# How Sentiment Impacts the Success of Blockchain Startups –
# An Analysis of Social Media Data and Initial Coin Offerings

| | | |
|---|---|---|
| Simon Albrecht | Bernhard Lutz | Dirk Neumann |
| University of Freiburg, Germany | University of Freiburg, Germany | University of Freiburg, Germany |
| simon.albrecht@is.uni-freiburg.de | bernhard.lutz@is.uni-freiburg.de | dirk.neumann@is.uni-freiburg.de |

## Abstract

*The blockchain technology is subject to intensive discourse on social media. One example is the promotion of Initial Coin Offerings (ICO) as a new financial instrument for startups. As ICOs are generally low-regulated, potential investors and entrepreneurs are faced with decision-making under uncertainty. This uncertainty might be partially offset by social factors. Based on established behavioral, economic and information systems theory, we investigate how different social media channels are linked to the capitalization of blockchain startups. We present empirical evidence from online search trends, financial data and a corpus of 231,758 tweets related to 524 ICOs. Our results suggest that higher search volume, positive sentiment and the increased use of emotive language on Twitter are linked to a high capitalization. The results are relevant for startups and potential ICO investors.*

## 1. Introduction

The advent of blockchain technology – and with it disintermediation, empowerment and the rethinking of value chains – is expected to pave the way for innovative applications and business models. Yet there are only a few market-ready blockchain services, as blockchain is considered to be an innovation still searching for a use case [1]. Finance is considered to be among the main industry sectors to be disruptively affected by blockchain [2], as this technology has the potential to change the nature of monetary and legal interactions [3], an influence observed in the case of Initial Coin Offerings (ICO), the distributed financing tool for blockchain startup firms. ICOs are becoming increasingly popular; in the first months of 2018, blockchain projects raised more money through ICOs than they had over the course of the previous year [4].

Since ICOs are not regulated by conventional legislation, the level of market uncertainty is pronounced and lawmakers are concerned that ICOs provide opportunities for financial fraud [5]. For instance, about 10% of ICOs can be classified as scams [6]. Accordingly, an objective evaluation of blockchain and ICOs has been identified as a relevant contribution in terms of both research and practice. In particular, the assessment of value propositions and risks is currently under-researched, a shortcoming that can be addressed with empirical investigations of strategies and tactics applied by companies working on new blockchain solutions [7]. Such assessments pose a challenge to analysts, as no regulatory authority is currently responsible for overseeing identities, criminal records or financial backgrounds. Thus, hard facts are sparsely available on most ICOs. To investors, it is therefore unclear whether an ICO will become successful or not. Entrepreneurs, likewise, are unsure of how to position their startup in a market characterized by uncertainty. Accordingly, it is relevant to investigate the underlying determinants of blockchain startups' success.

As blockchain, much like other innovations, can be assumed to pass through fashion cycles [8] with excessive discourse [9] often expressed as emotive or polarized speech – and blockchain is among those technological innovations accompanied by significant coverage in social media – we suspect that trends and discourse in social media will be utilizable as a data source for such an investigation.

To address the identified research gap, we investigate the social component of blockchain companies' strategies for raising funds with the following research question: *How is the financial success of firms' ICOs linked to their social media utilization and the overall blockchain trend?* This question is addressed by analyzing Twitter feeds with Natural Language Processing (NLP) tools and utilizing the output for a regression model in combination with online search statistics and financial ICO data. Doing so, we enhance researchers' and investors' ability to assess successful ICOs with social components. This approach contributes to the existing scientific literature by furthering the understanding of social mechanisms

HⓘCSS

of the current blockchain phenomenon. In addition, our study contributes to practice by explaining social features that are characteristic of startups' capitalization for retail investors. To the best of our knowledge, this is the first paper to use big data and NLP approaches to evaluate social media sentiment with regard to the financial success of ICOs.

The paper is structured as follows. Section 2 provides essential background information concerning the blockchain technology and ICOs. Section 3 lays the theoretical foundation for our research hypotheses, followed by Section 4, which describes the utilized methodology and data. In Section 5, the results of our model are discussed. Section 6 concludes and provides an outlook on future research.

## 2. Background

In this section, we origins of ICOs, the technical background, and the investment process.

### 2.1. Blockchain as a means of capitalization

During the last two years, ICOs have emerged as an efficient way to allocate capital for blockchain startup businesses as an increasingly popular alternative to conventional funding strategies such as venture capital and Initial Public Offerings (IPOs). During the observation period of this study, the cumulative ICO volume amounted to more than eight billion USD, as depicted in Figure 1.

In an ICO, a startup company issues a "token", which can be defined either as a digital asset or good, to investors who are willing to exchange fiat money in return (which usually must first be exchanged for an established cryptocurrency used for regular trade). Tokens being used as assets are called *equity* or *investment tokens*. Tokens that grant buyers access to an existing or future service or digital infrastructure, such as decentralized Apps (dApps), are called *utility tokens*. While the value of utility tokens is only determined by the perceived demand for the use of the respective network's infrastructure, equity tokens represent the current valuation of a legal firm that can (in theory) be claimed, since these tokens directly assign ownership of company shares to the token holder. Equity tokens are subject to a number of currently nebulous rules from regulators such as the U.S. Securities and Exchange Commission (SEC) and the European Securities and Markets Authority (ESMA) [10]. Potential regulatory steps to control ICOs include mandatory compliance programs, regulating cryptocurrency exchanges, establishing security standards, exercising legal force to combat fraud, an implementing government-led recommendation programs for end customers and retail investors. However, it appears as though most regulators are waiting for other countries' governments to take exemplary legislative steps which they can then emulate.

ICOs are consequently used by blockchain projects and startups due to the historical development of blockchain technology as a payment system in the form of Bitcoin [11] and its underlying distributed infrastructure, which allows for an immutable record with a single point of truth and consensus between network participants. The possibility to precisely display infinitesimal changes in ownership and a seamless record of transactions render the technology especially suitable for use in finance. While in principle, ICOs can also be utilized as a financial instrument by companies that are not proposing a blockchain-based service, a startup's proposal for the utilization of blockchain technology should be a logical solution in response to a real problem in order to avoid scams [12].

The advent of tokens as a digital and distributed tool for value allocation was introduced through the development of the Ethereum platform [13]. While the most prominent Bitcoin blockchain is solely a cryptographic distributed ledger, Ethereum allows for the intrinsic execution of smart contracts and dApps in its application layer. Smart contracts are conditionally self-executing applications, e.g. in the environment of the Ethereum Virtual Machine, that can, for instance, be triggered by external data sources (oracles). Smart contracts are the technological foundation for the creation, distribution, and tracking of tokens. In Ethereum's case, a smart contract contains the account addresses of all token holders and their respective balances. During a token transaction, the smart contract updates these balances (transactions can contain the cryptocurrency Ether, data or tokens). While it may appear that many blockchain startups are developing or operating their own blockchain platform, in reality, many ICOs, such as Aragon, OmiseGo and TenX, are issuing tokens that are being used on the Ethereum platform. These tokens are based on the popular ERC-20 smart contract template [14], a simple standard for the description of necessary information for token smart contracts.

### 2.2. The ICO process

When a blockchain startup aims to raise capital through an ICO, one of the first steps is to write a whitepaper in which the general idea of the project, technical foundations (platform and wallet design, consensus mechanism, routing protocols, etc.) and
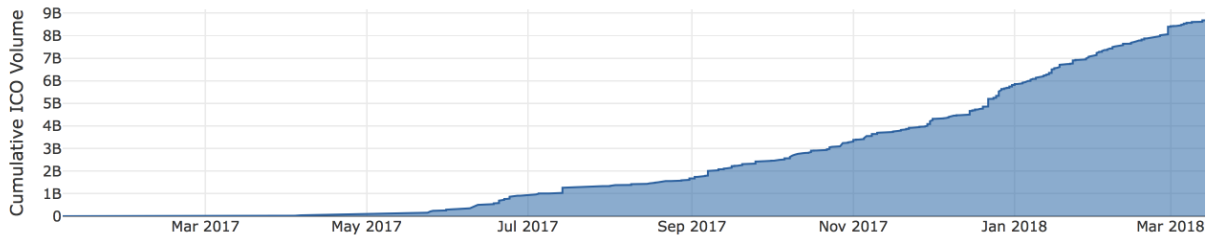
**Figure 1. Cumulative ICO volume for observation period based on our data (in billion USD).**

milestones are presented. Beyond the whitepaper, a description of the token contract and legal design is useful for potential investors. A transparent and open communication strategy ensures the broad participation of many investors. During the ICO, the wallet addresses of the designated investors are stored on the smart contract of the token. After the ICO's specified end date has been reached, and the target sum gathered, the tokens are transferred to the listed accounts [15].

A sound marketing strategy is expected to be a major determinant for an ICO's success since distributed business models depend on network effects. The value of a network depends on the number of other people already connected to it. Realizing network effects equates to positive feedback cycles in which market success can be achieved through a critical mass of more and more users who find an innovation useful [16]. Such community building can, for instance, be done through social media channels such as Reddit, Telegram, and Twitter, conduits through which potential users receive relevant information regarding the value proposition of the blockchain project. In addition to the most prevalent social media platforms, ICOs can provide additional information on designated ICO rating websites [15].

These kind of rating websites are among the first information sources in terms of reputability for retail investors seeking to differentiate between trustworthy blockchain projects and other ICOs that are likely to be scams. However, while basic information, such as the end date, target volume and linked social media accounts, are accessible through these websites, more detailed analyses are generally not available [12].

### 2.3. Concurrent research on successful ICOs

While few high-quality publications have released articles addressing the contemporary ICO phenomenon, various researchers have investigated the topic in working papers, conference proceedings and journals with limited impact. A recent study examined descriptive data from an ICO rating website (ending

in December 2017). The underlying dataset included background information on 450 tokens. Using the derived criteria and a logit model, the results suggest that geographic location is a determinant of success, while other variables (e.g. team size) are not relevant [17]. An econometric study found that even though ICOs are rather heterogeneous, their success rates are generally high. While providing additional information has differing impacts (e.g. high for code, low for whitepapers), presale initiatives with smaller tokens targeted at specific groups appear to have a positive effect on ICOs' success [18].

In a recent study, the social media activity levels of ICOs were found to vary greatly, with the top ICOs being highly active on Twitter, LinkedIn, and Facebook [19]. Subsequently, another study found the daily Twitter activity to be associated with positive market returns for ICOs [20]. In a qualitative study, researchers conducted semi-structured interviews to investigate signals during an ICO. The results suggest the relevance of social media discourse within the blockchain community due to the distributed nature of the community's philosophy. Surprisingly, some investors use sentiment as the only criterion for decision-making [21].

Another approach used financial and social media data from an ICO-listing website to model the prediction of successful ICOs. The descriptive data of 400 reviewed Twitter accounts, such as the number of followers and number of tweets, exhibited a weak correlation with the financial ICO result. Furthermore, reviewing the writing style of whitepapers also yielded a weak correlation with the ICO result [22].

### 3. Theory development

In this section, we review previous studies that analyze the characteristics of a high market capitalization. Subsequently, we derive our hypotheses based on behavioral, economic and information systems theory.

### 3.1. Economic determinants of successful capitalization

The (perceived) trustworthiness and legitimacy of a firm are crucial for its financial performance. Behavioral theories suggest that investors seek as much information as possible in order to make an informed decision [23]. In markets where inefficiencies are prevalent due to asymmetric information [24], signaling mechanisms, such as entrepreneurs' behavior or announced capital expenditures, help potential investors recognize the future value of firms. Previous work has yielded evidence that IPO investors consider signaling indicators, such as experience with the development of other products, to be determinants for the future financial success of firms. Empirical results suggest that high values in terms of these indicators significantly increase the raised capital throughout the IPO process [25].

To strengthen their business potentials, new firms can use signaling mechanisms to influence the perception of potential investors in order to indicate legitimacy and to overcome market newness. Research suggests that the relevance of non-financial information, as opposed to financial information (prices, returns, cash flows), has steadily increased for investors. Moreover, non-financial information is becoming an important factor in determining a firm's value. For instance, indicators like social capital among board members – and especially investors' perceptions of these indicators – improve a firm's market performance [26].

Signaling legitimacy can be done by writing reports, press releases or social media status updates. Investigating specific indicators, such as text length or active/passive voice, yielded evidence in favor of the importance of writing style [27]. Research on SEC-regulated document filings prior to the initiation of IPO processes has found evidence for the impact of tonality as a representation of a firm's business strategy on investors' valuation of the firm. IPOs with high levels of uncertainty in their writing, such as weak modal and negative word frequencies, yield higher returns on the first day preceding the IPO [28]. Other work focusing on stock markets has shown that the sentiment polarity of social media (Twitter) status updates significantly affects the direction of abnormal returns in the short term [29].

As the ICO mechanism can be considered similar to that of crowdfunding [18], economic behavior in ICOs has been reviewed through the lens of economic agents' options for influencing other stakeholders. For instance, marketing research has investigated crowdfunding as an instrument by which to realize network effects.

In particular, the establishment of trustful relations between entrepreneurs and investors potentially builds long-term relationships with the option to raise additional capital in the future [30]. Furthermore, a recent model analysis has shown that ICOs may allow entrepreneurs to incentivize buyer competition to reveal consumer value [31]. Accordingly, research on crowdfunding has shown the impact of entrepreneurs' signaling on investors' perception of uncertainty and risks. An empirical study that investigated firms' characteristics and eventual funding levels found that providing detailed information acts as signaling and can strongly impact funding success [32].

### 3.2. The effectiveness of initial price setting in ICOs

While classical financial paradigms state that investors make rational decisions, during the last several decades, research in behavioral economics has yielded evidence for inconsistent price patterns that do not fit with traditional models [33]. Biases in financial decision-making include optimism and disregard for statistical data. For instance, people make inconsistent choices in alternative formulations of the same problem [34]. In the blockchain community, entrepreneurs have been shown to exploit such behavioral patterns through strategic price setting in presale initiatives, targeting narrow groups of potential investors [18]. However, such an exploitation can only be successful when targeted subjects are manipulated in a specific direction. For the actual sales of tokens with initial prices, investor groups are inherently heterogeneous as this market is defined by low entry barriers. Accordingly, predictable projections of price developments cannot be assumed due to the lack of available instruments and because promising tokens may sell at the same price as underwhelming tokens [24]. While incentives for entrepreneurs to engage in strategic pricing behavior, such as over- or underpricing, have theoretically been modeled in prior research [31], we suspect that such strategies might not have a significant effect on the success of an ICO when applied to the initial token price. Accordingly, we propose the following hypothesis:

*Hypothesis 1. The initial price setting of tokens has no major impact on the financial success of an ICO.*

### 3.3. Bandwagon phenomena and market success throughout fashion cycles

Fashion cycles are innovation trends grounded in Abrahamson's management fashion theory [35]. In a fashion cycle, discourse is amplified and replicated by

champions of a fashionable innovation, leading to the recruiting of new champions. In such a cycle, the discourse of the innovation is disseminated outward towards new potential users. The upswing of a cycle is assumed to reflect emotional sentiment, while its downswing is assumed to exhibit rational sentiment [9]. Potential users of an innovation might be pushed towards conformity by observing fashion cycles in which a high adoption rate of an innovation is perceived [9, 36].

A valid tool by which to quantify fashion discourse is online search statistics, measuring the number of search queries for a given item. Prior research has used online search trends as an indicator of public interest in companies or stocks, in order to develop trading strategies [37]. Empirical results in these studies suggest that online search activity is significantly related to trading activity and the attention of retail investors [38, 39] and, in particular, uninformed individuals [38]. Furthermore, online search trends predict short-term increases in stock prices [39]. While keywords from ICO whitepapers have been used before to predict the ICO result [22], we aim to investigate a relationship between general societal trends and specific ICO results. This is analogous to the concept of bandwagon effects [40], in which trends spread on a large scale, build positive feedback loops [35], and pressure individuals to conformity. In the context of search trends, this is characterized by uninformed investors' pressure as a product of the fear of missing out (FoMO), the anxiety of being absent from potentially rewarding experiences [41]. Based on these theoretical and empirical foundations, we propose the following hypothesis:

*Hypothesis 2. The online search trends regarding blockchain technology and ICOs as a financial instrument is positively linked to the volume raised.*

### 3.4. Twitter sentiment and its relation to market success

Previous studies have pointed out the relevance of sentiment in economics: A recent study has found that negative sentiment in a column of the *Wall Street Journal* predicts downward pressure on market prices, while high sentiment values in either direction predict high trading volumes [42].

Twitter is a common data source for sentiment analysis [43], particularly for the prediction of financial markets, e.g., [44, 45]. These studies use the overall sentiment with respect to specific hashtags or company names to predict future stock market developments. In this study, we focus on the Twitter account of the company issuing cryptographic coins or tokens.

In contrast to IPOs, where companies are forced to publish a large amount of information in advance, including a company prospectus [28], ICOs are unregulated and heterogeneous [22, 15, 18]. Accordingly, there are no reliable information sources for potential investors. Companies will likely adapt and use alternative channels for signaling. Domain experts specifically suggest using social media for ICO marketing, including Twitter [46]. In addition, [22] found that among other social media channels, Twitter has the highest relevance for the success of an ICO. Hence, the Twitter account of the issuing company can be seen as an important source of information for investors. Concordant with the related literature, we expect positive sentiment to correspond to positive future developments. Accordingly, our hypothesis states:

*Hypothesis 3. An increase in the average sentiment of tweets from a startup's account preceding the ICO is positively linked to the volume raised.*

### 3.5. Time-variant emotionality and market success

In the context of IPOs, unemotional and conservative language has been found to be related to underpricing of stocks and lower short-term returns [47]. In addition, emotionally-charged messages typically strengthen opinions in a one-sided direction and have the potential to distract from relevant factual information [48]. Moreover, we expect ICO investors to be driven by emotionally-charged content and non-informative wording [49].

According to management fashion theory, emotionally-charged, enthusiastic, and unreasoned discourse characterizes the upswings of management fashion waves, while more reasoned, unemotional, and qualified discourse characterizes their downswings [35]. Since investment in blockchain technology was in the upswing phase of the underlying fashion cycle during the observation period [22], we expect to observe increasingly emotive language towards the ICO end date. We, therefore, hypothesize the following:

*Hypothesis 4. A continuous increase in the fraction of emotive language in the tweets with the approach of the ICO end date is positively linked to the volume raised.*

### 4. Data and methods

In this section, we describe the research process depicted in Figure 2. Our dataset originates from the website icodata.io in the version of March 15, 2018. This website contains a table with information about

recent ICOs. We collect the data of all completed ICOs since 2017. We exclude any previous data since according to Figure 1, major public interest in ICOs started in the early months of 2017. For each ICO, we collect the following information: (i) the name of the coin, (ii) the final date of the ICO, (iii) the raised volume in dollars, (iv) the initial price of the coin and (v) the name of the corresponding Twitter account. We remove all ICOs that did not create at least one tweet preceding their ICO date and all ICOs with a raised volume of zero. We account for outliers by removing ICOs with a raised volume in the highest percentile.
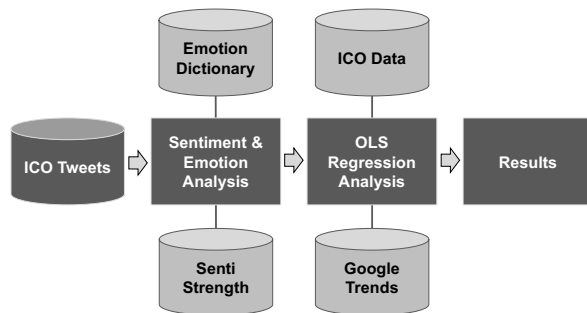


**Figure 2. Research process.**

As a next step, we download the timelines of these Twitter accounts and extract the number of followers and friends. In some cases, the timeline cannot be downloaded because some accounts have been deleted or set to private. The Twitter API limits the retrieval of historic tweets to 3,200 per user[1]. However, we assume that this limit does not affect our results as the average total number of tweets per account (including all tweets before and after the ICO end date) is only 444.83 with a standard deviation of 591.78. These accounts have created between 3 and more than 3,200 tweets prior to the date of the respective ICO. The mean number of tweets during the ICO process is 271.98 with a standard deviation of 437.66.

Finally, we determine the Google Search Trends[2] of the terms "Initial Coin Offering", "cryptocurrency" and "blockchain" starting from 2017. This data is scaled between 0 and 100 and is obtained on a weekly basis.

In the following, we describe the methods for creating the features required to test our hypotheses.

---

[1]https://developer.twitter.com/en/docs/tweets/timelines/api-reference/get-statuses-user_timeline.html, Twitter Developer Documentation, last lookup on September 15, 2018.

[2]https://trends.google.de/trends/

## 4.1. Feature extraction using NLP tools

In order to evaluate our hypotheses 3 and 4, we calculate several linguistic features for each tweet in our dataset. First, we use *SentiStrength* [50] to calculate the sentiment of each tweet. We opted for SentiStrength due to its simplicity and practicability. This tool is a common choice for sentiment analysis of short texts, e.g. [51, 52, 53, 54, 55]. Moreover, it has been shown to outperform supervised learning models [50]. SentiStrength determines the sentiment of short texts on two scales: a positive scale between 1 (neutral) and 4 (very positive), as well as a negative scale between -1 (neutral) and -4 (very negative). We follow the approach described by [53] of summing up the positive and the negative score. The resulting sentiment score contains values between -3 (very negative) and 3 (very positive). Second, we use the frequently-employed NRC dictionary [56] to determine the fraction of emotive words in each tweet. This value is located between 0 (no emotive words) and 1 (emotive words only).

In a next step, we calculate the average sentiment and average fraction of emotive words for each day preceding the ICO end date. We use this time series to determine the slope of the regression line going through the means of emotive words. The coefficient of the time variable indicates whether the use of emotive words is increasing or decreasing as the ICO's end date approaches. In the following, we describe how we integrate these variables into our regression model.

## 4.2. Regression analysis

Our goal is to explain the relationship between social media and the success of an ICO. Specifically, we aim to explain the amount of money (in USD) raised during the ICO process. In order to compensate for the skew in the distribution of raised volumes, we perform a log transformation on our target variable. This transformation also allows us to make more reasonable statements about the regression coefficients since potential effects are now describing level changes in the raised volume instead of changes in the absolute amount.

We use an OLS regression model to explain $Ln(RaisedVol)$ as described in Equation (1). Following previous research [26, 25, 22], we incorporate several control variables into our model: (i) the age of the Twitter account with respect to the ICO date ($AccountAge$), (ii) the number of followers ($NFollowers$), (iii) the number of friends ($NFriends$) and (iv) the total number of tweets during the ICO process ($NTweets$).

$$Ln(RaisedVol) = \beta_0 + \beta_1\, AccountAge \qquad (1)$$
$$+\, \beta_2\, InitialPrice + \beta_3\, NFollowers$$
$$+\, \beta_4\, NFriends + \beta_5\, NTweets$$
$$+\, \beta_6\, MGoogleTrend$$
$$+\, \beta_7\, MSentiment$$
$$+\, \beta_8\, SlopeEmotion + \epsilon.$$

The key variables in our model are the average Google trend for all three keywords ($MGoogleTrend$), the mean sentiment in all tweets preceding the ICO ($MSentiment$) and the slope of the linear regression line going through the mean fractions of emotive words ($SlopeEmotion$). We standardize all input variables in order to make a reasonable comparison between different effect sizes.

## 5. Discussion of results

We use the above regression design from Equation (1) to measure the impact of our suspected determinants on ICOs' success in terms of raised volume. The respective results are depicted in Table 1.

The regression results for our **baseline model** are shown in column $a$) of Table 1. The number of people following a blockchain startup on Twitter ($NFollowers$) is highly significant and positively linked to the volume raised during the ICO process. This finding is expected and in concordance with the theoretical foundation. The number of accounts the blockchain project is following ($NFriends$) and the total number of tweets ($NTweets$) are both statistically significant at the 5% level, which corresponds to $NFollowers$ as indicators for the overall social media strategy of a company and was therefore expected beforehand. The variable $AccountAge$, representing the creation date of the account, is not significant in the baseline model.

In order to test **Hypothesis 1**, we look at the coefficient of $InitialPrice$. In contrast to neoclassical economics which strictly stresses the relevance of price setting strategies, we expect the initial token prices to be negligible in the blockchain community. Our hypothesis is supported as the coefficient of $InitialPrice$ is not significant in any of the four model stages. This finding is in concordance with behavioral economics [24, 34], modern finance [26] and the first ICO publications [17].

Subsequently, we test **Hypothesis 2**. To that end, we add the variable $MGoogleTrend$ to our model. The results of this model are shown in column $b$) of Table 1. The coefficient of $MGoogleTrend$ is significant and positive ($\beta = 0.548, p < 0.001$). This means, that an increase of one standard deviation in the Google trend is positively linked to an increase in the raised volume of $e^{0.548} - 1 \approx 72.94\%$. Building on top of established work [40, 36] and concurrent research [38, 37, 22], these results support our hypothesis that fashion cycle upswings [9] are represented in blockchain projects' financial success.

In order to test **Hypothesis 3**, we add the variable $MSentiment$ to our model. The results of this model stage are shown in column $c$). The coefficient of $MSentiment$ is significant and positive ($\beta = 0.408, p < 0.001$). This means that an increase of one standard deviation in the average sentiment prior to the ICO date is positively linked to an increase in the raised volume of 50.34%. These results support our assumptions that the findings from IPO analysis [29, 28] can be transferred to the specific blockchain ICO community, as qualitative research has formerly indicated [21].

Finally, we test **Hypothesis 4**. We hypothesize that emotive language becomes more prevalent towards the end of the ICO. For this purpose, we add the variable $SlopeEmotion$ to our model. The results of this model are shown in column $d$). The coefficient of $SlopeEmotion$ is significant and positive ($\beta = 0.243, p < 0.05$). This means that an increase of one standard deviation is positively linked to an increase of 27.51% in the raised volume. This finding is supported by [35, 44].

The results of our four hypotheses are robust throughout all model stages. While the analysis of the full sample prior to the outlier removal yielded similar results with stronger effect sizes, we decided to focus on a more representative sample to safeguard generalizability. Furthermore, we standardized all predictor variables for our regression, which does not change the results in general but facilitates the interpretation of the model's effect sizes.

Our analysis is limited by a number of missing values in the Twitter dataset. Some blockchain projects do not follow a comprehensive social media strategy and either neglect the utilization of this medium altogether or else engage in limited activities. Startups located in countries with comparatively few English speakers (e.g. Japan) often use social media only in their native language. In addition, the observation period of this study was too short to investigate mid-term success. Another caveat is the dependence on a single ICO-listing website, which may be incomplete. We, therefore, aim to maximize the Twitter dataset by a) improving our web crawler to consider multiple ICO listing sites and b) enhance the dataset through a manual web search for missing entries.

**Table 1. Regression Results: Raised ICO volume on Google Trend, Twitter sentiment and emotionality.**

| Variable | (a) | (b) | (c) | (d) |
|---|---|---|---|---|
| *Intercept* | 14.462*** (0.116) | 14.462*** (0.114) | 14.462*** (0.113) | 14.462*** (0.112) |
| *AccountAge* | 0.173 (0.130) | 0.217 (0.127) | 0.218 (0.126) | 0.214 (0.126) |
| *InitialPrice* | −0.029 (0.116) | −0.020 (0.114) | −0.004 (0.113) | −0.008 (0.113) |
| *NFollowers* | 0.809*** (0.117) | 0.935*** (0.118) | 0.912*** (0.117) | 0.905*** (0.117) |
| *NFriends* | −0.303* (0.122) | −0.236 (0.121) | −0.236* (0.119) | −0.236* (0.119) |
| *NTweets* | 0.326* (0.135) | 0.226 (0.135) | 0.265* (0.134) | 0.257 (0.133) |
| *MGoogleTrend* | | 0.548*** (0.119) | 0.545*** (0.118) | 0.537*** (0.117) |
| *MSentiment* | | | 0.408*** (0.114) | 0.403*** (0.113) |
| *SlopeEmotion* | | | | 0.243* (0.113) |
| Observations | 524 | 524 | 524 | 524 |
| $R^2$ | 0.111 | 0.146 | 0.167 | 0.174 |
| Adjusted $R^2$ | 0.102 | 0.136 | 0.155 | 0.161 |

*Stated:* coeffficient and std. dev in parenthesis      $^*p<0.05$; $^{**}p<0.01$; $^{***}p<0.001$

These results can help retail investors and entrepreneurs to understand the social factors, such as fashion cycles and social media signals, related to successful ICOs. Entrepreneurs planning to start an ICO should be aware, that technological excellence and sound business plans may not have the same relevance for ICOs as for conventional markets. Accordingly, neglecting social media may have a negative impact on capitalization success. The same applies to price setting strategies, which may not yield the expected results. For retail investors, analyzing social media sentiment can be helpful in approximating whether an ICO will be successful or not. However, as the ICO market is characterized by high uncertainty, it is hardly predictable. For instance, in conventional markets, an increasing level of emotive speech might be considered untrustworthy and associated investment opportunities may suffer a negative impact. An increasing level of emotive speech during ICOs, on the other hand, was found to have a positive financial impact. Due to this uncertainty, ICOs may qualify for a certain degree of governance or international regulation from the perspective of legislators. However, reducing this uncertainty does not necessarily have to be the responsibility of government agencies. This niche may provide the opportunity for new business models regarding the evaluation of ICOs, which can cover the analysis of prices, social media sentiment or entrepreneurs' backgrounds. While countless websites already offer similar services, winning the trust of retail investors is hardly easy for an unknown web-service. Instead, established rating agencies and consultancies may have an opportunity to collaborate with analytics service providers to position their own brand within this market.

## 6.   Conclusion

In this paper, we investigated how the financial success of firms' ICOs are linked to social factors. In doing so, we addressed blockchain research gaps regarding value proposition and associated risks identified in previous research. For this purpose, we have created web crawlers to mine financial data on 524 completed ICOs and a total of 231,758 status updates from their Twitter accounts between January 2017 and March 2018. Using established NLP tools, we extracted features in order to measure the sentiment and emotionality of the mined tweets. We then enhanced out dataset with online search statistics. Finally, we used

an OLS regression model to assess the link between volume raised during an ICO and four behavioral factors: The sentiment of social media activities in polarity and emotiveness, strategic price setting and trend developments.

We found evidence of significant relationships between startups' raised volume and a) the general blockchain discourse as measured by search trends; b) their average Twitter sentiment; c) increasing emotionality in their tweets towards the ICO end date. While other intuitively expected controls, such as number of followers or friends, yielded significant results, the initial price of the issued token had no significant impact on the volume raised.

These results support the assumptions derived from our theoretical and empirical foundation, according to which we assumed blockchain technology and ICOs as a financial instrument to be subject to fashion cycles in an uncertain market environment in which startups utilize the underlying discourse as a signaling vehicle.

Our findings are relevant for retail investors, ICO entrepreneurs, legislators and analytics service providers. ICO entrepreneurs are advised to learn the targeted utilization of sentiment as part of their business strategy. Retail investors are advised to cautiously consider all available information sources, including social media. As most investors may not have access to the respective assessment tools, service providers and consultancies can help to reduce market uncertainty. Finally, legislators may be incentivized to impose further regulations to reduce the necessity of relying on alternative data sources for ICO assessments. In future research, we aim to address the aforementioned limitations by enhancing the dataset in terms of a longer observation period and more social media platforms.

# References

[1] F. Glaser, "Pervasive decentralisation of digital infrastructures: a framework for blockchain enabled system and use case analysis," in *Proceedings of the 50th Hawaii International Conference on System Sciences*, 2017.

[2] D. Tapscott and A. Tapscott, *Blockchain Revolution: How the technology behind bitcoin is changing money business, and the world*. New York, USA: Penguin Publishing Group, 2016.

[3] M. Nofer, P. Gomber, O. Hinz, and D. Schiereck, "Blockchain," *Business & Information Systems Engineering*, vol. 59, no. 3, pp. 183–187, 2017.

[4] D. Floyd, "2018 ICO Funding Has Passed 2017's Total - CoinDesk," 2018.

[5] J. Li and W. Mann, "Initial Coin Offering and Platform Building," *SSRN Electronic Journal*, 2018.

[6] U. W. Chohan, "Initial Coin Offerings (ICOs): Risks, Regulation, and Accountability," *SSRN Electronic Journal*, 2017.

[7] M. Risius and K. Spohrer, "A Blockchain Research Framework - What We (don't) Know, Where We Go from Here, and How We Will Get There," *Business & Information Systems Engineering*, vol. 59, no. 6, pp. 385–409, 2017.

[8] Gartner, "Hype Cycle for Emerging Technologies," 2017.

[9] R. L. Baskerville and M. D. Myers, "Fashion Waves in Information Systems Research and Practice," *MIS Quarterly*, vol. 33, no. 4, p. 647, 2009.

[10] P. Hacker and C. Thomale, "Crypto-Securities Regulation: ICOs, Token Sales and Cryptocurrencies under EU Financial Law," 2017.

[11] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System." 2008.

[12] F. Hartmann, X. Wang, and M. I. Lunesu, "Evaluation of initial cryptoasset offerings: the state of the practice," in *International Workshop on Blockchain Oriented Software Engineering*, pp. 33–39, 2018.

[13] V. Buterin, "Ethereum White Paper - A Next-Generation Smart Contract and Decentralized Application Platform." 2013.

[14] F. Vogelsteller and V. Buterin, "ERC-20 Token Standard," 2017.

[15] C. Hahn and A. Wons, *Initial Coin Offering (ICO)*. Springer, 2018.

[16] C. Shapiro and H. R. Varian, *Information rules: a strategic guide to the network economy*. Harvard Business Press, 1998.

[17] G. Fenu, L. Marchesi, M. Marchesi, and R. Tonelli, "The ICO phenomenon and its relationships with ethereum smart contract environment," in *International Workshop on Blockchain Oriented Software Engineering*, pp. 26–32, 2018.

[18] S. Adhami, G. Giudici, and S. Martinazzi, "Why do businesses go crypto? An empirical analysis of Initial Coin Offerings," *Journal of Economics and Business*, 2018.

[19] M. Fiedler and P. Sandner, "Identifying Leading Blockchain Startups on a Worldwide Level," *FSBC Working Paper*, 2017.

[20] L. Kostovetsky and H. Benedetti, "Digital Tulips? Returns to Investors in Initial Coin Offerings," *SSRN*, 2018.

[21] Y. Mohit, "Exploring signals for investing in an Initial Coin Offering (ICO)," *SSRN Electronic Journal*, p. 14, 2017.

[22] S. Jin, R. Ali, and A. V. Vlasov, "Cryptoeconomics: Data Application for Token Sales Analysis," 2017.

[23] T. D. Wilson, "Models in information behaviour research," *Journal of Documentation*, vol. 55, no. 3, pp. 249–270, 1999.

[24] G. A. Akerlof, "The market for lemons: Quality uncertainty and the market mechanism," in *Uncertainty in Economics*, pp. 235–251, Elsevier, 1978.

[25] D. L. Deeds, D. Decarolis, and J. E. Coombs, "The impact of firmspecific capabilities on the amount of capital raised in an initial public offering: Evidence from the biotechnology industry," *Journal of Business Venturing*, vol. 12, no. 1, pp. 31–46, 1997.

[26] S. T. Certo, "Influencing Initial Public Offering Investors with Prestige: Signaling with Board Structures," *The Academy of Management Review*, vol. 28, no. 3, p. 432, 2003.

[27] T. Loughran and B. McDonald, "Regulation and financial disclosure: The impact of plain English," *Journal of Regulatory Economics*, vol. 45, no. 1, pp. 94–113, 2014.

[28] T. Loughran and B. McDonald, "IPO first-day returns, offer price revisions, volatility, and form S-1 language," *Journal of Financial Economics*, vol. 109, no. 2, pp. 307–326, 2013.

[29] G. Ranco, D. Aleksovski, G. Caldarelli, M. Grčar, and I. Mozetič, "The Effects of Twitter Sentiment on Stock Price Returns," *PLOS ONE*, vol. 10, no. 9, p. e0138441, 2015.

[30] S. A. Macht, "Reaping Value-Added Benefits from Crowdfunders: What Can We Learn from Relationship Marketing?," *Strategic Change*, vol. 23, no. 7-8, pp. 439–460, 2014.

[31] C. Catalini and J. S. Gans, "Initial coin offerings and the value of crypto tokens," tech. rep., 2018.

[32] G. K. C. Ahlers, D. Cumming, C. Günther, and D. Schweizer, "Signaling in Equity Crowdfunding," *Entrepreneurship Theory and Practice*, vol. 39, no. 4, pp. 955–980, 2015.

[33] J. R. Nofsinger, *The psychology of investing*. Routledge, 2016.

[34] D. Kahneman and M. W. Riepe, "Aspects of investor psychology," *Journal of Portfolio Management*, vol. 24, no. 4, pp. 52–65, 1998.

[35] E. Abrahamson and G. Fairchild, "Management Fashion: Lifecycles, Triggers, and Collective Learning Processes," *Administrative Science Quarterly*, vol. 44, no. 4, p. 708, 1999.

[36] P. J. DiMaggio and W. W. Powell, "The Iron Cage Revisited: Institutional Isomorphism and Collective Rationality in Organizational Fields," *American Sociological Review*, 1983.

[37] T. Preis, H. S. Moat, and H. E. Stanley, "Quantifying trading behavior in financial markets using Google Trends," *Scientific Reports*, vol. 3, p. srep01684, 2013.

[38] M. Bank, M. Larch, and G. Peter, "Google search volume and its influence on liquidity and returns of German stocks," *Financial Markets and Portfolio Management*, vol. 25, no. 3, p. 239, 2011.

[39] Z. Da, J. Engelberg, and P. Gao, "In search of attention," *The Journal of Finance*, vol. 66, no. 5, pp. 1461–1499, 2011.

[40] H. Leibenstein, "Bandwagon, Snob, and Veblen Effects in the Theory of Consumers' Demand," *The Quarterly Journal of Economics*, vol. 64, no. 2, p. 183, 1950.

[41] A. K. Przybylski, K. Murayama, C. R. DeHaan, and V. Gladwell, "Motivational, emotional, and behavioral correlates of fear of missing out," *Computers in Human Behavior*, vol. 29, no. 4, pp. 1841–1848, 2013.

[42] P. C. Tetlock, "Giving content to investor sentiment: The role of media in the stock market," *The Journal of Finance*, vol. 62, no. 3, pp. 1139–1168, 2007.

[43] A. Pak and P. Paroubek, "Twitter as a corpus for sentiment analysis and opinion mining," in *Language Resources and Evaluation*, vol. 10, pp. 1320–1326, 2010.

[44] J. Bollen, H. Mao, and X. Zeng, "Twitter mood predicts the stock market," *Journal of Computational Science*, vol. 2, no. 1, pp. 1–8, 2011.

[45] J. K.-S. Liew and G. Z. Wang, "Twitter sentiment and ipo performance: A cross-sectional examination," *Journal of Portfolio Management*, vol. 42, no. 4, p. 129, 2016.

[46] K. Shilov, "What should your ICO marketing plan look like in 2018?," 2018.

[47] S. P. Ferris, Q. Hao, and M.-Y. Liao, "The effect of issuer conservatism on IPO pricing and performance," *Review of Finance*, vol. 17, no. 3, pp. 993–1027, 2012.

[48] K. Sweeny, D. Melnyk, W. Miller, and J. A. Shepperd, "Information avoidance: Who, what, when, and why.," *Review of General Psychology*, vol. 14, no. 4, p. 340, 2010.

[49] N. Pröllochs, S. Feuerriegel, and D. Neumann, "Is human information processing affected by emotional content? understanding the role of facts and emotions in the stock market," in *Proceedings of the 37th International Conference on Information Systems*, 2016.

[50] M. Thelwall, K. Buckley, G. Paltoglou, D. Cai, and A. Kappas, "Sentiment strength detection in short informal text," *Journal of the American Society for Information Science and Technology*, vol. 61, no. 12, pp. 2544–2558, 2010.

[51] E. Guzman, D. Azócar, and Y. Li, "Sentiment analysis of commit comments in GitHub: an empirical study," in *Proceedings of the 11th Working Conference on Mining Software Repositories*, pp. 352–355, ACM, 2014.

[52] M. Thelwall, K. Buckley, and G. Paltoglou, "Sentiment strength detection for the social web," *Journal of the Association for Information Science and Technology*, vol. 63, no. 1, pp. 163–173, 2012.

[53] S. Stieglitz and L. Dang-Xuan, "Emotions and information diffusion in social mediasentiment of microblogs and sharing behavior," *Journal of Management Information Systems*, vol. 29, no. 4, pp. 217–248, 2013.

[54] L. Dang-Xuan, S. Stieglitz, J. Wladarsch, and C. Neuberger, "An investigation of influentials and the role of sentiment in political communication on Twitter during election periods," *Information, Communication & Society*, vol. 16, no. 5, pp. 795–825, 2013.

[55] F. Calefato, F. Lanubile, and N. Novielli, "How to ask for technical help? Evidence-based guidelines for writing questions on Stack Overflow," *Information and Software Technology*, vol. 94, pp. 186–207, 2018.

[56] S. M. Mohammad and P. D. Turney, "Emotions evoked by common words and phrases: Using mechanical turk to create an emotion lexicon," in *Proceedings of the NAACL Workshop on Computational Approaches to Analysis and Generation of Emotion in Text*, pp. 26–34, 2010.