Journal of
Risk and Financial Management

MDPI

Article

An Entrepreneurial Definition of the Blockchain Technology and a Stacked Layer Model of the ICO Marketplace Using the Text Mining Approach

Ahmed Gomaa * and Yibai Li

Operations and Analytics Department, Kania School of Management, The University of Scranton, Scranton, PA 18510, USA
* Correspondence: ahmed.gomaa@scranton.edu

Abstract: The landscape of ICOs and its underlying Blockchain technology needs more clarity, given that several overlapping and opposing views exist from governmental institutions, institutional investors, economists, and academia. Those positions stem from confusion, bias, and vested interest. Having consensus from the pioneer entrepreneurs who define Blockchain technology usage, and its marketplace address this need. Furthermore, an agreement on the problems blockchain is solving from the industry perspective would further the understanding of the technology direction and its "raison d'être." or "reason of existence". The paper analyzes 4367 businesses that requested funding using ICO whitepapers and raising more than $20 billion US dollars during the most active ICO period. Using Latent Semantic Analyses (LSA), the paper identifies a one-factor solution that explains 98.15% of all ICOs. The paper conducts a second-order analysis that generates an 18-factor solution. Through the empirical analysis, the paper presents its findings as an ICO marketplace stacked layer model. The model is comprised of four layers: (1) Trust; (2) Value exchange; (3) Automation; and (4) Applications to enable value exchange, and an era of new business models. The paper then presents an unbiased, unified entrepreneurial definition of the Blockchain technology usage.

Keywords: initial coin offering (ICO); cryptocurrencies; peer-to-peer networks (P2P); blockchain; natural language processing (NLP); latent semantic analyses (LSA); entrepreneurial finance; value exchange

Citation: Gomaa, Ahmed, and Yibai Li. 2022. An Entrepreneurial Definition of the Blockchain Technology and a Stacked Layer Model of the ICO Marketplace Using the Text Mining Approach. *Journal of Risk and Financial Management* 15: 557. https://doi.org/10.3390/jrfm15120557

Academic Editors: Robert Brooks and Shigeyuki Hamori

Received: 1 July 2022
Accepted: 23 November 2022
Published: 28 November 2022

1. Introduction

The Initial Coin Offering (ICO) market is a new form of raising funds (Winotoatmojo et al. 2022). In a short period, ICOs raised tens of billions of dollars, and investors started to look seriously into this funding mechanism (Fisch 2019). Businesses that change their standard supply chain model by updating their intermediaries' using cryptocurrencies can raise funds using ICO (Filipov 2018). The underlying technology that supports ICOs is Blockchain. Blockchain allows anonymous entities to trust each other by using a consensus algorithm, thus supporting new business models (Gomaa 2018). The significant impact of the ICO in a short period brought the attention of the regulating agencies. The Internal Revenue Service (IRS) defined cryptocurrencies as a property for taxation purposes (Prentis 2015). In contrast with the IRS, the Security Exchange Commission (SEC) considers cryptocurrencies as either currencies or securities, according to the ICO specific conditions, and existing regulated fund-raising mechanisms. The confusion of being mislabeled as a security instead of a currency in the SEC eyes made the majority of the ICOs ban US investors in an attempt to become foreign offerings, thus not within the SEC reach (Lockaby 2018). The contrasts are event more on a global level (Karpenko et al. 2021).

During the early years of the rise and decline of ICO from 2016–2018, with 2017 as the highest year for ICO capital raising (Florysiak and Schandlbauer 2022), there was no clear landscape or clarity on the future usage of the ICO technology, making it like the

beginning of the "world wild web" (www), which was called the "world wild west" in reference to the ambiguity of its landscape while recognizing it as a disruptive technology. In addition to the confusing definitions from the different regulatory bodies, to date, not everyone shares the same views about ICOs and its underlying blockchain technology (De Andrés et al. 2022). Institutional investors, economist, and researchers have different views among themselves and each other when it comes to ICOs and cryptocurrencies (Zetzsche et al. 2017; Halaburda 2018; Zheng et al. 2018). The views vary from believing in the technology to considering the technology a total hoax, where hype and the fear of missing out are the driving forces for the widespread of the technology, especially that many ICOs failed in achieving their potential. The reason for the chaotic, overlapping positions across the spectrum is that the technology was still evolving, immature, and with considerable uncertainty (Cunningham et al. 2019; Varma 2019). Another reason for the unclear landscape is the misconception of the blockchain technology and its different layers (Halaburda 2018). Adding to the confusion is the own technology limitations, including potential subjective information fraud, and attacks like Sybil attack and camouflage attack, and the possible network fraud in 2017 to name a few (Zhu and Cai 2016; Swartz 2022). Other challenges exist, including electricity usage and speed (Zheng et al. 2018). In 2019, the financial time reported on the Germany central bank experience with blockchain that it is slower and more expensive (Kaminska 2019).

To address the chaos and lack of clarity around the ICOs landscape and blockchain from different constituents perspective numerous studies attempted to evaluate the ICOs success factors (Ahmad et al. 2021; Cai and Gomaa 2019) and post ICO performances (Lansky 2019; Jongsub et al. 2022), it is essential to recognize who determines the landscape of any technology. Once millions of people use technology, it no longer belongs to the inventor; instead, it belongs to the masses. It is the entrepreneurs, who collectively defines the ICO technology landscape. This paper aims to establish the technology landscape and its application from the entrepreneur's early adopter's perspective. After identifying the landscape, the paper generates a market definition of Blockchain. This would help in enhancing the understanding of how the industry perceives the contribution of blockchain to its current potentials. The paper uses the entire population of the 4367 ICOs from 2016 to 2018, including both successful and unsuccessful ICOs to extract the entrepreneurs understanding of the blockchain and the ICO marketplace. It provides a comprehensive view of how the market understands ICO and its underlying Blockchain technology. Compared to the many definitions already available in the literature (Nakamoto 2008; Morrison and Sinha 2016; Viriyasitavat and Hoonsoponb 2019), the paper is taking a bottom-up approach through a text mining approach named Latent Semantic Analysis (LSA). The paper generates a stacked layer ICO market model and definition of how the market—in its entirety—understands blockchain. This definition shed an unbiased, unified light on the technology as it enables a better understanding of the cryptocurrencies' potentials from the market own lens. Besides, the paper identifies how entrepreneurs understands the solutions provided using blockchain. The paper depicts the 18 most prominent themes related to the blockchain marketplace. Like the Transmission Control Protocol/Internet Protocol (TCP/IP) defining the technologies and rules for information exchange within networks (Clark 1988), The paper is the first to empirically identifies the four layers that define the technologies and rules to exchange value between networks. The paper reveals. a comprehensive view of how the market defines blockchain, its layers, and its mainstream applications.

The study contributes to the literature by introducing an ICO Marketplace stacked layer model, and a market generated blockchain definition, from the entire market understanding of blockchain according to the published ICOs descriptions using a bottom-up approach. The findings have important implications to both researchers and practitioners, as it explains how the entrepreneurs understand Blockchain technology and its potentials, while considering the different perspectives and speculation documented in the literature. It identifies the industry most essential stacked layers used to shape its understanding of

blockchain. The paper second section discusses the literature. The third section discusses the research objectives, data sources, and research framework. The fourth section interprets the generated tables and presents the resulting ICO stacked layer model, and the blockchain market definition. The fifth section concludes the paper.

2. Literature Review

The paper literature review highlights the history of the blockchain technology, and its origins, the different types of cryptocurrencies within the ICO landscape, and the current attempts from academia to employ data mining techniques, as a methodology to address challenges within the ICO and blockchain space.

2.1. Blockchain History

It is necessary to return to the origins of blockchain to understand how it evolved. Like most inventions, usually, an ideal is the driving force behind them. One may link the emergence of cryptocurrencies with the movement "free software," initiated in the 1980s, as well as with the "cypherpunk" community, that used encryption technologies to create a digital currency and guarantee anonymous transactions. As shown in Figure 1, first attempts took place in the 1990s, including how time stamp digital documents (Haber and Stornetta 1990). In addition to a few digital currencies attempt such as "digicash," "b-money," "bitgold," and "hashcash" to bring a decentralized digital currency and to lower spa (Back 2002). Those technologies enabled the creation of a framework that allows trusting third parties (Curran and Honan 2006). At that time, existing technologies were not mature enough to propose a functional digital currency. In specific, the digital currency domain, at that time, could not address double spending or other fault tolerance problems, which fiat currencies and current banking systems address them (Cao et al. 2005). This made centralized systems, with intermediaries the only practical path to generate a reliable digital transaction. Bitcoin provides the answer to the double spending problem. It describes the operation of a tamper-proof protocol using a peer-to-peer network—the blockchain—as a layer of Bitcoi (Nakamoto 2008). Blockchain is the first full spread implementation of public key infrastructure. The Public Key Infrastructure (PKI) is an ecosystem that enables trust online using digital signature (Ellison and Schneier 2000). It allows for transactions accountability and trust. Using blockchain infrastructure, any two parties, in a peer-to-peer network, can conduct a transaction without an intermediary that knows both parties. In a blockchain, a consensus mechanism between the blockchain participants is required. In November 2013, Vitalik Buterin introduced Ethereum, a platform that allows the automation of business logic, in a transparent manner, which marked the beginning of smart contracts to allow ICO to take plac (Buterin 2014). Currently Non-Fungible Tokens, followed by Security Token Offering (STO) and Initial Exchange Offerings (IEOs) are the market current state (Ante 2022; Takahashi 2020). A few entrepreneurs and researchers have hypothesis that blockchain creates and exchange value, instead of only exchanging information (Qiang and Lin 2019; Pazaitis et al. 2017; Zutshi et al. 2021; Yousif 2022). Our finding supports their hypothesis and provide further evidence that entrepreneurs where clear that Blockchain and the ICO marketplace create a new era of the internet, with value exchange instead of information exchange is the new direction.

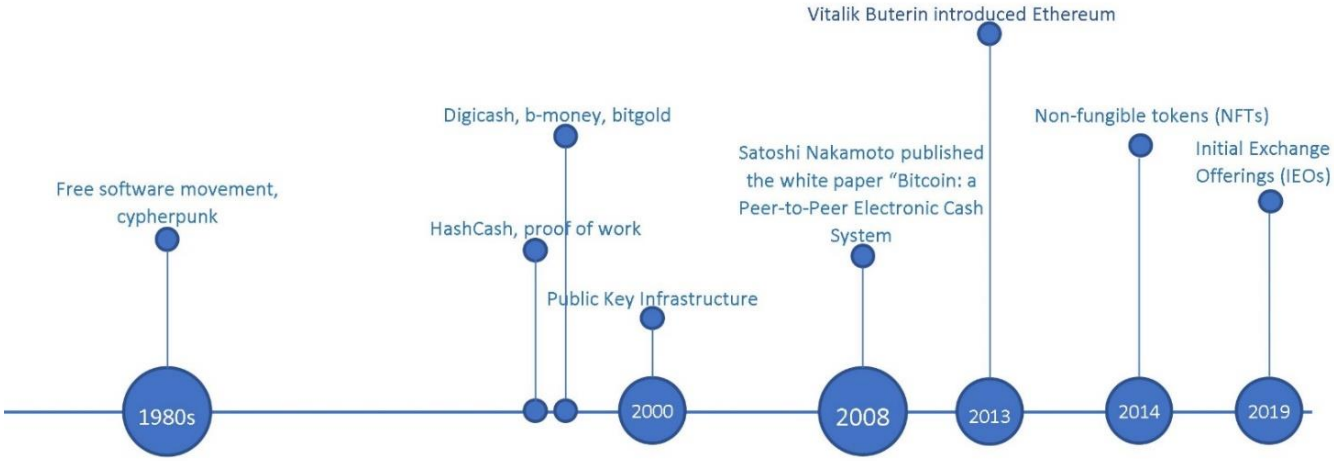

Figure 1. The Important Points on the Timeline of the Blockchain Evolvement.

2.2. Types of Cryptocurrencies

Cryptocurrencies are the only type of currencies with the following three features: ensuring pseudo anonymity, independence from central authority and double spending attack protection (Lansky 2018). There are different types of cryptocurrencies or tokens based on their usage. A currency token is like Bitcoin. Those are digital currencies designed to store and exchange values (Howell et al. 2018). An asset-backed token is a token that has a real-world representation (Cong and He 2019). A security token is a token that is subject to the Securities and Exchange Commission (SEC) (Senderowicz et al. 2018). Security tokens are like company stocks. A company can sell a utility token for allowing access to services they are developing (Lemchuk 2016). ICOs have different attributes; including the number of countries, they restrict dealing with, the amount of exposure on social media sites and the ICO executive team to name a few (Catalini and Gans 2018). Once the company finalizes their whitepaper, which explains the opportunity using blockchain technology and how it plans to use the funds raised, it usually invites whitelisted investors to join a pre-sale of the ICO on several phases, with incremental price points as an incentive to early bird investors (Fisch 2019). As shown in Figure 2, the ICO event is the following step. The companies usually implement their vision on a blockchain platform in phases as published on their whitepapers. Investors use the coins on cryptocurrencies exchanges if they wish to exit their investment. The market started to seeing growth in 2016. 2017 was the largest year that reflected the largest raised capital. A decline was visible by the third quarter of 2018. Since this paper focuses on how the early adopters understood Blockchain potentials and presented their ideas to investors directly, the paper uses information from 4367 identified ICOs during the study period. Figure 1 presents an overview of the marketplace From July 2016 to September 2018. From 4367, companies showing interest in ICO, 3472 successfully launched an ICO. Investors decided not to invest in 2135 companies. The remaining 1337 companies raised $20.8 Billion. 50% of those companies were trading on cryptocurrencies exchanges during 2018 (ICOBench.com 2022).

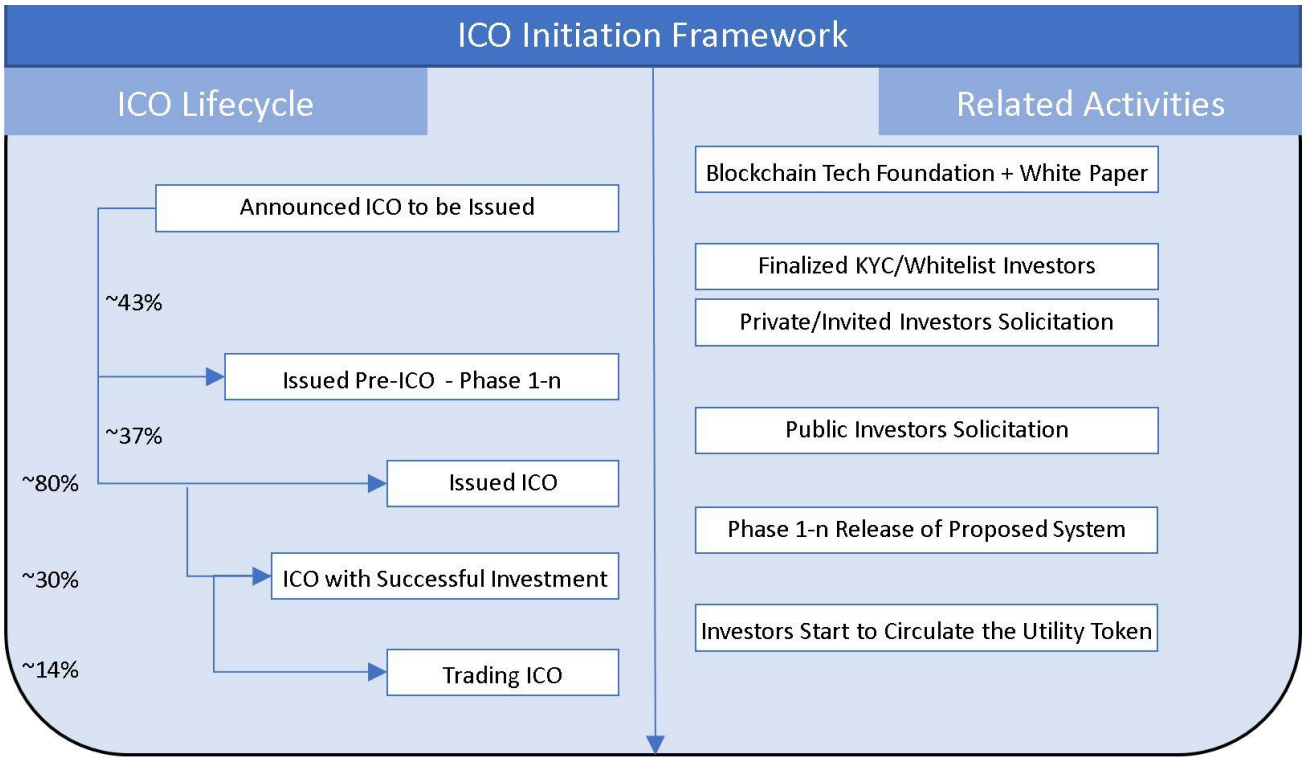

Figure 2. ICO overview of the marketplace. Source: From the verified dataset retrieved from ICOBench.com.

2.3. Using Data Mining with Blockchain

In the past few years, several researchers used mining to gain insights on the cryptocurrencies space. (Chitsazan et al. 2022) identified 78 empirical studies using data mining on ICO related data. Stenqvist and Lonno achieved a 79% accuracy in predicting the Bitcoin price by analyzing sentiment from twitter feeds (Stenqvist and Lönnö 2017). The researchers in (Yu et al. 2019) analyze the relationship between return volatility and user opinion difference and user interest. Instead of performing sentiment analysis on all social media content posted, Laskowski and Kim, use a prototype to analyze the Twitter feeds related to blockchain and could draw general insights about the increased demand in the market (Laskowski and Kim 2016). Mai et al. recognized that posts have a more significant impact than tweets on the Bitcoin price (Mai et al. 2018). In (Bian et al. 2018) they used latent Dirichlet allocation (LDA) model and manually created ten (10) labels for topic categorization. the author's focus is on identifying fraudulent ICOs. The researchers in (Cai and Gomaa 2019) analyze the ICO market and provide evidence that investor sentiment and investor awareness are determinants of the amounts raised in an ICO. Using data mining, the researchers in (Sun et al. 2019) achieved an 80% accuracy in de-anonymizing the Bitcoin Blockchain users by using Supervised Machine Learning to predict the type of yet-unidentified entities. (Chuanjie et al. 2019) also use LDA model, to categorize ICO whitepapers via topic modeling. In (Zhang et al. 2019), they compared the whitepapers readability and return. Sapkota and Grobys identify 37 predictor variables related to ICO whitepapers to determine the rational of ICO investors (Sapkota and Grobys 2021). (Thewissen et al. 2022) team suggests that the ICO white paper content are useful performance indicators. They produce semantically meaningful set of topics, with a high degree of similarity to this paper, despite the different sample size, and period covered in contrast with this paper.

The current literature shows the value of using text mining on ICO whitepapers. It does not provide a consensus of how the market understands the ICO marketplace, and the market definition of blockchain technology usage.

3. Research Objective

Given the different dimensions related to ICOs and the blockchain technology, we need a more comprehensive definition of how the market views blockchain. Specifically, we aim to answer two research questions:

RQ1: What are the main layers and core components of the ICO marketplace from the unified and unbiased entrepreneurs' perspective?

RQ2: What is the consensus of the industry on the definition of blockchain and its usage from the entrepreneur's perspective?

In contrast with the current literature, this paper provides a comprehensive evaluation of the ICO most relevant keywords representing the market. After preparing the complete market data for text mining, the paper uses Latent Semantic Analysis (LSA) to identify factors that define the blockchain and the ICO eco-system from an empirical perspective. This creates a new dimension to the literature of a comprehensive analysis of how the market defines the ICO marketplace layers and the blockchain technology using the collective wisdom of early ICO entrepreneurs

3.1. Data Sources, Collection, and Cleansing

To answer the identified research questions, we collect data from the issued ICO whitepapers to invite investors to purchase their coins. The paper collects ICO descriptions from 4367 companies that raised more than 20 billion dollars from the period of July 2016 to September 2018.

The paper first identifies five aggregators, namely, foundico.com, coinmarketcap.com, icomarks.com, icowatchlist.com, and icobench.com. From the empirical analysis, ICOBench.com was the most complete aggregator of issued ICOs and is used in this study. Data cleansing after validation takes place to ensure a consistent data format. For example, some data points representing the money raised is not numerical and is in Ethereum. Other data points are not correctly formatted as an integer and had to be manually corrected. In those instances, the authors validated the entries manually from the Ethereum blockchain directly. From the clean data set, a total of 4367 companies shows interests in ICOs, 3472 successfully launch an ICO. Investors decide not to invest in 2135 companies. The remaining 1337 companies raised $20.8 billion, and almost 50% of those companies are trading on cryptocurrency exchanges. The paper does not cover ICOs from November 2013 to June 2016; it covers ICOs from July 2016 to August 2018. According to (Kostovetsky and Benedetti 2018), there are 96 issued ICOs from 2013 to 2016. The paper focuses solely on the textual description provided by the companies issuing the ICO, regardless of their success in their ICO or the post ICO performance. The paper intentionally did not use any other financial data or contextual data attached to the ICO. The paper evaluates how entrepreneurs understands the blockchain value proposition.

3.2. Data Analysis

To achieve the research objectives of this study, the paper focuses on the textual data of the data set, especially the description of each ICO. To attract more investors, ICO companies must try their best to compile a concise description, usually in only a couple of sentences, which perfectly summarizes the core selling point of their ICOs. This makes the description of an ICO an ideal source for text mining because, with limited space, ICO companies are overly cautious in word selection. From the text mining perspective, it means this data source has very little noise and irrelevant information. Therefore, the description of ICOs become the input of the text mining analysis. Text mining represents a collection of techniques derived from areas such as natural language processing (NLP), information retrieval, information extraction, machine learning, and statistics. It refers to "the machine supported analysis of the text" (Hotho et al. 2005).

In recent years, text mining has been increasingly used for knowledge discovery from the text (Sidorova et al. 2008; Lin et al. 2017; Jensen et al. 2006). The advantage of text

mining is that it allows researchers to analyze a large amount of textual data quantitatively with little researchers' bias.

The specific text mining technique adopted in this study is a well-known technique named Latent Semantic Analysis (Han et al. 2011). Latent semantic analysis is an algebraic-statistical method that can detect the underlying topical structure of a document corpus and extract the hidden semantic structures of words and sentences (Evangelopoulos 2011). Latent Semantic Analysis was first proposed by a group of computer scientists at Bell Communication Research, University of Chicago, and the University of Western in 1988 (Dumais et al. 1988). This technique is also called latent semantic indexing (LSI). In computer science or library science, this method is often referred to as LSI. In social or behavioral science, the method is often referred to as LSA.

This method is initially developed for retrieving information from text. It is proposed to overcome the limitations of the conventional vector retrieval method or vector space model (Lochbaum and Streeter 1989). Different from the vector space model in which text documents are mapped in a multi-dimensional vector space constructed literally from the vocabulary of the text documents, LSA simulates the way the human brain distills meaning from text (Sidorova et al. 2008). It is because multiple words may share the same meaning, and one word may mean different things in different contexts. This approach aims to overcome the problem of variability in human word choices in the vector space model. It is one of the most widely accepted topic modeling methods (Crain et al. 2012).

This study follows the well-established procedures of latent semantic analysis demonstrated in prior studies (Sidorova et al. 2008; Evangelopoulos et al. 2012). A total of 4367 ICO descriptions were consolidated and loaded to RapidMiner 9.0. RapidMiner, formerly known as YALE, is an open-source software tool released under AGPL. It is one of the most widely accepted tools for machine learning, data mining, text mining, and predictive analytics (Altalhi et al. 2017). The following paragraphs describe the details and configurations for each step of pre-processing, term reduction, term frequency matrix transformation, and singular value decomposition.

Like other text mining techniques, latent semantic analysis requires some preprocessing of the textual data. We followed the well-established preprocessing and term reduction procedures (Evangelopoulos et al. 2012). All the letters in ICO descriptions were transformed into lower case. Each description was tokenized with non-letter separators. The stopwords such as "and," "the," "is," "a," "an" and so on were removed. All the tokens that are less than two letters (i.e., "s," "x," and so on) were removed. The tokens that appeared only in one ICO description were removed because these tokens are associated only with the specific ICO. However, the objective of this step is to identify the common themes among all the ICOs in our dataset. Term stemming was applied. The words that share the same root were considered one token. For example, "collaborate," "collaborating," "collaboration," and "collaborative" will be regarded as a single token, the "collabor." Like most other data mining and text mining methods, latent semantic analysis is also subject to the "curse of dimensionality" (Keogh and Mueen 2011). The basic rationale behind these procedures was to reduce the dimensionality of the matrix to get better results, and to reduce the noise and computation complexity in later steps of LSA.

Then, each ICO description was represented in a vector space model (Salton et al. 1975). This process is also referred to as the term frequency matrix transformation (Sidorova et al. 2008). Based on the terms or the tokens derived in the previous steps, the paper computes this matrix. Instead of using the absolute term frequency matrix directly, the frequency is transformed using TF-IDF (term frequency-inverse document frequency) weighting method (Han et al. 2011) Term frequency ($tf_{i,j}$) is calculated as $tf_{i,j} = \frac{n_{i,j}}{n_j}$, where n_i is the number occurrence of token i in document j, and $n_{i,j}$ is the total number of tokens in an ICO description j. The inverse document frequency is calculated as $idf_i = log\left(\frac{N}{df_i}\right)$, where N is the number of ICO descriptions in the database, and df_i is the frequency of ICO descriptions that have the token i. Finally, the weight of each token in each ICO description is calculated as $w_{i,j}$, where $w_{i,j} = tf_{i,j} \times idf_i$. This weighting method is widely adopted in text mining

studies and LSA studies (Sidorova et al. 2008) The matrix then went through singular value decomposition (SVD). SVD is a dimension reduction algorithm. SVD decomposes the TF-IDF weighted matrix, denoted matrix A into the production of three matrices—an orthogonal matrix U, a diagonal matrix S, and the transpose of an orthogonal matrix V. That is, $A_{mn} = U_{mm}S_{mn}V_{nn}^T$ where $U^TU = I$ and $V^TV = I$. The columns of V are orthonormal eigenvectors of A^TA; The columns of U are orthonormal eigenvectors of AA^T, and S is a diagonal $m \times n$ matrix containing the square roots of eigenvalues from U or V in descending order. The computation of singular value decomposition is discussed in the prior literature (Golub and Reinsch 1970).

The next step is factor reduction. LSA can provide a different level of abstraction by reducing the number of factors. We can restrict the matrix S to relatively higher eigenvalues and get a simplified matrix $\hat{S}_{kk}$ by deleting rows and columns from S. In order for the matrix multiplication to go through, we have to eliminate the corresponding row vectors of U and corresponding column vectors of V^T and get another two matrices $\hat{U}$ and $\hat{V}^T$. The result is as follows:

$$\hat{A}_{mn} = \hat{U}_{mk}\hat{S}_{kk}\hat{V}_{kn}^T$$

where $k < \min(m, n)$, $\hat{A}$ is an approximation the matrix A. That is why SVD is also a popular data compression method. $\hat{A}$ is a compression of the matrix A. Choosing different k value gives a different compression ratio.

This study identified a one-factor solution to be the optimal solution at the highest abstract level. This factor is denoted as F_0. With only one factor, F_0 is able to explain 98.15% of all the ICO descriptions and 91.35% of all the terms mentioned in these descriptions. Further, we conducted a second-order analysis for F_0, which resulted in an 18-factor solution.

After the determination of k, it was the factor interpretation and labelling. The meaning of a factor is collectively defined by the terms and ICO descriptions that load to the factor. Following the guidance of prior studies (Tirunillai and Tellis 2014), the paper focuses on the unique terms and ICO descriptions for each factor in the interpretation and labeling process. The labels for the 18 factors along with their top loading terms, top loading documents that were used for interpretation and labeling are included in the Appendix A.

4. Results and Discussion

The LSA generated one first-order factor, the F_0, and eighteen second-order factors, F_1 to F_{18}. These factors were not derived from a small sample through some types of a sampling method. Rather they were based on the entire population of the 4367 ICOs in the ICO marketplace. Therefore, these factors provide a very comprehensive view of the landscape of the ICO marketplace. Given that each ICO description can be considered as a definition for each ICO's underlying blockchain technology, F_0, which explains 98.15% of the population, can be considered a good definition for the blockchain technology collectively defined by the entire marketplace. If someone would like to understand what is going on in the ICO marketplace, she or he can look at the eighteen second-order factors, which depict the major structure of the landscape of ICO marketplace and form the basis of the ICO marketplace stacked layer model. The model is comprised of four layers: (1) Trust; (2) Value exchange; (3) Automation; and (4) Applications. where each layer host different factors.

4.1. ICO Marketplace Stacked Layers Model

Running the LSA on the entire market identified 18 high loading factors, in addition to Factor 0, which provides the market definition of the Blockchain technology, as shown in Table 1.

Table 1. F_0 The Market Definition of the Blockchain Technology.

Terms	Loadings	Top Loading ICO Descriptions
peer	0.882	
advertis	0.096	
bet	0.094	• ZPER is a "decentralized P2P (peer-to-peer) financial ecosystem" using smart contract technology.
economi	0.086	• AirSwap is a decentralized, peer-to-peer token trading network built on the Ethereum blockchain.
estat	0.071	• The Pareto Network is the first peer to peer financial content marketplace.
content	0.067	• Stater is creating three mobile applications and two market platforms promoting P2B (peer to business) and P2P (peer to peer) transactions.
bank	0.066	• Cryptocurrency betting. On-chain & peer-to-peer.
world's	0.061	• A Decentralized peer to peer energy exchange platform.
internet	0.061	• Cryptolancers platform is a decentralized peer-to-peer freelancing platform, with its tokens for online payment services.
freelanc	0.06	• ODYSSEY®'s mission is to build the next-generation decentralized sharing economy & Peer to Peer Ecosystem.
bui	0.058	• "BitherCash is a highly secure peer to peer decentralized cryptocurrency."
decentralis	0.057	• The future is here. Wanna Cab? Redcab offers a decentralized Peer2Peer transportation solution on the Blockchain.
monei	0.056	
integr	0.054	
sell	0.054	
investor	0.053	
energi	0.053	
fulli	0.052	
consum	0.05	
work	0.05	

To get a more intuitive understanding of the overall LSA result, we conducted some post hoc analysis by visualizing the data. Figure 3 presents an overview of the results of the LSA analysis. Detailed tables are presented in the supplementary data. The paper converts each factor label and high loading term into a network of nodes and edges, where each factor label is connected to its related terms with an edge. The result is a graph, with the notation G = (N, E). G is a set of nodes, N or terms and labels and set of edges, E that reflects the connection between the nodes. G can be represented as an adjacency matrix A such that $A_{ij} = I((i, j) \in E)$. If A_{ij} does not include a 0, then there is an edge between the nodes and J. If more than one edge exists between the same vertices that will reflect the edge weight. The Yifan Hu algorithm is applied (Hu 2005). Then, the modularity to restructure the network is enforced. The results appear in Figure 3. In general, P2P, and PKI are leveraged in innovative ways to create value in different domains. The business logic is embedded in smart contracts and executed programmatically. We manually grouped the different industry sectors as Application Program Interface (API). We labeled the types of tokens such as ERC-20 and supportive consensus models as (Snart Contracts). We also labeled the terms referring to transactions as (Monetary Solutions).

2 out of 18 Factors, namely F_1 and F_{16}, high loading factors identifies P2P as a central component of the Blockchain. Looking at F_5, and F_{17} identify that a dominant component of the Blockchain technology is addressing the Trust, Transparency, and Fairness (F_{17}), especially when it comes to content creation (F_5). Those are due to the public Key Infrastructure. This creates the foundation layer of the ICO marketplace model, which can be labeled as "Trust Layer" where secure, and distributed trust technologies exist.

The factors that are addressing the digital economy dimension of the ICO are F_7, F_8, F_9, F_{10}, F_{13}. The factors are monetary solutions. They cover E-commerce (F_7), Digital economy (F_8), Digital Wallets (F_9), investment (F_{10}), and Alternatives to Fiat money (F_{13}). This creates the second layer of the ICO marketplace model, which can be labeled as "Value exchange Layer" where technologies that present the consensus to record and exchange of value exists.

F_{14} emphasizes the need to have an execution tool to enable execution of agreements that are beyond the transfer of value. This is where Ethereum Blockchain emerged and is now one of the most extensive platforms to enable smart contracts. This creates the third layer of the ICO marketplace model, which can be labeled as "Automation Layer" where the execution of the business logic related technologies exists.

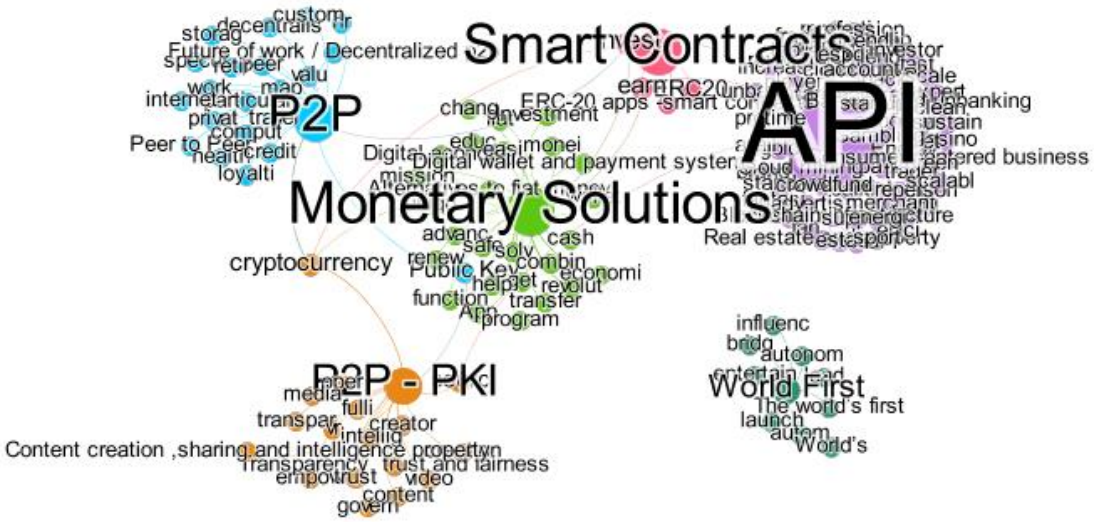

Figure 3. An overview of the LSA results.

The LSA analysis also reveals the dominant industries that the ICO initiators believe they may benefit from. Factors F_2, F_3, F_4, F_6, F_{11}, F_{15}, and F_{18} are Real estate and property, Banking, Cloud mining, Gambling, Energy, Consumer-centered businesses, and Blockchain platforms. Those are industry-specific applications. This creates the fourth and last layer of the ICO marketplace model, which can be labeled as "Application Layer" where the application technologies exist.

Analyzing the entire market space using LSA reveals exciting views on the market; first, it explains the driving force making Blockchain a phenomenon. One interesting finding that was not in the paper initial research objectives, that it highlighted that the current problems identified in the P2P community, namely security concerns and the problem of distributing copyrighted materials (Masood et al. 2018), were already solved by the blockchain community several years earlier. The initiated ICO on the market level identified via 4 high loading factors (F_1, F_5, F_{16}, F_{17}) on the ICO marketplace Trust layer, that Blockchain is providing a solution built on top the P2P networks (F_1, F_{16}), that addresses P2P current identified challenges of disseminating copyrighted content (F_5) securely and safely (F_{17}).

The LSA analysis also reveals that Blockchain emerged from the need to create a payment model, and with its introduction to the double spending, solution using a consensus model (F_7, F_8, F_9, F_{10}, F_{13}) as shown in the second layer of the ICO marketplace model, the "Value Exchange" layer The industry sees Blockchain as a gate opener that enables numerous value exchange solutions, and paves the way to a new era in the digital economy, where value, not only information as in the case of the internet may be exchanged. The industry position, based on the LSA analysis, confirms that Blockchain bridged the gap between the "Proof of Work" concept established because of HashCash in 1997 (Back 2002) and a functional digital currency as a foundation of a new digital economy due to the consensus mechanism presented that addresses the double spending problem (Nakamoto 2008). F_{12} is reinforcing that the Blockchain is an innovative technology, and the remaining factors F_2, F_3, F_4, F_6, F_{11}, and F_{18} presents the different industry segments highly considered by the market.

From this analysis, in parallel with the TCP/IP layered model that governs information exchange over networks, the paper presents a stacked layer model of the ICO marketplace that governs the value exchange over networks in Figure 4. The ICO staked layer model is the first empirical confirmation of how the ICO marketplace with blockchain as its foundation is viewed by the market, confirming that the ICO marketplace is a value exchange system, a new era of the internet from the typical information exchange system.

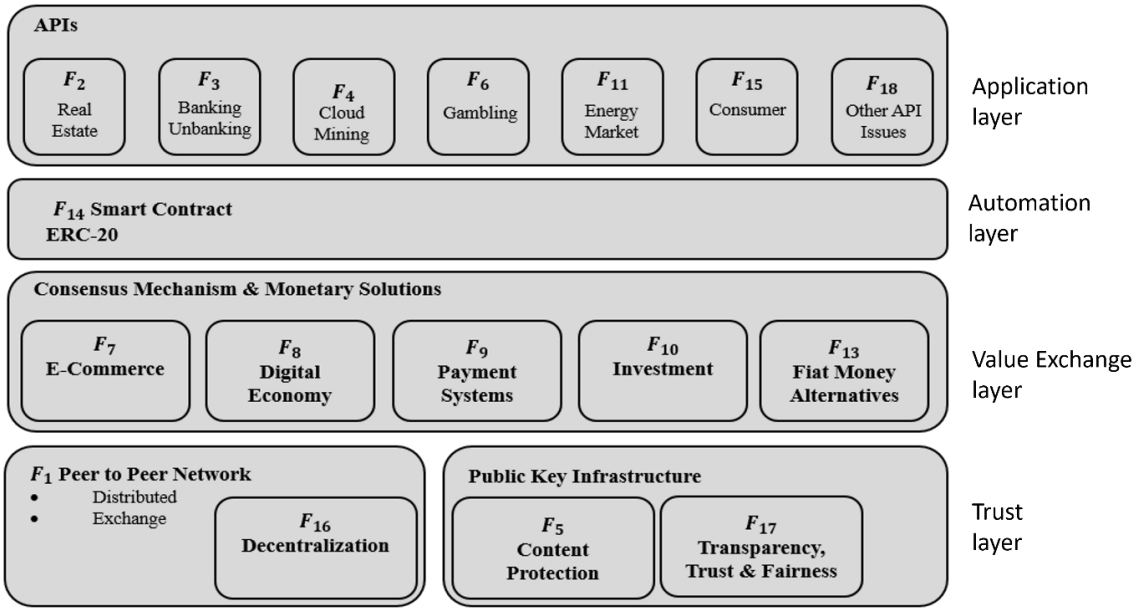

Figure 4. A Stacked layer model of the ICO marketplace.

Mapping Figures 3 and 4 to Table 1 and to the 18 high loading factor tables in Appendix A, the paper groups the different keywords into representative labels. For instance, in Figure 3, the terms related to Application Program Interface (API) in the Application layer of the ICO marketplace model are the list of keywords that connect to different industry segments. For example, Real Estate, Advertising, Banking, and Energy. The paper is doing the same with the Smart Contracts label in the Automation layer of the ICO marketplace model, where it groups the terminologies related to Smart contracts like ERC20 under the label Smart Contract. Accordingly, Figure 3 depicts how different concepts are connected.

In Figure 4, each loading factor table shows as a box. For instance, in the fourth layer, one would see multiple loading factor tables, including cloud mining, gambling, and banking, to name a few. The paper labels the tables using the semantics of the identified high loading keywords. Subsequently, it groups all the related tables under one label on the layer level, which is API for the fourth layer. The same is done for the other labels, for instance, the smart contract label is created based on Factor F_{14}, which groups the keywords related to ERC20, the most used format to create smart contracts in the promoted business models within the analyzed ICOs.

In summary, the ICO marketplace four stacked layers model started with the foundation trust layer which includes Peer-to-Peer Networks and a Public Key Infrastructure. On top of this foundation, there is a value exchange layer to enable the exchange of value using different consensus and payment mechanisms. Then, an Automation layer, as the third layer to execute the approved transactions according to the business logic. The top layer is the Application layer that exposes the business logic to the outside world for use case implementations.

4.2. Blockchain Technology Definition

Table 1 provides the information for F_0, which includes the top loading terms, loadings of these terms, and top ICO descriptions loaded to F_0. The paper reviews the top ICO descriptions of F_0 and finds that the token "peer" represented the term "peer-to-peer." The token "to" does not appear in the result because it is a stop word. As mentioned earlier, F_0 is a market definition for the blockchain technology, "peer-to-peer," the term that has the highest loading, explains 78% of this definition (0.882^2). This means that, from the market perspective, the majority of blockchains, 78% of it, are built on the foundation of peer-to-peer (P2P) technology.

Furthermore, blockchain is more than just P2P. The other 22% of the definition spreads out in a wide variety of industry-specific applications such as advertising (0.9%), gambling (0.9%), and economy (0.7%), real estate (0.5%) and so on. The applications of blockchain are so wide that each specific application only accounts for less than one percent of the definition. F_0 also reveals the practical and application orientation of the market on the blockchain technology. It essentially answers two more concerned questions of the market: "what is blockchain technology usage" and "what an ICO marketplace is comprised of".

According to Factor 0, the paper definition of the Blockchain technology using the market wisdom as depicted on the high loading terms, where the terms in bold are from the high loading factors in Table 1, one may articulate the definition as follows:

Definition 1. *Blockchain technology usage: Blockchain is a decentralized Peer to Peer technology that uses Proof of Work, among other models to reach consensus and enable new economic business models in different industries including Advertising, Banking, and Energy to name a few.*

Mapping the definition to Table 1, the definition uses a sample of the keywords identified to create a market definition. The paper derives the definition of what the market thinks is correct. It does not have to be how investors reacted to those offerings, but it shows where the market is moving in general, using their own market words.

5. Conclusions

As highlighted in the literature review, at least 78 empirical studies used data mining on ICO whitepapers and related information sources adding up to thirty-seven (37) predictor variables in one study to answer different questions including fraud risks, success rates, and future returns. A few researchers used LDA for ICO whitepaper topic classification as a step to help improve their predictions accuracy. This paper used a different latent method, namely the LSA method which is usually used when the goal is dimensionality reduction, not topic modeling.

To address the contradicting definitions from different stakeholders and governing bodies, the paper uses a bottom-up approach and text mining, namely LSA, to identify the market consensus and to provide clarity to entrepreneurs on how to best define blockchain and understand the ICO marketplace.

This paper contributes to the literature by being the only paper that provides a consensus of how the market understands the ICO marketplace, and the market definition of blockchain technology usage. In contrast with the current literature, this paper focuses on summarizing the wisdom of the masses with no bias and in a unified fashion to understand blockchain technology usage, and its ecosystem.

First, the paper analyzed 4367 businesses that published ICO whitepapers. The paper groups the LSA second order identified 18 factors (F_1 to F_{18}) in an ICO marketplace stacked model. This paper is the first to empirically identify the ICO marketplace layers of Trust, Value Exchange, Automation and Application. Second, the paper uses LSA to explain 98.15% of the analyzed business as depicted in F_0. This also generates the first empirical unified and unbiased definition of blockchain technology usage by entrepreneurs. P2P represented 78% of the explanation, enforcing that P2P is in the heart of all ICO value propositions of creating new business models and cutting intermediaries.

The paper contributes to the literature by introducing an ICO Marketplace stacked layer model, and a market generated blockchain technology usage definition using a bottom-up approach to provide clarity to entrepreneurs by entrepreneurs on how to look at Blockchain and its ecosystem, instead of the conflicting definitions issued by the different stakeholders with different objectives.

In summary, the paper identifies the "raisons d'être" or the "reason of existence" for Blockchain technology and the ICO marketplaces. It is the exchange of value between distributed computer networks as it provides a mechanism that enables trust between anonymous entities, without a central authority. This paves the way to address interoper-

ability, and the typical silos of information. It also enables new information flows between entities. With this new mechanism, new business models will be now possible. This paper identifies the major components within this space using empirical analysis. The varied factors shaped the ICO marketplace four layers model to enable the exchange of value, a new area from a previous internet era of exchanging information.

Author Contributions: Methodology, A.G.; data curation, Y.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Acknowledgments: The authors would like to thank the editor, and the anonymous reviewers for their critical reading, clear recommendations, and suggestions to improve the manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A

Table A1. Peer to Peer.

Label	Top Loading Terms	Top Loading ICO Descriptions
F_1 Peer to Peer	peer endors map retir jet hr specul particular pow	• A blockchain platform for effective freelancing and secure peer-to-peer exchanges. • The global freelance peer to peer ecosystem that shares revenue. • Sagittae introduces peer-to-peer financial protection based on blockchain technology. • As a peer-to-peer decentralized network, MEVU allows people to bet on anything, anytime, against anyone. • Dinerocoin is a Peer-to-Peer Transactions for both fiat and cryptocurrencies, which aims to revolutionize the financial market. • Peer-to-peer solution for trade, payment, and transaction of business for agricultural products and services used in the production of alcoholic spirits. • Peer-2-peer decentralized payment & remittance gateway. • Arcade City is a distributed community of peer-to-peer services a blockchain-based sharing economy". • ODYSSEY's mission is to build the next-generation decentralized sharing economy & Peer to Peer Ecosystem. • Teambrella is a peer-to-peer alternative to insurance. Teambrella platform is already live.

Table A2. Real Estate.

Label	Top Loading Terms	Top Loading ICO Descriptions
F_2 Real estate and property	estat advertis properti back rent rental fraction	• Invest in real estate all over the world through blockchain. • Blockchain powered real estate investment platform. • Mijn Vastgoe aims to it possible for everyone to invest in real estate. • Devolve Developer is the first blockchain-based global real estate platform that will disrupt the real estate industry. • RedFund provides access to Real Estate Development investments through an asset-backed. • World's Best Blockchain Real Estate Ecosystem. • The Decentralized Real Estate Platform for Property Marketing. • Relex is the world's first cryptocurrency-based real estate development investment opportunity. • The "X Real Estate Development" project (XRED) aims to develop a technological and organisational model for the real estate development market. • BrickEX is an Open Decentralized Real Estate Marketplace that enables fractional real estate investment and trading by means of asset backed tokens.

Table A3. Banking and Unbanking.

Label	Top Loading Terms	Top Loading ICO Descriptions
F_3 Banking and Unbanking	bank lend tradit unbank account card friendli	• ORCA is the first Open Banking platform designed for crypto users. • ImpalaCoin is the crypto currency that is powering the Impala crypto bank or Financial network. • The world's first global blockchain-powered small business banking. • The Coin of the WorldÂ's first Crypto Bank. The crypto bank for your future. • The Bytemine ICO is a Blockchain Ecosystem for Mining, Assets and Banking. • CryptobankConnecting Blockchain Technology and Traditional Banking to bank the unbanked. • Alux Bank is a digital bank built under the blockchain tecnhnology, with the potential of smart contracts and the power of our own Trading Bot. • MyCryptoBank â€" is an online bank allowing any client registered in the electronic bank system to make a full range of bank operations. • World Bit Bank project will create a group of cryptocurrency banks by acquiring existing banks in different countries to service cryptocurrency customers. • Blockchain-Friendly British Bank.

Table A4. Cloud Mining.

Label	Top Loading Terms	Top Loading ICO Descriptions
F_4 Cloud Mining	mine profit gold crowdfund scale	• Mining technologies for everyone. • MCAP token is a mining and ICO fund. • MinedBlock offers the opportunity for investors to take advantage of using the resources from a large scale mining operation. • The Kerberos ICO is an opportunity for prospective investors to purchase a profit share of an industrial scale, Crypto mining company. • Gramaton is a crowdfunded token with an adjacent large-scale mining operation. • FiveStarMiningNetwork is the first cryptocurrency mining network that will provide guaranteed profit to investors each month. • At Veritas Mining Company, we believe in turning renewable green energy into cryptocurrency by means of crypto mining. • We are provide a cryptocurrency mining service via our large cloud mining pool. • ORET Token is one of the first options in the world of cryptocurrencies to invest and share the profits of a Gold Mining Company. • Innovative Technologies of Cloud Mining.

Table A5. Digital Content Protection.

Label	Top Loading Terms	Top Loading ICO Descriptions
F_5 Digital Content Protection	content media creator video vr intellig monet artifici creation	• Narrative is a Content Economy, generating rewards for content creators. • REOS is the world's first blockchain-based cryptographic exchange for digital content, empowering content creators and consumers. • Next-Generation Content Distribution. • We offer a powerful gaming platform that will be easily integrated by game content providers via an API. • Democratizing Influence Marketing on Blockchain integrated in an existing omnichannel marketplace and shoppable content platform as-a-Service. • ASQ Protocol is a decentralized solution for publishing, storing, sharing, engaging with and monetizing content that empowers a new content economy. • PROPS is Decentralizing the Digital Media Economy. • Bitacium is bringing in sensational innovation to blockchain and media services, integrating business and social activities into one network. • SocialX is a decentralized social media platform allowing users to distribute content seamlessly. • AllSpark Chain is a public blockchain where content creators, advertisers and social media users co-exist a reciprocal sharing economy.

Table A6. Gambling.

Label	Top Loading Terms	Top Loading ICO Descriptions
F_6 Gambling	bet sport event esport gambl fantasi standard casino revolutionari player fan	• An innovative online sports betting platform. • RoBET is a smart-contract decentralized, block-chain based ecosystem for sports betting. • EtherSport is a platform where people all over the world can place bets on sports events. • vDice is a blockchain-based betting game for the Ethereum network. • The Decentralized, Smart Contract Global Betting Exchange. • Transparent betting platform for sporting and other events with revenue sharing using the decentralized smart contracts on the Ethereum blockchain. • Our aim is to become the gold standard for online betting, with bets ranging from sports, eSports to social events. • Introducing BETR—the betting crypto-currency that will establish truly decentralised sports betting on the internet. • PlayChip will deliver the worlds first fully-integrated, blockchain enabled global fantasy sports, online sports betting and gaming ecosystem. • FansUnite is providing a free and open protocol to facilitate fully decentralized sports betting DApps.

Table A7. E-commere.

Label	Top Loading Terms	Top Loading ICO Descriptions
F_7 E-commerce	integr commerc sell bui suppli enterpris sector	• Vanig is the world's first integrated E-Commerce platform and Supply Chain ecosystem powered by the blockchain. • Blockchain crypto-currency for business integrated platform—MIR COIN. • World's First Fiat-Crypto Exchange with Integrated Digital Payment Solutions. • The Lokalize Platform is a fully integrated, 4-in-1 system. • DatEat is the world's first ever dating platform that integrated the power of Blockchain technology, cryptocurrencies and AI engine. • Data integrity for supply chain operations, powered by blockchain technology. • Enterprise Data Integrity Powered by Blockchain. • A smart contract platform with direct commerce integration. • A Chain built with one purposeâ€¦To Scale and Integrate Real-World applications. • A blockchain-based integrative data management system for the global renewable energy industry.

Table A8. Digital Economy.

Label	Top Loading Terms	Top Loading ICO Descriptions
F_8 Digital Economy	economi uniqu get mission revolut	• A blockchain platform for the global digital economy. • The Next Generation Distributed Smart Network Blockchain for the Sharing Economy. • The Compositecoin platform brings together innovative projects for the real economy and their potential investors. • Based on the new money we build a new economy and a better world. • The Token Fund provides an opportunity for investors looking for the simplest way to get into the decentralised economy. • "CryptoDime provides an opportunity for investors looking for the simplest way to get into the decentralised economy." • Revolutionize the digital economy. • PATRON is the sharing economy for influencers. • Starflow is a blockchain-powered ecosystem for the influencer economy. • ICONOMI's mission is to provide an opportunity for investors looking for high profits, not possible in the old economy.

Table A9. Payment systems.

Label	Top Loading Terms	Top Loading ICO Descriptions
F_9 Payment Systems	app wallet combin advanc function solv program	• Mazzuma is a mobile money payment system that utilizes a distributed secure infrastructure and cryptocurrency to enable seamless payments.

Table A10. Investment.

Label	Top Loading Terms	Top Loading ICO Descriptions
F_{10} Investment	investor trader opportun profession level avail expert time tool startup person	• Ecosystem for crypto-investors, traders and funds. • Rublix is developing a unique ecosystem for investors, traders and market professionals. • We will create and operate Abele Trust, one of the world's first fully digital custodians for ALL investors. • RAISON provides you with the investment opportunities which were available only to professional investors. • OpenAssets is an open ecosystem for more effective interaction with the ICO with funds, investors, experts, and traders. • Rewarding investors by building cryptocurrency infrastructure and securing the blockchain. • Give investors the opportunity to own shares in an Ethereum mine. • For creating an open investment fund available for investors of any level. • Social platform for independent developers and investors based on Ethereum. • The Otcrit platform provides traders and investors with a unique combination of an advanced cryptocurrency exchange and research marketplace.

Table A11. Energy Market.

Label	Top Loading Terms	Top Loading ICO Descriptions
F_{11} Energy Market	util energi green sustain increas clean effici	• BlockChain based trade systems make energy trade transparent. • EtainPower is a blockchain-based Energy Ecosystem trading platform. It builds a bridge between investors and green energy. • We are a Utility Payment Blockchain platform. • Bittwatt is a decentralized B2B and B2C energy trading platform. • The first decentralized energy exchange platform powered by renewable energy. • A platform for decentralized crowdfunding and trading of renewable energy assets, making green energy accessible to everyone. • The platform utilizes smart contracts on the Ethereum blockchain to increase transparency and AI automation. • A New Utility Token for Social Commerce. • Peoples Token is an erc20 utility token. Peoples token aims to create a fully community driven and decentralized token. • Decentralizing global energy markets by rewarding energy efficient behavior.

Table A12. The World First.

Label	Top Loading Terms	Top Loading ICO Descriptions
F_{12} The world's first	world's lead autonom autom bridg entertain influenc launch	• The world's first Community Crypto Exchange Coin. • Based On Blockchain Financial System World's First BCH Marketplace. • Snovio is the world's first decentralized lead generation service. • The world's best cryptocurrency-based autonomous marketplace of services. • Bolt Global, the world's first open entertainment economy powered by Zilliqa. • FANFARE is the world's first Social Commerce ecosystem powered by Blockchain technology. • Zuum is the world's first decentralized on-demand services platform that allows consumers and service providers to interact without intermediaries. • This is the world's first socially contributing token ecosystem, with the SocialGood cryptocurrency. • The world's first blockchain-based marketplace for the global $3 Trillion IT products and services Industry. • Bitdollar Fund is the world's first collectively managed fund.

Table A13. Fiat Money Alternatives.

Label	Top Loading Terms	Top Loading ICO Descriptions
F_{13} Fiat Money Alternatives	monei earn fiat help renew transfer easi educ chang safe cash	• Stash: Digital Money For Everyone. • We build a decentralized platform where you can earn money with your data. • A better way to manage your money. Transfer, exchange and spend your money all over the world. An easy way. • Earn money by watching, sharing and interacting with digital content. • HexanCoin is a smart crypto-trading platform where users will be able invest their fiat money on cryptocurrencies. • The next generation of advanced solution for global money transaction. • Telegraf Money allows members to communicate with each other, instantly send funds and cryptocurrencies, issue/receive loans online, and earn money. • The Mammoth is an innovative payment network and a new kind of money. • We will develop fair and completely transparent StreamDesk service for the cryptocurrency-fiat money exchange without intermediaries. • We protect money invested in ICOs and cryptocurrencies.

Table A14. Smart Contract.

Label	Top Loading Terms	Top Loading ICO Descriptions
F_{14} Smart Contract	erc	• DAR is a modern utility token based on Ethereum blockchain ERC-20 protocol. • Skyhope Coin (SHC) is an ERC-20 token ethereum-based platform that will be an utility token used for freelancer secure and fast payment. • NuMoney Exchange is launching a utility token to burn 66.6% of trading fees. • The hirefreehands token ($HFT) is an ERC20 utility token designed to be used by clients and freelancers within the platform. • The World's First Internet Service Provider utilizing decentralized blockchain technology. • MATOX is a fully audited Ethereum based ERC20 utility token that is used for exchange fees on the MATOX Exchange. • Qravity is a decentralized content production and distribution platform where creators own and profit from their work. • The FUCK Token is a new ERC20 Token built on the Ethereum blockchain.

Table A15. Consumer-centered Business.

Label	Top Loading Terms	Top Loading ICO Descriptions
F_{15} Consumer-centered Business	consum driven manufactur brand retail merchant	• A decentralized ecosystem directly connecting consumers and businesses. • Multicryptocurrency payment platform for businesses, consumers & community. • Crypto Kafe is the decentralized consumer ecosystem powered by blockchain. • VR platform for global integration of developers, consumers and business. • Global Decentralized Network Directly Connecting Manufacturers, Designers, and Consumers. • MaxData is building a protocol to utilize anonymous user data to save billions for consumers and companies. • A decentralized social currency for Influencers, Companies and Consumers. • Atlas will provide a smart-contract based infrastructure that connects all consumers and businesses in the travel industry. • The world's first decentralized, AI-driven marketplace providing real-time solutions to retailers, manufacturers and consumers. • Your premier seat on top of the global consumer network—make money helping other save.

Table A16. Decentralization.

Label	Top Loading Terms	Top Loading ICOs Descriptions
F_{16} Decentralization	decentralis work privat custom loyalti internet travel credit store comput valu storag health	• Future of Work Decentralise work with Blockchain. • Sociall is a new generation of networking. It is a secure and private decentralised social network for all. • Pally is the first decentralised social travel ecosystem in the world. • Fern is a smart contract protocol for decentralising, and increasing trust and security in private blockchain networks. • Unleashing the value of medical data for the new decentralised healthcare economy. • Golem is the new way the Internet will work. • Designed for Decentralised Applications that require SPEED and PRIVACY when sharing content. • World's first decentralised, open source and device independent computer. • Blocknubie is a decentralised Ecosystem for Blockchain Startups. • The new digital decentralised address platform.

Table A17. Transparency, Trust and Fairness.

Label	Top Loading Terms	Top Loading ICOs Descriptions
F_{17} Transparency, Trust and Fairness	transpar oper sourc trust fulli fair empow govern	• PrimeLend is a transparent and decentralized investment platform built on the mighty Ethereum blockchain. • A global & transparent blockchain-based data exchange ecosystem. • The Mercury Protocol is a decentralized, transparent, and open source protocol built on the Ethereum blockchain. • System for trusted financial transactions between people based on transparency and blockchain technology. • COINDAQ â€œCDAQâ€ brings transparency, regulation and security to digital currency and assets through. • The TaTaTu Platform will be a fair and transparent social media and entertainment blockchain-powered platform. • Cryptaur aims to create a decentralized ecosystem with a focus on establishing a fully transparent platform for consumers. • Global casino industry transformation. Transparent. Fair. Empowered by the Etherium blockchain. • Utile Network is an infrastructure that allows you to crowdsource blockchain ecosystem data in a transparent environment. • Enzym is a dApp and the ZYM is an ERC20 token. We use the blockchain to bring transparency and security to the user.

Table A18. Blockchain Infrastructure Applications.

Label	Top Loading Terms	Top Loading ICOs Descriptions
F_{18} Blockchain Infrastructure Applications	infrastructur sourc support scalabl realiti virtual free multi softwar financ dapp cloud fast	• The Most Powerful Infrastructure for Decentralized Applications. • A protocol for decentralized fund infrastructure. • The Neverdie ICO will finance the development of a multi-platform API for monetized gaming infrastructure to support a virtual goods economy. • Lambda is a high-speed, secure and scalable blockchain infrastructure project. • Cross-industry data infrastructure. • The Apollon Blockchain project aims to develop and support all the related business applications with the underlying Blockchain infrastructure technology. • Enterprise Blockchain Infrastructure for Decentralized Internet. • REMME is building an open source distributed Public Key Infrastructure protocol with a set of DApps, enabling passwordless authentication for humans. • TravelerToken—A Blockchain based infrastructure and payment system for the travel industry. • Mobile ad hoc blockchains, an open-source infrastructure for distributed mobile peer-to-peer applications.

References

Ahmad, Muhammad Farooq, Oskar Kowalewski, and Paweł Pisany. 2021. What determines initial coin offering success: A cross-country study. *Economics of Innovation and New Technology*, 1–24. [CrossRef]

Altalhi, Abdulrahman H., José María Luna, M. A. Vallejo, and Sebastián Ventura. 2017. Evaluation and comparison of open source software suites for data mining and knowledge discovery. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery* 7: e1204. [CrossRef]

Ante, Lennart. 2022. Non-fungible token (NFT) markets on the Ethereum blockchain: Temporal development, cointegration and interrelations. *Economics of Innovation and New Technology 1–19*, 1–19. [CrossRef]

Back, Adam. 2002. Hashcash—A Denial of Service Counter-Measure. Available online: http://www.hashcash.org/papers/hashcash.pdf (accessed on 22 November 2022).

Bian, Shuqing, Zhenpeng Deng, Fei Li, Will Monroe, Peng Shi, Zijun Sun, Wei Wu, Sikuang Wang, William Yang Wang, Arianna Yuan, and et al. 2018. IcoRating: A Deep-Learning System for Scam ICO Identification. *arXiv* arXiv:1803.03670.

Buterin, Vitalik. 2014. A Next-Generation Smart Contract and Decentralized Application Platform. Available online: https://cryptorating.eu/whitepapers/Ethereum/Ethereum_white_paper.pdf (accessed on 22 November 2022).

Cai, Jinghan, and Ahmed Gomaa. 2019. Initial Coin Offering to Finance Venture Capital: A BehavioralPerspective. *The Journal of Private Equity* 22: 93–101. [CrossRef]

Cao, Tianjie, Dongdai Lin, and Rui Xue. 2005. A randomized RSA-based partially blind signature scheme for electronic cash. *Computers and Security* 24: 44–49. [CrossRef]

Catalini, Christian, and Joshua S. Gans. 2018. *Initial Coin Offerings and the Value of Crypto Tokens*. Cambridge: National Bureau of Economic Research.

Chitsazan, Hasti, Afsaneh Bagheri, and Mahdi Tajeddin. 2022. Initial coin offerings (ICOs) success: Conceptualization, theories and systematic analysis of empirical studies. *Technological Forecasting and Social Change* 180: 121729. [CrossRef]

Chuanjie, Fu, Andrew Koh, and Paul Griffin. 2019. Automated theme search in ICO whitepapers. *The Journal of Financial Data Science* 1: 140–58. [CrossRef]

Clark, David. 1988. The design philosophy of the DARPA Internet protocols. In *Symposium Proceedings on Communications Architectures and Protocols*. New York: Association for Computing Machinery, pp. 106–14.

Cong, Lin William, and Zhiguo He. 2019. Blockchain disruption and smart contracts. *The Review of Financial Studies* 32: 1754–97. [CrossRef]

Crain, Steven P., Ke Zhou, Shuang-Hong Yang, and Hongyuan Zha. 2012. Dimensionality reduction and topic modeling: From latent semantic indexing to latent dirichlet allocation and beyond. In *Mining Text Data*. Boston: Springer, pp. 129–61.

Cunningham, William Michael, Bobby Henebry, and Hunter Horsley. 2019. BLOCKCHAIN: A Primer for Planners. *Journal of Financial Planning* 32: 23–25.

Curran, Kevin, and John Honan. 2006. Eliminating the Volume of Spam E-Mails Using a Hashcash-Based Solution. *Information Systems Security* 15: 22–41. [CrossRef]

De Andrés, Pablo, David Arroyo, Ricardo Correia, and Alvaro Rezola. 2022. Challenges of the market for initial coin offerings. *International Review of Financial Analysis* 79: 101966. [CrossRef]

Dumais, Susan T., George W. Furnas, Thomas K. Landauer, Scott Deerwester, and Richard Harshman. 1988. Using latent semantic analysis to improve access to textual information. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York: Association for Computing Machinery, pp. 281–85.

Ellison, Carl, and Bruce Schneier. 2000. Ten risks of PKI: What you're not being told about public key infrastructure. *Computer Security Journal* 16: 1–7.
Evangelopoulos, Nicholas. 2011. Citing Taylor: Tracing Taylorism's technical and sociotechnical duality through latent semantic analysis. *Journal of Business and Management* 17: 57–74.
Evangelopoulos, Nicholas, Xiaoni Zhang, and Victor R. Prybutok. 2012. Latent semantic analysis: Five methodological recommendations. *European Journal of Information Systems* 21: 70–86. [CrossRef]
Filipov, Nadezhda. 2018. Blockchain—An Opportunity For Developing New Business Models. *Business Management* 2018: 75–92.
Fisch, Christian. 2019. Initial coin offerings (ICOs) to finance new ventures: An exploratory study. *Journal of Business Venturing* 34: 1–22. [CrossRef]
Florysiak, David, and Alexander Schandlbauer. 2022. Experts or charlatans? ICO analysts and white paper informativeness. *Journal of Banking & Finance* 12: 106456.
Golub, G. H., and C. Reinsch. 1970. Singular value decomposition and least squares solutions. *Numerische Mathematik* 14: 403–20. [CrossRef]
Gomaa, Ahmed. 2018. The Creation of a Cryptocurrency to Support Disintermediation. *Operatoins Management Education Review (OMER)*, 12. [CrossRef]
Haber, Stuart, and W. Scott Stornetta. 1990. How to time-stamp a digital document. In *Conference on the Theory and Application of Cryptography*. Berlin: Springer, pp. 437–55.
Halaburda, Hanna. 2018. Blockchain Revolution without the Blockchain? *Communications of the ACM* 61: 27–29. [CrossRef]
Han, Jiawei, Jian Pei, and Micheline Kamber. 2011. *Data Mining: Concepts and Techniques*. Amsterdam: Elsevier.
Hotho, Andreas, Andreas Nürnberger, and Gerhard Paaß. 2005. A brief survey of text mining. *Ldv Forum* 20: 19–62.
Howell, Sabrina T., Marina Niessner, and David Yermack. 2018. *Initial Coin Offerings: Financing Growth with Cryptocurrency Token Sales*. Cambridge: National Bureau of Economic Research.
Hu, Yifan. 2005. Efficient, high-quality force-directed graph drawing. *Mathematica Journal* 10: 37–71.
ICOBench.com. 2022. September 17. Available online: https://icobench.com/ (accessed on 22 November 2022).
Jensen, Lars Juhl, Jasmin Saric, and Peer Bork. 2006. Literature mining for the biologist: From information retrieval to biological discovery. *Nature Reviews Genetics* 7: 119–29. [CrossRef] [PubMed]
Jongsub, Lee, Tao Li, and Donghwa Shin. 2022. The Wisdom of Crowds in FinTech: Evidence from Initial Coin Offerings. *The Review of Corporate Finance Studies* 11: 1–46.
Kaminska, Izabella. 2019. Blockchain officially confirmed as slower and more expensive. *Financial Times*. May 29. Available online: https://ftalphaville.ft.com/2019/05/29/1559146404000/Blockchain-officially-confirmed-as-slower-and-more-expensive/ (accessed on 22 November 2022).
Karpenko, Oksana A., Tatiana K. Blokhina, and Lali V. Chebukhanova. 2021. The Initial Coin Offering (ICO) Process: Regulation and Risks. *Journal of Risk and Financial Management* 14: 599. [CrossRef]
Keogh, Eamonn, and Abdullah Mueen. 2011. Curse of dimensionality. In *Encyclopedia of Machine Learning*. Boston: Springer, pp. 257–58.
Kostovetsky, Leonard, and Hugo Benedetti. 2018. Digital Tulips? Returns to Investors in Initial Coin Offerings. *SSRN Electronic Journal*. [CrossRef]
Lansky, Jan. 2018. Possible state approaches to cryptocurrencies. *Journal of Systems Integration* 9: 19–31. [CrossRef]
Lansky, Jan. 2019. Cryptocurrency survival analysis. *The Journal of Alternative Investments* 22: 55–64. [CrossRef]
Laskowski, M., and H. M. Kim. 2016. Rapid Prototyping of a Text Mining Application for Cryptocurrency Market Intelligence. Paper presented at IEEE 17th International Conference on Information Reuse and Integration (IRI), Pittsburgh, PA, USA, July 28–30.
Lemchuk, Stephanie A. 2016. Virtual Whats: Defining Virtual Currencies in the Face of Conflicting Regulatory Guidances. *Cardozo Public Law, Policy, and Ethics Journal* 15: 319.
Lin, Xiaolin, Yibai Li, and Xuequn Wang. 2017. Social commerce research: Definition, research themes and the trends. *International Journal of Information Management* 37: 190–201. [CrossRef]
Lochbaum, Karen E., and Lynn A. Streeter. 1989. Comparing and combining the effectiveness of latent semantic indexing and the ordinary vector space model for information retrieval. *Information Processing & Management* 25: 665–76.
Lockaby, C. Daniel. 2018. The Sec Rides into Town: Defining an Ico Securities Safe Harbor in the Cryptocurrency Wild West. *Georgia Law Review* 53: 335–66.
Mai, Feng, Zhe Shan, Qing Bai, Xin (Shane) Wang, and Roger H. L. Chiang. 2018. How Does Social Media Impact Bitcoin Value? A Test of the Silent Majority Hypothesis. *Journal of Management Information Systems* 35: 19–52. [CrossRef]
Masood, Saleha, Muhammad Alyas Shahid, Muhammad Sharif, and Mussarat Yasmin. 2018. Comparative Analysis of Peer to Peer Networks. *International Journal of Advanced Networking and Applications* 9: 3477–91.
Morrison, Alan, and Subhankar Sinha. 2016. Blockchain and smart contract automation: Blockchains defined. In *The Future of Law and Technologies*. Edited by Tanel Kerikmae and Addi Rull. Berlin and Heidelberg: Springer.
Nakamoto, Satoshi. 2008. Bitcoin: A Peer-to-Peer Electronic Cash System. Available online: https://bitcoin.org/bitcoin.pdf (accessed on 22 November 2022).
Pazaitis, Alex, Primavera De Filippi, and Vasilis Kostakis. 2017. Blockchain and value systems in the sharing economy: The illustrative case of Backfeed. *Technological Forecasting and Social Change* 125: 105–16. [CrossRef]
Prentis, Mitchell. 2015. Digital Metal: Regulating Bitcoin as a Commodity. *Case Western Reserve Law Review* 66: 609–38.

Qiang, Minqian, and Fei Lin. 2019. The Challenges of Existence, Status, and Value for Improving Blockchain. *IEEE Access* 7: 7747–58.
Salton, Gerard, Anita Wong, and Chung-Shu Yang. 1975. A vector space model for automatic indexing. *Communications of the ACM* 18: 613–20. [CrossRef]
Sapkota, Niranjan, and Klaus Grobys. 2021. Fear Sells: Determinants of Fund-Raising Success in the cross-section of Initial Coin Offerings. *SSRN*. [CrossRef]
Senderowicz, Jeremy I., K. Susan Grafton, Timothy Spangler, Kristopher D. Brown, and Andrew J. Schaffer. 2018. SEC focuses on initial coin offerings: Tokens may be securities under federal securities laws. *Journal of Investment Compliance* 19: 10–14. [CrossRef]
Sidorova, Anna, Nicholas Evangelopoulos, Joseph S. Valacich, and Thiagarajan Ramakrishnan. 2008. Uncovering the intellectual core of the information systems discipline. *MIS Quarterly* 32: 467–82. [CrossRef]
Stenqvist, Evita, and Jacob Lönnö. 2017. *Predicting Bitcoin Price Fluctuation with Twitter Sentiment Analysis*. Stockholm: KHT.
Sun, Yin Hao Hua, Klaus Langenheldt, Mikkel Harlev, Raghava Rao Mukkamala, and Ravi Vatrapu. 2019. Regulating Cryptocurrencies: A Supervised Machine Learning Approach to De-Anonymizing the Bitcoin Blockchain. *Journal of Management Information Systems* 36: 37–73. [CrossRef]
Swartz, Lana. 2022. Theorizing the 2017 blockchain ICO bubble as a network scam. new media & society. *New Media & Society* 24: 1695–713.
Takahashi, Koji. 2020. Prescriptive Jurisdiction in Securities Regulations: Transformation from the ICO (Initial Coin Offering) to the STO (Security Token Offering) and the IEO (Initial Exchange Offering). *Ilkam Law Review* 20: 31–50. [CrossRef]
Thewissen, James, Prabal Shrestha, Wouter Torsin, and Anna M. Pastwa. 2022. Unpacking the black box of ICO white papers: A topic modeling approach. *Journal of Corporate Finance* 75: 102225. [CrossRef]
Tirunillai, Seshadri, and Gerard J. Tellis. 2014. Mining marketing meaning from online chatter: Strategic brand analysis of big data using latent dirichlet allocation. *Journal of Marketing Research* 51: 463–79. [CrossRef]
Varma, Jayanth Rama. 2019. Blockchain in Finance. *The Journal for Decision Makers* 44: 1–11. [CrossRef]
Viriyasitavat, Wattana, and Danupol Hoonsoponb. 2019. Blockchain characteristics and consensus in modern business processes. *Journal of Industrial Information Integration* 13: 32–39. [CrossRef]
Winotoatmojo, Hugo Prasetyo, Idris Gautama, Elidjen, and Meyliana. 2022. Initial Coin Offering (ICO) an Innovation Funding for Digitalpreneur. *I T A L I E N I S C H* 12: 1–9. [CrossRef]
Yousif, Ahmed. 2022. Blockchain: The Protocol of Value. Makkah, Saudi Arabia. June 1. Available online: https://bsvblockchain.org/news/blockchain-the-protocol-of-value/ (accessed on 22 November 2022).
Yu, Ju Hyun, Juyoung Kang, and Sangun Park. 2019. Information availability and return volatility in the bitcoin Market: Analyzing differences of user opinion and interest. *Information Processing & Management* 56: 721–31.
Zetzsche, Dirk A., Ross P. Buckley, Douglas W. Arner, and Linus Föhr. 2017. The ICO Gold Rush: It's a scam, it's a bubble, it's a super challenge for regulators. *University of Luxembourg Law Working Paper* 11: 17–83. [CrossRef]
Zhang, Shuyu, Walter Aerts, Liping Lu, and Huifeng Pan. 2019. Readability of token whitepaper and ICO first-day return. *Economics Letters* 180: 58–61. [CrossRef]
Zheng, Zibin, Shaoan Xie, Hong-Ning Dai, Xiangping Chen, and Huaimin Wang. 2018. Blockchain challenges and opportunities: A survey. *International Journal of Web and Grid Services* 14: 352–75. [CrossRef]
Zhu, Dan, and Yuanfeng Cai. 2016. Fraud detections for online businesses: A perspective from blockchain technology. *Financial Innovation* 2: 20.
Zutshi, Aneesh, Antonio Grilo, and Tahereh Nodehi. 2021. The value proposition of blockchain technologies and its impact on digital platforms. *Computers & Industrial Engineering* 155: 107187.