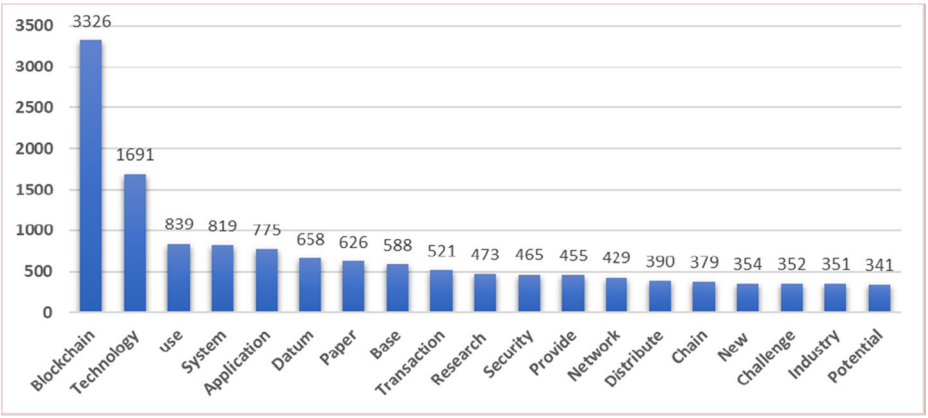

Fig. 6 Top words from corpus with frequency

- Use of *mallet* for choosing T

For optimizing the hyperparameters, a java-based NLP package named mallet is used. The deployment of mallets accomplished the task of the number of topics. According to [106], there is no established measure to defend the optimal number of solutions. However, heuristic parameters suggested by ([9, 23] can be applied to find the optimal range of topic solutions. The choice of the topic solution has been influenced by the heuristics and findings of the studies ([9, 21, 23, 106].

- Algorithmic-base choice of topic

The choice of several topics is also performed algorithmically., An algorithm-based selection has been accomplished using the k means clustering algorithm; the optimal number of topic solutions for identifying research trends has been chosen. The deployment of k means clustering depicts the optimal choice of several topics for depicting the core research areas is five. Thus, five topic solution has been selected optimistically. Also, the researchers have intended to explore for a high ranging value of ten so that the research trends could be more discovered more evidently.

The role of coherence score is also significant while deciding the number of topic solutions. Ideally, the coherence score ranging within the values 0.3 to 0.6 [114] is considered good to have been used. In the stances where high coherence scores have been achieved, the corresponding values of k have been used to choose several topic solutions. In the current scenario, the coherence score obtained for two topic solution is suggestive as 0.49, and the score obtained for five topics optimal solution is 0.51. And for the ten topic solution, the coherence value has been acquired as 0.46. Thus, five topic solution is the optimally best choice.

6.5 Labeling process and high loading articles

The labeling process has been performed after extensive hard work, and every member of our review team has contributed well to the formations of labels for each formulated topic. It is one of the crucial tasks in the process of topic modeling. This review study exceptionally focused on high loading articles under each topic's solutions. The extensive process of labeling each topic solution has been performed after several rounds of meetings and discussions among the research team. This labeling process is accomplished to reach out to conclusive and depictive labels. High loading articles corresponding to the each topic label is represented in Table 4.

7 Result analysis

The result analysis is one of the most highlighting sections of this review paper. Here, the dominance of every topic solution is based on the count of articles extracted under its head. There is a decent possibility that the different topic solutions may likely depict similar research areas. The articles under every heading of topic solution can be favorably overlapping in different label headers. Thus, semantic mapping has been performed to discover the common and overlapping sections of the topic solution (refer Fig. 7). The relationship between the different topic extracted in represented in Table 7.

Table 4 High loading research articles for topic solutions

Topic ID	Key Terms	Topic Label	High Loading Paper	Contribution(%)
2.1	system, datum, transaction, network, security, base, IoT, propose, paper, distribute, privacy, secure, provide, application, trust, bitcoin, record, consensus, solution, smart	Taxonomy and Architecture of Blockchain	[66]	89.72
			[62]	89.47
			[28]	89.42
			[40]	89.1
			[77]	89.06
			[104]	88.89
2.2	application, research, paper, industry, chain, potential, study, supply, business, challenge, provide, financial, case, energy, discuss, base, area, current, future, review	Blockchain implementation and integration in Industry	[88]	89.66
			[89]	89.11
			[29]	89.04
			[14]	88.94
			[93]	88.85
			[105]	88.84
5.1	transaction, system, network, consensus, block, performance, base, propose, bitcoin, model, node, database, distribute, paper, process, protocol, design, time, proof, trust	Protocols and Modeling of Blockchain Technology	[138]	89.76
			[63]	89.74
			[68]	89.4
			[61]	89.25
			[90]	89.07
5.2	datum, system, record, base, data, patient, privacy, secure, share, health, medical, propose, storage, access, healthcare, paper, provide, information, solution, store	Blockchain in Healthcare	[131]	89.18
			[10]	88.98
			[132]	87.71
			[66]	87.09
			[39]	84.61
			[37]	84.26
			[67]	83.89
5.3	application, financial, potential, transaction, system, bitcoin, digital, distribute, contract, trust, business, ledger, sector, smart, market, paper, include, create, provide, industry	Blockchain in the Finance Industry	[64]	89.46
			[119]	89.34
			[117]	89.09
			[71]	88.87
			[22]	88.3
			[22]	87.57
5.4	research, chain, supply, study, paper, industry, application, literature, review, case, provide, adoption, base, management, business, potential, current, identity, benefit, analysis	Blockchain in Supply Chain Management	[89]	83.76
			[33]	83.55
			[46]	82.6
			[127]	82.4
			[129]	81.08
5.5	iot, security, application, internet, energy, smart, challenge, system, thing, privacy, device, network, base, solution, datum, present, paper, provide, research, secure	Privacy and Security of IoT using Blockchain	[43]	85.5
			[12]	83.03
			[78]	81.47
			[130]	87.6
			[44]	82.34
			[17]	79.91
10.1	energy, industry, financial, market, sector, system, peer, trading, article, service, power, social, base, distribute, information, government, efficiency, current, electricity, sustainable	Blockchain technology in Energy Sector	[6]	88.5
			[130]	87.03
			[1]	86.47
			[140]	83.6
10.2	transaction, system, record, peer, application, distribute, ledger, network, base, digital, trust, block, decentralize, chain, store, paper, user, secure, proof, bitcoin	Security and Privacy on Blockchain	[31]	89.52
			[3]	87.75
			[3]	87.52
			[92]	87.49
			[24]	86.22
			[53]	85.33
10.3			[100]	89.61

Table 4 (continued)

Topic ID	Key Terms	Topic Label	High Loading Paper	Contribution(%)
	iot, security, device, network, internet, thing, smart, base, datum, system, challenge, solution, secure, privacy, model, propose, architecture, resource, service, provide	IoT Privacy and Security Challenges for Smart Cities	[17] [136]	88.61 87.14
10.4	transaction, system, network, performance, consensus, propose, node, block, base, model, protocol, paper, time, database, number, work, bitcoin, security, attack, proof	Consensus mechanism in Blockchain Technology	[61] [90] [138]	89.79 88.50 88.32
10.5	bitcoin, financial, business, application, change, currency, digital, potential, service, world, market, trust, impact, transaction, industry, internet, information, include, technological, model	Cryptocurrency and Blockchain	[118] [116] [71] [122] [32] [22]	89.59 88.58 88.44 88.32 87.13
10.6	application, security, paper, issue, challenge, research, provide, system, survey, privacy, discuss, present, future, distribute, solution, internet, recent, review, consensus, industry	Security and Privacy Issues in	[83] [92] [20] [139]	89.79 88.23 88.14 87.13
10.7	chain, supply, management, process, business, food, traceability, logistic, industrial, industry, improve, benefit, product, cost, paper, provide, adoption, transparency, operation, efficiency	Efficiency of Suppychain Management Systems using Blockchain	[19] [36] [30]	89.77 89.71 88.35
10.8	datum, patient, system, base, data, medical, storage, share, privacy, record, health, healthcare, information, propose, provide, access, provider, secure, management, security	Security and Privacy of electronic health record systems using Blockchain Technology	[72] [37] [108] [111]	89.44 89.21 88.99 88.56
10.9	smart, contract, application, system, process, distribute, governance, base, design, paper, platform, trust, software, public, share, information, case, economy, implementation, create	Smart Contract System using Blockchain	[80] [42] [60]	89.92 88.45 88.02
10.10	research, study, paper, industry, potential, literature, application, case, analysis, review, identify, adoption, base, future, provide, implementation, approach, framework, present, business	Strategic Management Frameworks using Blockchain	[52] [57] [127]	89.72 89.34 88.9

Table 5 summarizes the count of publications corresponding to each topic solution.

7.1 Topic labelling

The core research zones that have been explored and discovered based on the two topic solution has been depicted by the topic T2.1 and T2.2. Lets us discuss how this labeling has

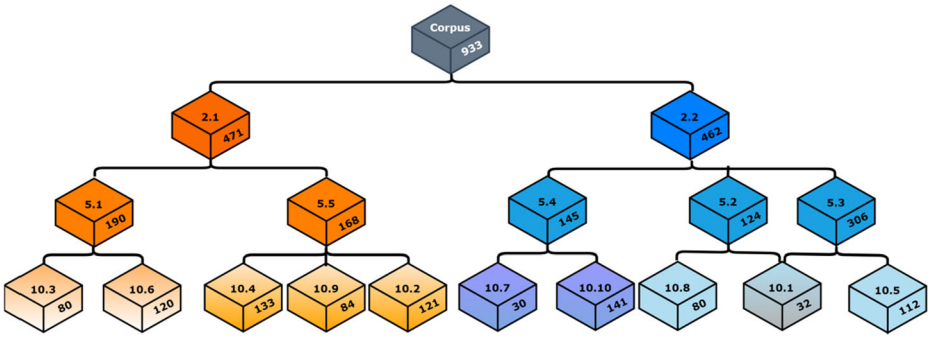

Fig. 7 Semantic mapping of topic solution

been performed. While implementing LDA on two topic solution, the keywords and their loading has been extracted. The extraction results of LDA depict the high loading articles per topic and also the high loading terms or keywords per topic. The labeling process is based on the high-loading keywords that have been collected. Thus, in Table 6, the labeling per topic solution has been performed corresponding to the terms that have been extracted under the heads T2.1, T2.2, and so on; it goes for five and ten topic solutions.

Table 5 Yearwise publication analysis for 2, 5, and 10 topic solutions

T-ID	Topic Name	<2015	2015	2016	2017	2018	2019	2020	2021	Total
2.1	Texonomy and Architecture of Blockchain	3	9	37	95	156	107	60	4	471
2.2	Blockchain implementation and integration in Industry	0	6	32	90	117	137	78	2	462
5.1	Protocols and Modeling of Blockchain Technology	1	3	18	31	67	46	22	2	190
5.2	Blcokchain in Healthcare	0	3	9	31	40	26	14	1	124
5.3	Blockchain in the Finance Industry	2	7	32	88	80	66	31	0	306
5.4	Blockchain in Supply Chain Management	0	0	4	14	32	58	36	1	145
5.5	Privacy and Security of IoT using Blockchain	0	2	6	21	54	48	35	2	168
10.1	Blockchain technology in Energy Sector	0	0	0	13	8	5	6	0	32
10.2	Security and Privacy on Blockchain	1	5	17	28	40	23	7	0	121
10.3	IoT Privacy and Security Challenges for Smart Cities	0	0	4	15	32	14	14	1	80
10.4	Consensus mechanism in Blockchain Technology	1	2	11	21	49	29	19	1	133
10.5	Cryptocurrency and Blockchain	1	3	12	32	27	23	14	0	112
10.6	Security and Privacy Issues in	0	1	6	17	26	46	23	1	120
10.7	Efficiency of Suppychain Management Systems using Blockchain	0	0	1	1	8	12	8	0	30
10.8	Security and Privacy of electronic health record systems using Blockchain Technology	0	0	6	19	22	22	10	1	80
10.9	Smart Contract System using Blockchain	0	4	9	19	30	15	6	1	84
10.10.	Strategic Management Frameworks using Blockchain	0	0	3	20	31	55	31	1	141

Table 6 High loading terms for 2,5 and 10 topic solution

T-ID	Topic Labels	Topic Terms	Count of Studies
2.1	Taxonomy and Architecture of Blockchain	system, datum, transaction, network, security, base, IoT, propose, paper, distribute, privacy, secure, provide, application, trust, bitcoin, record, consensus, solution, smart	471
2.2	Blockchain implementation and integration in Industry	application, research, paper, industry, chain, potential, study, supply, business, challenge, provide, financial, case, energy, discuss, base, area, current, future, review	462
5.1	Protocols and Modeling of Blockchain Technology	transaction, system, network, consensus, block, performance, base, propose, bitcoin, model, node, database, distribute, paper, process, protocol, design, time, proof, trust	190
5.2	Blockchain in Healthcare	datum, system, record, base, data, patient, privacy, secure, share, health, medical, propose, storage, access, healthcare, paper, provide, information, solution, store	124
5.3	Blockchain in the Finance Industry	application, financial, potential, transaction, system, bitcoin, digital, distribute, contract, trust, business, ledger, sector, smart, market, paper, include, create, provide, industry	306
5.4	Blockchain in Supply Chain Management	research, chain, supply, study, paper, industry, application, literature, review, case, provide, adoption, base, management, business, potential, current, identity, benefit, analysis	145
5.5	Privacy and Security of IoT using Blockchain	iot, security, application, internet, energy, smart, challenge, system, thing, privacy, device, network, base, solution, datum, present, paper, provide, research, secure	168
10.1	Blockchain technology in Energy Sector	energy, industry, financial, market, sector, system, peer, trading, article, service, power, social, base, distribute, information, government, efficiency, current, electricity, sustainable	32
10.2	Security and Privacy on Blockchain	transaction, system, record, peer, application, distribute, ledger, network, base, digital, trust, block, decentralize, chain, store, paper, user, secure, proof, bitcoin	121
10.3	IoT Privacy and Security Challenges for Smart Cities	iot, security, device, network, internet, thing, smart, base, datum, system, challenge, solution, secure, privacy, model, propose, architecture, resource, service, provide	80
10.4	Consensus mechanism in Blockchain Technology	transaction, system, network, performance, consensus, propose, node, block, base, model, protocol, paper, time, database, number, work, bitcoin, security, attack, proof	133
10.5	Cryptocurrency and Blockchain	bitcoin, financial, business, application, change, currency, digital, potential, service, world, market, trust, impact, transaction, industry, internet, information, include, technological, model	112
10.6	Internet Security and Privacy Issues	application, security, paper, issue, challenge, research, provide, system, survey, privacy, discuss, present, future, distribute, solution, internet, recent, review, consensus, industry	120
10.7	Efficiency of Suppychain Management Systems using Blockchain	chain, supply, management, process, business, food, traceability, logistic, industrial, industry,	30

Table 6 (continued)

T-ID	Topic Labels	Topic Terms	Count of Studies
		improve, benefit, product, cost, paper, provide, adoption, transparency, operation, efficiency	
10.8	Security and Privacy of electronic health record systems using Blockchain Technology	datum, patient, system, base, data, medical, storage, share, privacy, record, health, healthcare, information, propose, provide, access, provider, secure, management, security	80
10.9	Smart Contract System using Blockchain	smart, contract, application, system, process, distribute, governance, base, design, paper, platform, trust, software, public, share, information, case, economy, implementation, create	84
10.10	Strategic Management Frameworks using Blockchain	research, study, paper, industry, potential, literature, application, case, analysis, review, identify, adoption, base, future, provide, implementation, approach, framework, present, business	141

7.1.1 Two topic solution: Broader concept of Blockchain technology

In this study firstly the corpus is viewed as two topic solution which is broader view of Blockchain technology. Two topic titled as "Taxonomy and Architecture of Blockchain"

Table 7 Semantic mapping of 2 topic, 5 topic and 10 topic

2 Topic Solution		5 Topic Solution		10 Topic Solution	
Topic ID	Topic label	Topic ID	Topic label	Topic ID	Topic Label
2.1	Taxonomy and Architecture of Blockchain	5.5	Privacy and Security of IoT using Blockchain	10.3	IoT Privacy and Security Challenges for Smart Cities
				10.6	Internet Security and Privacy Issues
		5.1	Protocols and Modelling of Blockchain Technology	10.4	Consensus mechanism in Blockchain Technology
				10.9	Smart Contract System using Blockchain
				10.2	Security and Privacy on Blockchain
2.2	Blockchain implementation and integration in Industry	5.4	Blockchain in Supply Chain Management	10.7	Efficiency of Supply chain Management Systems using Blockchain
				10.10	Strategic Management Frameworks using Blockchain
		5.2	Blockchain in Healthcare	10.8	Security and Privacy of electronic health record systems using Blockchain Technology
				10.1	Blockchain technology in Service Industry
		5.3	Blockchain in the Finance Industry	10.5	Cryptocurrency and Blockchain
				10.1	Blockchain technology in Service Industry

(T2.1) and "Blockchain implementation and integration in Industry"(T2.2). These are two major labels that depict the core research areas that have been extensively explored by the researchers. These topics are the broader concept used in blockchain technology which comprises how the block is implemented in the industry and various architecture of blockchain technology.

7.1.2 Research area of Blockchain technology

To get a better insight into the research areas, the core research areas depicted in two topic solutions have been further widened, and five topic solutions have been explored. In the five-topic solution, keywords again played a significant role in the nomenclature of the topic solution. As per the keywords extracted, the labeling is accomplished. After being labeled based on the corresponding keywords' loading values, the identified topics become ready in the absolute state to predict and depict the major research areas that have been extensively researched in the domain of blockchain technology. The identified research areas are presented in topics (T5.1), (T5.2), (T5.3), (T5.4), and (T5.5). Research areas corresponds to this study entiteled as "Protocols and Modeling of Blockchain Technology" (T5.1), "Blockchain in Healthcare" (T5.2), "Blockchain in the Finance Industry" (T5.3), "Blockchain in Supply Chain Management" (T5.4), and "Privacy and Security of IoT using Blockchain" (T5.5). Each of the identified research areas' roles can be distinctly visualized in Fig. 5 based on the count of studies or documents fetched under each topic. Further, the researchers have endeavored to map the identified stigmas under the appropriated core research zones (T2.1) and (T2.2) in Table 5. Figure 8 illustrates the active research areas in the field of blockchain technology.

7.1.3 Recent trends for researcher: Blockchain technology

Quantifying the number of topics extracted by the LDA topic model becomes necessary to get a detailed analysis of what research patterns can be observed through the surfaced keywords. The ten topic solution further resulted in comprehensive research trends in blockchain technology. In ten topic solution, more research trends appeared, namely, "Blockchain technology in Energy Sector" (T10.1), "Security and Privacy on Blockchain" (T10.2), "IoT Privacy and Security Challenges for Smart Cities" (T10.3), "Consensus mechanism in Blockchain Technology" (T10.4), "Cryptocurrency and Blockchain"(T10.5), "Internet

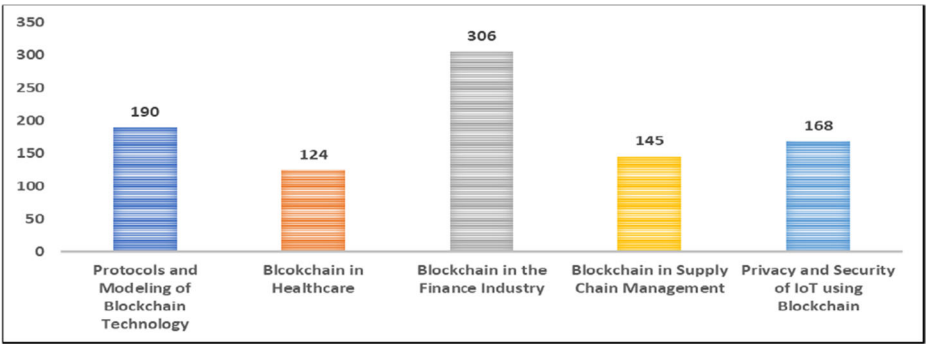

Fig. 8 Trending research area

Security and Privacy Issues"(T10.6), "Efficiency of Suppychain Management Systems using Blockchain"(10.7), "Security and Privacy of electronic health record systems using Blockchain Technology"(T10.8), "Smart Contract System using Blockchain"(T10.9), and "Strategic Management Frameworks using Blockchain"(T10.10). High-loading terms and top five articles for ten topic solutions are depicted in Table 6.

8 Research questions and discussions

In the current study, the researchers aim to summarize the research trends in blockchain technology based on the selected corpus of 993 articles. The corpus formed after the filtration process includes articles from bibliographic databases from the Year 2000 to the Year 2021. We have tried to compile the entire research literature from the year 2000 to the day 2021. Analysis of the corpus for n topic solution has been performed by applying LDA to find, extract and identify the latent research patterns and trends. In this section, an extensive technical discussion has been carried out for each formulated research question. Every answer below presents the gist summary of all the extractions after applying LDA. Thus, all of the three research questions have been discussed given the literature dataset findings, and further future research opportunities have been explored and identified.

8.1 Research question 1

8.1.1 Which research areas have been explored mostly by the researchers?

Taxonomy and blockchain architecture has been prominent areas of research since the beginning. In 2008, after the first bitcoin scheme share by Satoshi Nakamoto, the trend to find out new and prominent systems for cryptocurrencies had been started. The integration of this technology with other technologies was the next phase of the research. After the boom in cryptocurrencies, research has tried to explore the architecture of this technology [120] [134] and implemented it over insecure areas [11]. The decentralized architecture has allowed the researchers to implement it into various segments of the industry. The table shows that the topic labeled 2.1 is further linked to topics 5.1 and 5.5. The privacy and security of IoT networks using the blockchain [38] is the key focal area of research in the initial level of this technology. The integration of both technologies has new challenges [34], but some of the security concerns are addressed by this amalgamation. In addition to the second area of research for blockchain technology implementation and integration of the three prominent areas are identified of technology integration to Supply Chain Management [45], Healthcare [15], and Financial Services [87]. Topic 2.2 is further mapped with three areas which are labels as 5.2, 5.3, and 5.4 in the table.

8.2 Research question 2

8.2.1 What have the researcher's used research models to provide security to different domains using Blockchain?

With the implementation of blockchain over various industrial practices, so many protocols and models have been used to provide strong security mechanisms. There are so many protocols and models that have been implemented to make blockchain more robust and

efficient. Researchers have impended various protocols over various simulators [41][51]. These models and protocols answer questions like keeping the failure probability for a given sharing protocol smaller than a predefined threshold. Most of the courties have started considering the blockchain technology as a reliance machine for accelarting the growth in other technologies also in terms of security and robustness [103]. Many countirs have stared the mission of Smart Cities for the ease of govenance and livelihood of the citizens; as being on the internet this technology need some robust security frameworks for data privacy and security.

8.3 Research question 3

8.3.1 Which research areas demand greater attention from researchers?

With the evolution of internet technologies and increasing role of smart devices in it, the security has been become a critical issue to the researchers. Verious cryptographic techniques are been invented and implemented over internet to make it more secure but still there are various issues which need the attention. Blochchain technology is growing as a platform to provide robust security mechanism for data sharing over the network, that is why it became the favorite technology for researchers to be implemented in Internet of Things (IoT), Finance and Banking, Supply Chain Maanagement, Healthcare Sector and even in food industry. The topics labels from 10.1 to 10.10 are listed as the most demanded areas of research in blockchain technology. The Blockchain technology in Service Industry [58] [54] on SCM in the Indian service industry has given new dimensions to the security services in the industry. There are certain issues associated with privacy and security measures [59] with this technology. Smart City Project has attracted many researchers to it due to its viability in IoT Technology. The security challenges were adhered to by the technology of blockchain [81]. The basic principles and characteristics of the consensus algorithms have allowed them to be sued in smart cities [79]. The first version of blockchain technology was for cryptocurrencies. This is the unprevailing truth behind the technology of blockchain. The cryptocurrency [83] had introduced blackchin technology to the world. The Internet has been the backbone of every technology in computer science, and it has been the most exposed thing in terms of security [34]. The arrival of blockchain technology has brought new opportunities to the field of the internet. The decentralized mechanism made this technology more secure and transparent to its users. The next research area was researchers have shown their interest is “Blockchain Technology in Supply Chain management”. The efficiency [2] of this integration is an important aspect where researchers have to focus upon. With the concept of tempered proof, the digital document has given a new perspective to electronic health record systems [108].

9 Conclusion and future scope

This paper the researchers have considered various articles related to blockchain available on google scholar, IEEE, and ACM libraries for LDA modeling. The study shows that the most common research area in the domain of blockchain is “Taxonomy and Architecture of Blockchain”, followed by “Blockchain Imlementeation and Integtation over various technologies”. These two major areas of research are further linked to the prominent sub areas dealing with other segment of the industry. The year wise progress of this technology has also been

discussed by the researchers alongwith the diffrent generations of the blockchain technology like Blockchain 1.0, Blockchain 2.0, Blockchain 3.0 and 4.0 to show the relation among all the prominent research areas. The generations of blockchain technologies will also help the forthcoming researchers and academician in this field to choose a veracious research pathway. The main aim of the study is to provide an abstract view of the research areas to the exixting as well as prospective researchers. Being an prominent and emerging domain, its integration with many other domains has been seen and it is predicted that this technology will play an important role in the field of security, hance it is recommended by the researchers of this paper that the issues related to the domentic as well as commercial applicactions of this technology need to be adhere critically by academic, research and industrial communities. More than 15 research directions are discussed in this paper which can guide the researchers to find the research gaps in various technologies which are associated to blockchain technology.

References

1. Ahl A, Yarime M, Tanaka K, Sagawa D (2019) Review of blockchain-based distributed energy: implications for institutional development. Renew Sust Energ Rev 107:200–211
2. Aich S, Chakraborty S, Sain M, Lee H, Kim H-C (2019) A review on benefits of IoT integrated blockchain based supply chain management implementations across different sectors with case study. In: 2019 21st international conference on advanced communication technology (ICACT), pp 138–141
3. Ali MS, Vecchio M, Pincheira M, Dolui K, Antonelli F, Rehmani MH (2018) Applications of blockchains in the internet of things: a comprehensive survey. IEEE Commun Surv Tutor 21(2):1676–1717
4. Alketbi A, Nasir Q, Talib MA (2018) Blockchain for government services—use cases, security benefits and challenges. In: 2018 15th learning and technology conference (L \ T), pp 112–119
5. Alvi ST, Uddin MN, Islam L (2020) Digital voting: a Blockchain-based E-voting system using biohash and smart contract. In: 2020 third international conference on smart systems and inventive technology (ICSSIT), pp 228–233
6. Andoni M, Robu V, Flynn D, Abram S, Geach D, Jenkins D, McCallum P, Peacock A (2019) Blockchain technology in the energy sector: a systematic review of challenges and opportunities. Renew Sust Energ Rev 100:143–174
7. Androulaki E, Barger A, Bortnikov V, Cachin C, Christidis K, De Caro A, Enyeart D, Ferris C, Laventman G, Manevich Y et al (2018) Hyperledger fabric: a distributed operating system for permissioned blockchains. Proceedings of the Thirteenth EuroSys Conference:1–15
8. Angelis J, da Silva ER (2019) Blockchain adoption: a value driver perspective. Business Horizons 62(3): 307–314
9. Arun R, Suresh V, Madhavan CEV, Murty MN (2010) On finding the natural number of topics with latent Dirichlet allocation: some observations. Lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics), 6118 LNAI(PART 1), 391–402. https://doi.org/10.1007/978-3-642-13657-3_43
10. Azaria A, Ekblaw A, Vieira T, Lippman A (2016) Medrec: using blockchain for medical data access and permission management. In: 2016 2nd international conference on open and big data (OBD), pp 25–30
11. Ballandies MC, Dapp MM, Pournaras E (2018) Decrypting distributed ledger design–taxonomy, classification and Blockchain community evaluation. ArXiv Preprint ArXiv:1811.03419
12. Banafa A (2017) IoT and blockchain convergence: benefits and challenges. IEEE Internet of Things J 9
13. Bansod S, Ragha L (2020) Blockchain technology: applications and research challenges. In: 2020 international conference for emerging technology (INCET), pp 1–6
14. Beck R, Müller-Bloch C (2017) Blockchain as radical innovation: a framework for engaging with distributed ledgers as incumbent organization. In: Proceedings of the 50th Hawaii international conference on system sciences
15. Bell L, Buchanan WJ, Cameron J, Lo O (2018) Applications of blockchain within healthcare. Blockchain Healthc Today 1(8):10.30953/bhty.v1.8

16. Bhadoria RS, Arora Y, Gautam K (2020) Blockchain hands on for developing genesis block. In: Kim S, Deka G (eds) Advanced applications of blockchain technology. Studies in Big Data, vol 60. Springer, Singapore. https://doi.org/10.1007/978-981-13-8775-3_13
17. Biswas K, Muthukkumarasamy V (2016) Securing smart cities using blockchain technology. 2016 IEEE 18th international conference on high performance computing and communications. In: IEEE 14th international conference on Smart City; IEEE 2nd international conference on data science and systems (HPCC/SmartCity/DSS), pp 1392–1393
18. Blei DM, Ng AY, Jordan MI (2003) Latent dirichlet allocation. J Mach Learn Res 3:993–1022
19. Bosona T, Gebresenbet G (2013) Food traceability as an integral part of logistics management in food and agricultural supply chain. Food Control 33(1):32–48
20. Bozic N, Pujolle G, Secci S (2016) A tutorial on blockchain and applications to secure network control-planes. In: 2016 3rd smart cloud networks \ systems (SCNS), pp 1–8
21. Bradford RB (2008) An empirical study of required dimensionality for large-scale latent semantic indexing applications. In: Proceedings of the 17th ACM conference on information and knowledge management, pp 153–162
22. Brandon D (2016) The blockchain: the future of business information systems. Int J Acad Bus World 10(2):33–40
23. Cao J, Xia T, Li J, Zhang Y, Tang S (2009) A density-based method for adaptive LDA model selection. Neurocomputing 72(7–9):1775–1781
24. Carlozo L (2017) What is Blockchain? Here's a primer on the potentially transformative digital ledger technology. J Account 224(1):29
25. Cermeño JS (2016) Blockchain in financial services: regulatory landscape and future challenges for its commercial application. BBVA Research Paper 16:20
26. Chamola V, Hassija V, Gupta V, Guizani M (2020) A comprehensive review of the COVID-19 pandemic and the role of IoT, drones, AI, blockchain, and 5G in managing its impact. IEEE Access 8:90225–90265
27. Chang SE, Luo HL, Chen Y (2020) Blockchain-enabled trade finance innovation: a potential paradigm shift on using letter of credit. Sustainability 12(1):188
28. Chen Y, Li H, Li K, Zhang J (2017) An improved P2P file system scheme based on IPFS and Blockchain. IEEE Int Conf Big Data (Big Data) 2017:2652–2657
29. Chen G, Xu B, Lu M, Chen N-S (2018) Exploring blockchain technology and its potential applications for education. Smart Learn Environ 5(1):1–10
30. Cole R, Stevenson M, Aitken J (2019) Blockchain technology: implications for operations and supply chain management. Supply Chain Manag: An International Journal 24:469–483
31. Crosby M, Pattanayak P, Verma S, Kalyanaraman V et al (2016) Blockchain technology: beyond bitcoin. Appl Innov 2(6–10):71
32. De Filippi P (2017) What blockchain means for the sharing economy. Harv Bus Rev:1–5
33. Dobrovnik M, Herold DM, Fürst E, Kummer S (2018) Blockchain for and in logistics: what to adopt and where to start. Logistics 2(3):18
34. Dorri A, Kanhere SS, Jurdak R (2016) Blockchain in internet of things: challenges and solutions. ArXiv Preprint ArXiv:1608.05187
35. Dorri A, Kanhere SS, Jurdak R (2017) Towards an optimized blockchain for IoT. IEEE/ACM second international conference on internet-of-things design and implementation (IoTDI) 2017:173–178
36. Duan J, Zhang C, Gong Y, Brown S, Li Z (2020) A content-analysis based literature review in blockchain adoption within food supply chain. Int J Environ Res Public Health 17(5):1784
37. Dubovitskaya A, Xu Z, Ryu S, Schumacher M, Wang F (2017) Secure and trustable electronic medical records sharing using blockchain. AMIA Ann Symp Proc 2017:650–659
38. Dwivedi AD, Srivastava G, Dhar S, Singh R (2019) A decentralized privacy-preserving healthcare blockchain for IoT. Sensors 19(2):326
39. Ekblaw A, Azaria A, Halamka JD, Lippman A (2016) A case study for Blockchain in healthcare: "MedRec" prototype for electronic health records and medical research data. Proc IEEE Open Big Data Conf 13:13
40. Eyal I, Gencer AE, Sirer EG, Van Renesse R (2016) Bitcoin-ng: a scalable blockchain protocol. 13th $\{$USENIX$\}$ symposium on networked systems design and implementation ($\{$NSDI$\}$ 16), 45–59
41. Faria C, Correia M (2019) BlockSim: Blockchain simulator. IEEE International Conference on Blockchain (Blockchain) 2019:439–446
42. Felin T, Lakhani K (2018) What problems will you solve with blockchain? MIT Sloan Manag Rev 60(1): 32–38
43. Fernández-Caramés TM, Fraga-Lamas P (2018) A review on the use of Blockchain for the internet of things. IEEE Access 6:32979–33001

44. Ferrag MA, Derdour M, Mukherjee M, Derhab A, Maglaras L, Janicke H (2018) Blockchain technologies for the internet of things: research issues and challenges. IEEE Internet Things J 6(2):2188–2204
45. Francisco K, Swanson D (2018) The supply chain has no clothes: technology adoption of blockchain for supply chain transparency. Logistics 2(1):2
46. Ge L, Brewster C, Spek J, Smeenk A, Top J, van Diepen F, Klaase B, Graumans C, de Wildt MM (2017) Blockchain for agriculture and food: findings from the pilot study (issues 2017–112). Wageningen Economic Research
47. Ghag KV, Shah K (2015) Comparative analysis of effect of stopwords removal on sentiment classification. In: 2015 international conference on computer, communication and Control (IC4), pp 1–6
48. Golosova J, Romanovs A (2018) The advantages and disadvantages of the blockchain technology. In: 2018 IEEE 6th workshop on advances in information. Electronic and Electrical Engineering (AIEEE), pp 1–6
49. Haber S, Stornetta WS (1990) How to time-stamp a digital document. In: Conference on the theory and application of cryptography, pp 437–455
50. Hertz L (2019) "Hackernoon," 2019. https://medium.com/hackernoon/what-is-blockchain-as-a-service-28667754d6dc. Accessed 26 Mar 2021
51. Hafid A, Hafid AS, Samih M (2019) New mathematical model to analyze security of sharding-based blockchain protocols. IEEE Access 7:185447–185457
52. Halaburda H (2018) Blockchain revolution without the blockchain? Commun ACM 61(7):27–29
53. Halpin H, Piekarska M (2017) Introduction to security and privacy on the Blockchain. In: 2017 IEEE European symposium on security and privacy workshops (EuroS \ PW), pp 1–3
54. Helebrandt P, Bellus M, Ries M, Kotuliak I, Khilenko V (2018) Blockchain adoption for monitoring and management of enterprise networks. In: 2018 IEEE 9th annual information technology, Electronics and Mobile Communication Conference (IEMCON), pp 1221–1225
55. Joung J, Kim HM (2021) Automated keyword filtering in latent Dirichlet allocation for identifying product attributes from online reviews. J Mech Des 143(8):84501
56. Junruo S, Huancheng S, Haifeng L, Jiabin C, Xi L et al (2020) Software defect classification approach based on the modified latent Dirichlet allocation topic model considering the domain characters. In: 2020 7th international conference on dependable systems and their applications (DSA), pp 478–482
57. Kamble SS, Gunasekaran A, Gawankar SA (2018) Sustainable industry 4.0 framework: a systematic literature review identifying the current trends and future perspectives. Process Saf Environ Prot 117:408–425
58. Karamchandani A, Srivastava SK, Srivastava RK (2020) Perception-based model for analyzing the impact of enterprise blockchain adoption on SCM in the Indian service industry. Int J Inf Manag 52:102019
59. Karame G, Capkun S (2018) Blockchain security and privacy. IEEE Secur Priv 16(04):11–12
60. Khatoon A (2020) A blockchain-based smart contract system for healthcare management. Electronics 9(1): 94
61. Kiayias A, Russell A, David B, Oliynykov R (2017) Ouroboros: a provably secure proof-of-stake blockchain protocol. In: Annual international cryptology conference, pp 357–388
62. Kiyomoto S, Rahman MS, Basu A (2017) On blockchain-based anonymized dataset distribution platform. In: 2017 IEEE 15th international conference on software engineering research, management and applications (SERA), pp 85–92
63. Kraft D (2016) Difficulty control for blockchain-based consensus systems. Peer Peer Netw Appl 9(2):397–413
64. Lee L (2015) New kids on the blockchain: how bitcoin's technology could reinvent the stock market. Hastings Bus LJ 12:81
65. Li H, Cryer S, Raymond J, Acharya L (2020) Interpreting atomization of agricultural spray image patterns using latent Dirichlet allocation techniques. Artif Intell Agric 4:253–261
66. Liang X, Shetty S, Tosh D, Kamhoua C, Kwiat K, Njilla L (2017a) Provchain: a blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability. In: 2017 17th IEEE/ACM international symposium on cluster, cloud and grid computing (CCGRID), pp 468–477
67. Liang X, Zhao J, Shetty S, Liu J, Li D (2017b) Integrating blockchain for data sharing and collaboration in mobile healthcare applications. In: 2017 IEEE 28th annual international symposium on personal, indoor, and Mobile radio communications (PIMRC), pp 1–5
68. Lo SK, Xu X, Chiam YK, Lu Q (2017) Evaluating suitability of applying blockchain. In: 2017 22nd international conference on engineering of complex computer systems (ICECCS), pp 158–161
69. Madaan A, Sharma V, Pahwa P, Das P, Sharma C (2018) Hadoop: solution to unstructured data handling. In: Big data analytics. Springer, pp 47–54
70. Maesa DDF, Mori P (2020) Blockchain 3.0 applications survey. J Parallel Distrib Comput 138:99–114

71. Manski S (2017) Building the blockchain world: technological commonwealth or just more of the same? Strateg Chang 26(5):511–522
72. Marbouh D, Abbasi T, Maasmi F, Omar IA, Debe MS, Salah K, Jayaraman R, Ellahham S (2020) Blockchain for COVID-19: review, opportunities, and a trusted tracking system. Arab J Sci Eng 45:9895–9911
73. Mattila J, Seppälä T, Naucler C, Stahl R, Tikkanen M, Bådenlid A, Seppälä J (2016) Industrial blockchain platforms: An exercise in use case development in the energy industry, vol No. 43. ETLA Working Papers
74. Mavridis T, Symeonidis AL (2014) Semantic analysis of web documents for the generation of optimal content. Eng Appl Artif Intell 35:114–130
75. McGhin T, Choo K-KR, Liu CZ, He D (2019) Blockchain in healthcare applications: research challenges and opportunities. J Netw Comput Appl 135:62–75
76. Mearian L (2021) "Computer World." 2021. [Online]. Available: https://www.computerworld.com/article/3340373/jp-morgan-to-launch-a-us-dollar-backed-cryptocurrency.html. Accessed 25 Mar 2021
77. Meshkov D, Chepurnoy A, Jansen M (2017) Short paper: Revisiting difficulty control for blockchain systems. In: Data privacy management, cryptocurrencies and blockchain technology. Springer, Cham, pp 429–436
78. Michelman P (2017) Seeing beyond the blockchain hype. MIT Sloan Manag Rev 58(4):17
79. Mingxiao D, Xiaofeng M, Zhe Z, Xiangwei W, Qijun C (2017) A review on consensus algorithm of blockchain. In: 2017 IEEE international conference on systems, man, and cybernetics (SMC), pp 2567–2572
80. Mohanta BK, Panda SS, Jena D (2018) An overview of smart contract and use cases in blockchain technology. In: 2018 9th international conference on computing, communication and networking technologies (ICCCNT), pp 1–4
81. Mohanta BK, Jena D, Satapathy U, Patnaik S (2020) Survey on IoT security: challenges and solution using machine learning, artificial intelligence and blockchain technology. Internet of Things 11:100227
82. Control M (2019) Bitcoin SV to raise its block cap to 2GB. https://www.moneycontrol.com/news/technology/bitcoin-sv-to-raise-its-block-cap-to-2gb-3857881.html. Accessed 26 Mar 2021
83. Monrat AA, Schelén O, Andersson K (2019) A survey of blockchain from the perspectives of applications, challenges, and opportunities. IEEE Access 7:117134–117151
84. Möser M, Böhme R, Breuker D (2013) An inquiry into money laundering tools in the Bitcoin ecosystem. In: 2013 APWG ECrime Researchers Summit, pp 1–14
85. Nakamoto S (2008) Bitcoin whitepaper. URL: https://bitcoin.org/bitcoin.pdf-(17.07.2019)
86. Nakamoto S, Bitcoin A (2008) A peer-to-peer electronic cash system. Bitcoin–URL: https://bitcoin.org/bitcoin.pdf, 4
87. Nguyen QK (2016) Blockchain-a financial technology for future sustainable development. In: 2016 3rd international conference on green technology and sustainable development (GTSD), pp 51–54
88. Notheisen B, Hawlitschek F, Weinhardt C (2017) Breaking down the blockchain hype–towards a blockchain market engineering approach. In: Proceedings of the 25th European Conference onInformation Systems (ECIS), Guimarães, Portugal, June 5-10, 2017. Research Papers, pp 1062–1080 http://aisel.aisnet.org/ecis2017_rp/69
89. Nowiński W, Kozma M (2017) How can blockchain technology disrupt the existing business models? Entrepreneurial Bus Econ Rev 5(3):173–188
90. Pass R, Seeman L, Shelat A (2017) Analysis of the blockchain protocol in asynchronous networks. In: Annual international conference on the theory and applications of cryptographic techniques, pp 643–673
91. Pedrosa AR, Pau G (2018) ChargeltUp: on blockchain-based technologies for autonomous vehicles. In: Proceedings of the 1st workshop on cryptocurrencies and Blockchains for distributed systems, pp 87–92
92. Pilkington M (2016) Blockchain technology: principles and applications. In: Research handbook on digital transformations. Edward Elgar Publishing, pp 225–253
93. Pisa M, Juden M (2017) Blockchain and economic development: hype vs. reality. Center for Global Development Policy Paper 107:150
94. Pixelplex (2021) Pixelplex. https://pixelplex.io/work/ethereum-based-platform-for-crowdfunding-and-icos/. Accessed 20 Mar 2021
95. Plisson J, Lavrac N, Mladenic D et al (2004) A rule based approach to word lemmatization. Proc IS 3:83–86
96. Pop C, Cioara T, Anghel I, Antal M, Salomie I (2020) Blockchain based decentralized applications: technology review and development guidelines. ArXiv Preprint ArXiv:2003.07131
97. Popper N (2014) Hal Finney, cryptographer and Bitcoin Pioneer, dies at 58. NYTimes Archived from the original on 3
98. Porter MF (2001) Snowball: a language for stemming algorithms

99. Puthal D, Malik N, Mohanty SP, Kougianos E, Yang C (2018) The blockchain as a decentralized security framework [future directions]. IEEE Consum Electron Mag 7(2):18–21
100. Qiu J, Liang X, Shetty S, Bowden D (2018) Towards secure and smart healthcare in smart cities using blockchain. In: 2018 IEEE international smart cities conference (ISC2), pp 1–4
101. Saberi S, Kouhizadeh M, Sarkis J, Shen L (2019) Blockchain technology and its relationships to sustainable supply chain management. Int J Prod Res 57(7):2117–2135
102. Saghafi F, Pakyari M, Rezaei M (2019) Prioritizing capabilities of Blockchain Technology in Telecommunication for promoting customer satisfaction. In: 16th international conference on information technology-new generations (ITNG 2019), pp 499–503
103. Salimitari M, Chatterjee M (2018) A survey on consensus protocols in blockchain for iot networks. ArXiv Preprint ArXiv:1809.05613
104. Samaniego M, Jamsrandorj U, Deters R (2016) Blockchain as a service for IoT. In: 2016 IEEE international conference on internet of things (IThings) and IEEE green computing and communications (GreenCom) and IEEE cyber, physical and social computing (CPSCom) and IEEE smart data (SmartData), pp 433–436
105. Scott B, Loonam J, Kumar V (2017) Exploring the rise of blockchain technology: towards distributed collaborative organizations. Strateg Chang 26(5):423–428
106. Sehra SK, Brar YS, Kaur N, Sehra SS (2017) Research patterns and trends in software effort estimation. Inf Softw Technol 91:1–21. https://doi.org/10.1016/j.infsof.2017.06.002
107. Shah D, Shah P (2020) Soil study of coastal hyperspectral data using K-means and LDA (latent Dirichlet allocation). In: 2020 4th international conference on electronics, communication and aerospace technology (ICECA), pp 1258–1262
108. Shahnaz A, Qamar U, Khalid A (2019) Using blockchain for electronic health records. IEEE Access 7: 147782–147795
109. Sharma S, Kishore K (2017) Data dissemination algorithm using cloud services: a proposed integrated architecture using IoT. In: 2nd international conference on innovative research in engineering science and technology (IREST-2017). Eternal University, Baru sahib, Sirmour (HP), pp 7–8
110. Sharma S, Kishore K (2017) Internet of things (IoT): a review of integration of precedent, existing & inevitable technologies. AGU Int J Eng Technol 4:442–2455
111. Shi S, He D, Li L, Kumar N, Khan MK, Choo K-KR (2020) Applications of blockchain in ensuring the security and privacy of electronic health record systems: a survey. Comput Secur 97:101966
112. Sims A (2019) Blockchain and Decentralised Autonomous Organisations (DAOs): The evolution of companies? Available at SSRN
113. Slattery T (2014) Taking a bit out of crime: Bitcoin and cross-border tax evasion. Brook J Int'l L 39:829
114. Stack Overflow (2019) Evaluation of topic modeling: How to understand a coherence value / c_v of 0.4, is it good or bad? https://stackoverflow.com/questions/54762690/evaluation-of-topic-modeling-how-to-understand-a-coherence-value-c-v-of-0-4
115. Su C-W, Qin M, Tao R, Umar M (2020) Financial implications of fourth industrial revolution: can bitcoin improve prospects of energy investment? Technol Forecast Soc Chang 158:120178
116. Subramanian H (2017) Decentralized blockchain-based electronic marketplaces. Commun ACM 61(1): 78–84
117. Swan M (2017) Anticipating the economic benefits of blockchain. Technol Innov Manag Rev 7(10):6–13
118. Tapscott A, Tapscott D (2017a) How blockchain is changing finance. Harv Bus Rev 1(9):2–5
119. Tapscott D, Tapscott A (2017b) How blockchain will change organizations. MIT Sloan Manag Rev 58(2): 10
120. Tasca P, Tessone CJ (2017) Taxonomy of blockchain technologies. Principles of identification and classification. ArXiv preprint ArXiv:1708.04872
121. Tikhomirov S, Voskresenskaya E, Ivanitskiy I, Takhaviev R, Marchenko E, Alexandrov Y (2018) Smartcheck: static analysis of ethereum smart contracts. In: Proceedings of the 1st international workshop on emerging trends in software engineering for Blockchain, pp 9–16
122. Trautman LJ (2016) Is disruptive blockchain technology the future of financial services?. 69 The Consumer Finance Law Quarterly Report 232(2016). Available at SSRN: https://ssrn.com/abstract=2786186
123. Ulieru M (2016) Blockchain 2.0 and beyond: adhocracies. In: Banking beyond banks and money. Springer, pp 297–303
124. Valenta M, Sandner P (2017) Comparison of ethereum, hyperledger fabric and corda. Frankfurt School Blockchain Center 8:101966
125. Wang Y, Xu W (2018) Leveraging deep learning with LDA-based text analytics to detect automobile insurance fraud. Decis Support Syst 105:87–95

126. Wang X, McCallum A, Wei X (2007) Topical n-grams: phrase and topic discovery, with an application to information retrieval. In: Seventh IEEE international conference on data mining (ICDM 2007), pp 697–702
127. Wang H, Chen K, Xu D (2016) A maturity model for blockchain adoption. Financial Innov 2(1):1–5
128. Webster JJ, Kit C (1992) Tokenization as the initial phase in NLP. In: COLING 1992 volume 4: the 15th international conference on computational linguistics
129. White GRT (2017) Future applications of blockchain in business and management: a Delphi study. Strateg Chang 26(5):439–451
130. Wu J, Tran NK (2018) Application of blockchain technology in sustainable energy systems: an overview. Sustainability 10(9):3067
131. Xia QI, Sifah EB, Asamoah KO, Gao J, Du X, Guizani M (2017a) MeDShare: trust-less medical data sharing among cloud service providers via blockchain. IEEE Access 5:14757–14767
132. Xia Q, Sifah EB, Smahi A, Amofa S, Zhang X (2017b) BBDS: Blockchain-based data sharing for electronic medical records in cloud environments. Information 8(2):44
133. Xu X, Weber I, Staples M, Zhu L, Bosch J, Bass L, Pautasso C, Rimba P (2017) A taxonomy of blockchain-based systems for architecture design. In: 2017 IEEE international conference on software architecture (ICSA), pp 243–252
134. Xu X, Weber I, Staples M (2019) Architecture for blockchain applications. Springer
135. Yang W, Garg S, Raza A, Herbert D, Kang B (2018) Blockchain: trends and future. Pacific Rim Knowledge Acquisition Workshop:201–210
136. Yang L, Elisa N, Eliot N (2019) Privacy and security aspects of E-government in smart cities. In: Smart cities cybersecurity and privacy. Elsevier, pp 89–102
137. Yun J, Geum Y (2020) Automated classification of patents: a topic modeling approach. Comput Ind Eng 147:106636
138. Zamani M, Movahedi M, Raykova M (2018) Rapidchain: scaling blockchain via full sharding. In: Proceedings of the 2018 ACM SIGSAC conference on computer and communications security, pp 931–948
139. Zheng Z, Xie S, Dai H, Chen X, Wang H (2017) An overview of blockchain technology: architecture, consensus, and future trends. IEEE International Congress on Big Data (BigData Congress) 2017:557–564
140. Zhu S, Song M, Lim MK, Wang J, Zhao J (2020) The development of energy blockchain and its implications for China's energy sector. Res Policy 66:101595