Journal Pre-proof

Blockchain Technology Forecasting by Patent Analytics and text mining

Seyed Mojtaba Hosseini Bamakan, Alireza Babaei Bondarti, Parinaz Babaei Bondarti, Qiang Qu

PII: S2096-7209(21)00014-2

DOI: https://doi.org/10.1016/j.bcra.2021.100019

Reference: BCRA 100019

To appear in: *Blockchain: Research and Applications*

Received Date: 25 March 2021

Revised Date: 22 May 2021

Accepted Date: 20 June 2021

Please cite this article as: S.M. Hosseini Bamakan, A.B. Bondarti, P.B. Bondarti, Q. Qu, Blockchain Technology Forecasting by Patent Analytics and text mining, *Blockchain: Research and Applications* (2021), doi: https://doi.org/10.1016/j.bcra.2021.100019.

Blockchain Technology Forecasting by Patent Analytics and Text Mining

Seyed Mojtaba Hosseini Bamakan [a,c1], Alireza Babaei Bondarti[a], Parinaz Babaei Bondarti[b], Qiang Qu[c]

[a] Department of Industrial Management, Yazd University, Yazd, 89195-741, Iran
[b] Department of Library and Information Science, Kharazmi University, Tehran, Iran
[c] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, 518055, China

Abstract

Information technologies (ITs) have been playing an important role in improving our society, and the fast evolution of ITs creates a competitive environment not only for companies but also for regions. Hence, recognizing the future trend of technologies can be effective in decision-making with regard to technology selection and investment. Blockchain technology with its vast and impressive applications has received considerable attention from researchers, investors, and public agencies. The purpose of this research is to investigate blockchain technology to explore its trends according to their classification by use of the World Intellectual Property Organization (WIPO) database. Furthermore, we particularly evaluate the registered patents in the world's most well-known patent databases such as the USA patent database. We drew the current technology trends in blockchain patents by applying the text mining and clustering approach. The results represent that the registered patents in the USA patent database have been achieved in the growth phase. That means, attention to the blockchain is rising nowadays and most patents focused the cryptocurrencies and their application in finance. However, blockchain technology is in the emergence phase and is evolving by researchers and inventors.

Keywords: Technology forecasting, blockchain technology, patent Analysis, text mining, technology life cycle

1 Introduction

Blockchain technology is an emerging technology that has been designed for managing a distributed ledger of transactions (Qu, Nurgaliev, Muzammal, Jensen, & Fan, 2019; Xu, Chen, & Kou, 2019; Yaga, Mell, Roby, & Scarfone, 2019; Zhang & Wen, 2015). This technology was traced back to 2008 while bitcoin was introduced by Satoshi Nakamoto (Chernov & Chernova, 2018). Blockchain is a sequence of blocks in which each block contains some transactions that are validated by an encryption mechanism. The information maintained in the blocks does not remove and change (Yaga et al., 2019). The blocks are adding continually with each other and making a chain of blocks (Zheng, Xie, Dai, Chen, & Wang, 2017). From 2014, attention to cryptocurrencies has increased, as far as the time of this writing, 7300 cryptocurrencies are listed in CoinMarketCap. Generally, blockchain technology can be regarded as a combination of sciences such as software engineering, encryption knowledge, game theory in economics, and distributed computing. Eliminating intermediaries, reducing transaction costs, enhanced security, immutability, and transparency are among the influencing factors in the great development and expansion of blockchain technology.

The use case of blockchain is not limited to cryptocurrencies, rather today blockchain technology is applied in a vast range of domains including financial services (Foroglou & Tsilidou, 2015; F. Gao et al., 2018; Peters, Panayi, & Chapelle, 2015), smart contracts (Kosba, Miller, Shi, Wen, & Papamanthou, 2016), Internet of Things (IoT)

[1] Corresponding author
E-mail address: smhosseini@yazd.ac.ir
Postal address: University Blvd , Safayieh, Yazd,Iran, Tel, +9835-8123 7515 POST Box: 8915818411

(Akins, Chapman, & Gordon, 2014; Kravitz & Cooper, 2017), identification services (Zhang & Wen, 2015), security services (Noyes, 2016), healthcare services (Kuo, Kim, & Ohno-Machado, 2017; Mettler, 2016; Witchey, 2019), supply chain (Abeyratne & Monfared, 2016; Korpela, Hallikas, & Dahlberg, 2017; Moosavi, Naeni, Fathollahi-Fard, & Fiore, 2021; Tian, 2016; Turk & Klinc, 2017) and so forth. Fig. 1 shows a wide range of blockchain applications in both financial and non-financial domains.

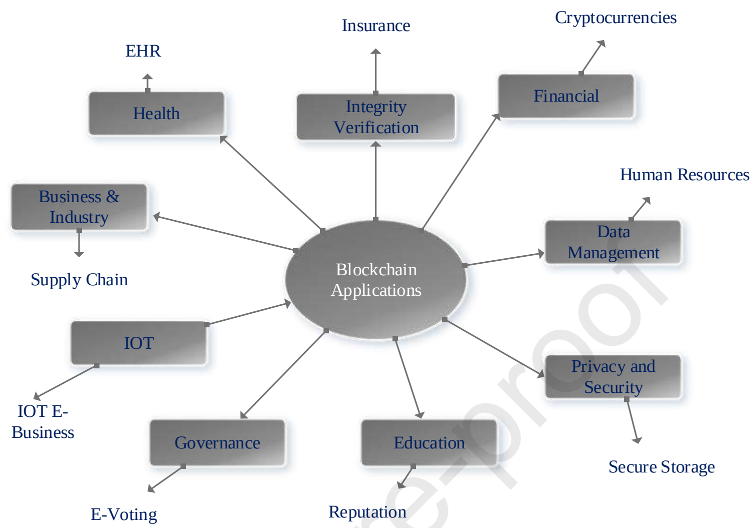


Fig. 1. Classification of Blockchain Applications (Casino, Dasaklis, & Patsakis, 2019).

Despite the substantial growth in the development of blockchain technology and its application, this technology has faced some limitations as scalability, consensus mechanism, blockchain interoperability, power consumption, low transaction throughput, and limited storage volume of each block (Meva, 2018; Monrat, Schelén, & Andersson, 2019; Muzammal, Qu, & Nasrulin, 2019; Zheng, Xie, Dai, Chen, & Wang, 2018; Zheng et al., 2017). Consequently, various studies have been done to address these shortcomings that mostly are registered in patent registration databases.

Nowadays, the complexity and variety of digital transformations have caused high-technology to play a critical role in businesses as a competitive advantage. Hence, forecasting the emerging technologies and their future trends at the right time is a substantial force for business continuity and its compliance (L. Gao et al., 2013; Li, Xie, Daim, & Huang, 2019; Watts & Porter, 1997). Although, due to the extremely rapid changes, the future of technologies is accompanied by uncertainty and ambiguity (Cho & Daim, 2013), by forecasting future trends it's possible to decrease this indeterminacy (Firat, Woon, & Madnick, 2008) and increase the awareness and certainty in the decision-making for technology investment (Altuntas, Dereli, & Kusiak, 2015). As a result, for investing in new technology, its current position on the product life cycle should be determined. Technology forecasting is a regular and systematic process whose objective is technology monitoring for trend forecasting and to determine future profits or losses made by a company (Adamuthe & Thampi, 2019). Technology forecasting would help companies to make an informed decision, priority setting for R&D unit, and knowledge exploitation (Abbas, Zhang, & Khan, 2014).

A vast range of methods for technology forecasting and trend analysis has been designed and classified into a variety of categories (Adamuthe & Thampi, 2019). Some categories of these methods are presented in (Cho & Daim, 2013; Firat et al., 2008; Yoon & Park, 2007). One of these methods is patent analysis in which has been considered by many researchers. Patents are powerful information resources for technology. Innovation is the foundation of global competition and technological advancement. Patents, as the most important source of information from technologies, are the main driver of the innovative process. Patents provide new knowledge through the disclosure of technological information, and their analysis avoids the repetitive efforts of researchers in a technological field and enables researchers to create technological subjects. The extensive literature on their use in the field of technology analysis in creating various indicators in research and development, reviewing the quality of technologies, comparing technologies, different policies of technology adoption, identifying technological gaps, or reviewing the life cycle of technologies has been done (Barcelon Yang, 2012; Baumann et al., 2021; Dou, Leveillé, Manullang, & JM Jr, 2005). In 1997, Ernest (Ernst, 1997) suggested that patents and their analysis can be used for technology forecasting and drawing their future trends. Many studies have been executed with the use of patent analysis by different tools to forecast different kinds of technologies that could be mentioned to flash memory, remote communication (Altuntas et al., 2015), green energy technology (Lee, Kim, Kim, Lee, & Oh, 2014), biotechnology (Jun, Park, & Jang, 2012), RFID technology (Trappey, Wu, Taghaboni-Dutta, & Trappey, 2011), file storage (Daim, Ploykitikoon, Kennedy, & Choothian, 2008), Nanoceramics materials (Cheng & Chen, 2008) and computerized numerical control technology (Ernst, 1997). Patents have detailed information about technologies that indicate the boundaries of legal uses of an invention. The trend of technologies and their competitors would be identified in the market by using patent analysis (Trappey et al., 2011). Patents are unique sources of various technological information and their information is not published in any other source of information. Given the growing importance of blockchain technologies and predicting their bright future, leading research has analyzed patents in this field. Since patents contain structured and unstructured information about technologies, and due to their large number, text analysis and clustering approaches have been used to analyze them. Sothe main objective of this research is to examine blockchain technology patents to identify its trends as an emerging technology by applying the text mining and clustering techniques. Moreover, we provide a forecasting model base on life cycle technology to achieve potential opportunities in research and development sector.

The remainder of the paper is organized as follows. Section 2 provides a brief description of the necessary background including patent analysis, clustering analysis, self-organizing map, and technology life cycle. The research methodology is presented in section 3. The results of blockchain patent analysis and technology life cycle prediction are discussed in Section 4. We provide possible challenges to the adoption of blockchain technology in Section 5. Finally, Section 6 concludes the paper with discussions.

2 Background

2.1 Patent analysis

Based on World Intellectual Property Organization (WIPO) a patent is defined as a document that shows a new product or a new process for product development and/or a new solution for a problem (WIPO). This document is an exclusive right and legal protection for inventors over a period of time that plays an important role in fair technology development and distribution (Adamuthe & Thampi, 2019; Altuntas et al., 2015).

Patents have more detailed information on technologies rather than academic papers. Analyzing the information extracted from patents has resulted in strategic planning, technology management, competitor analysis, and R&D unit management (Abbas et al., 2014; Altuntas et al., 2015; Firat et al., 2008). Therefore, patent analysis has become a strategic tool in technology trend analysis and its forecasting area. Patents have updated information on technology areas and are the best tools for technology forecasting and technology decision making (Altuntas et al., 2015; Campbell, 1983; Jaranyagorn & Chansa Ngavej, 2012). Hence, patent analysis has been executed with a vast range of methods to achieve a wide range of objectives. For example, organizations are interested in exploring patent innovation, trend analysis of technology, technology development forecasting in a specific area, strategic planning, drawing the road map of technology, and competitor identification (Abbas et al., 2014).

Since patents have structured and unstructured data, it is required analytical tools for better understanding (Abbas et al., 2014). Unlike there is no accurate classification of analytical tools and methods of patent analysis some researchers suggested a general category of patent analysis methods through literature review. For example, patent analysis methods are divided into quantitative (network-based) and qualitative (keyword-based) tools. Accordingly, the qualitative analysis focuses on patents' concepts and text mining techniques. It is applied in keyword extraction of the patents. On the other hand, the quantitative analysis uses the patents bibliographic information such as inventors, publication date, applicants, and other specifications for metadata of patents which can be achieved to competitors' analysis and a new exploration of technology opportunities (Abbas et al., 2014; J. Choi, Im, Kim, & Hwang, 2012; Y. Choi & Hong, 2020; Li et al., 2019).

In another category, patent analysis methodologies have been divided into two classes of text mining and visualization methods (such as clustering). Patent analysis can be useful in trend analysis and technology forecasting, drawing technology roadmap, and leading countries in innovative technology (Abbas et al., 2014).

2.2 Clustering analysis

We consider the patent analysis as a knowledge extraction process from a high volume of structured and unstructured patent data by text mining techniques. Text mining results and their interpretations would help technology investment decision-making to be more wisely (Abbas et al., 2014; Tseng, Lin, & Lin, 2007). Clustering is an effective text mining technique (Allahyari et al., 2017; Chen & Hu, 2006), is an unsupervised technique, its goal is to put similar objects in the groups that their members have more similarities with each other, however, has more dissimilarities with the other groups' members (Han, Pei, & Kamber, 2011).

Clustering is a continuous process that includes collecting data, determining a similarity criterion between data, selecting an appropriate clustering method, evaluating the performance of the selective method, and finally interpreting the results of clustering (Sarstedt & Mooi, 2014). In this study, we adopted K- Means Self-Organizing Map (SOM), to do the task of patent clustering analysis.

2.3 Self-Organizing Map (SOM)

Self-Organizing Map (SOM) is one of the intelligent computing methods and a powerful tool in data analysis. SOM is a neural artificial network that is introduced in 1982 by Kohonen (Asan & Ercan, 2012; Cottrell, Olteanu, Rossi, & Villa-Vialaneix, 2018). Generally, SOM is a method to analyze and visualize the current pattern in the high-dimensional data sets (Asan & Ercan, 2012; Kohonen, 2012). SOM neural network has been made of two internal and external layers that are connected directly to each other without a hidden layer. The network's internal layer is an *m*-dimensional vector whose elements encompass a set of data or characteristics (Asan & Ercan, 2012). The external layer indicates a low dimensional visualization of data. The node numbers in the external layer specify that the maximal number of clusters, which affects SOM method precision (Asan & Ercan, 2012).

In general, the network's topology is either rectangular or hexagon. In the rectangular topology, the internal node has four neighbors. However, this number is six in the hexagonal topology. Fig. 2 illustrates the SOM method's structure. SOM method receives internal data sets with a high dimension and illustrates it in a low dimensional space with internal data topology.

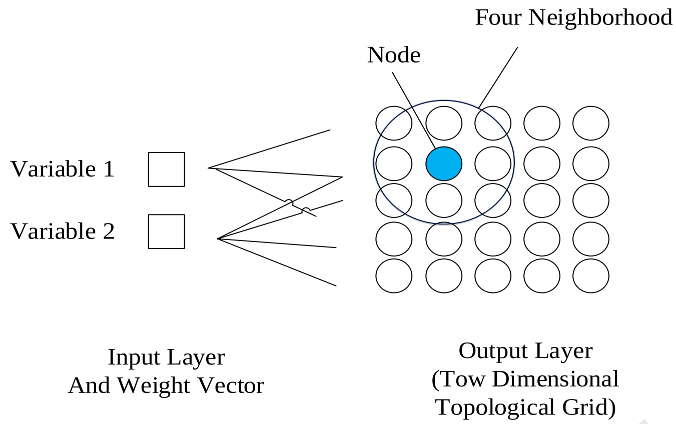


Fig. 2. Self-Organizing Maps topology.

2.4 Technology life cycle

Each technology life cycle encompasses three stages of presentation, growth, and saturation as presented in Fig. 3. In the presentation stage, the technology has recently entered the market and therefore the number of patents is so limited. After technology's representation, the patents' numbers have been increasing and the technology has been entered the growth stage. In the decline stage, the patents' numbers have been achieved to the maximal value. Therefore, the technology may be replaced with another technology in this stage (Adamuthe & Thampi, 2019; Altuntas et al., 2015; Trappey et al., 2011). Investing in technology requires analyzing technology life cycle stages (Altuntas et al., 2015).

"S-shaped curve" is among the most well-known methods for technology forecasting by drawing the technology life cycle. To illustrate an S-shaped curve, there are a variety of models that the logistic model is the most appropriate one. One characteristic of this model is its symmetry with regard to the inflection point. If the inflection point has occurred, the technological forecasting path maintenance will be simple (L. Gao et al., 2013). Eq. 1shows the logistic model that is applied in this research.

$$y_t = \frac{k}{1+e^{-\left(\frac{t-a}{b}\right)}} \quad (1$$

Where t is time, y_t indicates the sum of patents in the time. a is an inflection point and b is a parameter of the model which controls the curved shapes K is the upper limit of patent numbers and is the highest degree of y_t (L. Gao et al., 2013). These parameters can be achieved by the use of Least-Square Fitting computations with aid of statistical software such as Sigma Plot. After achieving these parameters for determining technology life cycle stages, if $\frac{y_t}{K} < 10\%$, then technology is in presentation stage, if $10\% \leq \frac{y_t}{K} < 50\%$, then technology is in the growth stage if $50\% \leq \frac{y_t}{K} < 90\%$, then technology is in the maturity stage, and finally if $90\% \leq \frac{y_t}{K}$, then technology is in the saturation stage (Trappey et al., 2011).

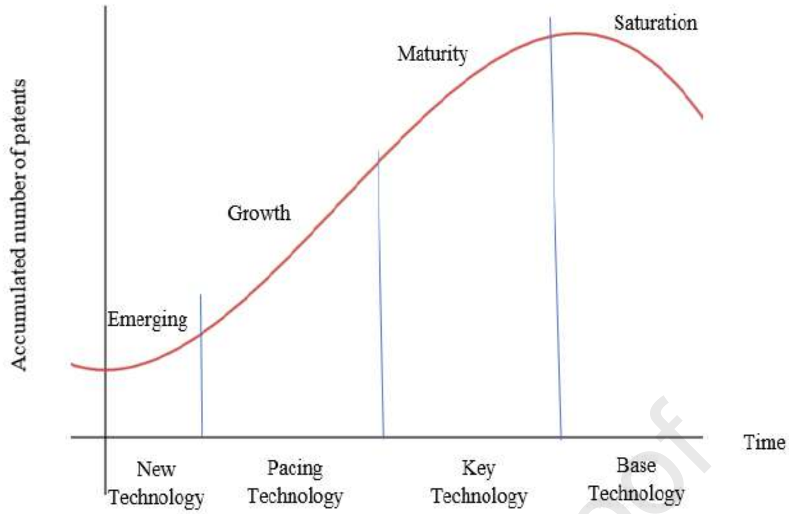


Fig. 3. S-Curve for Stage Technology Forecasting (Altuntas et al., 2015; L. Gao et al., 2013; Liu & Wang, 2010).

Based on the S-curve of the technology life cycle, researchers believe that investing in technology must occur when the technology is in the growth stage (Altuntas et al., 2015).

3 Research Methodology

In this section, we describe the whole process of patent analysis utilized in this research including; patent documents text mining and keywords extraction, patent documents clustering, and finally cluster analysis as is illustrated in Fig. 4.

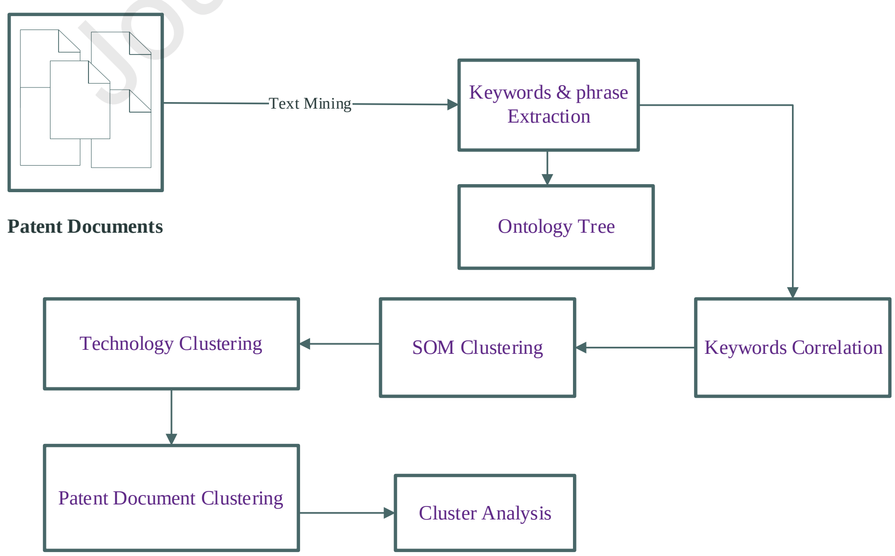


Fig. 4. Research framework.

3.1 Keywords Extraction

In the first step, some general blockchain-related keywords are used for patent extraction and construct a database of patents. After the appearance of this database, text mining has been performed on this dataset to extract the most important specific keywords. Text mining process and consequently keywords retrieval requires transforming the text data into the standard format, namely text pre-processing process that encompasses the following steps:

— *Tokenization* is responsible for breaking down and dividing an array of characters and converting them to alphabets or words. In other words, tokenization determines the words' boundaries in the textual data and transforms the text into an array of words.

— *Stop words* are applied in the documents to remove some of the non-relevant words. For example, the removal of the most duplicate words in the text that hasn't any useful information.

— *Lemmatization* is responsible for morphological analysis of the words. In other words, it draws infinitive tens form for verbs and one form for a noun.

— *Stemming* is used to discover the stem of words in the text by removing the prefixes and suffixes that disregard their role in the text (Allahyari et al., 2017).

3.2 Ontology tree

The objective of the Ontology tree is to put the extracted patents' keywords into the categories with regard to experts' viewpoints. In fact, this tree represents relevant concepts to extracted keywords. Blockchain technology Ontology encompasses the most important keywords in blockchain technology and their classifications as discussed in (Tasca, Thanabalasingham, & Tessone, 2017).

3.3 Keywords correlation

Following the key phrases extraction, we need to set up a logical relationship between them. The authors in (Hsu, Trappey, Trappey, Hou, & Liu, 2006), suggested a four stages algorithm for keywords analysis. At the first point, a keywords vector achieves by patent analysis. In the second phase, the keywords are updated by removing the redundant phrases and adding desired phrases. In the third phase, the keywords repetition matrix is achieved and finally, the correlation between the keywords are calculated with the use of Eq. 2:

$$R_{ij} = \frac{\sum_{k=1}^{N_D} X_{ik}X_{jk} - N_D\,\overline{X_i}\,\overline{X_j}}{\sqrt{\left(\sum_{k=1}^{N_D} X_{ik}{}^2 - N_D\overline{X_i^2}\right)\left(\sum_{k=1}^{N_D} X_{jk}{}^2 - N_D\overline{X_j^2}\right)}} \quad (2$$

N_D is the total number of patents and $X_{ik}$is the repetition number of keywords i in the patent D_k.

Keywords clustering is calculated by the correlation between them based on their frequencies in the patents.

Keywords \ Keywords	K_1	K_2	K_3	$\cdots$	K_m
K_1	CoR_{11}	CoR_{12}	CoR_{13}	$\cdots$	CoR_{1m}
K_2	CoR_{21}	CoR_{22}	CoR_{23}	$\cdots$	CoR_{2m}
K_3	CoR_{31}	CoR_{32}	CoR_{33}	$\cdots$	CoR_{3m}
$\vdots$			$\vdots$		
K_m	$CoR_{\mathrm{m}1}$	$CoR_{\mathrm{m}2}$	CoR_{m3}	$\cdots$	CoR_{mm}

As a result, the above correlation matrix is considered as an input of the clustering phase, and then each cluster of keywords is labeled as a technology cluster.

3.4 Patent document clustering

Despite the technology clustering, which represents the technology clusters by using the keywords; patent documents are classified according to their similarities. For patent document clustering, technology clusters that are produced by the correlation matrix, are applied as key variables for patent document clustering (Trappey et al., 2011).

Tech Cluster \ Patents	P_1	P_2	P_3	$\cdots$	P_n
TC_1	N_{11}	N_{12}	N_{13}	$\cdots$	N_{1m}
TC_2	N_{21}	N_{22}	N_{23}	$\cdots$	N_{2m}
TC_3	N_{31}	N_{32}	N_{33}	$\cdots$	N_{3m}
$\vdots$			$\vdots$		
TC_k	N_{n1}	N_{n2}	N_{n3}	$\cdots$	N_{kn}

N_{ij} is the sum of repetition of the keyword in patent j which places in cluster i.

4 Blockchain Patent Analysis

4.1 Patents Database construction

Initially, a database for relevant patents of blockchain was constructed. Building such a database requires searching in the World Intellectual Property Organization (WIPO) to retrieve patents. It is a strong database for patent searching. The important phase of searching patents is choosing appropriate and comprehensive sets of keywords. In this research, we used keywords to searching patent as below:

> "Blockchain" OR "Distributed Ledger" OR "Cryptocurrency" OR "Virtual Currency" OR "Digital ledger" OR "cryptographic Ledger" OR "Cryptocoin" OR "Digital Currency" OR "Smart Contract"

The total number of retrieved documents as the result of this search pattern was approximately 14000 patents worldwide. The highest number of patents are registered in China. The US has the second rank with 26% of patent registration and 19 % of the patents have registered through the Patent Cooperation Treaty (PCT) and WIPO. The 10 best countries in the field of patent registrations have been shown in Fig. 5.

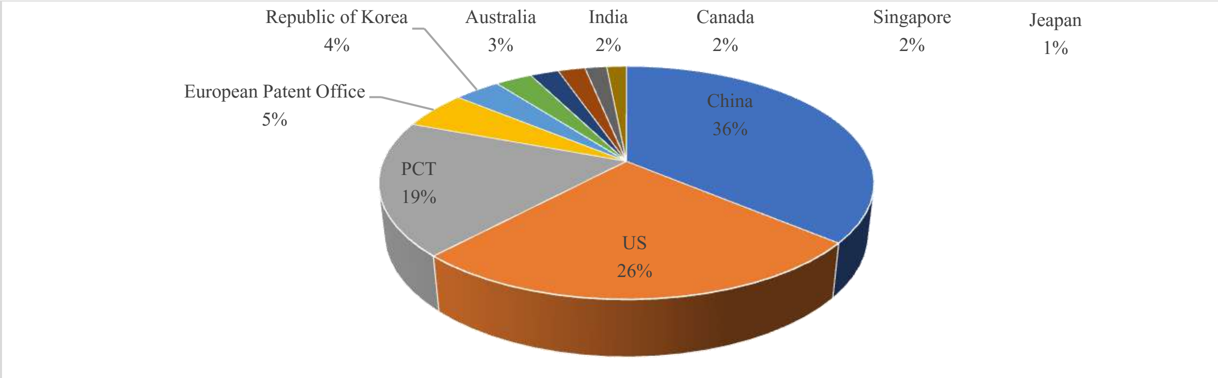


Fig. 5. Distribution of blockchain patent registration by countries.

The best owners of patents in the field of blockchain are the E-commerce giant ALIBABA, IBM technology firm, Nchain Holding protocol builders, Chinese Technology Company Tencent, and Master Card Multi-national finance services. The main owners or applicants of the blockchain patents have been shown in Fig. 6.

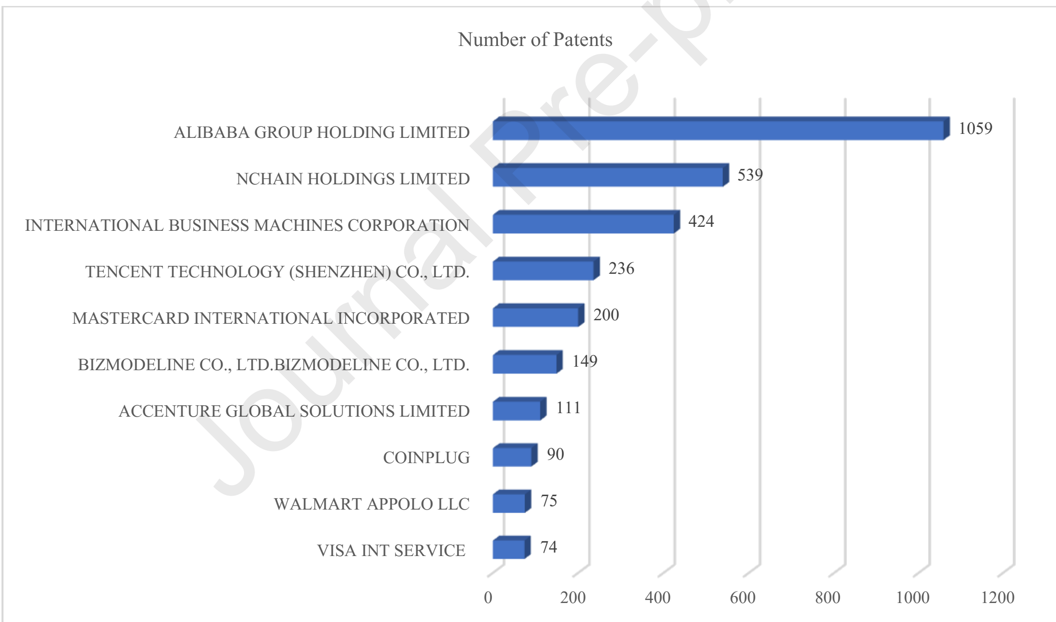


Fig. 6. The number of registered patents by companies.

The other results of blockchain patent analysis are the best inventors, which have the highest number of patents as presented in Fig. 7. WRIGHT CRAIG STEVEN (an Australian scientist and the pioneer of bitcoin production) with 286 patents is in the first rank. The second rank is SAVANAH STEPHANE, an information technology consultant, and the blockchain expert in London with 258 patents. KIM JAE HYUNG and KWON BONG KI are two Korean inventors who are placed in the third and fourth rank with 243 patents. The next rank is JI JIANXUN, Chinese inventor and Chief Technology Officer (CTO) member of the insight chain with 176 patents. UHR JOON SUM (a Korean technologist and the pioneer of Coin plug startup) with 147 and ZHANG WENBIN (a Chinese researcher and a member of the blockchain platform in the IBM Center) with 140 patents are the best inventors in the field of blockchain.

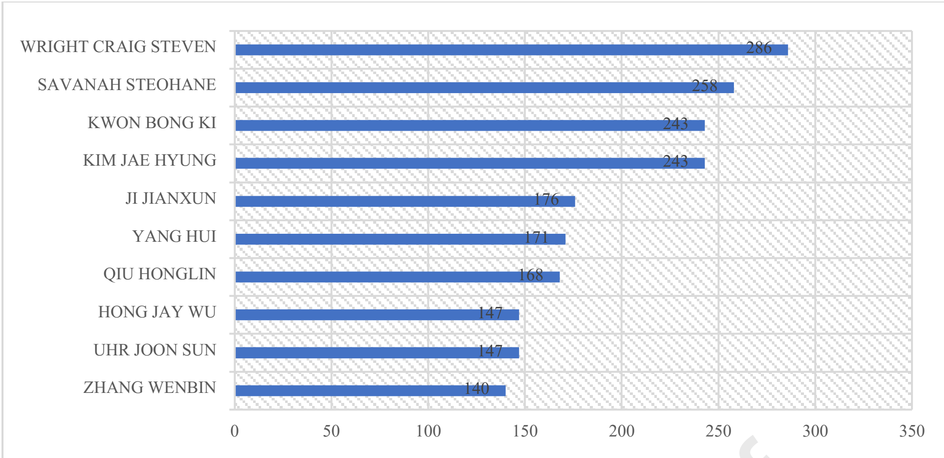


Fig. 7. Inventors in the field of blockchain.

The review of retrieved patents represented that blockchain technologies have been considered since 2012 by 50 patents. The total number of patents from 2000 to 2011 is 120. The number of patents has significantly increased in 2019. Fig. 8 shows a comparison between the numbers of patents in various years. These patent documents can be grouped into three categories including Patent application, Granted patent, and other patents (such as Search Report and Limited Patent). This information is available at https://www.lens.org/.

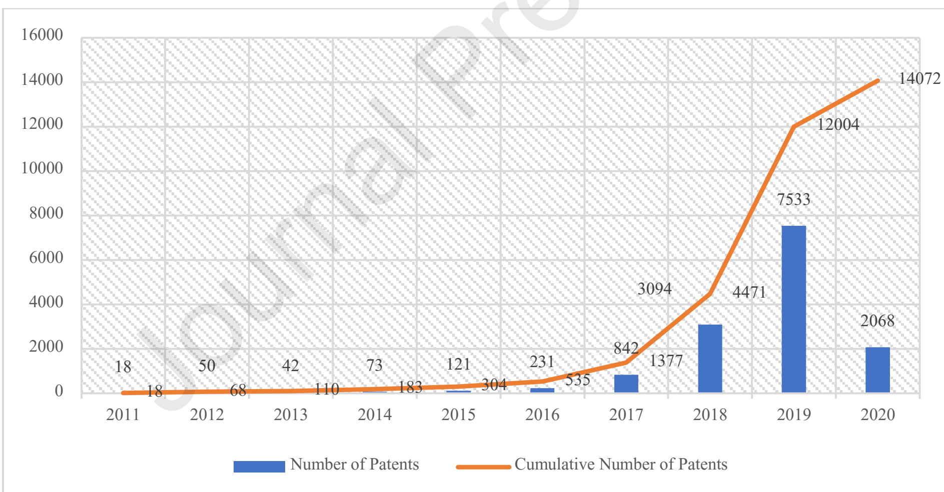


Fig. 8. The number of patents in each year from 2011 to 2020.

The patents are classified according to an international and standard system to evaluate and search quickly. One of these classifications is the International Patent Classification (IPC) system that is under the surveillance of WIPO. The IPC classification has been represented in Table 1.

Table 1. IPC groups for blockchain technology.

IPC Code	Description
G06Q	Data Processing Systems or Method's
H04L	Transmission of Digital Information, e.g. Telegraphic Communication
G06F	Electric digital data processing
H04W	Wireless communication networks
G06K	Recognition of Data; Presentation of Data, Record Carries, Handling Records Carriers

G06N	Computer systems based on specific computational models
A63F	Card, board, or roulette games; indoor games using small moving playing bodies; video games; games not otherwise provided for
G16H	Healthcare Informatics, i.e. Information and Communication Technology [ICT] Specially Adapted for the Handing or Processing of Medical or Healthcare Data
G07C	The time or Attendance Registers; Registering or Indicating the Working of Machines; Generating Random Numbers; Voting or Lottery Apparatus;
G07F	Coin- Freed or Like Apparatus

Fig. 9 represents the contribution of each category in the IPC classification from the retrieved patents.

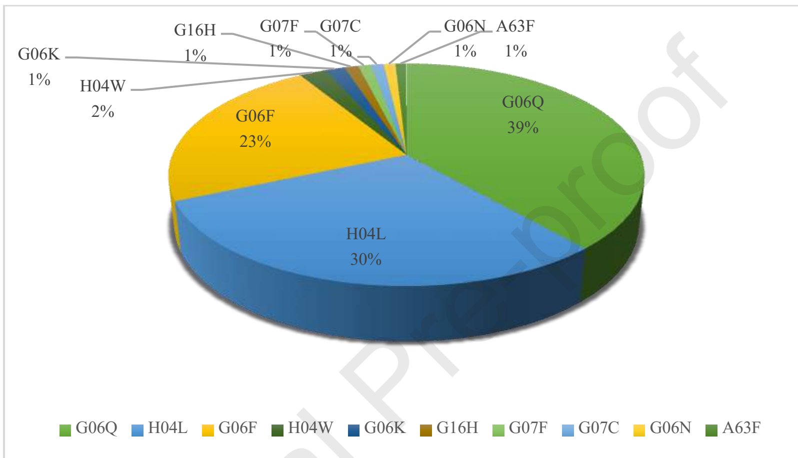


Fig. 9. Distribution of blockchain patents based on IPC.

In order to patent analysis and investigate blockchain technologies, we considered the US patents. As the US has the most remarkable market that the inventors can register their patents in this market, the US has a powerful database for patents and invention registrations (Lehtovirta, 2019). Therefore, we searched our keywords in the American Database to retrieve patents. In this phase, the patent search process has been performed through WIPO.

Following the search process, 3600 patents have been retrieved. Finally, after excluding the irrelevant and duplicate patents, 3261 patents have been gained from 2001 until May 2020. The frequency and the total number of patents in each year have been illustrated in Fig. 10. According to this figure, the number of patents, i.e. the development of blockchain technology inventions has been started in 2015 and increased in 2019, and would continue in 2020.

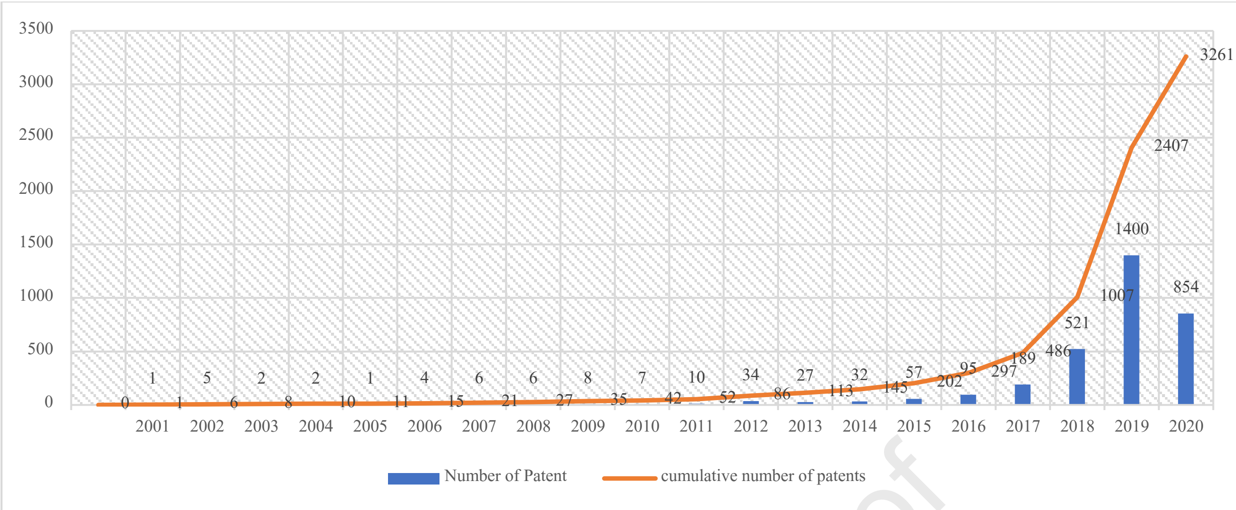


Fig. 10.The frequency and the total number of patents in each year.

Most applicants of blockchain patents in the United States of America have been shown in Table 1.

Table 2. Applicants of blockchain patents in the USA.

Applicant Name	Countries	Number of Patent application
IBM	US	337
ALIBABA	China	194
Bank of America	US	89
MasterCard	US	75
NCHAIN	China	75
ACCENTURE	Ireland	64
CAPITAL ONE	US	42
WALMART	US	39
MICROSOFT	US	36
CISCO	US	33

The technology lifecycle of each class of blockchain technology is mapped by utilizing the IPC. As already mentioned in the IPC classification, each patent can assign to more than one class. In other words, one patent can have more than one IPC classification code. We utilized Sigma Plot 14.0 software to investigate the life stage of technologies in each class. The achieved results have been demonstrated in Table 3.

Table 3. The life stage of technology is based on IPC classification.

Class	Number of Patents (to 2020)	Estimated maximum patent	Share of the upper limit (%)	Stage of the technology life cycle
H04L	2265	2454	92	saturation
G06Q	1983	3335	59	maturity
G06F	1509	1704	86	maturity
H04W	173	187	93	saturation
G06N	131	142	92	saturation
G06K	122	185	66	maturity
G07F	92	197	50	maturity
A63F	79	110	72	maturity

G16H	61	-	-	-	-
G07C	58	-	-	-	-

According to Table 3, we have not achieved a result for the last two classes including G16H and G07C because of the few numbers of patents in these groups. The technology lifecycle represents whether the specific technology is in which stage of its lifecycle from its introduction to decline. According to extracted patents from the USA patent registration website and selected keywords, Table 3 indicates that technology related to H04L with 2265 patents is in the saturation stage until 2020. The saturation stage is the last stage of a technology life that technology is replaced by other technology.

4.2 Text Mining

The text mining process output is gaining a set of keywords or key phrases out of the patents which is called key phrases vector. It is an important step as the other stages of patent analysis including clustering and technology forecasting and finally, its results depend on this extracted term vector. The selection of key phrases vector has been achieved by patent studies and by experts' judgment. Fig. 11 illustrates the word cloud of extracted key phrases from the patent text mining.

Fig. 11. Word cloud of Blockchain words

By implementing patent text mining, 100 keywords were extracted which has been applied in the other stages of this study.

4.3 Ontology Tree

The blockchain technology's ontology tree is built upon keywords and phrases extracted from patent text mining. One of the most important applications of the ontology tree is that it can be applied in recognizing technological gaps. It also provides a perspective for future research. The blockchain ontology tree consists of three segments according to Fig. 12. *Blockchain structure* includes the main components or body of the blockchain. *Blockchain features* such as high security, authentication, blockchain types, or decentralized distribution office are another segment of this tree. The third part of the ontology tree is related to the *application of blockchain* in various fields such as finance, treatment, insurance, etc.

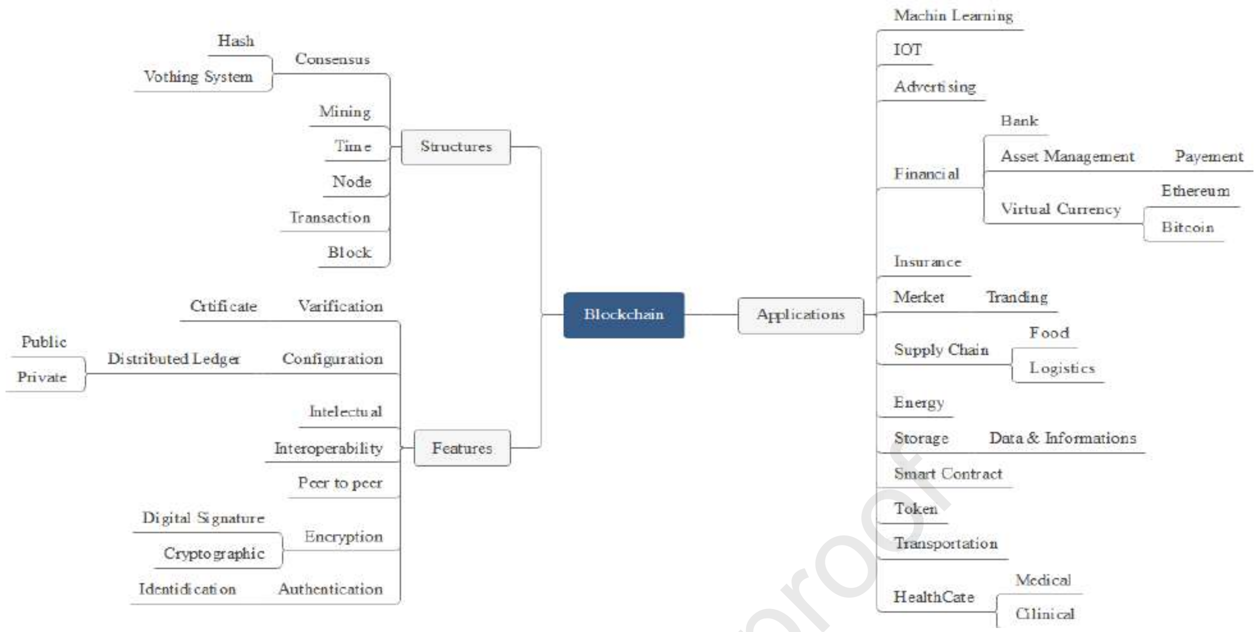


Fig. 12. Blockchain ontology tree.

The main branches of Figure 12 are based on the keywords extracted from text mining into three categories inclusing, words related to blockchain structure, words related to attributes, and finally words related to the use of blockchain in different sections. The sub-branches of the tree are formed due to their proximity to the main branch. For example, words like consensus are related to the structure of the blockchain, which is also related to the hash and vote of the word. Words like virtual currencies refer to the use of blockchain technology, such as bitcoin or atrium. Blockchain decentralization and digital authentication also refer to the features of blockchain technology.

4.4 Technology Clustering

4.4.1 Correlation between keywords

In the next stage, we calculate the correlation between the words with regard to the number of repetitions in each patent document. Hence, we need the repetition matrix of words in each patent. We used SPSS statistics software to compute the Pearson correlation that a part of its result is represented in Table 4.

Table 4: Correlation between keywords.

	ADVERTISING	ASSET	AUTHENTICATION	AUTHORIZATION	BANK	BITCOIN
ADVERTISING	1	-0.011	-0.006	-0.008	-0.006	-0.003
ASSET	-0.011	1	0.017	-0.009	-0.002	0.000
AUTHENTICATION	-0.006	0.017	1	0.030	-0.015	-0.024
AUTHORIZATION	-0.008	-0.009	0.030	1	-0.002	-0.015
BANK	-0.006	-0.002	-0.015	-0.002	1	-0.011
BITCOIN	-0.003	0.000	-0.024	-0.015	-0.011	1

4.4.2 Keywords Clustering

The objective of keyword clustering is recognizing the technologies that exist in the patents. Each cluster is called technology clusters. SOM clustering technique and Viscovery SOMine software have been applied in keyword clustering. One of the most challenging issues in the clustering process is determining the optimal number of clusters. Although the experts' opinions can be utilized to determine the number of clusters, there are some methods to verify the optimal number of clusters such as the Variation ratio criterion of Calinski and Harabasz, Dunn Index, Hartigan Method, Davis- Bouldin Index, Krzanowski and Lai Method, Gap Method of Tibshirani and Silhouette Index of Kaufman and Rousseeuw (Yan, 2005).

In this research, we adopted Davis-Bouldin Clustering Index that is based on the similarity between two clusters (R_{ij}) and is defined by distribution(s_i) of class (C_i) . $\|C_i\|$ is the number of clusters i elements and v_i is its centroid, the extent of distribution of this cluster is computed by use of Eq. 3:

$$s_i = \frac{1}{\|C_i\|} \sum_{x \in C_i} d(x, v_i) \qquad (3$$

Where d is a distance function (e.g. Euclidean Distance). Therefore, the similarity extent between cluster C_i and C_j achieve by Eq. 4:

$$R_{ij} = \frac{s_i + s_j}{d(v_i, v_j)} \qquad i, j = 1,2, \cdots k, \quad i \neq j \qquad (4$$

Where k is the sum of all clusters. Finally, if:

$$R_i = \max_{\substack{j=1,2,\cdots,k \\ j \neq i}} R_{ij} \qquad i = 1,2, \cdots, n \qquad (5$$

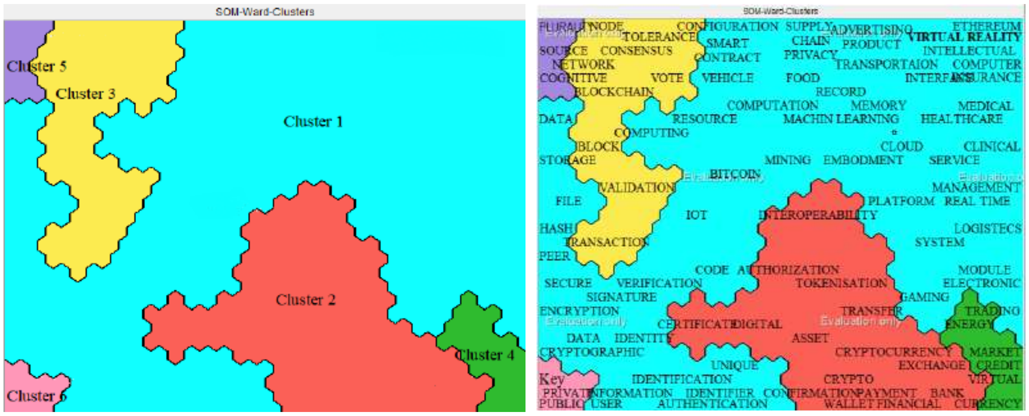
Then the Davis-Boulding Index is calculated by Eq. 6:

$$DVB = \frac{1}{k} \sum_{i=1}^{k} R_i \qquad (6$$

This index computes the similarity average between each cluster with the most similar cluster to it. According to the clustering objective which is declining the differences inside the cluster and the most dispersal between the clusters, the optimal number of clusters in the clustering process is achieved when the DVB index value matches the minimum (Lamirel, Dugué, & Cuxac, 2016). By replicating the SOM algorithm and different clusters, Davis-Bouldin Index has been calculated according to Table 5. By experts' confirmation, the keywords have been replaced in 6 clusters. The clusters are represented in Fig. 13-a and the distribution of words in each cluster is shown in Fig. 13-b.

Table 5. The number of clusters and their Davis-Bouldin Index.

Number of Clusters	3	4	5	6	7
Index DVS	3.4	3.2	2.7	2.3	3.1
Quantization Error	0	0	0	0	0



a) b)

Fig. 13. The number of clusters obtained by SOM techniques, a) technology clusters b) distribution of words in each cluster.

4.5 Patent Clustering

Patent document clustering is a method that classifies patents based on technology similarity (Trappey et al., 2011). After keyword clustering, each cluster of keywords represents the special technology that is classified in a group of patents. For instance, the first cluster that has the largest number of keywords represents blockchain technologies that have been applied in different areas such as healthcare, Internet of Things, Supply Chain, Smart Contract, and the third cluster especially refers to the virtual currency technology and its commerce. The patents that have technological similarities are grouped in one cluster. Table 5 represents the results of the patent document clustering. The clustering of words is attained by using the SOM method and based on the correlation between keywords. However, the patent document clustering is done based on the conceptual similarity between patents as presented in Table 5.

Table 6. Patent document clustering by text mining.

Cluster	Number of Patents	Cluster representative words	Description
Cluster 1	1545 (48%)	Digital Asset, Financial, Charity and donation, wagers on sports, gambling, Public events, Virtual Currency, Validating the transaction record, Game smart contract, In-game digital assets, Networked video game system, Scheduling and distributing user-generated content, Distributed journal, Healthcare, Energy, Supply chain, Logistics, Insurance	Smart contact applications
Cluster 2	364 (11%)	Intellectual property, distributed ledger, eco-system, trust models, licensing royalty, fraud detection, recordation, Data privacy, Internet of Things, Reliable, information sharing, terminal processor, road safety, data security, data layer, adaptive control unit, cognitive computing unit, artificial intelligence unit, machine-learning unit.	Privacy-preserving and Intellectual property
Cluster 3	287 (9%)	Digital certificate, System, Authorization, Code, Configuration, Validation, certificate revocation, Authority, Identifier, identity querying, Plurality, Hash, Device, Encryption, Certificate, Signature, Ownership, Verification, educational offering, wallet authority	Certificate issuance and verification
Cluster 4	350 (11%)	Token, Identity, Configuration, fractional, ownership, entity, security token, securitization, utility, assets, Access, circuit, liquidization, fungible,	Tokenization

Cluster 5	533 (17%)	Currency transactions, conversion, interaction, swap platform, subsequently, second blockchain, authentication, furnisher, cryptologic committal, exchange, cross-currency	Interoperability and consensus in cross-blockchain
Cluster 6	162 (5%)	AR display, platform, GPS system, virtual hub systems, graphical user interface, GPS map routing, mixed reality, transport, console, video game, dynamic conversion, reward, virtual asset, player, gambling, social gaming platform, instant messaging	Virtual reality and gaming

4.6 Results of technology life cycle prediction

The results of patent document clustering are applied for predicting the future trends of technologies in each cluster. The results are obtained by the logistic model and Sigma Plot Software as presented in Table 6.

Table 7. Prediction of future trends of technologies.

	Number of Patents (2001-2019)	Estimated Maximum Patent Number	Year of Upper Limit	Share of Upper limit (%)	Stage of Technology Life Cycle
Cluster1	1545	3026	2034	0.51	maturity
Cluster2	364	486	2028	0.73	maturity
Cluster3	287	645	2026	0.44	growth
Cluster4	350	792	2031	0.44	growth
Cluster5	553	1142	2026	0.48	growth
Cluster6	162	351	2031	0.46	growth
Total	3261	5234	2037	0.62	maturity

The first cluster includes approximately half of the available patents in the current database. This cluster is achieved by its representative words (combined of the extracted keywords in text mining and other words are achieved by the manual investigation of patents), indicates applications of blockchain in different areas including smart contract, supply chain, financial, and cryptocurrency, insurance, healthcare, etc. In other words, half of the patents are referring to the usability of blockchain. The total number of these patents from 2001 to May 2020 is 1545. Table 6 shows that this cluster is at the beginning of the maturity stage and the number of patents of this cluster will be reached from 3026 to 2034. This cluster encompasses those patents relates to blockchain application in various areas. This cluster can be examined in-depth and separately studied the trend of patents in each domain.

The second cluster, which contains 11 percent of all patents, refers to the security and privacy in a distributed ledger and Internet of things. The patents of this cluster have been registered since 2008 and its growth is started by 54 patents since 2017, in which its total number is 364 until May 2020. According to Table 6, the trend of technology available in the patents of this cluster has been entered the maturity stage and it will be reached to the largest number (486) by 2028.

The number of patents in the third cluster has been registered since 2012 and it can be concluded that its growth has been started since 2017 by 27 patents. Totally, by May 2020, it has been registered 287 patents in this cluster. The technology of this cluster relates to the security of the blockchain network and the transaction validity and accuracy and users' identification. Table 6 indicates that the trend of this cluster's patents is in the growth stage.

Tokenization technology is used in the financial processes and its objective is transforming a chunk of data into a random string of known characters, namely Token, has not mathematical relationships with real data that indicate them (Morrow & Zarrebini, 2019; Smith, 2019). The fourth cluster with 11 percent frequency refers to this technology. Table 6 shows that the trend of these patents' clusters is in the growth stage and it is expected that it will be increased to 792 patents by 203.

Another significant technology in blockchain technology is the consensus algorithm. There are different methods to achieve consensus in the blockchain which are analyzed and investigated in (Bamakan, Motavali, &

Bondarti, 2020). This cluster with 533 patents until May 2020, contains 17 percent of all total patents. The registered patents in this field hold patents since 2016. The growth trend of the patent numbers has been 293 in 2019. Table 6 shows that it can be expected that the patent numbers of this cluster will be reached to 1142 patents by the end of 2026 and the patents of this cluster are in the growth stage. It is noted that forecasting technology trends in each cluster require an adequate frequency of patents in different several years. In this cluster, the forecasting has been performed based on approximately 5 years (from 2016 until May 2020). Thus, this cluster is in the growth stage.

Finally, the sixth cluster is referring to the application of blockchain in the games and virtual world. Five percent of patents are in this cluster and Table 6 indicates that the trend of patents technology in this cluster is in the growth stage and it is predicted that the number of patents of this cluster will be reached to 351 by 2031. In the last row of Table 6, the forecasting process has been performed for all available patents in the database. As it is indicated, the trend of blockchain patent production is in the maturity stage from 2001 until May 2020 and it can be expected that it will be reached to 5234 by approximately the future 17 years. The growth curve of the patents in each cluster has been shown in Fig. 14. According to Table 6, almost all patent technologies in different clusters have the potential for development and investment.

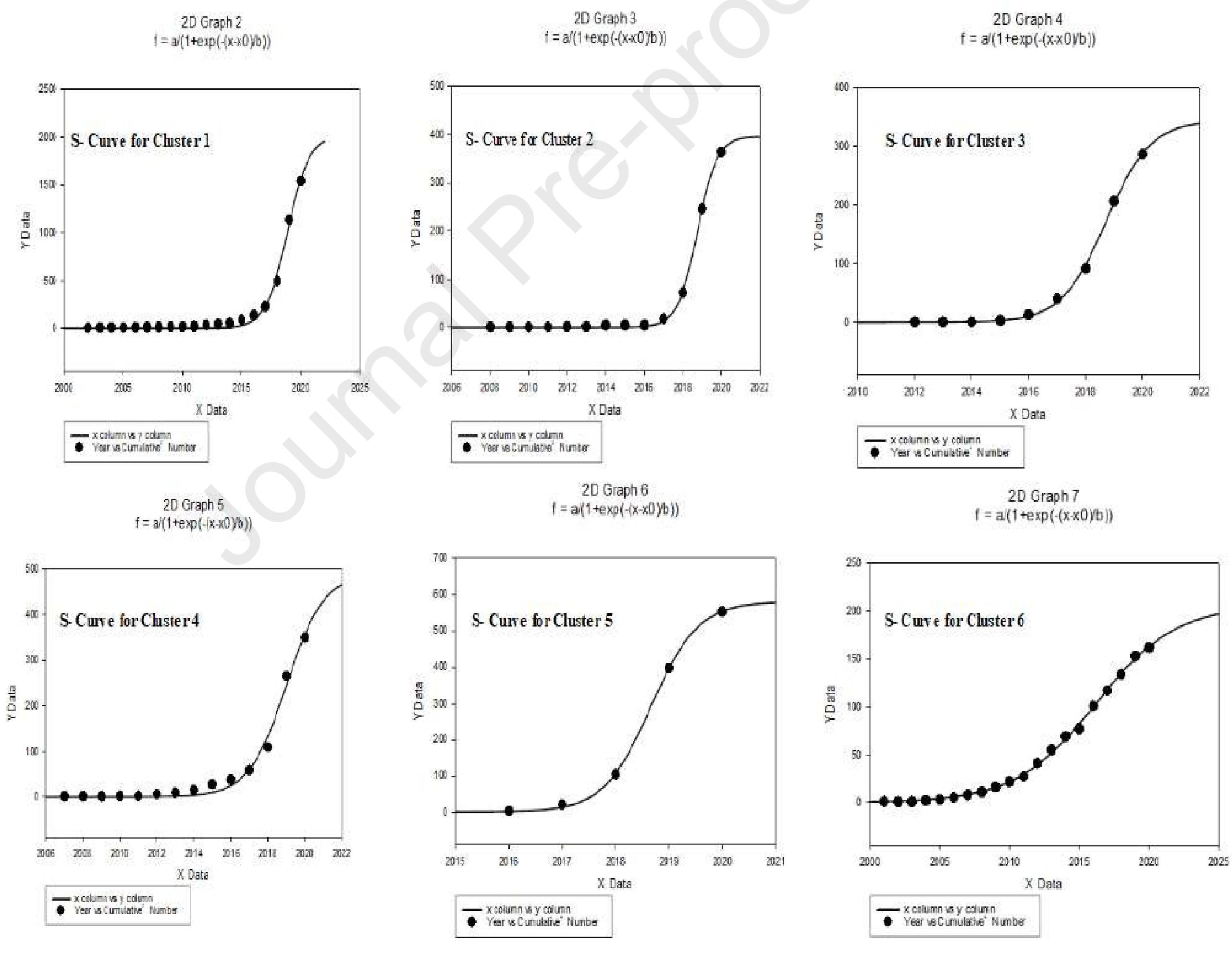


Fig. 14. Growth curve of blockchain technology patents in six clusters.

5 Discussion and challenges

Blockchain technologies have been developed for almost a decade, even though, it is considered more precisely in the last five years (Bhatt, Kumar, Lu, Cho, & Lai, 2020). Researchers have published their studies of blockchain patents as academic articles and technical reports. Furthermore, a simple search with the keyword "blockchain" in the Scopus database demonstrates that the publications on this topic have increased from 10 in 2014 to 5806 in 2019. It is evidenced that the popularity of this technology has risen more and more amongst researchers and enthusiasts. Hunting for the papers in the blockchain dataset has resulted in 13833 papers whose keywords' relationship map is illustrated in Fig. 15 by VOSviewer software.

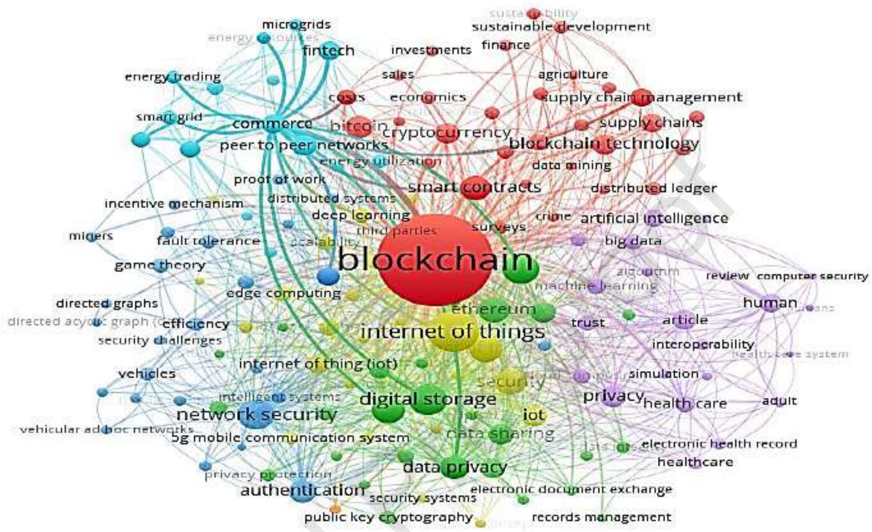


Fig. 15. The relationship mapping between the keywords of studies on blockchain technology.

According to Figure 15 and keyword clustering terms like IoT, Cryptocurrencies, Network Security, Consensus and Digital Storage have the most frequencies. Fig. 16 shows the researcher's partnership from various countries that represents the more impressive attendance of American and Chinese researchers.

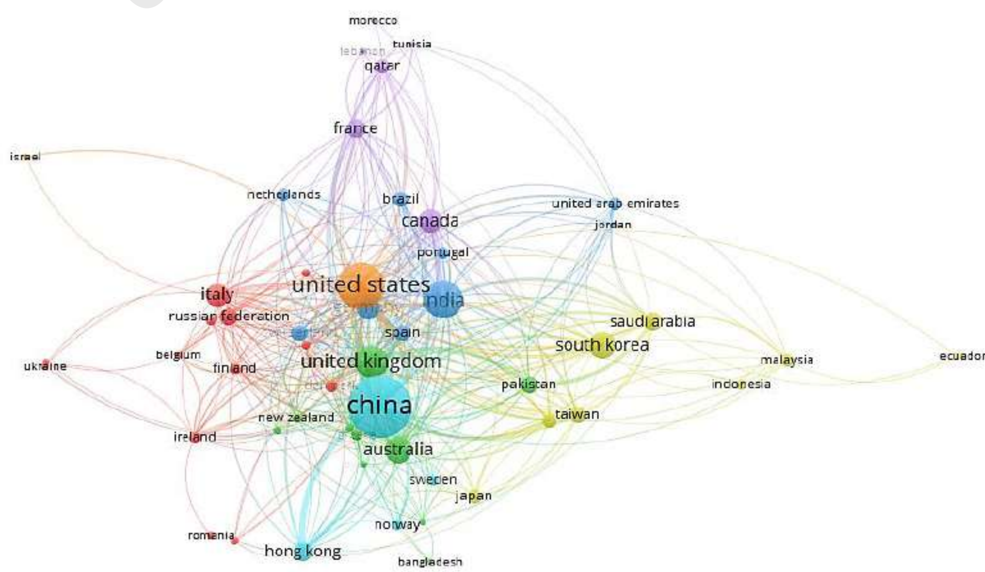


Fig. 16. Researcher's cooperation from different countries on the blockchain domain.

The popularity of blockchain according to the high volume of papers and patents shows that the blockchain is not only applicable for cryptocurrencies, but all are seeking its more extended applications in other areas and recognizing this technology's potentials (Filippova, Scharl, & Filippov, 2019). Blockchain is still an emerging technology right now and its applications are enquired and analyzed by many researchers and inventors (Bhatt et al., 2020). E-commerce, the Internet of Things, digital signature, and cryptocurrencies are the most significant topics amongst the researchers (Filippova et al., 2019). Furthermore, the capacities of blockchain technology by integrating with other technologies like Artificial Intelligence (AI), Cloud Computing, and the Internet of Things is growing, where it has resulted in the investment in blockchain (Batubara, Ubacht, & Janssen, 2018; Berryhill, Bourgery, & Hanson, 2018; Bhatt et al., 2020; Grody, 2018; Wijaya, Liu, Suwarsono, & Zhang, 2017).

There are many open challenges in the adoption of blockchain technology that need to be addressed. In the case of cryptocurrencies, the challenges including the speed of transactions, the cost of transactions, and the miners' award are regarded (Bamakan et al., 2020; Filippova et al., 2019; Meva, 2018). How to achieving the consensus is one of the most significant challenges that a vast variety of studies and patents has noticed (Bamakan et al., 2020). Scalability, privacy, selfish extraction, security, fork issues, affirmation timeline, partnership competency, regular issues, and energy consumption are the other challenges that the researchers have referred to them (Bamakan et al., 2020; Batubara et al., 2018; Zheng et al., 2017).

Blockchain Security and Privacy is a key concept and important concern that about Transaction in blockchain and includes protecting the information and data contained in the blocks against various attacks. The most important security issues that could put Blockchain at risk are Liveness Attack, Double Spending Attacks, 51% Vulnerability Attack, Private Key Security Attack, Transaction Privacy Leakage, Selfish Mining Attack, Dao Attack, BGP Hijacking Attack, Sybil Attack, Mining pool attacks, Criminal activity, Transaction privacy leakage (Bamakan, Moghaddam, & Manshadi, 2021; Gupta, 2020; Singh, Hosen, & Yoon, 2021). Security and validation of a block is an important task performed by a special mechanism in blockchain called consensus algorithms. Reaching consensus in a distributed system is a challenge. The Chinese block does not have a central node to control and verify transactions. Therefore, it is necessary to design protocols to verify transactions. Hence, consensus algorithms are the core of blockchain technology. According to these protocols, the new transaction must be approved by all members of the network (nodes) and thus added to the previous block. So far, more than 15 consensus algorithms have been designed, the most important of which is Proof of Work, Proof of Stake, Proof of Elapsed Time, or Practical Byzantine Fault Tolerance (Bamakan et al., 2020; Zheng et al., 2017).

Another challenge that blockchain technology faces is energy consumption. According to the algorithms used in the Chinese block, its energy consumption is high. Therefore, institutions and companies are trying to find a way to solve the problem of energy consumption in the blockchain and the use of renewable energy for sustainable development and the prevention of environmental hazards (Ghosh & Das, 2019). Another challenge of blockchain technology is the waste of energy resources. This concern raises two important questions. The first question is how to reduce energy consumption and the second is how computational power can be used to process user's data (Bamakan, Faregh, & ZareRavasan, 2021; Casino et al., 2019). One of the issues in the blockchain is scalability and usability, which play an important role in its development as the number of blockchain users increases. Every new technology faces these challenges. Scalability in the blockchain is the processing of a large number of transactions per second. Usability in any technology refers to the degree of ease of use by users (Hafid, Hafid, & Samih, 2020; Moniruzzaman, Chowdhury, & Ferdous, 2020). Fig. 17 shows a summary of the most important blockchain challenges.

Fig. 17. The most important blockchain technology challenges.

6 Conclusion

The continuous progress of science and technology has developed a competitive environment, and technology has become a significant competitive tool. An organization would be in a competitive circumstance by implementing the technology when it is applied at an appropriate time. Technology trend or technology lifecycle begins with the initiation phase. Then, it enters the maturity phase and finally the saturation phase. If technology enters the saturation phase, it will be removed or replaced by another technology. Therefore, the investment in technologies in this phase could not be so logical. Hence, technology forecasting is one of the significant tools in research and development to determine its future trend.

In this research, we attempted to investigate the trends of blockchain technologies based on registered patents. Initially, the registered patents in the WIPO database have been searched and extracted, then, classified based on IPC code and their trend diagram is drawn according to this code. The results manifested that the patent registry trend has increased since 2014 and a total of 14072 patents has been registered until May 2020. Among the patents, 39% of patents are assisted to G06Q class and 40% to H04L. Furthermore, the exploration of patent trends by use of Sigma Plot represented that technologies of H04L and G06Q are in the maturity and saturation phase.

In order to further analyze, we concentrate on the registered patent in the US patent database and we applied a text mining approach for patent classification (in contrast to the previous section that was based on IPC code). Finally, the patents were extracted based on keywords and were clustering based on co-occurrence between them. The obtained clusters are containing the keywords in the patents that are called technology clusters and are applied for patents' classification. The results indicated that the main topic of most patents is in the various applications of blockchain technology including cryptocurrencies, financial, SCM, healthcare, gaming, security, and privacy. The clusters of most technologies are approximately in the maturity phase. However, blockchain technology is in the initialization phase and there are many challenges and limitations, which need to be addressed by researchers.

Data Availability Statements

The data underlying this article are available in https://github.com/smhbamakan/Blockchain_patent_data.

Acknowledgments

This work has been partially supported by CAS President's International Fellowship Initiative, China [grant number 2021VTB0002, 2021].

References

Abbas, A., Zhang, L., & Khan, S. U. (2014). A literature review on the state-of-the-art in patent analysis. *World Patent Information, 37*, 3-13.

Abeyratne, S. A., & Monfared, R. P. (2016). Blockchain ready manufacturing supply chain using distributed ledger. *International Journal of Research in Engineering and Technology, 5*(9), 1-10.

Adamuthe, A. C., & Thampi, G. T. (2019). Technology forecasting: A case study of computational technologies. *Technological Forecasting and Social Change, 143*, 181-189.

Akins, B. W., Chapman, J. L., & Gordon, J. M. (2014). A whole new world: Income tax considerations of the Bitcoin economy. *Pitt. Tax Rev., 12*, 25.

Allahyari, M., Pouriyeh, S., Assefi, M., Safaei, S., Trippe, E. D., Gutierrez, J. B., & Kochut, K. (2017). A brief survey of text mining: Classification, clustering and extraction techniques. *arXiv preprint arXiv:1707.02919*.

Altuntas, S., Dereli, T., & Kusiak, A. (2015). Forecasting technology success based on patent data. *Technological Forecasting and Social Change, 96*, 202-214.

Asan, U., & Ercan, S. (2012). An introduction to self-organizing maps *Computational Intelligence Systems in Industrial Engineering* (pp. 295-315): Springer.

Bamakan, S. M. H., Faregh, N., & ZareRavasan, A. (2021). Di-ANFIS: an integrated blockchain–IoT–big data-enabled framework for evaluating service supply chain performance. *Journal of Computational Design and Engineering, 8*(2), 676-690.

Bamakan, S. M. H., Moghaddam, S. G., & Manshadi, S. D. (2021). Blockchain-enabled pharmaceutical cold chain: applications, key challenges, and future trends. *Journal of Cleaner Production*, 127021.

Bamakan, S. M. H., Motavali, A., & Bondarti, A. B. (2020). A survey of blockchain consensus algorithms performance evaluation criteria. *Expert Systems with Applications*, 113385.

Barcelon Yang, C. (2012). Role of patent analysis in corporate R&D. *Pharmaceutical patent analyst, 1*(1), 5-7.

Batubara, F. R., Ubacht, J., & Janssen, M. (2018). *Challenges of blockchain technology adoption for e-government: a systematic literature review.* Paper presented at the Proceedings of the 19th Annual International Conference on Digital Government Research: Governance in the Data Age.

Baumann, M., Domnik, T., Haase, M., Wulf, C., Emmerich, P., Rösch, C., . . . Weil, M. (2021). Comparative patent analysis for the identification of global research trends for the case of battery storage, hydrogen and bioenergy. *Technological Forecasting and Social Change, 165*, 120505.

Berryhill, J., Bourgery, T., & Hanson, A. (2018). Blockchains unchained.

Bhatt, P. C., Kumar, V., Lu, T.-C., Cho, R. L.-T., & Lai, K. K. (2020). *Rise and Rise of Blockchain: A Patent Statistics Approach to Identify the Underlying Technologies.* Paper presented at the Asian Conference on Intelligent Information and Database Systems.

Campbell, R. S. (1983). Patent trends as a technological forecasting tool. *World Patent Information, 5*(3), 137-143.

Casino, F., Dasaklis, T. K., & Patsakis, C. (2019). A systematic literature review of blockchain-based applications: current status, classification and open issues. *Telematics and informatics, 36*, 55-81.

Chen, Y.-L., & Hu, H.-L. (2006). An overlapping cluster algorithm to provide non-exhaustive clustering. *European Journal of Operational Research, 173*(3), 762-780.

Cheng, A.-C., & Chen, C.-Y. (2008). The technology forecasting of new materials: the example of nanosized ceramic powders. *Romanian Journal of Economic Forecasting, 4*, 88-110.

Chernov, A., & Chernova, V. (2018). Global Blockchain Technology Market Analysis-Current Situations And Forecast. *Economic and Social Development: Book of Proceedings*, 143-152.

Cho, Y., & Daim, T. (2013). Technology forecasting methods *Research and Technology Management in the Electricity Industry* (pp. 67-112): Springer.

Choi, J., Im, N., Kim, H., & Hwang, Y.-S. (2012). *A study on the patent analysis for effective technology forecasting.* Paper presented at the 2012 8th International Conference on Information Science and Digital Content Technology (ICIDT2012).

Choi, Y., & Hong, S. (2020). Qualitative and quantitative analysis of patent data in nanomedicine for bridging the gap between research activities and practical applications. *World Patent Information, 60*, 101943.

Cottrell, M., Olteanu, M., Rossi, F., & Villa-Vialaneix, N. (2018). Self-OrganizingMaps, theory and applications.

Daim, T. U., Ploykitikoon, P., Kennedy, E., & Choothian, W. (2008). Forecasting the future of data storage: case of hard disk drive and flash memory. *foresight*.

Dou, H., Leveillé, V., Manullang, S., & JM Jr, D. (2005). Patent analysis for competitive technical intelligence and innovative thinking. *Data science journal, 4*, 209-236.

Ernst, H. (1997). The use of patent data for technological forecasting: the diffusion of CNC-technology in the machine tool industry. *Small business economics, 9*(4), 361-381.

Filippova, E., Scharl, A., & Filippov, P. (2019). *Blockchain: an empirical investigation of its scope for improvement.* Paper presented at the International Conference on Blockchain.

Firat, A. K., Woon, W. L., & Madnick, S. (2008). Technological forecasting–A review. *Composite Information Systems Laboratory (CISL), Massachusetts Institute of Technology*.

Foroglou, G., & Tsilidou, A.-L. (2015). *Further applications of the blockchain.* Paper presented at the 12th student conference on managerial science and technology.

Gao, F., Zhu, L., Shen, M., Sharif, K., Wan, Z., & Ren, K. (2018). A blockchain-based privacy-preserving payment mechanism for vehicle-to-grid networks. *IEEE network, 32*(6), 184-192.

Gao, L., Porter, A. L., Wang, J., Fang, S., Zhang, X., Ma, T., . . . Huang, L. (2013). Technology life cycle analysis method based on patent documents. *Technological Forecasting and Social Change, 80*(3), 398-407.

Ghosh, E., & Das, B. (2019). *A study on the issue of blockchain's energy consumption.* Paper presented at the International Ethical Hacking Conference.

Grody, A. D. (2018). Rebuilding financial industry infrastructure. *Journal of Risk Management in Financial Institutions, 11*(1), 34-46.

Gupta, N. (2020). A Deep Dive Into Security and Privacy Issues of Blockchain Technologies *Handbook of Research on Blockchain Technology* (pp. 95-112): Elsevier.

Hafid, A., Hafid, A. S., & Samih, M. (2020). Scaling blockchains: A comprehensive survey. *IEEE Access, 8*, 125244-125262.

Han, J., Pei, J., & Kamber, M. (2011). *Data mining: concepts and techniques*: Elsevier.

Hsu, F.-C., Trappey, A. J., Trappey, C. V., Hou, J.-L., & Liu, S.-J. (2006). Technology and knowledge document cluster analysis for enterprise R&D strategic planning. *International Journal of Technology Management, 36*(4), 336-353.

Jaranyagorn, P., & Chansa Ngavej, C. (2012). *Technology forecasting methods with emphasis on bibliographic analysis and curve fitting: a photocatalytic case example.* Paper presented at the Advanced Materials Research.

Jun, S., Park, S. S., & Jang, D. S. (2012). Patent management for Technology Forecasting: A case study of the Bio-Industry.

Kohonen, T. (2012). *Self-organizing maps* (Vol. 30): Springer Science & Business Media.

Korpela, K., Hallikas, J., & Dahlberg, T. (2017). *Digital supply chain transformation toward blockchain integration.* Paper presented at the proceedings of the 50th Hawaii international conference on system sciences.

Kosba, A., Miller, A., Shi, E., Wen, Z., & Papamanthou, C. (2016). *Hawk: The blockchain model of cryptography and privacy-preserving smart contracts.* Paper presented at the 2016 IEEE symposium on security and privacy (SP).

Kravitz, D. W., & Cooper, J. (2017). *Securing user identity and transactions symbiotically: IoT meets blockchain.* Paper presented at the 2017 Global Internet of Things Summit (GIoTS).

Kuo, T.-T., Kim, H.-E., & Ohno-Machado, L. (2017). Blockchain distributed ledger technologies for biomedical and health care applications. *Journal of the American Medical Informatics Association, 24*(6), 1211-1220.

Lamirel, J.-C., Dugué, N., & Cuxac, P. (2016). *New efficient clustering quality indexes.* Paper presented at the 2016 International Joint Conference on Neural Networks (IJCNN).

Lee, S., Kim, W., Kim, Y. M., Lee, H. Y., & Oh, K. J. (2014). The prioritization and verification of IT emerging technologies using an analytic hierarchy process and cluster analysis. *Technological Forecasting and Social Change, 87*, 292-304.

Lehtovirta, V. (2019). Patent analysis for social usage of blockchain.

Li, X., Xie, Q., Daim, T., & Huang, L. (2019). Forecasting technology trends using text mining of the gaps between science and technology: The case of perovskite solar cell technology. *Technological Forecasting and Social Change, 146*, 432-449.

Liu, C.-Y., & Wang, J.-C. (2010). Forecasting the development of the biped robot walking technique in Japan through S-curve model analysis. *Scientometrics, 82*(1), 21-36.

Mettler, M. (2016). *Blockchain technology in healthcare: The revolution starts here.* Paper presented at the 2016 IEEE 18th international conference on e-health networking, applications and services (Healthcom).

Meva, D. (2018). Issues and challenges with blockchain: A survey. *International Journal of Computer Sciences and Engineering, 6*, 488-491.

Moniruzzaman, M., Chowdhury, F., & Ferdous, M. S. (2020). *Examining Usability Issues in Blockchain-Based Cryptocurrency Wallets.* Paper presented at the International Conference on Cyber Security and Computer Science.

Monrat, A. A., Schelén, O., & Andersson, K. (2019). A survey of blockchain from the perspectives of applications, challenges, and opportunities. *IEEE Access, 7*, 117134-117151.

Moosavi, J., Naeni, L. M., Fathollahi-Fard, A. M., & Fiore, U. (2021). Blockchain in supply chain management: a review, bibliometric, and network analysis. *Environmental Science and Pollution Research*, 1-15.

Morrow, M. J., & Zarrebini, M. (2019). Blockchain and the Tokenization of the Individual: Societal Implications. *Future Internet, 11*(10), 220.

Muzammal, M., Qu, Q., & Nasrulin, B. (2019). Renovating blockchain with distributed databases: An open source system. *Future Generation Computer Systems, 90*, 105-117.

Noyes, C. (2016). Bitav: Fast anti-malware by distributed blockchain consensus and feedforward scanning. *arXiv preprint arXiv:1601.01405*.

Peters, G., Panayi, E., & Chapelle, A. (2015). Trends in cryptocurrencies and blockchain technologies: A monetary theory and regulation perspective. *Journal of Financial Perspectives, 3*(3).
Qu, Q., Nurgaliev, I., Muzammal, M., Jensen, C. S., & Fan, J. (2019). On spatio-temporal blockchain query processing. *Future Generation Computer Systems, 98*, 208-218.
Sarstedt, M., & Mooi, E. (2014). A concise guide to market research. *The Process, Data, and, 12*.
Singh, S., Hosen, A. S., & Yoon, B. (2021). Blockchain security attacks, challenges, and solutions for the future distributed iot network. *IEEE Access, 9*, 13938-13959.
Smith, S. S. (2019). Blockchain, Tokenization, and Implications for Financial Services Practitioners. *International Journal of Accounting and Financial Reporting, 9*.
Tasca, P., Thanabalasingham, T., & Tessone, C. J. (2017). Ontology of blockchain technologies. Principles of identification and classification. *SSRN Electronic Journal, 10*.
Tian, F. (2016). *An agri-food supply chain traceability system for China based on RFID & blockchain technology.* Paper presented at the 2016 13th international conference on service systems and service management (ICSSSM).
Trappey, C. V., Wu, H.-Y., Taghaboni-Dutta, F., & Trappey, A. J. (2011). Using patent data for technology forecasting: China RFID patent analysis. *Advanced Engineering Informatics, 25*(1), 53-64.
Tseng, Y.-H., Lin, C.-J., & Lin, Y.-I. (2007). Text mining techniques for patent analysis. *Information processing & management, 43*(5), 1216-1247.
Turk, Ž., & Klinc, R. (2017). Potentials of blockchain technology for construction management. *Procedia engineering, 196*, 638-645.
Watts, R. J., & Porter, A. L. (1997). Innovation forecasting. *Technological Forecasting and Social Change, 56*(1), 25-47.
Wijaya, D. A., Liu, J. K., Suwarsono, D. A., & Zhang, P. (2017). *A new blockchain-based value-added tax system.* Paper presented at the International conference on provable security.
Witchey, N. J. (2019). Healthcare transaction validation via blockchain, systems and methods: Google Patents.
Xu, M., Chen, X., & Kou, G. (2019). A systematic review of blockchain. *Financial Innovation, 5*(1), 27.
Yaga, D., Mell, P., Roby, N., & Scarfone, K. (2019). Blockchain technology overview. *arXiv preprint arXiv:1906.11078*.
Yan, M. (2005). *Methods of determining the number of clusters in a data set and a new clustering criterion.* Virginia Tech.
Yoon, B., & Park, Y. (2007). Development of new technology forecasting algorithm: Hybrid approach for morphology analysis and conjoint analysis of patent information. *IEEE Transactions on Engineering Management, 54*(3), 588-599.
Zhang, Y., & Wen, J. (2015). *An IoT electric business model based on the protocol of bitcoin.* Paper presented at the 2015 18th international conference on intelligence in next generation networks.
Zheng, Z., Xie, S., Dai, H.-N., Chen, X., & Wang, H. (2018). Blockchain challenges and opportunities: A survey. *International Journal of Web and Grid Services, 14*(4), 352-375.
Zheng, Z., Xie, S., Dai, H., Chen, X., & Wang, H. (2017). *An overview of blockchain technology: Architecture, consensus, and future trends.* Paper presented at the 2017 IEEE international congress on big data (BigData congress).

Investigating blockchain technology to explore its trends by IPC classification.

Examining and clustering blockchain technology patents by text mining and clustering techniques.

Predicting blockchain technology life cycle by registered patents in WIPO.

Smart contact applications, privacy-preserving, and tokenization are the main groups of patents.

Blockchain Technology Forecasting by Patent Analytics

Conflict of Interest

We have no conflict of interest to declare.

Sincerely,

Dr. Seyed Mojtaba Hosseini Bamakan[a,b]
Alireza Babaei Bondarti[a]
Parinaz Babaei Bondarti[b]
Qiang Qu[c]

[a] Data Science Research Center, Yazd University, Yazd, 89195-741, Iran
[b] Department of Library and Information Science, Kharazmi University, Tehran, Iran
[c] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen, 518055, China