ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Legal Analysis of the Legal Effect of Smart Contracts in Commercial Transactions

Qianyu Chen[1,*]

[1] College of Law, Central South University of Forestry and Technology, Changsha, Hunan, 410004, China

SUMMARY: *Blockchain technology has a significant impact on smart contracts. In practice, they reflect the redistribution and organization of the rights and obligations of the contract parties, but the existing legal regulations regarding their functioning are still underdeveloped. With this regard, it is highly important to consider the legal effect of smart contracts in commercial transactions. The current research uses the methods of text-mining such as Chinese word segmentation, keyword extraction, co-occurrence network analysis, and LDA topic clustering to derive the factors that contribute to the legal effect of smart contracts in commercial transactions and then explores the relationships between those factors using association-rule mining in order to reveal the particularities of their legal effect. The experimental outcomes demonstrate that there are four critical areas of concern related to smart contracts: validity determination, transaction security, responsibility definition, and risk prevention and control. There are altogether 11 LDA topics and 33 dispute focus points of the legal effect of smart contracts in business transactions. Of them, the most significant share belongs to the disputes based on professional ethics and responsibility boundaries, which amount to 52.83 percent. The obtained result indicates that the professional ethics and the demarcation of responsibilities are very important aspects of the development of the legal effect of smart contracts.*

KEYWORDS: *text mining technology; association rules; smart contract; legal effect*

1 Introduction

The notion of a “contract” is present in nearly every legal system. Although its formal definition may vary, the essential foundation remains consistent, namely, a legal norm directly governed by the relationship between claims and debts, that is, the configuration of rights and obligations between the parties [1-3]. Literature [4] discusses the importance of contracts in commercial relations and notes that contract administration in enterprises often appears relatively monotonous, while the conclusion of a contract in the legal sense is intended to ensure expected outcomes, fulfillment of obligations, and assistance under unforeseen circumstances. Literature [5] evaluates the effectiveness of contract law in dealing with the legal complexity and practical challenges of business transactions, emphasizing that contract law is an important field for learning and engaging in commercial affairs. With the rapid growth of the digital economy, smart contracts, as one of the core applications of blockchain technology, are reshaping the traditional model of contract formation and performance through the self-executing mechanism of “code as law” [6-8].

*znlchenqianyu@163.com
https://doi.org/10.65102/is2026383
Published: 30 April, 2026

Smart contracts are able to enforce contract terms and conditions automatically without relying on intermediaries, thereby enabling transactions that are credible, dependable, and secure [9, 10]. In commercial dealings, they can significantly improve contractual transparency, precision, and efficiency, while also facilitating transaction development and innovation in transactional methods [11, 12]. The most notable distinction between traditional contracts and smart contracts lies in the latter's ability to establish mutual trust between parties through decentralization [13, 14]. Owing to the immutability and decentralized structure of blockchain technology, smart contracts can prevent contracts from being altered or deleted and can ensure that parties carry out transactions according to the agreed terms and conditions [15, 16]. Literature [17] investigated sustainable business models supported by smart contracts, aiming to identify the key elements involved in applying smart contracts to realize SBMs, and, based on case studies, showed that smart contracts can reduce transaction costs and enhance social trust so as to support the development of new sustainable business models. Literature [18 reviewed the implementation of smart contracts in enterprises through blockchain technology to improve performance and described both the applications and the benefits and shortcomings associated with such contracts. Literature [19] analyzed the concept of smart contracts and their relevance to the business sector on the basis of blockchain technology and offered insights into both the advantages and the current limitations of their applicability. Literature [20] examined blockchain technology from the perspective of legal regulation and the possibility of using smart contracts, treating it as one of the application areas of this technology and as a simulation of civil circulation contracts grounded in legal analysis. Literature [21] pointed out that smart contracts have attracted considerable attention because of their unique features of automatic execution, transparency, and resistance to tampering in blockchain environments, and it further verified their characteristics, benefits, and potential advantages from the procurement perspective. Literature [22], based on a literature review, explored the potential areas that can effectively benefit from smart-contract implementation, emphasized that smart contracts are a viable instrument, and carried out related testing. These studies collectively demonstrate the applications, effects, strengths, and limitations of smart contracts. Even so, the legal validity of this new contract form, which depends on blockchain technology, still raises a collision between technological characteristics and legal rules. For example, does the automatic execution of code amount to consent in the legal sense [23]? Do the technical characteristics of non-tampering restrict the right to modify and revoke the contract [24]? How can ambiguity in the subject of liability be resolved within the existing legal framework [25]?

These issues concern not only the limits of applying smart contracts in the commercial sphere, but also directly affect the protection of the rights and interests of participants in commercial transactions [26]. Regarding the legal effect of smart contracts, literature [27] introduces their characteristics, highlights their ability to improve business efficiency, reduce transaction and legal costs, and realize anonymous transactions, and also analyzes the legal and practical enforceability issues that may arise when smart contracts are applied from the perspectives of civil law and common law. Literature [28] emphasizes that smart contracts represent a new contractual form capable of automatic and self-execution in whole or in part, and it regards them as a development and innovation of traditional contracts from a legal point of view. Literature [29] describes the use of smart contracts in financial innovation and, through systematic analysis, shows that the legal effects of smart contracts can be clarified in the context of the rapid development of financial technology. Therefore, to clarify the legal effect of smart contracts in commercial transactions, it is necessary to combine the "credibility of the code" with the "binding force of the legal text," and to examine the legal basis and practical difficulties relating to their formation, performance, evidence, remedies, and legal application [30-32].

The current article initially performs a thorough analysis of the fundamental technology behind the smart contract and subsequently outlines the techniques and processes of text mining and association-rule mining. Next, information about smart contracts in business transactions is treated as the corpus of text mining and the TF-IDF algorithm is used to derive the text keywords and identify the distribution pattern of the high-frequency words on time. Then, the LDA topic model is employed to conduct topic mining on smart-contract case in commercial transactions to determine what causes the legal effect of smart-contracts and offer data backup in the process of mining correlations among the relevant components. Following this, the internal correlation structure of the variables that influence the legal effect of smart contracts in commercial transactions is further investigated using association rules. Lastly, the paper examines the unique characteristics of the legal effect of smart contracts as a type of contract based on four aspects.

2 Data mining of the legal effects of smart contracts in commercial transactions

2.1 Smart Contracts

The implementation of smart contracts in the blockchain network might be viewed as a digital version of the paper-based agreements in the real-world transactions. Upon linking a node to the blockchain, it must perform particular tasks based on the smart contracts implemented within the network. Such deployed smart contracts become user-specific, which assists in maintaining the confidentiality of the contract, and developers can alter and update them at any time. Simultaneously, smart contracts minimize human control during the execution of the contract. To give an example, it can compute the required payment of one party under a contract and subsequently issue the payment instruction. Consequently, there is no need to involve a third-party organization, and no manual intervention is necessary to make a transaction happen.

Smart contracts integration with blockchain technology enables storing the traditional contracts on computers in a distributed way to ensure decentralization, resistance to tampering and high reliability. Codes of smart-contracts include those terms that were agreed between both parties. After such conditions are met, the contract is executed automatically, solving the issue of decentralization and reducing the cost of trust. Moreover, smart contracts offer the functional property of non-tampering to the blockchain system.

In terms of coalition-chain use cases, a smart contract is just a collection of logical rules that represents a business case and tends to manifest itself at the simulated transaction level prior to reaching consensus. The smart contract is prepared and tested by developers before the simulated transaction step, and once checked and signed by various participants, it is installed on the blockchain. That is the moment when all activities occurring on the blockchain can be tracked. As soon as an endorsing node identifies the fact that a smart contract was called and that the predefined contract conditions were met, the contract becomes active and performs the respective operation. There can also be some conditional function interfaces which call other smart contracts. Figure 1 shows how the smart contract works.

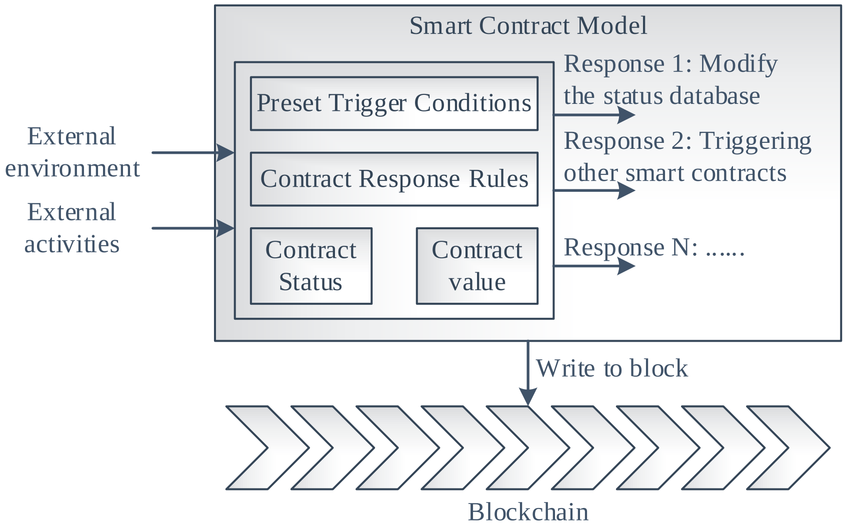


Figure 1: operation principle of smart contract

2.2 Text Mining Related Technologies

2.2.1 Web crawling techniques

Web crawler has a set rules strategy to follow, which automatically crawls and downloads Internet pages, and subsequently processes the information of the downloaded pages based on relevant rules and algorithms, including data extraction and indexing. In terms of gathering information about the websites, one needs to have some idea of how the target website network operates. In Figure 2, the web-crawling process is illustrated.

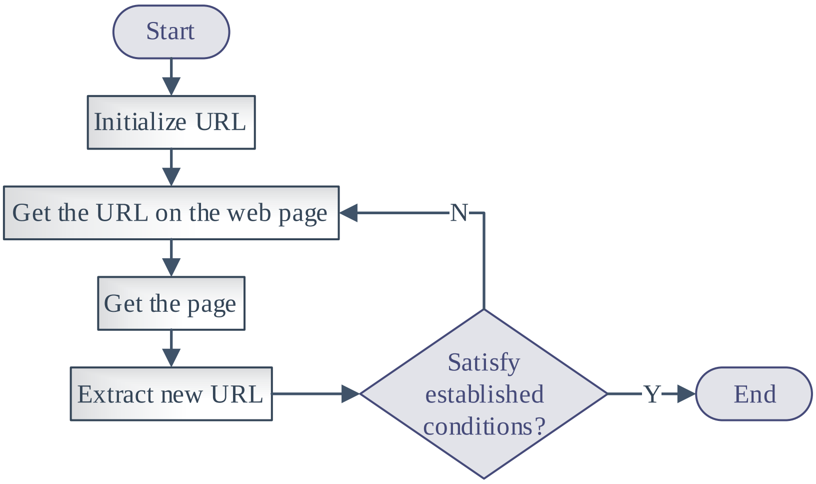


Figure 2: The process of web crawl

2.2.2 Chinese word separation

Word segmentation in Chinese means dividing a series of Chinese characters to form separate words. Unlike in English where spaces serve as separators, Chinese words do not have natural separators, which makes Chinese text segmentation harder than English text segmentation. According to the realization principles and specific features of the Chinese segmentation methods, the Chinese word-segmentation algorithms may be divided into three general groups:

string-matching-based algorithms, statistical algorithms and semantics-based algorithms. This paper has chosen Jieba as its segmentation tool.

(1) Segmentation Algorithm

(a) String Matching Based Segmentation Algorithm

The central concept behind the string-matching segmentation algorithm is that, initially, a matching rule is specified and then, based on that rule, the Chinese character sequence to be segmented is compared against all of the words in the so-called rich enough participle dictionary until a valid match is found, which leads to segmentation. Based on various search directions, this approach may be classified into three types, including forward matching, reverse matching, and both-way matching. Because the basic use of string-based matching algorithms has numerous constraints, it is frequently required to combine them with other algorithms to enhance the quality of segmentation.

(b) Statistical Based Segmentation Algorithm

The statistical segmentation algorithm is based on probabilistic combinations. It assumes that the frequency with which neighboring Chinese characters appear together can, to some extent, represent the probability that they form words. If the algorithm is expressed mathematically, then for a string Y, there may exist m kinds of segmentation results:

$$\begin{matrix} \{X_{11} & X_{12} & \dots & X_{1n_1}\} \\ \{X_{21} & X_{22} & \dots & X_{2n_2}\} \\ \{\vdots & \vdots & \vdots & \vdots\} \\ \{X_{m1} & X_{m2} & \dots & X_{mn_n}\} \end{matrix} \tag{1}$$

where $n_i (i = 1, 2, \dots m)$ denotes the number of words in the i th segmentation result. If we want to obtain the expected optimal segmentation result j, the corresponding distribution probability of that segmentation result should be the largest among all possible segmentation results, i.e.:

$$j = \arg\max_i P\left(X_{i1}, X_{i1}, \dots, X_{im_i}\right) \tag{2}$$

Because the joint distribution of n_i words is involved in $P\left(X_{i1}, X_{i1}, \dots, X_{im_i}\right)$, direct calculation is difficult. Therefore, in practical solutions, a Markov assumption is often introduced, namely, that the next word is related only to one or several preceding words.

If the first n th word is related to the previous 1 word, the distribution probability of the segmentation result is given by Equation (3):

$$P\left(X_{i1}, X_{i1}, \dots, X_{im_i}\right) = P\left(X_{i1}\right) P\left(X_{i2} \mid X_{i1}\right) \dots P\left(X_{in_i} \mid X_{i(n_i-1)}\right) \tag{3}$$

If the nth word is related to the previous 2 words, the distribution probability of the partition result is shown in Equation (4):

$$P\left(X_{i1}, X_{i1}, \dots, X_{im_i}\right) = P\left(X_{i1}\right) P\left(X_{i2} \mid X_{i1}\right) P\left(X_{i3} \mid X_{i1}, X_{i2}\right) \dots P\left(X_{in,} \mid X_{i(n_i-2)}, X_{i(n_i-1)}\right) \tag{4}$$

Equation (3) is called the binary model, whereas Equation (4) is referred to as the ternary model, and this framework can be generalized further to the N -meta model.

(c) Semantic-based Segmentation Algorithm

The segmentation algorithm based on semantics essentially uses the division of text into words using Chinese grammar and contextual information. Since Chinese linguistic rules and semantics are very complex, this human-assisted semi-supervised learning model needs a significant volume of linguistic knowledge and information. That is why any semantic-based segmentation system has not gone beyond the phase of constant experimentation.

(2) Jieba Participle

Jieba has the principle of creating a prefix dictionary based on a statistical dictionary. Based on this approach, the prefix dictionary can be used to scan the input sentence and derive all possible slicing outcomes. Depending on the cut locations, a directed acyclic graph (DAG) is generated, and the highest probability path is determined using dynamic programming to give the final segmentation result. Jieba offers three modes of segmentation: full mode, precise mode, and search-engine mode. Default is full mode, which eliminates all segmentation possibilities in the sentence. Precise mode makes an attempt to divide the sentence as accurately as possible and is best applicable in the text analysis. Search-engine mode is a modification of the precise mode which will keep dividing the long words until it is finished.

2.2.3 TF-IDF algorithm

Term frequency-inverse document frequency, or TF-IDF, is an unsupervised statistical technique that can be applied to determine how important a term is to the document in question within its corpus. The significance of a particular word is directly proportional to the number of occurrences of the word in the given document and inversely proportional to the number of documents in the corpus that contain the word.

TF, or term frequency, represents the number of times a word appears in a given document, and this value is usually normalized to avoid TF bias toward long documents. If any word is defined as T_i, then the TF value of word T_i is calculated as:

$$TF_{i,j} = \frac{N_{i,j}}{\sum_k N_{k,j}} \tag{5}$$

In Eq. (5), $N_{i,j}$ is the number of occurrences of the word T_i in the document D_j, and $\sum_k N_{k,j}$ is the sum of occurrences of all words in the document D_j.

IDF, i.e. Inverse Document Frequency, represents the distribution of documents containing a word in the corpus. The IDF of the word T_i is calculated as:

$$IDF_i = \log \frac{|D|}{1+\left|\left\{ j : T_i \in D_j \right\}\right|} \tag{6}$$

In Equation (6), $|D|$ denotes the total number of documents in the corpus, and $1+\left|\left\{ j : T_i \in D_j \right\}\right|$ represents the number of documents containing the word, where the added 1 is used to avoid a denominator of zero.

TF-IDF is then obtained as shown in Equation (7), and it can be seen that the larger the TF value is, the larger the corresponding TF-IDF value will be.

$$TFIDF_{i,j} = TF_{i,j} \times IDF_i = \frac{N_{i,j}}{\sum_k N_{k,j}} \times \log \frac{|D|}{1 + \left| \left\{ j : T_i \in D_j \right\} \right|} \tag{7}$$

2.2.4 LDA Subject Modeling

(1) Introduction of LDA Topic Model

The topic model is a type of text representation model that includes the semantic association by extracting hidden topics in the text, which will be used to identify semantic correlations in the text. Another name of the LDA model is a three-layer Bayesian probabilistic model consisting of three structural levels: documents, topics, and words.

LDA uses the bag of words model to treat every document as a word-frequency vector to convert textual data into a simple digital form without taking into account the sequential nature of the words. The probability distribution of a document consists of various topics and the probability distribution of a topic consists of various words. With bag-of-words modeling, the probability of a given word appearing in a certain document can be expressed through this formula:

$$p(\text{Words} \mid \text{Documentation}) = \sum_{\text{Topic}} p(\text{Words} \mid \text{Topic}) \times p(\text{Topics} \mid \text{Documentation}) \tag{8}$$

where $p\,(\text{word} \mid \text{topic})$ denotes the probability of occurrence of each word under a given topic, and $p\,(\text{topic} \mid \text{document})$ denotes the probability of occurrence of each topic in a given document. The probability formula can also be illustrated by a matrix diagram, and the probability of occurrence of each word in the LDA model is shown in Figure 3.

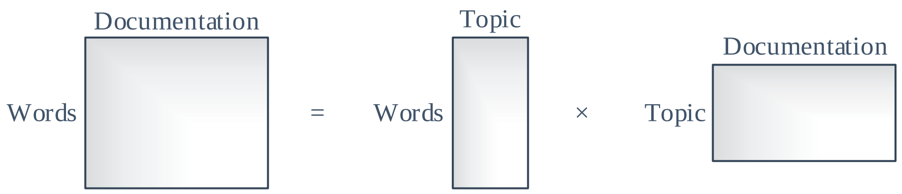


Figure 3: Generated probability of each word in LDA model

The structure of the LDA probability map is shown in Figure 4.

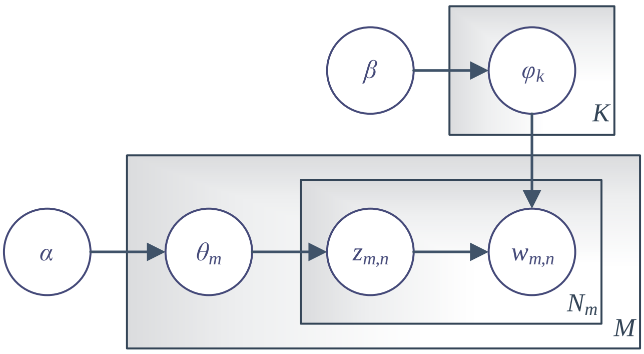


Figure 4: Probability graphical model of LDA

Based on the above probability plot, the joint distribution of the variables can be obtained:

$$p(w,z|\alpha,\beta) = p(w|z,\beta)p(z|\alpha) = \prod_{k=1}^{K}\frac{\Delta(\varphi_k+\beta)}{\Delta(\beta)}\prod_{m=1}^{M}\frac{\Delta(\theta_m+\alpha)}{\Delta(\alpha)} \tag{9}$$

(2) LDA parameter solving

Let $i=(m,n)$ denote a two-dimensional subscript. Then $z_{m,n}$ may be written as z_i, $w_{m,n}$ may be written as $w_i, \neg i$. Here, i refers to a word that does not contain the subscript i. Under Gibbs sampling, if the current word is assumed to be $w_i = t$ the probability that this word is assigned to a certain topic during one iteration can be obtained from Bayes' law as:

$$p(z_i = k \mid z_{\neg i}, w) \propto p(z_i = k, w_i = t \mid z_{\neg i}, w_{\neg i}) \tag{10}$$

After the iteration ends, the topic-word matrix φ and the document-topic matrix θ are output.

$$\varphi_{kt} = \frac{n^{(t)}{}_{k,\neg i} + \beta_t}{\sum_{t=1}^{V}\left(n^{(t)}{}_{k,\neg i} + \beta_t\right)} \tag{11}$$

$$\theta_{mk} = \frac{n^{(k)}{}_{m,\neg i} + \alpha_k}{\sum_{k=1}^{K}\left(n^{(k)}{}_{m,\neg i} + \alpha_k\right)} \tag{12}$$

2.3 Association Rule Mining Methods

2.3.1 Basic concepts

The association rule approach, i.e., through data mining techniques, explores the strong correlations contained in a large number of smart contracts, reflecting certain dependencies or correlations hidden between a certain element in the data set and elements other than it.

(1) Item and item set, transaction

Assume that $I = \{i_1, i_2, i_3, \cdots, i_n\}$ denotes the collection of all items in the information database of a smart contract; this collection is called the item set I, where each i_k represents one item. If the item set I contains k items, it is regarded as a k-itemset. Let $T = \{t_1, t_2, t_3, \cdots, t_m\}$ represent all transactions in the transaction database. In other words, each transaction-processing record contains all items involved in a single processing event, and each tit_iti is a subset of I.

(2) Association rules

An association rule describes a logical implication relationship between item sets. For example, $A \Rightarrow B$ is an association rule, where A belongs to I, B belongs to I, and $A \cap B \neq \varnothing$. Here, A, B are called the antecedent event and the consequent event, respectively. This implication means that when item set A appears, there is a certain probability that item set B will also be triggered. The strength of such a rule is measured through two indicators, namely support and confidence.

(3) Support

Support reflects how frequently itemset A and itemset B occur together, that is, the proportion of transactions containing both A and B in the total number of transactions T. It is used to characterize the occurrence frequency of an association rule, as shown in Eq. (13):

$$S(A \Rightarrow B) = P(A \cup B) \tag{13}$$

(4) Confidence

Confidence represents the conditional probability that itemset B appears when itemset A has already occurred, that is, the proportion of transactions containing both A and B in the total number of transactions containing A. This index is used to describe the reliability level of the association rule, as given in Eq. (14).

$$C(A \Rightarrow B) = P(B \mid A) \tag{14}$$

(5) Frequent itemset

Set the threshold according to the application scenario, set the minimum support threshold as S_{min}, any non-empty itemset I whose support is not less than this threshold (S_{min}) can be regarded as a frequent itemset, denoted as I_i.

(6) Strong association rule

Set the minimum confidence threshold as C_{min}, if an association rule $A \Rightarrow B$ satisfies the threshold condition of C_{min} while the support is not less than S_{min}, the rule is called a strong association rule, and vice versa is regarded as a weak association rule.

(7) Enhancement

The lift measure is defined as the ratio between the probability that itemset B occurs under the condition that itemset A occurs and the overall probability of occurrence of itemset B. It is used to indicate the degree of association between elements before and after the rule, namely, the influence of itemset A on the appearance of itemset B, as shown in Eq. (15):

$$L(A \Rightarrow B) = \frac{C(A \Rightarrow B)}{S(A \Rightarrow B)} = \frac{P(B \mid A)}{P(B)} \tag{15}$$

When $L > 1$, a positive correlation exists between the item sets before and after the rule. The larger the lift value is, the more important the rule is in correlation analysis of smart contracts. When $L < 1$, the rule indicates a negative correlation between the antecedent and consequent items, meaning that an inhibitory effect exists.

2.3.2 Association Rule Mining Steps and Apriori Algorithm

(1) Association rule mining steps

The main steps of association rule mining are to generate frequent itemsets and filter strong association rules.

Frequent Itemset Generation Step 1. Enter the smart-contract data and use a suitable algorithm to determine the support of all candidate itemsets. All candidate itemsets with support that is not below the specified threshold are considered frequent itemsets.

Step 2: Filter strong association rules.

(2) Apriori algorithm

Apriori algorithm is a traditional approach to association rule technology in data mining information relevance. Its main concepts are based on the prior principle and the anti-monotonicity principle and it primarily utilizes the steps of scanning, self-connection, pruning, and others. It is also a bottom-up search strategy because it starts with less complicated search patterns and ends with more complicated search patterns. Prior principle implies that after an itemset I is determined as a frequent itemset, all its non-empty subsets should be satisfied as well with the frequent-itemset property. Conversely, the anti-monotonicity principle asserts that when an itemset I is infrequent, then any superset which contains it will likewise be infrequent. The algorithm works on these two principles as follows:

Step 1: Traverse the database to generate candidate 1-itemsets I_1, count the number of occurrences of each item, and then calculate the support of each item.

Step 2: According to the minimum support threshold, retain the members satisfying the condition as frequent itemsets and generate the frequent 1-itemset I_1. Then perform self-connection on the resulting frequent itemset I_1 to generate candidate 2-itemset I_2.

Step 3: Traverse the database again and calculate the support of each member in I_2. The items that fail to satisfy the minimum support threshold are removed by pruning, whereas those satisfying the condition are retained and filtered to obtain the frequent 2-itemset I_2.

Step 4: Repeat the above process to generate candidate k-itemset I_k and frequent k-itemset I_k until the candidate $k+1$-itemset I_{k+l} becomes empty. In other words, when the support of all items in the candidate itemset produced by self-connection of I_k does not meet the minimum support threshold, no higher-order candidate itemset can be generated, and the cycle ends. In this process, the frequent itemset obtained after screening is the result of association-rule mining, where I_k is the highest-order frequent kkk-itemset.

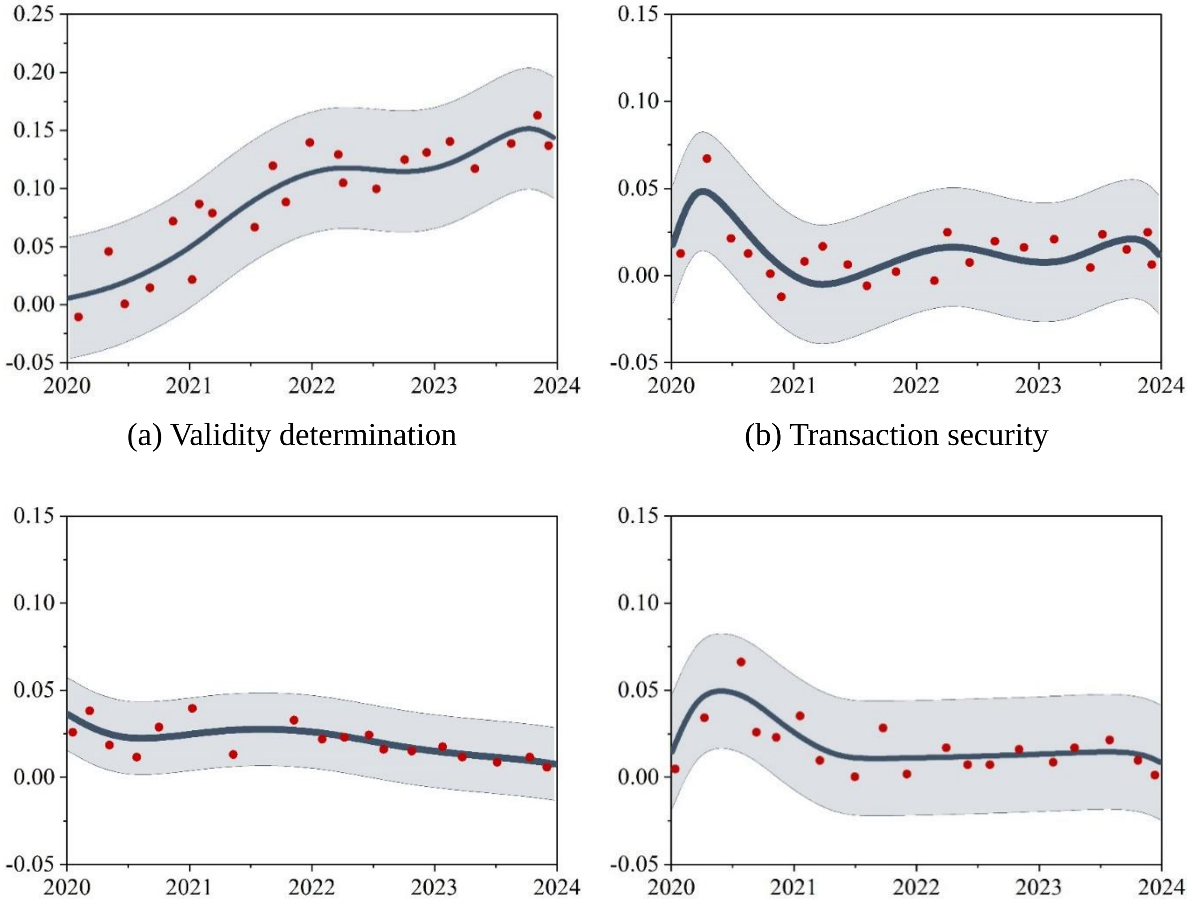
3 Analysis of results

3.1 Cluster Analysis of Smart Contract Themes

3.1.1 Dynamic Evolution Law of High Frequency Words in Smart Contracts

Using semantic analysis technology to extract shallow semantic information from the textual information of smart contracts in business transactions, and use this to carry out word division technology, based on the TF-IDF algorithm textual keyword extraction and word frequency statistics, after eliminating meaningless conjunctions, dummy words and pronouns, the statistics obtained 1259655 original keywords and frequency, totaling 9871 paragraphs. The keywords are plotted according to the descending order of the frequency ranking, and then the ranking and word frequency are fitted on the double logarithmic coordinate system. The keywords in the smart contract have the phenomenon of power law distribution, which obeys Zipf's law.

In order to deeply explore the evolution law of smart contracts over time, the first 4 main high-frequency words in TF-IDF are extracted, and the change trend of high-frequency dispute problems over time is plotted, and the dynamic evolution river diagram of the high-frequency words is shown in Fig. 5, and Figs. a~d represent the validity determination, transaction security, responsibility definition and risk prevention and control, respectively. From the figure, it can be seen that most of the problems show a slow growth from 2020, and reach the maximum value in 2024, and the risk behind the problems should not be ignored. Meanwhile, the graph also indicates that "validity determination", "transaction security", "liability definition" and "risk prevention and control" are the focuses of smart contracts. In addition, “validity determination” and “security of the transaction” have been occurring steadily over time and are trending upwards, which implies that attention should be paid to the security of financial transactions in commercial transactions and to whether or not a smart contract is recognized as a valid contract in the context of a specific commercial transaction.



(a) Validity determination

(b) Transaction security

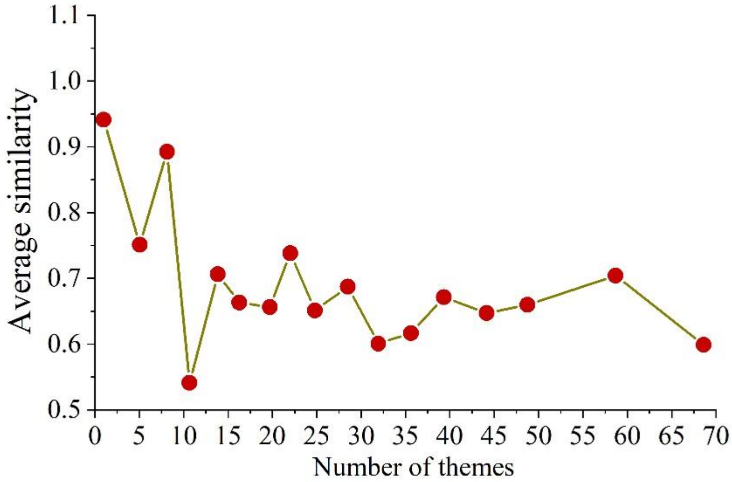
(c) Definition of responsibility

(d) Risk prevention and control

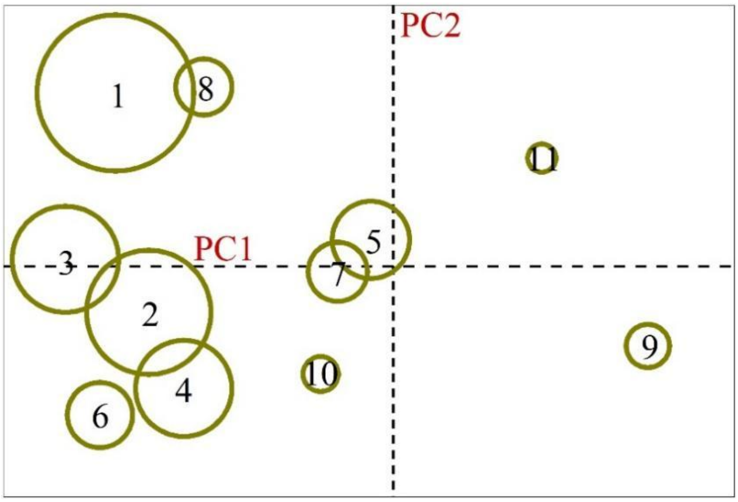
Figure 5: Dynamic evolution river diagram of high-frequency words

3.1.2 Smart Contract Theme-High Probability Feature Word Distribution

As opposed to a high-frequency word analysis based on TF-IDF, topic modeling is applied in this case to classify smart-contract data into particular topics and provide more insight into the links between various topics. This experiment produces a total of 365,955 word pairs of which 285,698 valid word pairs remain after selecting those that have co-occurrence exceeding 0.25. On the basis of the formula that is applied to compute the mean similarity among topics, the best number of topics is established. The obtained results are depicted graphically in Fig. 6. Fig. 6(a) depicts the average similarity between topics and Fig. 6(b) depicts the distance between topics. It can be seen in Fig. 6(a) that at the topic number of 11, the average-similarity curve enters the lowest valley and then the further data behaviour becomes relatively stable. Simultaneously, Fig. 6(b) shows that the topic overlap is relatively low when K=11 and the classification performance is higher. Thus, the last topic number is chosen as K=11 and the diameter of the circle denotes the strength of the topic.



(a) The average similarity between topics



(b) Distance graph between themes

Figure 6: The number of topics is visualized

Through the LDA topic model classification to get the smart contract focus content of the 11 topics and word distribution, the topic corresponding to the top 15 high-probability feature map for combing,topic-high probability feature word distribution as shown in Table 1.

Table 1: Theme - Distribution of High-probability feature words

Topic 1		Topic 2		Topic 3		Topic 4	
Legal adaptation		The termination of intelligent contracts and the design of legal intervention interfaces		The legal personality of the autonomous organization as the subject of the trading subject		Intellectual property trading and intelligent contracts	
Legislative philosophy	0.165	Emergency suspension	0.087	Limited liability	0.178	Digital copyright management	0.146
Formal equality	0.098	Multiple signature	0.082	Unlimited liability	0.161	The royalties are automatically divided	0.161
Essential function	0.093	Court command	0.093	Legal formulation	0.111	License chain	0.108
Acknowledgement of law	0.066	Agent contract model	0.019	Member responsibility	0.072	Conditional coding	0.094
Electronic signature	0.04	Legal event trigger	0.048	Asset lock	0.059	Intellectual property pledge	0.051
Topic 5		Topic 6		Topic 7		Topic 8	
The intellectual contract legal risk map in supply chain finance		Insurance technology and parameterized insurance contract		Intelligent contract as the integrity, authenticity and relevance of electronic evidence		The compliance conflict between cross-border data flows and intelligent contracts	
Accounts receivable	0.254	Trigger claim	0.137	Hash fixation	0.1	Data localization	0.092
Right of claim	0.088	Predictive machine data source	0.068	Timestamp service	0.073	Exit safety assessment	0.004
Multistage circulation	0.06	Explicit obligation	0.06	Certificate of storage platform	0.062	Personal sensitive information	0.064
Automatic cleaning	0.08	Data privacy	0.039	Raw data extraction	0.046	Default privacy design	0.06
Core enterprise credit penetration	0.055	Regulatory approval	0.093	Clean environment	0.056	Zero proof	0.039
Topic 9		Topic 10		Topic 11			
The central bank's digital currency and programmable payments		The ability of legal practitioners to transition and the rise of legal engineers		Legislative race			
Programmable currency	0.105	Cross boundary knowledge	0.101	Demonstration method	0.114		
Conditional payment	0.088	Code review	0.09	Technical authority	0.053		
Fund orientation	0.074	Logical mapping	0.069	Industry standard	0.038		
Smart Treasury	0.05	Demand translation	0.065	interoperability	0.014		
Performance automation	0.064	System architecture understanding	0.09	Cross-border judicial recognition	0.042		

3.1.3 Classification Analysis of Smart Contract Controversy Focus Topics

Based on the above smart contract theme analysis and controversy focus summary, its content can be summarized into a table and its coding is sequentially analyzed by co-occurrence network. The set of dispute focus causes for the legal effect of smart contracts is shown in Table 2. There are a total of 11 themes and 33 controversial focus items of smart contracts and the weights that each occupies. As can be seen from the table, the highest proportion of professional ethics and responsibility boundary disputes is the most likely to cause disputes, and its weight is 52.83%.

Table 2: The set of reasons for the focus of dispute

Theme category	Focus of dispute	Coding	Weight(%)
Topic 1	Principle conflict	A1	6.17
	Judicial discretion boundary	A2	0.31
	Legislative lag	A3	3.5
Topic 2	The fundamental contradiction between decentralization and centralization intervention	B1	0.56
	Risk of technical security and legal abuse	B2	0.08
	International problems	B3	7.81
Topic 3	The ultimate problem of responsibility	C1	5.26
	The service and execution of the legal procedure	C2	0.75
	The legal standards identified by members	C3	0.02
Topic 4	Conflict of code license and legal permit	D1	0.78
	Ambiguity of ownership mapping	D2	3.91
	The automation paradox of tort	D3	14.81
Topic 5	Data authenticity challenge	E1	7.74
	Traditional ticket and chain digital assets conflict	E2	6.98
	The risk conduction after the core enterprise credit "upper chain"	E3	5.93
Topic 6	Formalized dilemma of "explicit obligation"	F1	11.26
	Loss of data source failure	F2	11.72
	The transfer and activation of moral hazard	F3	18.22
Topic 7	The authenticity of the "upper chain" is a problem	G1	5.04
	Technical black box and judge cognitive gap	G2	14.22
	Anonymity and relevance	G3	2.11
Topic 8	Irreversibility and irreconcilable rights	H1	9.38
	The conflict between judicial jurisdiction and data access	H2	5.6
	The legal effect of privacy computing technology is pending	H3	14.28
Topic 9	Monetary programmable versus free fundamental limits	I1	13.05
	The legal confirmation of the final sex of intelligent contracts and payment settlement	I2	1.25
	The legal responsibility of double payment risk is transferred	I3	4.66
Topic 10	Occupational ethics and responsibility boundary	J1	52.83
	The path of law education reform	J2	6.25
	The contention of the industry	J3	0.84
Topic 11	The paradox of regulation and the supervision of the highlands	K1	8.34
	The coupling problem of technical standards and legal rules	K2	2.17
	The limits of the sovereignty	K3	1

Based on the above, the co-occurrence network coding is carried out to explore the changes of causal factors of the legal effect of smart contracts in commercial transactions based on the co-occurrence relationship between words. The co-occurrence network of dispute focus is shown in Figure 7. From the figure, it can be seen that there are co-occurrence relationships among different topics and they can be transmitted to each other. As shown in the red part of the figure, D1, D2-D3-E1-E2-E3-F1 correspond to the specific content: The conflict between code licensing and legal licensing - the ambiguity of ownership mapping - the paradox of automation in infringement - the challenge of data authenticity - the conflict between traditional negotiable instruments law and on-chain digital assets - the accelerated risk transmission after core enterprise credit is "on-chain" - the formalization predicament of the "obligation to clearly explain". This co-occurrence rule passes to show that in commercial transactions, the legal effects of smart contracts need to fully consider issues such as liability and legal risk. The dispute focus co-occurrence network relationship can clearly see that the factors are related to each other and influence each other.

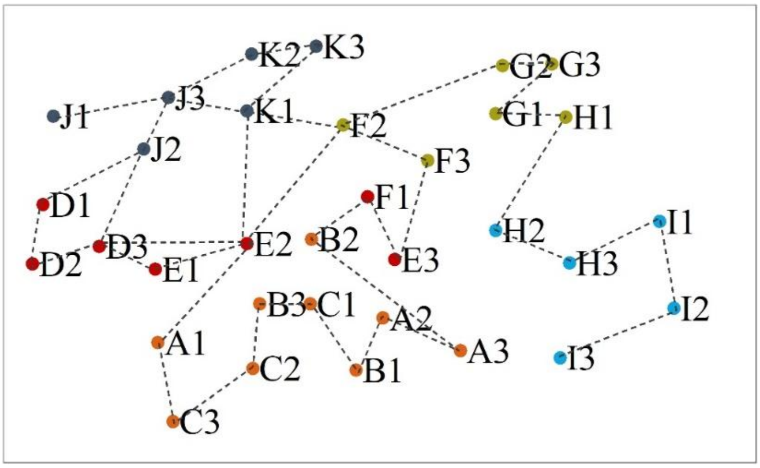


Figure 7: The controversial focus has emerged as an online woman

3.2 Strong association rule mining

3.2.1 Construction of data sets

In this paper, based on 300 smart contract data collected, we construct the association rules dataset of factors affecting the legal effect of smart contracts in commercial transactions, and the object of association analysis is 33 smart contract dispute factors through text mining, and the factor coding follows the numbering set during the categorization of smart contract influence factors. The association rule data set of smart contract dispute factors adopts Boolean structured data representation, if the smart contract impact factor is in this information, it is represented by the number 1, if the smart contract impact factor does not appear, it is represented by the number 0. Therefore, it is necessary to convert the vector space model matrix into a Boolean matrix by using the formula, and the association rule data set is shown in Table 3.

Table 3: Association rule dataset

	EF1	EF2	EF3	EF4	EF5	EF6	MF1	MF2	MF3	…	CF11
1	0	0	0	1	0	0	1	0	1	…	1
2	0	0	0	0	0	0	0	0	0	…	0
3	0	0	0	0	0	0	0	0	1	…	1
4	0	0	0	0	0	0	0	0	1	…	1
5	0	0	0	1	0	0	0	0	0	…	0
6	0	0	0	1	0	0	0	0	0	…	1
7	0	0	0	0	0	0	0	0	0	…	0
8	0	0	0	0	0	0	0	0	0	…	1
9	0	0	0	0	0	0	0	0	0	…	1
…	…	…	…	…	…	…	…	…	…	…	…
300	0	0	0	1	0	0	0	0	1	…	1

3.2.2 Minimum support and confidence settings

The co-occurrence network graph can only show whether there are co-occurrence relationships between the smart-contract dispute factors, and what level of co-occurrence has been observed. It does not decide whether the factors that co-occur are actually related in a strong manner. Consequently, it is required to continue analyzing the data using the Apriori algorithm and check the amount of strongly correlated rules (Table 4).Given the fact that the legal effects of smart contracts in commercial transactions are affected by numerous factors, certain factors might also play a role in the creation of legal effects, regardless of their frequency. That is why the factors influencing the legal effect of smart contracts are identified each time they arise, and the support threshold is defined in the following way: support = (X Y) >= 0 percent.

In case of setting the confidence level to 60 per cent the number of related elements varies slightly and it means that at the confidence level of no less than 60 percent more reliable association rules are produced on the influencing factors of the legal effect of the smart contract. The number of strongly associated items is 212 at this threshold. Consequently, the least confidence level is 60, which can be written as Confidence(X-Y)60%.

As the influence connections between the factors that determine the legal impact of smart contracts are mostly promoting ones instead of inhibiting ones, it is accepted that the relationship between the factors will be considered effective only when the lift value between them is more than 1. When the factor effect is below 1, the offsetting relationship between factors in the context of smart-contract disputes is not considered. In this way, the action threshold is defined as Lift:(X Y) >1.

Table 4: Analysis of the number of association rule items

	1%	10%	20%	30%	40%	50%	60%	70%	80%	90%
Support rate: 0%	1245	974	813	652	478	348	212	151	86	13
Support rate: 5%	1086	837	653	550	389	262	191	87	51	45
Support rate: 10%	897	667	555	454	310	220	139	86	34	34
Support rate: 20%	736	543	450	343	259	159	91	72	54	8

3.2.3 Strong association rule generation

The factor association rules are shown in Table 5. In the strong correlation rule of the factors affecting the legal effect of smart contracts in commercial transactions, there are 33 pairs of correlation factors with confidence level greater than 90%, indicating that the factors have strong influence, among which the correlations such as “A1”, “A2”, and “A3” have high confidence but low support.

Table 5: Factor association rule

Serial number	Factor	Support level	Confidence level	Degree of effectiveness.
1	A1	1.1022	100	2.77893
2	A2	1.10083	100	1.74256
3	A3	1.89389	100	1.03899
4	B1	2.44395	100	1.76551
5	B2	2.46162	100	1.01077
6	B3	2.97913	100	1.06091
7	C1	6.78011	100	1.05711
8	C2	7.06982	100	1.32609
9	C3	7.06809	100	1.01065
10	D1	13.30929	100	1.02417
11	D2	14.13517	100	1.03634
12	D3	27.44635	100	1.02191
13	E1	52.18076	100	1.00647
14	E2	35.30162	99.2401	1.0312
15	E3	69.0299	99.23652	1.00964
16	F1	33.69629	99.17358	1.02608
17	F2	30.41949	99.09679	0.99598
18	F3	27.73728	99.01031	1.01128
19	G1	23.64912	98.84662	1.04081
20	G2	40.49594	98.65324	1.02506
21	G3	38.59625	98.57873	1.02439
22	H1	18.75555	98.56958	1.01272
23	H2	50.28579	98.38015	1.02361
24	H3	15.47447	98.24287	1.03489
25	I1	76.35734	98.22961	0.99874
26	I2	40.48423	97.98369	1.2916
27	I3	13.29611	97.99002	1.93859
28	J1	25.27635	97.85122	1.01188
29	J2	21.67762	96.65941	1.13164
30	J3	39.07356	96.35757	0.98646
31	K1	40.38989	96.22541	0.88564
32	K2	21.1309	96.12455	0.93654
33	K3	44.76316	96.00248	0.96881

The statistics of the number of strong correlation rules affecting the factors are shown in Table 6. From the table, it can be seen that the relationship between the factors is not a simple one-to-one correspondence, such as “C3”, “K1” associated with a large number of rules, indicating that these factors have a strong influence in the legal effect of smart contracts.

Table 6: Statistics on the number of strongly associated rule items of influencing factors

Serial number	Factor	Rule term	Serial number	Factor	Rule term
1	A1	11	18	F3	10
2	A2	6	19	G1	10
3	A3	7	20	G2	3
4	B1	4	21	G3	5
5	B2	2	22	H1	6
6	B3	8	23	H2	8
7	C1	8	24	H3	8
8	C2	7	25	I1	11
9	C3	36	26	I2	8
10	D1	5	27	I3	7
11	D2	10	28	J1	8
12	D3	6	29	J2	9
13	E1	9	30	J3	6
14	E2	9	31	K1	34
15	E3	5	32	K2	23
16	F1	3	33	K3	3
17	F2	16			

4 Specificity of the legal effects of smart contracts

Whatever the subject, content, or implementation approach of a smart contract may be, it is safe to claim that a smart contract remains a type of contract. Despite similarities with traditional contracts, smart contracts have numerous unique properties, which also determine the application of traditional law to smart contracts.

4.1 Specificity of Smart Contract Formation and Content

Prior to exploring the particularities of the legal consequences of smart contracts, it is important to define the connection between smart contracts and electronic contracts. They are distinct in their forms of existence. At the same time, due to its peculiar form of expression, smart contracts are different compared to natural language-based traditional contracts. Computer language has a far lower ambiguity, which assists in lowering disputes between the contracting parties on the terms of the contract and enhances the performance efficiency. Nevertheless, even with this benefit, there are numerous possible risks. Like, smart contracts might not reflect the actual wishes of certain parties, therefore making smart contract interpretation harder than in conventional contracts.

4.2 Subjects of Smart Contracts and Contract Validity

One of the main attributes of contract validity is the ability of the parties to act. Traditional contracts can check the subject qualification of the parties to ascertain whether they have the contractual capacity. Nevertheless, in cyberspace, the identity of the parties to a contract cannot necessarily be established as it would be in the case of traditional contracts. Despite the fact that electronic contracts have been developed over more than ten years, it is still problematic to assess whether the identity of users accessing the network fulfills the subject-qualification conditions stipulated by Contract Law.

4.3 Recognition of Intentional Representation of Smart Contracts and Contractual Validity

Viewed through the prism of the decentralization and the de-trust nature of blockchain, the participants in a smart contract have to base their actions on information exchange between the nodes spanning the whole network to finalize the deal, and these nodes assist in ensuring the authenticity, accuracy and safety of the contract content by verifying its contents. Hence, the expression of intent confirmed by the whole network of nodes has a significant distinction with the traditional form of expressing intent in terms of contract validity.

4.4 Analysis of Smart Contract Performance Effectiveness

According to the performance mode view, the smart-contract execution is closed. After a smart contract has been validated by all the nodes in the network and recorded on the blockchain, it cannot be altered until the trigger conditions have been satisfied, at which point it will be executed automatically. It is different with electronic contracts, where, in case of modification or termination, the contract performance may be technically interrupted before the completion of the contract, e.g., in refund and return procedures related to the online shopping contract. Smart contract performance process, on the other hand, is isolated and immutable. Since execution is now coded, smart contracts cease to have the flexibility that was present in traditional contracts.

5 Conclusion

This paper analyzes and researches the legal effects of smart contracts in business transactions using data mining techniques. The main research results and conclusions of this paper are as follows:

(1) Through the exploration of the dynamic evolution laws of high-frequency words in smart contracts, it is found that "validity determination", "transaction security", "responsibility definition" and "risk prevention and control" are the focuses of smart contracts.

(2) Through the study of the focal causes of the legal effect of smart contracts. It is found that there are a total of 33 controversial focus items of 11 themes in smart contracts, among which the controversy of professional ethics and responsibility boundary accounts for the highest proportion, and its weight is 52.83%. This proves the importance of professional ethics and responsibility boundary in the legal effect of smart contracts.

(3) The legal effect of smart contract has special characteristics, which is reflected in the special characteristics of establishment and content, the special characteristics of subject and contractual effect, the special characteristics of the determination of meaning and contractual effect, and the special characteristics of performance effect.

About the Author

Qianyu Chen was born in Henan, China in 2005. She is currently studying at the Law School of Central South University of Forestry and Technology. The main research directions are civil and commercial law and jurisprudence.

References

[1] Möhring, M. M., & Finch, J. (2015). Contracts, relationships and innovation in business-to-business exchanges. Journal of Business & Industrial Marketing, 30(3/4), 405-413.

[2] Shiffrin, S. V. (2023). Hidden delegations: The assignment of contractual rights and consumer debt. The Modern Law Review, 86(1), 1-30.

[3] Sianipar, J. R., & Prasetyawati, E. (2024). Legal Protection for Debtors When There Is a Transfer Of Debt By Creditors Without The Debtor's Consent. International Journal of Social Sciences and Humanities, 2(2), 44-48.

[4] Grâu-Panţureac, M. (2022). The importance of contracts in business relationship. In Competitivitatea şi inovarea în economia cunoaşterii (pp. 471-476).

[5] Karati, M. (2023). Law of Contracts: Their presence in corporate transactions learning. European Journal of Economics, 7(2).

[6] Mik, E. (2017). Smart contracts: terminology, technical limitations and real world complexity. Law, innovation and technology, 9(2), 269-300.

[7] Meitinger, T. H. (2017). Smart contracts. Informatik-Spektrum, 40(4), 371-375.

[8] Liu, Y., He, J., Li, X., Chen, J., Liu, X., Peng, S., ... & Wang, Y. (2024). An overview of blockchain smart contract execution mechanism. Journal of Industrial Information Integration, 41, 100674.

[9] John, K., Kogan, L., & Saleh, F. (2023). Smart contracts and decentralized finance. Annual Review of Financial Economics, 15(1), 523-542.

[10] Zhang, Y., Kasahara, S., Shen, Y., Jiang, X., & Wan, J. (2018). Smart contract-based access control for the internet of things. IEEE Internet of Things Journal, 6(2), 1594-1605.

[11] Muneeb, M., Raza, Z., Haq, I. U., & Shafiq, O. (2021). Smartcon: A blockchain-based framework for smart contracts and transaction management. IEEE Access, 10, 23687-23699.

[12] Chikov, I. A., Koliadenko, S. V., Supryhan, V. A., Tabenska, O. I., Nitsenko, V. S., & Holinko, O. V. (2023). Smart contracts and business process automation: the technical aspect. Naukovyi Visnyk Natsionalnoho Hirnychoho Universytetu. 2023.№ 5. P. 186-192. DOI: 10.33271/nvngu/2023-5/186.

[13] McKinney, S. A., Landy, R., & Wilka, R. (2017). Smart contracts, blockchain, and the next frontier of transactional law. Wash. JL Tech. & Arts, 13, 313.

[14] Khan, S. N., Loukil, F., Ghedira-Guegan, C., Benkhelifa, E., & Bani-Hani, A. (2021). Blockchain smart contracts: Applications, challenges, and future trends. Peer-to-peer Networking and Applications, 14(5), 2901-2925.

[15] Taherdoost, H. (2023). Smart contracts in blockchain technology: A critical review. Information, 14(2), 117.

[16] Prause, G. (2019). Smart contracts for smart supply chains. IFAC-PapersOnLine, 52(13), 2501-2506.

[17] Dal Mas, F., Dicuonzo, G., Massaro, M., & Dell'Atti, V. (2020). Smart contracts to enable sustainable business models. A case study. Management Decision, 58(8), 1601-1619.

[18] Nzuva, S. M. (2019). Smart contracts implementation, applications, benefits, and limitations. Journal of Information Engineering and Applications.

[19] Allam, Z. (2018). On smart contracts and organisational performance: A review of smart contracts through the blockchain technology. Review of Economic and Business Studies (REBS), (22), 137-156.

[20] Deeva, T. V., Nikiporets-Takigawa, G., Lustina, T. N., Podsevalova, E. N., & Didenko, E. N. (2020). Blockchain technologies and smart contracts: New technological methods to regulate transactions and trade operations. Int. J, 8, 3659-3664.

[21] Xu, Y., Chong, H. Y., & Chi, M. (2021). A review of smart contracts applications in various industries: a procurement perspective. Advances in Civil Engineering, 2021(1), 5530755.

[22] Borgogno, O. (2019). Usefulness and dangers of smart contracts in consumer and commercial transactions. A modified version is forthcoming in “Smart Contracts and Blockchain Technology: Role of Contract Law”, L. DiMatteo, M. Cannarsa & C. Poncibo eds (Cambridge University Press, 2019).

[23] Sanz Bayón, P. (2019). Key legal issues surrounding smart contract applications. KLRI Journal of Law and Legislation, 9(1), 63-91.

[24] Sklaroff, J. M. (2017). Smart contracts and the cost of inflexibility. U. Pa. L. Rev., 166, 263.

[25] Cannarsa, M. (2018). Interpretation of contracts and smart contracts: smart interpretation or interpretation of smart contracts?. European review of private law, 26(6).

[26] Papantoniou, A. A. (2020). Smart contracts in the new era of contract law. Digital LJ, 1, 8.

[27] Giancaspro, M. (2017). Is a ‘smart contract’really a smart idea? Insights from a legal perspective. Computer law & security review, 33(6), 825-835.

[28] Liu, Y., & Huang, J. (2019, November). Legal creation of smart contracts and the legal effects. In Journal of Physics: Conference Series (Vol. 1345, No. 4, p. 042033). IOP Publishing.

[29] Li, X. (2020, August). Legal effect of smart contracts based on blockchain. In International Conference of Pioneering Computer Scientists, Engineers and Educators (pp. 166-186). Singapore: Springer Singapore.

[30] De Caria, R. (2018). The legal meaning of smart contracts. European Review of Private Law, 26(6).

[31] Kirillova, E. A., Bogdan, V. V., Lagutin, I. B., & Gorevoy, E. D. (2019). Legal status of smart contracts: features, role, significance. Juridicas Cuc, 15(1), 285-300.

[32] Savelyev, A. (2017). Contract law 2.0:‘Smart’contracts as the beginning of the end of classic contract law. Information & communications technology law, 26(2), 116-134.