ARTICLE

Check for updates

The dark web privacy dilemma: linguistic diversity, talkativeness, and user engagement on the cryptomarket forums

Zhicong Chen [1], Xiang Meng[2] & Cheng-Jun Wang [1✉]

The users of the Dark Web require a secure and highly anonymous environment to exchange information while protecting their online privacy, which presents a privacy dilemma. This paper examines the digital footprints of user behavior on the three most popular cryptomarket forums on the Dark Web, namely Silk Road 1, Silk Road 2, and Agora. The results indicate that users who engage in more conversations and employ a wider range of vocabulary are more likely to discontinue their participation on the forum. Intriguingly, no significant relationship is found between network characteristics and user engagement. These findings emphasize that the risk of exposure within anonymous communities primarily stems from the potency of information rather than social connections, which sheds light on the privacy dilemma inherent in the Dark Web and provides deeper insights into the online user behavior surrounding anonymity-granting technologies on the Internet.

[1] School of Journalism and Communication, Nanjing University, Nanjing, China. [2] Department of Media and Communication, City University of Hong Kong, Hong Kong SAR, China. ✉email: wangchengjun@nju.edu.cn

Introduction

The users of the Dark Web forums have a strong need to effectively communicate with other members while protecting their privacy. Theoretically, there is a tradeoff between online self-disclosure and privacy protection. Inspired by the seminal research of social penetration theory (Altman and Taylor, 1973) and communication privacy management theory (Petronio, 2002), as information exchanges increase, there is a higher possibility that user identities in real life could be revealed. Social interactions also play a competing role in privacy calculus (Jiang et al. 2013; Krasnova et al. 2010; Laufer and Wolfe, 1977). For example, Lewis et al. (2008) show that people's taste for privacy is shaped by both social influence and personal motives. When the cumulative risk is larger than the future benefit, the user has to leave the community. The question pertains to how they smell the danger. We argue that there exists a prospective theoretical puzzle in explaining anonymous users' withdrawal behaviors: which of the two factors, information intensity or network connectivity expose anonymous users? By tracking how users leave the forums on the Dark Web, we can take a close look on the mechanisms of social penetration and privacy management in a fully anonymous setting.

Prior research on the Dark Web has left notable gaps that this study aims to fill. First, the bulk of empirical investigations into the Dark Web has predominantly centered on its economic dimensions, particularly focusing on operations within Dark Web markets, as demonstrated by a range of studies (Aldridge and Decary-Hetu, 2016; Barratt and Aldridge, 2016; ElBahrawy et al. 2020; Hardy and Norgaard, 2016; Martin, 2014b; Paquet-Clouston et al. 2018). Nevertheless, there is a scarcity of studies that have approached the Dark Web from a communication standpoint. This fully anonymous environment offers a communication context markedly distinct from traditional online communities on the Internet, characterized by unique patterns of online interaction. Second, while qualitative and critical studies have explored and discussed the potential and social implications of the Dark Web technology (Bancroft and Reid, 2017; Maddox et al. 2015; Martin, 2014a; Munksgaard and Demant, 2016; Pace, 2016; Sotirakopoulos, 2018; Tzanetakis et al. 2016), very few studies have ventured into investigating online behavior with behavioral logs. Third, prevailing theories have predominantly centered on the social penetration of close relationships, but the intricacies of how individuals manage their privacy in fully anonymous social settings warrant more extensive exploration. The dynamics of how users on the Dark Web interact while remaining hidden in the shadows remain under-examined in the existing literature, offering potential theoretical contributions to the domains of anonymous communication and group socialization. Fourth, while existing theories have indeed identified various factors contributing to social penetration, such as culture, gender, motive, context, and the risk-benefit ratio, their relative importance has not been subjected to comprehensive comparative examination (Altman and Taylor, 1973; Petronio, 2002). We aim to fill these gaps and provide valuable insights into the Dark Web's communication dynamics, user interactions, and the different aspects of factors influencing them.

This study aims to bridge these research gaps by delving into the communication dynamics of the Dark Web and exploring online behavior through behavioral logs, offering a fresh perspective on this unique digital realm. First, by conceptualizing the cryptomarket forums as anonymous online communities, the study examines the sustainability of user activities within this context. Second, through investigating patterns and mechanisms within this unique online communication environment, we aim to enhance our understanding of user online behavior and the dynamics of group socialization in anonymous online communities. Third, utilizing online behavioral data from cryptomarket forums, the study investigates how language usage and online social networks influence the long-term usage of these forums. The overall findings of this study have substantial implications for understanding group dynamics and anonymous online communities.

In the following sections, we will begin by introducing the Tor network, the Dark Web, and the cryptomarket, which serve as the research contexts for this study. Subsequently, we will provide a comprehensive review of the theoretical perspectives on technical anonymity and anonymous communication in online communities. Furthermore, we will formulate two research questions pertaining to the potential factors influencing users' engagement in anonymous online communities. Using data collected from the digital traces of online user behavior, we will assess talkativeness, linguistic diversity, and online leadership within the online social network. Survival analysis will be employed to examine the relationships between these potential factors and user engagement in Dark Web marketplaces. Lastly, we will discuss the implications of the research findings.

Tor, dark web, and cryptomarket forums. The Tor (The Onion Router) network is a special type of network that is only accessible through the Tor browser. Just like the name "The Onion Router" goes, the structure of the Tor network is analogical to an onion, which consists of a group of relays operated by volunteer individual users or non-profit organizations. The relays work as a series of virtual tunnels for connecting the information provider (sender) and user (receiver). Therefore, the traffic on the Tor network could be hidden and encrypted by multiple layers and routed between different relays multiple times. With this careful design, user identities are fully anonymized and untraceable on the Tor network. In summary, the Tor anonymity network has provided a convenient way to be extremely anonymous in some special application contexts, which to some extent becomes an important alternative ecosystem on the Internet, usually known as the Dark Web. The primary emphasis of this study will center on user behavior within the Dark Web forums, with a specific focus on the forums associated with the cryptomarket.

Many anonymous services and websites hosted on the Dark Web contain unethical content (e.g., pedophilia and child pornography) and online criminal activity (e.g., drug markets, financial fraud, illegal weapons, and espionage) (Biryukov et al. 2014; Biryukov et al. 2013; Guitton, 2013; Soska and Christin, 2015). Among them, the marketplaces have been highly prosperous since 2011, formally known as *cryptomarkets* (Martin, 2014b). Cryptomarket is a special type of online market hidden on the "Dark Web". The technical basis of the Tor anonymity network has made anonymity a fundamental feature of how cryptomarkets are constructed and function as markets and as communities (Bancroft and Reid, 2017). In cryptomarkets, users are highly anonymized and well protected by the onion-routing technology behind the Tor network, which creates a very different space for communicating and trading with each other in many aspects of the communication process. Moreover, most cryptomarkets not only present and sell various products but also provide a forum for users to discuss valuable topics during the whole purchase process, such as the usage of the website, delivery of the products, and vendors' ratings. The market-affiliated forums have provided great opportunities for researchers to analyze user behavior in anonymous online communities. With the advanced privacy-enhancing technologies employed to protect user identities, cryptomarkets have shown great potential in changing the landscape of global drug distribution networks, leading to

worldwide opportunities for drug consumers and challenges for law enforcement (Nugent, 2019).

Starting from 2011, there have been several prosperous cryptomarkets, created and operated widely around the whole world through the Tor anonymity network. As one of the most popular types of services on the Tor network, in most cases, the cryptomarkets will provide affiliated forums for users to discuss the purchase process, the usage of the website, delivery of the products, vendors' ratings, etc. Topics discussed in the forums typically center around sense-making, technical difficulties, purchase processes, logistic issues, vendor and product reviews, and cyber security issues (Hazel Kwon and Shao, 2020; Porter, 2018). As the technical design of the Tor network has made anonymity a fundamental feature of the cryptomarket forums, users are highly anonymized and well protected by the onion routing technology behind the Tor network, which creates a very different space for communicating and trading with each other in many aspects of the communication process (Bancroft and Reid, 2017). Therefore, the cryptomarket-affiliated forums have provided great opportunities for researchers to analyze users' online behavior in a fully anonymous setting. In this study, Dark Web forums are examined as anonymous online communities.

Technical anonymity versus social anonymity in online interactions. Theoretical works in anonymous communication and organizational communication argue that there are several types and dimensions of anonymity, e.g., technical anonymity versus social anonymity, physical anonymity versus discursive anonymity, and self-anonymity versus other-anonymity (Rains and Scott, 2007; Scott, 1998; Scott et al. 2011). In the context of the Dark Web, it is especially important to clarify the technical anonymity and social anonymity provided by privacy-enhancing technologies. As previously introduced, technical anonymity is well ensured by the design of the onion routing network. However, this only provides part of the guarantees for user identity. Tor networks can't solve all anonymity problems. Tor network only focuses on protecting the data transmission procedure. There could still be potential risks caused by inappropriate online behavior by the user of the Tor network. Although technically well protected, making no chances for third-party tracking and surveillance, every user must be vigilant and careful about their self-disclosure when using the Tor network to ensure a high level of anonymity.

Technically, it is always possible to infer user identity from their social interactions and group socialization process with other people (Al Jawaheri et al. 2020; Biryukov et al. 2013). On the normal websites on the World Wide Web such as social networking sites, the group socialization process naturally occurs everywhere and never stops. Assisted by rapidly developed Web technologies, people can now share various types of content, e.g., text, photo, voice, video, etc. with others in a short time at almost any place. The enormous data about social interactions exposed on social media has already provided enough information for user identification and personality prediction, even without directly tracking users' IP addresses (Kosinski et al. 2013; Stachl et al. 2020). This could make a huge impact on Dark Web users. Any casual leak of personal information may lead to vital consequences, especially for some controversial or illegal activities. For example, the founder of Silk Road, Ross Ulbricht, got caught and arrested by law enforcement only because he had once accidentally exposed his personal email address on the Dark Web. Therefore, the Dark Web users are consistently faced with a social dilemma between self-disclosure (social interactions) and privacy protection (social anonymity) during the entire group socialization in anonymous online communities.

Much like conventional counterparts such as Reddit, which have attracted the attention of researchers across various disciplines in recent years (Medvedev et al. 2019; Parmentier and Cohen, 2019; Seering et al. 2019), the cryptomarket forums primarily serve as platforms for discussion and social interaction, enabling individuals with shared interests or opinions to engage over the Internet, all without face-to-face contact. Notably, Dark Web forums exhibit significant differences from mainstream online platforms like Reddit. Firstly, Dark Web users typically have limited real-life or offline social connections, leading to natural disconnection and separation among its users, unlike popular social networking sites such as Facebook and Twitter, which heavily rely on offline social relationships. Moreover, Dark Web access requires sophisticated technology and carries a high entry threshold, challenging ordinary users. Consequently, Dark Web users tend to be more motivated and active compared to users on mainstream social media platforms. The elevated entry requirements attract individuals with specific needs for a high degree of anonymity, often driven by economic incentives, which may be illegal or unethical. Lastly, although the Dark Web is perceived as an environment conducive to anonymous online interactions, achieving technical anonymity does not ensure complete social and perceived anonymity, as cautioned by the theory of anonymous communication. Users must exercise care in their online self-disclosure to prevent inadvertent disclosure or inference of private information during online social interactions.

In summary, the Dark Web forums differ from common online forums in terms of social connections, user motivation, and the need for anonymity. Understanding these distinctions is crucial for comprehending the dynamics of Dark Web communities and the unique challenges associated with online interactions in this context. Just like all communities, the first step for every user of anonymous online communities is to join the group. Then the new members need to learn the group norms through social interactions and communication with each other. Theoretically, the whole group socialization process can be categorized into five phases (Levine and Moreland, 1994; Moreland and Levine, 1982), namely investigation (entry), socialization, maintenance (acceptance), resocialization (divergence), and remembrance (exit). A successful socialization process could lead to good acceptance of the newcomers into the group, while a failed socialization process will just result in divergence of the group and the worst case is that the group member exit from the group. In this sophisticated process, the most important outcome and consequence is exit behavior, which completely ends all the social interactions between the group members.

User engagement in anonymous online communities. From the perspective of media use, one of the most important research questions is how and why users use and continuously use online communities as a media platform to communicate with others. As anonymity is an important affordance of computer-mediated communication, which could lead to substantial implications and consequences (Bancroft and Reid, 2017; Scott and Rains, 2020; Scott et al. 2011; Woo, 2006), this study pays special attention to the use and sustainable use of online communities in a fully anonymous setting, i.e., the anonymous online communities. Therefore, the exit behavior (or user sustainability in contrast way) on the Tor network is a focal variable in this study.

There has been a large body of research on continuance behavior in information systems. Bhattacherjee (2001) claims that users' continuance intention is determined by their satisfaction with the IS (information system) use and the perceived usefulness of continued IS use. Limayem and Cheung (2008) expanded Bhattacherjee's IS continuance model by adding a moderating

effect (IS habit) to IS continuance intention and IS continued usage. Bhaskar et al. (2019) decomposed the perceived usefulness into the perceived quality and perceived usability. In the context of open-source software (OSS) development, Wu et al. (2007) find that developers' feelings of satisfaction and their intentions to continue with OSS development were influenced by both helping behavior and economic incentives. Jin et al. (2010) find that users' continuance intention to participate in an online community is determined by both satisfaction and affective commitment.

On the Dark Web, Christin (2013) has analyzed the survivability of both users and products on the market of Silk Road 1 and found that about half of the sellers leave the site within 100 days of initial appearance; around a fifth of the sellers stay for less than three weeks. Most of the sellers will disappear within roughly three months since their arrival. Only a core of 112 sellers has been online throughout their whole observation interval. Similar findings were replicated on another 16 cryptomarkets from 2013 to 2015, including Silk Road, Agora, and Evolution (Soska and Christin, 2015) and also OpenBazaar (Arps and Christin, 2020). All the empirical observations indicate that vendors usually hold low stocks and operate primarily in the retail space, with small product quantities, low sales volumes, and a short lifetime in the cryptomarkets. These previous works have already paid attention to survival analysis on the cryptomarket ecosystems but are majorly limited to the analysis of products and vendors. Also, most of the work stops at describing the estimated Kaplan-Meier curve (the K-M curve), which only depicts the decaying process of user and vendor lifetime without further explorations on the potential explanations.

As most of the data collected on the cryptomarket forums are user behavior logs in the form of text on the webpage, it is necessary to develop various methods and strategies to characterize users' behavior according to the content they produced (Bagozzi and Dholakia, 2002; Chen and Liu, 2021). Previous studies have claimed that both language use and online social networks on the Dark Web demonstrate different patterns from the normal platforms on the Internet. For example, Choshen et al. (2019) find that the text related to legal and illegal activities on the Dark Web is different from each other in terms of not only word use but also shallow syntactic structure (represented through POS tags). Further, Zamani et al. (2019) examined the differences in the structure and dynamics of networks in both dark and public web forums. They find that the degree distribution in public forums is much more homogeneous than in Dark Web forums and even "semi-dark forums" (e.g., 8chan). Therefore, both language use and online social network characteristics should be considered when trying to understand user behavior in anonymous online communities.

Linguistic diversity and talkativeness. Language use can be used to reflect and infer many characteristics of the users, from the psychological process, e.g., affective processes and perceptual processes, to social processes, e.g., social concern and social support (Tausczik and Pennebaker, 2010). Among the psychological meanings of language use, Batikas and Kretschmer (2018) found that negative emotion is highly correlated to exit behavior in the Agora market. In other words, entrepreneurs with a high share of negative feedback are much more likely to exit the market. This finding is consistent with another work by Bhaskar et al. (2019), who also find that sellers with poor ratings are significantly more likely to exit. The finding is further confirmed across several different cryptomarkets, including Silk Road 2.0, Agora, Evolution, and Nucleus. All the previous findings suggest that online illegal markets function in a similar way to legal markets—the economic performance is significantly harmed by

negative ratings. Since negative ratings and comments under the vendor profiles in the anonymous online communities play essential roles in the group socialization process, we give special focus on the user sentiment when exploring the factors that explain the user survivability from textual features. Therefore, we propose the following research question:

RQ1: What is the relationship between the user's language use and exit behavior in anonymous online communities?

Online leadership in social network. Social network plays important roles in shaping markets, either online or offline (White, 1981). Previous studies on cryptomarkets show that buyers tend to exit the market instead of retaliation against sellers after negative experiences in social interactions with others (Norbutas et al. 2020). In many different types of online communities, social rewards such as enjoyment and reputation are important predictors of continuous engagement (Nov et al. 2010; Zhang et al. 2020). Also, social support (e.g., emotional support and informational support) gained from the online social network on the online communities could further motivate the user to keep staying in the online group, especially for health-related communities (Wang et al. 2017; Wang et al. 2012). From the perspectives of online influence and leadership, the more central position a user occupies in the network, the more likely he/she would become the leader and have more influence in the community, which could further enhance the commitment to the group (Huffaker, 2010; Johnson et al. 2015). Based on all the considerations, we propose the second research question:

RQ2: What is the relationship between the user's position in online social networks and exit behavior in anonymous online communities?

Methods
Data Collection. We analyze data from the forums affiliated to three popular cryptomarkets, Silk Road 1, Silk Road 2, and Agora. Silk Road 1 was the first online black market on the Tor network (Ormsby, 2014). It was initially established in February 2011 and later shut down by the Federal Bureau of Investigation (FBI) in October 2013 for large-scale illegal trade of drugs. After that, Silk Road 2 was established as a substitute for the previously closed market. At the same time, Agora was also launched, which lasted for another two years until August 2015. Agora even survived Operation Onymous in November 2014, when Silk Road 2 was seized by a joint law enforcement operation between the Federal Bureau of Investigation (FBI) and the European Union Intelligence Agency Europol. After Evolution closed in an exit scam in March 2015, Agora became the largest cryptomarket on the Dark Web.

All the data are collected from the publicly available Darknet Market Archives (Branwen et al. 2015). For the Silk Road 1 forum, the data range from the beginning of the forum (June 2011) to November 2013, including 39514 usernames and 81408 threads with 872961 posts. In Silk Road 2 forum, there are 39862 usernames and 29002 threads with 390086 posts from October 2012 to April 2014. Data collected on the Agora forum has a shorter time range from December 2013 to April 2014, containing 13590 users and 10666 threads with 84918 posts. Users are quite active and contribute substantially to the forum. For example, more than 3500 users (9%) on the Silk Road 1 forum wrote at least 50 posts during their lifetime. In Silk Road 2 and Agora forum, the percentages are 12 and 7%.

User engagement. Since all three cryptomarkets were taken down by exogenous forces (law enforcement), there is a censoring problem in the data collected, i.e., it is hard to identify whether it is an intended leave or due to the closure of the website. Therefore, we take a time window, e.g., 180 days for the Silk Road 1 forum, to define the survival status of a user (for the Silk Road 2 forum and Agora forum, we use 30 days as the period in the collected data is much shorter). For the whole observed period, if the number of days counting from the last post time to the end date of data collection, denoted as $T_{\text{lasti nactive}}$, is larger than 180 days, the user will be identified as a user who exits. Otherwise, the user will be labeled as right-censored. This can be written in a function as below:

$$\text{Status} = \begin{cases} \text{Exit, } T_{\text{last inactive}} > 180\,\text{days} \\ \text{Censored, } T_{\text{last inactive}} \leq 180\,\text{days} \end{cases}$$

The survival function $S(t)$, is the probability that a subject survives longer than time t. Let T be a random variable of the user lifetime, i.e., the length of time before a user exits from the forum. The survival function defines the probability that the death event has not occurred yet at time t, or equivalently, the probability of surviving past time t. $S(t)$ is theoretically represented by a continuous and smooth curve. Typically, it is approximated with the K-M curve using the Kaplan-Meier Estimator, which calculates the proportion of users surviving for a specific duration before the event takes place. Mathematically, it is defined as:

$$\widehat{S}(t) = \prod_{t_i \leq t} 1 - \frac{d_i}{n_i}$$

where t_i is the time when at least one event happened. d_i is the number of events (e.g., exit) at the time t_i and n_i is the number of individuals known to have survived (have not yet had an event or been censored) up to the time t_i.

Linguistic diversity and talkativeness. For the message-level characteristics, we measure two variables. The first is *Linguistic Diversity*, which is measured as the number of unique words found in a message, calculated by dividing the number of different words by the total number of words. This is also referred to as the *type/token ratio* (Huffaker, 2010) or *richness* (Spitters et al. 2015). The second is *Talkativeness*, which is the average length of messages contributed by each participant, i.e., the sum of the total words found in each message divided by the total number of messages.

Sentiment (positive/negative emotion). The valence of emotion in a piece of text is measured by the percentage of positive and negative words according to LIWC (Tausczik and Pennebaker, 2010). The emotional content of each user's posts is measured by calculating the average of both positive and negative emotions expressed in all of their messages.

Online leadership in social network. Except for the measurement of language use, we also seek measurements based on the online social network formed by user interactions. An edge is created between who replied to whom under each thread. For example, if user A (source) replies to user B (target), it will be considered as a directed edge in the network. In this network, various node-level structural measures can be calculated. To be specific:

Expansiveness is measured by outdegree centrality. This measure represents the number of outgoing links for each author (i.e., the number of users that this user has replied to), which is one of the fundamental factors that shape the formation of lasting identification as a member in a virtual community (Kozinets, 1999).

Reply Trigger is measured by in-degree centrality in the reply network, i.e., the total number of incoming links (i.e., the number of users that have replied to this user). This measure indicates the ability of a user to inspire and motivate other users to reply to a message.

Brokering in the reply network identifies which authors link the shortest paths or geodesics in the network. It is an indicator of to what extent a user acts as a "structural hole" (Burt, 2009) in the reply network.

Reciprocity is measured as the frequency of user interactions in a mutual dyad, i.e., both actors reply to one another, regardless of the order of the replies (Wasserman & Faust, 1994). When doing this, the networks were converted into symmetric relations, and the sum of reciprocal links was calculated for both authors.

User activity. Two variables concerning user activity were measured as control variables. The first one is the number of posts, measured by the total number of posts that a user contributes to the forum. In addition to that, we also measure the number of threads that are started by the user. Both measures indicate the commitment to the anonymous online community, which is expected to correlate to the final survival status (active or not) on the platform.

Analytical approach. Survival analysis is a widely used statistical model to analyze the duration of time until one or more events happen, such as a death in biological organisms and failure in mechanical systems. Usually, the event in survival analysis refers to death, disease occurrence, disease recurrence, recovery, etc. In this study, the exit behavior from the cryptomarket forums is considered as the event of interest for survival analysis.

The basic Cox proportional hazard model takes the log hazard ratio of an individual as a linear function of a group of time-independent covariates and a population-level baseline hazard that changes over time. Let $X = [X_1, ..., X_p]$ denote a group of covariates for an individual at time or duration t, and the log hazard ratio can be modeled by:

$$h(t|X) = h_0(t)exp(\beta X)$$

where $h_0(t)$ is the baseline probability that a user will exit at time t with $X = 0$, which reflects the change of hazard ratio over time. $e^{\beta i}$ represents the relative hazard, a proportionate increase or reduction in hazard, associated with X_i. With one unit increase in the covariate X_i (with all other covariates held constant), the probability that the user exits the forum at time t will increase by $e^{\beta j}$.

Results

The growth curves for both user and post activity in all three distinct cryptomarket forums are depicted in Fig. 1. The Silk Road 1 forum experienced the most rapid growth right from its beginning, signifying its widespread popularity within the 'Dark Web' community. The Silk Road 2 forum exhibited a growth pattern similar to Silk Road 1, with its rapid expansion following the closure of Silk Road 1 in October 2013. This abrupt shift in growth pattern underlines a strong substitution relationship between the Silk Road 1 and Silk Road 2 platforms. The Agora forum, another prominent cryptomarket that emerged after the shutdown of Silk Road 2, also displayed a growth pattern similar to that of the Silk Road 1 forum, despite its relatively short observation period of approximately four months since its inception. All three communities have undergone growth stages. Towards the end of the observed period, the growth of these
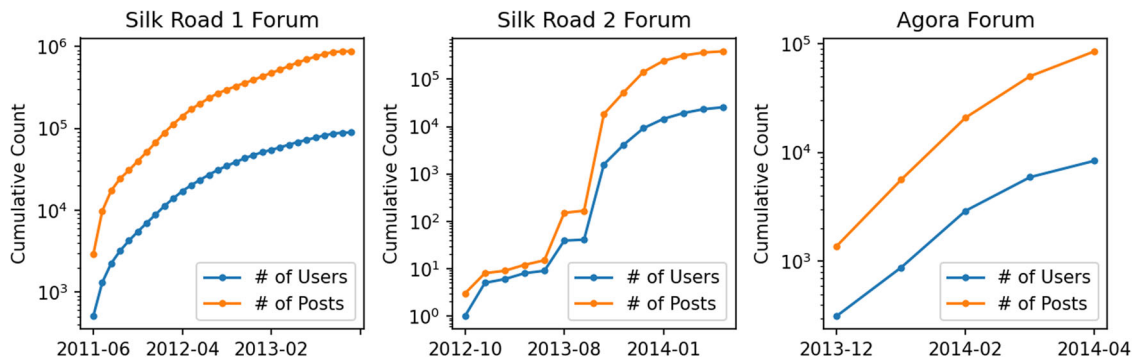
Fig. 1 S-curve of community growth on three cryptomarket forums. Three cryptomarket forums incudes the Silk Road 1 forum (left), the Silk Road 2 forum (middle), and the Agora forum (right). The blue line represents the number of users and the orange line represents the number of posts.
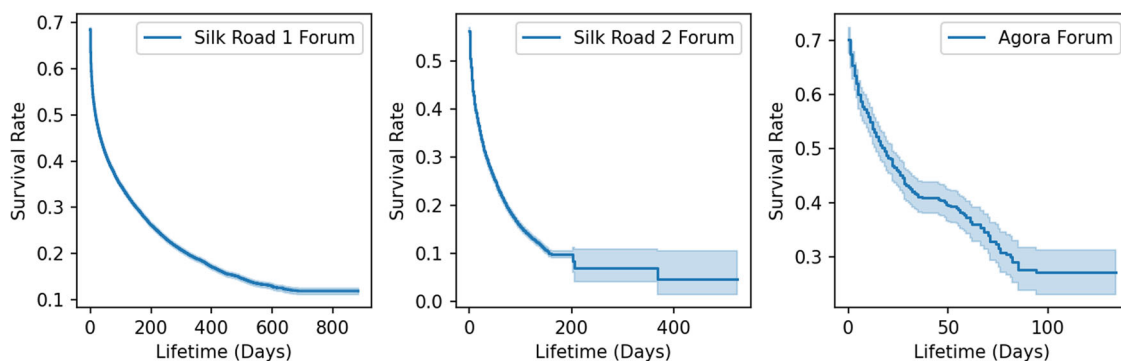
Fig. 2 Survival curves on three cryptomarket forums. Three cryptomarket forums includes the Silk Road 1 forum (left), the Silk Road 2 forum (middle), and the Agora forum (right).

communities has plateaued, leading to a flattened curve. This indicates that the critical mass of users has been achieved within the observed timeframe. These findings align well with the expected S-Curve as described in the Diffusion of Innovations theory (Rogers, 2003).

Figure 2 illustrates the survival curves of user lifetimes in the three forums. Users from all three platforms exhibit similar survival patterns. Roughly half of the users exit the platform within 30 days of joining, indicating a generally short user lifetime. Conversely, the number of users with longer lifetimes in anonymous online communities declines at a much slower rate. As user lifetimes increase, the survival curve becomes flatter. Among the three platforms, Silk Road 1 has the longest lifespan, with the most long-lasting user remaining active on the platform for over 800 days. Following law enforcement actions in 2013, variations in patterns have become more apparent in Silk Road 2 and the Agora forum.

Further, we investigate the relationship between overall lifetime linguistic diversity and survival status for each user across the three platforms. Cox proportional hazard regressions, as explained in the Methods part, are conducted. Both textual characteristics and user characteristics in the online social network are used as control variables. Table 1 presents the number of cases, means, and standard deviations for each variable calculated from the three platforms. The three platforms exhibit similarities in terms of mean and standard deviations, enabling examinations and comparisons across the platforms.

The regression results are shown in Tables 2, 3, and 4 for three forums, respectively. In the regressions, communication activities, i.e., the number of posts and number of threads started by the user are both negative predictors of the exit behavior. This result is very easy to understand as the survival status is defined by the posting activities. We recognize users who haven't posted for a

certain period as dead (leaving) users. Therefore, if a user has more posts or starts more threads, it will be less likely that he is a short-living user in the community. Besides, positive emotion ($B = -0.304$, 95% $C.I. = [-0.585, -0.023]$, $p < 0.05$) is a negative predictor of exit behavior. However, this relationship is found in neither Silk Road 2 nor the Agora forum.

Linguistic diversity and talkativeness turn out to be critical factors that explain the survival behavior in anonymous online communities. For regressions in the Silk Road 1 forum, linguistic diversity ($B = 2.523$, 95% $C.I. = [2.446, 2.600]$, $p < 0.001$) and talkativeness ($B = 0.154$, 95% $C.I. = [0.139, 0.168]$, $p < 0.001$) are both significant predictors that positively correlated to the proportional hazard, i.e., the probability to exit from the platform when we only control for activities. The relationship still exists even after the textual characteristics and network characteristics are controlled. The findings are well replicated on both Silk Road 2 and Agora forums. In the Silk Road 2 forum, linguistic diversity ($B = 3.311$, 95% $C.I. = [3.196, 3.425]$, $p < 0.001$) and talkativeness ($B = 0.921$, 95% $C.I. = [0.851, 0.991]$, $p < 0.001$) both significantly predicts the proportional hazard. The relationship still exists after controlling textual and network characteristics. In the Agora forum, linguistic diversity ($B = 2.634$, 95% $C.I. = [2.114, 3.154]$, $p < 0.001$) and talkativeness ($B = 1.698$, 95% $C.I. = [1.130, 2.267]$, $p < 0.001$) are both significant predictors to the proportional hazard, controlling for textual and network characteristics.

Discussion and conclusion
The analytical results in this paper yield several noteworthy findings. Firstly, rather than sentiment and online leadership, the analysis demonstrates that both linguistic diversity and

Table 1 Descriptive statistics on three cryptomarket forums.

Variables on Silk Road 1 Forum	N	Mean	SD	Min	Max
Variables on Silk Road 1 Forum					
Linguistic Diversity	39,249	0.609	0.234	0.0002	1.000
Talkativeness	39,249	0.086	0.300	0.00003	37.878
Positive Emotion	39,249	0.041	0.045	0.000	1.000
Negative Emotion	39,249	0.016	0.024	0.000	1.000
Expansiveness	37,638	10.097	34.271	0.000	1451.135
Reply Trigger	37,638	10.101	43.085	0.000	3741.255
Reciprocity	37,638	1.236	16.124	0.000	1,605.887
Brokering	37,638	0.008	0.034	0.000	1.000
Number of Posts	39,249	22.207	89.257	1	8404
Number of Threads Started	39,249	2.073	5.683	0	251
Variables on Silk Road 2 Forum					
Linguistic Diversity	13,960	0.583	0.237	0.002	1.000
Talkativeness	13,960	0.077	0.107	0.002	6.572
Positive Emotion	13,960	0.025	0.031	0.000	1.000
Negative Emotion	13,960	0.013	0.023	0.000	1.000
Expansiveness	13,017	9.440	29.977	0.000	1420.034
Reply Trigger	13,017	9.299	26.548	0.000	1010.448
Reciprocity	13,017	1.120	8.988	0.000	532.082
Brokering	13,017	0.019	0.066	0.000	1.000
Number of Posts	13,960	25.533	99.285	1	3836
Number of Threads Started	13,960	1.947	5.014	0	147
Variables on Agora Forum					
Linguistic Diversity	1421	0.568	0.232	0.016	1.000
Talkativeness	1421	0.077	0.087	0.002	1.586
Positive Emotion	1421	0.026	0.020	0.000	0.200
Negative Emotion	1421	0.012	0.016	0.000	0.384
Expansiveness	1421	1.632	4.658	0.000	92.650
Reply Trigger	1421	1.771	5.794	0.000	152.941
Reciprocity	1421	0.365	2.704	0.000	65.537
Brokering	1421	0.018	0.059	0.000	0.500
Number of Posts	1421	21.550	66.344	1	1301
Number of Threads Started	1421	1.856	4.666	0	75

Table 2 Cox proportional hazard model on Silk Road 1 forum.

	Model 1	Model 2	Model 3
Linguistic Diversity	2.523[***] (2.446, 2.600)		2.552[***] (2.472, 2.632)
Talkativeness	0.154[***] (0.139, 0.168)		0.399[***] (0.374, 0.423)
Positive Emotion	−0.304[*] (−0.585, −0.023)		−0.206 (−0.488, 0.077)
Negative Emotion	0.051 (−0.448, 0.550)		0.272 (−0.222, 0.766)
Expansiveness		−0.002 (−0.005, 0.002)	0.001 (−0.002, 0.004)
Reply Trigger		−0.005[***] (−0.007, −0.004)	−0.002[***] (−0.003, −0.001)
Brokering		0.036[***] (0.031, 0.041)	0.016[***] (0.012, 0.021)
Reciprocity		0.287 (−0.083, 0.658)	0.774[***] (0.414, 1.134)
Numberof Posts	−0.017[***] (−0.019, −0.016)	−0.031[***] (−0.033, −0.029)	−0.018[***] (−0.020, −0.016)
Numberof Threads Started	−0.010[**] (−0.017, −0.004)	−0.041[***] (−0.049, −0.032)	−0.010[*] (−0.017, −0.002)
Observations	39,125	37,609	37,609
R^2	0.348	0.282	0.353
Max. Possible R^2	1.000	1.000	1.000
Log Likelihood	−243,012.700	−234,178.800	−232,227.300
Wald Test	11,387.580[***] (df = 6)	5453.940[***] (df = 6)	11,409.340[***] (df = 10)
LR Test	16,760.060[***] (df = 6)	12,465.650[***] (df = 6)	16,368.660[***] (df = 10)
Score (Logrank) Test	14,982.520[***] (df = 6)	4350.028[***] (df = 6)	15,572.650[***] (df = 10)

Note: [*]$p < 0.05$; [**]$p < 0.01$; [***]$p < 0.001$.

talkativeness serve as significant predictors that correlate with user exit behavior. In regression analyses conducted on three cryptomarkets, only positive emotion exhibits significance on the Silk Road 1 forum, while neither positive nor negative emotions prove to be significant predictors of user exit behavior on the other two forums. Furthermore, centrality-based network measures (in-degree centrality, out-degree centrality, and brokering) are all found to be insignificant in predicting user exit behavior. This result is intriguing since network centrality is commonly considered an important indicator of social position and leadership in social networks. In summary, these findings indicate that the exposure risk in anonymous communities

Table 3 Cox proportional hazard model on Silk Road 2 forum.

	Model 1	Model 2	Model 3
Linguistic Diversity	3.311*** (3.196, 3.425)		3.259*** (3.140, 3.379)
Talkativeness	0.921*** (0.851, 0.991)		0.905*** (0.831, 0.978)
Positive Emotion	−0.649 (−1.306, 0.008)		−1.313** (−2.181, −0.445)
Negative Emotion	0.233 (−0.502, 0.968)		0.454 (−0.248, 1.156)
Expansiveness		−0.021*** (−0.024, −0.017)	−0.011*** (−0.014, −0.008)
Reply Trigger		−0.0003 (−0.002, 0.001)	0.002 (−0.0001, 0.003)
Brokering		0.024*** (0.015, 0.033)	0.008 (−0.003, 0.020)
Reciprocity		0.201 (−0.107, 0.509)	0.204 (−0.093, 0.500)
Numberof Posts	−0.004*** (−0.005, −0.003)	−0.009*** (−0.010, −0.007)	−0.001 (−0.002, 0.0003)
Numberof Threads Started	−0.038*** (−0.048, −0.029)	−0.093*** (−0.105, −0.082)	−0.048*** (−0.059, −0.038)
Observations	13,960	13,017	13,017
R^2	0.402	0.266	0.401
Max. Possible R^2	1.000	1.000	1.000
Log Likelihood	−93,154.150	−86,695.210	−85,365.200
Wald Test	6071.890*** (df = 6)	1896.790*** (df = 6)	5648.150*** (df = 10)
LR Test	7181.929*** (df = 6)	4018.018*** (df = 6)	6678.022*** (df = 10)
Score (Logrank) Test	7221.752*** (df = 6)	1712.071*** (df = 6)	6776.119*** (df = 10)

Note: **$p < 0.01$; ***$p < 0.001$.

Table 4 Cox proportional hazard model on Agora forum.

	Model 1	Model 2	Model 3
Linguistic Diversity	2.634*** (2.114, 3.154)		2.586*** (2.061, 3.111)
Talkativeness	1.698*** (1.130, 2.267)		1.668*** (1.093, 2.243)
Positive Emotion	−0.917 (−4.211, 2.376)		−0.990 (−4.287, 2.307)
Negative Emotion	2.173 (−2.555, 6.902)		2.119 (−2.563, 6.800)
Expansiveness		−0.151* (−0.292, −0.010)	−0.063 (−0.193, 0.067)
Reply Trigger		−0.035 (−0.100, 0.030)	−0.032 (−0.097, 0.033)
Brokering		0.351* (0.022, 0.679)	0.278 (−0.055, 0.611)
Reciprocity		−0.125 (−1.484, 1.235)	−0.298 (−1.668, 1.072)
Numberof Posts	−0.047*** (−0.058, −0.036)	−0.068*** (−0.084, −0.052)	−0.044*** (−0.058, −0.030)
Numberof Threads Started	−0.031 (−0.082, 0.020)	−0.089** (−0.152, −0.026)	−0.029 (−0.091, 0.033)
Observations	1421	1421	1421
R^2	0.396	0.359	0.398
Max. Possible R^2	0.999	0.999	0.999
Log Likelihood	−4903.077	−4946.219	−4901.457
Wald Test	400.870*** (df = 6)	242.180*** (df = 6)	404.440*** (df = 10)
LR Test	717.605*** (df = 6)	631.320*** (df = 6)	720.844*** (df = 10)
Score (Logrank) Test	597.275*** (df = 6)	194.346*** (df = 6)	598.499*** (df = 10)

Note: *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

primarily stems from the strength of information rather than social connections.

The users who engage in more extensive and varied conversations are more likely to leave the forum on the Dark Web. One possible explanation is that linguistic diversity and talkativeness collectively reflect the volume of information contained in the messages posted by users. It is important to consider the nature of cryptomarket forums in the analysis. The majority of discussions on these forums revolve around trading and beginner-oriented topics, such as learning how to successfully conduct purchases (Van Hout and Bingham, 2013). For instance, users need to familiarize themselves with using Bitcoin for transactions and understanding Escrow services to ensure product delivery (Lacson and Jones, 2016). In other words, every new user entering the cryptomarkets undergoes a knowledge accumulation process due to the distinct operations compared to conventional online markets (Chen et al. 2022). The slight increase from the 10 to 20% user life stage on the Silk Road 1

forum (as depicted in the figure) also suggests the presence of this learning process.

What motivates users to join and leave anonymous online communities? In a nutshell, social interactions are not the primary motivator for users in anonymous online communities; instead, they are driven by distinct and well-defined purposes, primarily centered around transactions between buyers and vendors in the cryptomarket forums (Tzanetakis et al. 2016). This differs from common online communities. With the progression of Web 2.0 technology, webpages have become more interactive, and social interactions have gained prominence across various online platforms (Bagozzi and Dholakia, 2002). Users tend to develop stronger attachments to online platforms, seeking information, social support, and engaging in conversations (Wang et al. 2017). Today, it is often assumed that people enter the online world to enhance their offline social relationships and daily lives (Kosinski et al. 2013). Nevertheless, the analysis of anonymous online communities presents a contrasting narrative

and reveals alternative possibilities for technological design in human-computer interactions.

To provide a more comprehensive insight into user dynamics on the Dark Web, we propose the Dark Web Privacy Dilemma, expanding upon and developing this concept in alignment with the existing concepts, such as the Dark Web Dilemma (Jardine, 2015), and the transparency paradox (Tzanetakis et al. 2016). The Dark Web poses a compelling privacy dilemma for its forum users, who strive to engage in effective communication while safeguarding their identities. In this realm, there exists an inherent tension between the desire for online self-disclosure and the imperative to protect one's privacy. As information exchanges intensify on Dark Web forums, the risk of revealing real-life identities from online conversations increases (Jardine, 2019). This raises concerns about the potential exposure and the subsequent consequences that may follow. When the cumulative risk outweighs the anticipated future benefits, users are compelled to make the difficult decision of leaving the community altogether.

Exploring the factors that prompt users to withdraw from Dark Web forums opens a window into the intricate interplay of privacy concerns and social interaction. It allows us to delve into the motivations, risk assessments, and decision-making processes that users undertake to safeguard their identities (Martin, 2014b). Understanding the privacy dilemma faced by Dark Web users not only contributes to theoretical advancements in the field but also has practical implications. The insights gleaned from this exploration can inform the design of privacy-enhancing technologies, policies, and educational initiatives that better align with users' needs for effective communication while preserving their privacy. One example of such an initiative is the creation of privacy discussion rooms (Doiciar and Cretan, 2021).

Overall, the contribution of this study is three-fold. First, the previous analysis mainly focuses on vendors and products (Christin, 2013; Christin, 2017; Soska and Christin, 2015). In this study, the unit of analysis is an individual user. This is the first work to examine the online social interactions between cryptomarket users with a focus on sustainable user behavior. Second, this study explored user continuity from a language and network perspective to the Cox model on the cryptomarkets. Most of the previous works merely plot the estimated Kaplan-Mayer curve to demonstrate the lifetime of either users or products. Third, the findings of the analysis imply a highly purpose-driven behavior in online anonymous communities, which is quite different from the commonly seen driving forces in normal online communities, where most of the users are intended to socially interact with each other.

The main shortcomings of this study come from data collection. Unlike common approaches in examining the sustainable use of media technology in the existing literature, such as surveys and experiments, all the measurements in this study are calculated from online behavior (posts), which makes it hard to collect individual-level demographics and perceptions. This limitation is largely constrained by the special design of the cryptomarkets, which forbid collecting user identity from the technical foundations. It is particularly hard to collect individual-level information using traditional survey due to the potentially strong social desirability as most of the behavior on cryptomarkets are illegal and unethical. Another limitation of this study is the measurement of information disclosure in online conversations. This study suggests that linguistic diversity and talkativeness together imply the volume of information delivered in the messages. The findings could be further confirmed by quantifying information through more refined methods.

Subsequent research endeavors may explore various promising directions. Firstly, acquiring additional data can enhance the robustness of findings and support more comprehensive cross-

validation. Future studies might combine behavioral log analysis of cryptomarket users with survey data, further fortifying the empirical foundation. Secondly, there is an opportunity to advance methodologies for assessing information disclosure, such as implementing measures like the information redundancy measure on social media platforms (Liang and Fu, 2016). This approach has the potential to notably enhance research findings. Thirdly, advancements in measurement techniques can contribute to a more precise operationalization of theoretical frameworks. For example, future investigations may employ advanced methodologies to calculate individualized risk-benefit ratios based on users' behavioral traces, providing more direct insight into factors influencing online forum participation.

In conclusion, this study investigates the digital traces of user actions on the Dark Web's three most popular cryptomarket forums: Silk Road 1, Silk Road 2, and Agora. The findings reveal that users who actively engage in discussions and use a diverse vocabulary are more inclined to discontinue their forum participation. Moreover, there is no significant correlation between network characteristics and user engagement. These results imply that the risk of exposure within anonymous communities primarily arises from the intensity of information rather than social networking. The research sheds light on the privacy predicament inherent in this hidden online realm and offers valuable insights into the dynamics of user behavior surrounding anonymity-enabling technologies on the Internet.

Data availability

The dataset and code employed in this study are available on Open Science Framework https://osf.io/cv4ef/. The original data is available on Darknet Market Archives https://gwern.net/dnm-archive.

References

Al Jawaheri, H., Al Sabah, M., Boshmaf, Y., & Erbad, A. (2020) Deanonymizing Tor hidden service users through Bitcoin transactions analysis. Computers & Security, 89. https://doi.org/10.1016/j.cose.2019.101684

Aldridge J, Decary-Hetu D (2016) Hidden wholesale: the drug diffusing capacity of online drug cryptomarkets. Int J Drug Policy 35:7–15. https://doi.org/10.1016/j.drugpo.2016.04.020

Altman I, Taylor DA (1973) Social penetration: The development of interpersonal relationships. Holt, Rinehart & Winston, New York

Arps, J. E., & Christin, N. (2020) Open Market or Ghost Town? The Curious Case of OpenBazaar. In International Conference on Financial Cryptography and Data Security. Springer International Publishing. https://doi.org/10.1007/978-3-030-51280-4_30

Bagozzi RP, Dholakia UM (2002) Intentional social action in virtual communities. Journal of Interactive Marketing 16(2):2–21. https://doi.org/10.1002/dir.10006

Bancroft A, Reid PS (2017) Challenging the techno-politics of anonymity: the case of cryptomarket users. Information Communication & Society 20(4):497–512. https://doi.org/10.1080/1369118x.2016.1187643

Barratt MJ, Aldridge J (2016) Everything you always wanted to know about drug cryptomarkets* (*but were afraid to ask). Int J Drug Policy 35:1–6. https://doi.org/10.1016/j.drugpo.2016.07.005

Batikas, M., & Kretschmer, T. (2018) Entrepreneurs on the Darknet Reaction to Negative Feedback. SSRN. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3238141

Bhaskar V, Linacre R, Machin S (2019) The economic functioning of online drugs markets. Journal of Economic Behavior & Organization 159:426–441. https://doi.org/10.1016/j.jebo.2017.09.022

Bhattacherjee A (2001) Understanding information systems continuance: An expectation-confirmation model. MIS Quarterly 25(3):351–370. https://doi.org/10.2307/3250921

Biryukov, A., Pustogarov, I., Thill, F., & Weinmann, R. P. (2014) Content and popularity analysis of Tor hidden services. 2014 Ieee 34th International

Conference on Distributed Computing Systems Workshops (Icdcsw), 188-193. https://doi.org/10.1109/Icdcsw.2014.20

Biryukov, A., Pustogarov, I., & Weinmann, R. P. (2013) Trawling for Tor Hidden Services: Detection, Measurement, Deanonymization. *2013 Ieee Symposium on Security and Privacy (Sp)*, 80-94. https://doi.org/10.1109/Sp.2013.15

Branwen, G., Christin, N., Décary-Hétu, D., Munksgaard, R., & Presidente, E. (2015) *Darknet Market Archives, 2011–2015.* https://www.gwern.net/DNM-archives

Burt RS (2009) Structural holes: The social structure of competition. Harvard University Press, Cambridge, MA

Chen Z, Jardine E, Fan Liu X, Zhu JJH (2022) Seeking anonymity on the Internet: The knowledge accumulation process and global usage of the Tor network. New Media & Society 0(0):14614448211072201. https://doi.org/10.1177/14614448211072201

Chen, Z., & Liu, X. (2021) *Why Do Users Leave Anonymous Online Communities? An Exploratory Study of the Silk Road Forum on the Dark Web* 71st Annual International Communication Association Conference (ICA21), Denver, Colorado, 27-31 May 2021

Choshen, L., Eldad, D., Hershcovich, D., Sulem, E., & Abend, O. (2019) The Language of Legal and Illegal Activity on the Darknet. *57th Annual Meeting of the Association for Computational Linguistics (Acl 2019)*, 4271-4279. https://doi.org/10.18653/v1/P19-1419

Christin, N. (2013) Traveling the Silk Road: A measurement analysis of a large anonymous online marketplace. *Proceedings of the 22nd international conference on World Wide Web*, 213-224. https://doi.org/10.1145/2488388.2488408

Christin, N. (2017) *An EU-focused analysis of drug supply on the anonymous online marketplace*. EMCDDA Commissioned Paper. https://www.emcdda.europa.eu/document-library/eu-focused-analysis-drug-supply-online-anonymous-marketplace-ecosystem_en

Doiciar, C., & Cretan, R. (2021) Pandemic populism: COVID-19 and the rise of the nationalist AUR party in Romania. *Geographica Pannonica*, 25(4). https://doi.org/10.5937/gp25-33782

ElBahrawy, A., Alessandretti, L., Rusnac, L., Goldsmith, D., Teytelboym, A., & Baronchelli, A. (2020) Collective dynamics of dark web marketplaces. *Scientific Reports*, 10(1). https://doi.org/10.1038/s41598-020-74416-y

Guitton C (2013) A review of the available content on Tor hidden services: The case against further development. Computers in Human Behavior 29(6):2805–2815. https://doi.org/10.1016/j.chb.2013.07.031

Hardy RA, Norgaard JR (2016) Reputation in the Internet black market: an empirical and theoretical analysis of the Deep Web. Journal of Institutional Economics 12(3):515–539. https://doi.org/10.1017/S1744137415000454

Hazel Kwon K, Shao C (2020) Communicative Constitution of Illicit Online Trade Collectives: An Exploration of Darkweb Market Subreddits. International Conference on Social Media and Society, Toronto, ON, Canada, 10.1145/3400806.3400815

Huffaker D (2010) Dimensions of leadership and social influence in online communities. Human Communication Research 36(4):593–617. https://doi.org/10.1111/j.1468-2958.2010.01390.x

Jardine, E. (2015) The Dark Web Dilemma: Tor, Anonymity and Online Policing. *Global Commission on Internet Governance Paper Series.* https://papers.ssrn.com/abstract=2667711

Jardine, E. (2019) Online content moderation and the Dark Web: Policy responses to radicalizing hate speech and malicious content on the Darknet. *First Monday*, 24(12). https://doi.org/10.5210/fm.v24i12.10266

Jiang Z, Heng CS, Choi BC (2013) Research note—privacy concerns and privacy-protective behavior in synchronous online social interactions. Information Systems Research 24(3):579–595. https://www.jstor.org/stable/42004283

Jin XL, Lee MKO, Cheung CMK (2010) Predicting continuance in online communities: model development and empirical test. Behaviour & Information Technology 29(4):383–394. https://doi.org/10.1080/01449290903398190

Johnson SL, Safadi H, Faraj S (2015) The Emergence of Online Community Leadership. Information Systems Research 26(1):165–187. https://doi.org/10.1287/isre.2014.0562

Kosinski M, Stillwell D, Graepel T (2013) Private traits and attributes are predictable from digital records of human behavior. Proceedings of the National Academy of Sciences 110(15):5802–5805. https://doi.org/10.1073/pnas.1218772110

Kozinets RV (1999) E-tribalized marketing?: the strategic implications of virtual communities of consumption. European Management Journal 17(3):252–264. https://doi.org/10.1016/s0263-2373(99)00004-3

Krasnova H, Spiekermann S, Koroleva K, Hildebrand T (2010) Online social networks: why we disclose. Journal of Information Technology 25(2):109–125. https://doi.org/10.1057/jit.2010.6

Lacson W, Jones B (2016) The 21st century darknet market: lessons from the fall of silk road. International Journal of Cyber Criminology 10(1):40–61. https://doi.org/10.5281/zenodo.58521

Laufer RS, Wolfe M (1977) Privacy as a concept and a social issue: a multi-dimensional developmental theory. Journal of Social Issues 33(3):22–42. https://doi.org/10.1111/j.1540-4560.1977.tb01880.x

Levine JM, Moreland RL (1994) Group socialization: theory and research. European Review of Social Psychology 5(1):305–336. https://doi.org/10.1080/14792779543000093

Lewis K, Kaufman J, Christakis N (2008) The taste for privacy: an analysis of college student privacy settings in an online social network. Journal of Computer-Mediated Communication 14(1):79–100. https://doi.org/10.1111/j.1083-6101.2008.01432.x

Liang H, Fu K-W (2016) Information overload, similarity, and redundancy: unsubscribing information sources on twitter. Journal of Computer-Mediated Communication 22(1):1–17. https://doi.org/10.1111/jcc4.12178

Limayem M, Cheung CMK (2008) Understanding information systems continuance: the case of Internet-based learning technologies. Information & Management 45(4):227–232. https://doi.org/10.1016/j.im.2008.02.005

Maddox A, Barratt MJ, Allen M, Lenton S (2015) Constructive activism in the dark web: cryptomarkets and illicit drugs in the digital 'demimonde'. Information, Communication & Society 19(1):111–126. https://doi.org/10.1080/1369118x.2015.1093531

Martin J (2014a) Drugs on the Dark Net: How Cryptomarkets are Transforming the Global Trade in Illicit Drugs. Palgrave Pivot, London, 10.1057/9781137399052

Martin J (2014b) Lost on the Silk Road: Online drug distribution and the 'cryptomarket. Criminology & Criminal Justice 14(3):351–367. https://doi.org/10.1177/1748895813505234

Medvedev AN, Delvenne JC, Lambiotte R (2019) Modelling structure and predicting dynamics of discussion threads in online boards. Journal of Complex Networks 7(1):67–82. https://doi.org/10.1093/comnet/cny010

Moreland RL, Levine JM (1982) Socialization in small groups: temporal changes in individual-group relations. Advances in experimental social psychology 15:137–192. https://doi.org/10.1016/S0065-2601(08)60297-X

Munksgaard R, Demant J (2016) Mixing politics and crime - The prevalence and decline of political discourse on the cryptomarket. International Journal of Drug Policy 35:77–83. https://doi.org/10.1016/j.drugpo.2016.04.021

Norbutas L, Ruiter S, Corten R (2020) Believe it when you see it: dyadic embeddedness and reputation effects on trust in cryptomarkets for illegal drugs. Social Networks 63:150–161. https://doi.org/10.1016/j.socnet.2020.07.003

Nov O, Naaman M, Ye C (2010) Analysis of participation in an online photo-sharing community: a multidimensional perspective. Journal of the American Society for Information Science and Technology 61(3):555–566. https://doi.org/10.1002/asi.21278

Nugent TJ (2019) Prosecuting dark net drug marketplace operators under the federal crack house statute. Fordham Law Review 88(1):345–379. https://ir.lawnet.fordham.edu/flr/vol88/iss1/9

Ormsby E (2014) Silk Road. Macmillan Publishers, Australia

Pace J (2016) Exchange relations on the dark web. Critical Studies in Media Communication 34(1):1–13. https://doi.org/10.1080/15295036.2016.1243249

Paquet-Clouston M, Decary-Hetu D, Morselli C (2018) Assessing market competition and vendors' size and scope on AlphaBay. Int J Drug Policy 54:87–98. https://doi.org/10.1016/j.drugpo.2018.01.003

Parmentier, A., & Cohen, R. (2019) Learning User Reputation on Reddit. *2019 Ieee/Wic/Acm International Conference on Web Intelligence (Wi 2019)*, 242-247. https://doi.org/10.1145/3350546.3352524

Petronio, S. (2002) *Boundaries of privacy: Dialectics of disclosure.* State University of New York Press

Porter K (2018) Analyzing the DarkNetMarkets subreddit for evolutions of tools and trends using LDA topic modeling. Digital Investigation 26:S87–S97. https://doi.org/10.1016/j.diin.2018.04.023

Rains SA, Scott CR (2007) To identify or not to identify: a theoretical model of receiver responses to anonymous communication. Communication Theory 17(1):61. https://doi.org/10.1111/j.1468-2885.2007.00288.x

Rogers EM (2003) Diffusion of innovations, 5th ed. Free Press, New York, NY

Scott CR (1998) To reveal or not to reveal: a theoretical model of anonymous communication. Communication Theory 8(4):381–407. https://doi.org/10.1111/j.1468-2885.2007.00288.x

Scott CR, Rains SA (2020) (Dis)connections in anonymous communication theory: exploring conceptualizations of anonymity in communication research. Annals of the International Communication Association 44(4):385–400. https://doi.org/10.1080/23808985.2020.1843367

Scott CR, Rains SA, Haseki M (2011) Anonymous communication unmasking findings across fields. Communication Yearbook 35(1):299–340. https://doi.org/10.1080/23808985.2011.11679120

Seering J, Wang T, Yoon J, Kaufman G (2019) Moderator engagement and community development in the age of algorithms. New Media & Society 21(7):1417–1443. https://doi.org/10.1177/1461444818821316

Soska, K., & Christin, N. (2015) Measuring the longitudinal evolution of the online anonymous marketplace ecosystem. 24th {USENIX} Security Symposium ({USENIX} Security 15)

Sotirakopoulos N (2018) Cryptomarkets as a libertarian counter-conduct of resistance. European Journal of Social Theory 21(2):189–206. https://doi.org/10.1177/1368431017718534

Spitters, M., Klaver, F., Koot, G., & van Staalduinen, M. (2015) Authorship Analysis on Dark Marketplace Forums. *2015 European Intelligence and Security Informatics Conference (Eisic)*, 1-8. https://doi.org/10.1109/Eisic.2015.47

Stachl C, Au Q, Schoedel R, Gosling SD, Harari GM, Buschek D, Völkel ST, Schuwerk T, Oldemeier M, Ullmann T, Hussmann H, Bischl B, Bühner M (2020) Predicting personality from patterns of behavior collected with smartphones. Proceedings of the National Academy of Sciences 117(30):17680–17687. https://doi.org/10.1073/pnas.1920484117

Tausczik YR, Pennebaker JW (2010) The Psychological Meaning of Words: LIWC and computerized text analysis methods. Journal of Language and Social Psychology 29(1):24–54. https://doi.org/10.1177/0261927x09351676

Tzanetakis M, Kamphausen G, Werse B, von Laufenberg R (2016) The transparency paradox. Building trust, resolving disputes and optimising logistics on conventional and online drugs markets. Int J Drug Policy 35:58–68. https://doi.org/10.1016/j.drugpo.2015.12.010

Van Hout MC, Bingham T (2013) Silk Road', the virtual drug marketplace: a single case study of user experiences. Int J Drug Policy 24(5):385–391. https://doi.org/10.1016/j.drugpo.2013.01.005

Wang X, Zhao K, Street N (2017) Analyzing and predicting user participations in online health communities: a social support perspective. J Med Internet Res 19(4):e130. https://doi.org/10.2196/jmir.6834

Wang Y-C, Kraut R, Levine JM (2012) To stay or leave? The relationship of emotional and informational support to commitment in online health support groups. Proceedings of the ACM 2012 conference on Computer Supported Cooperative Work - CSCW '12, Seattle, Washington, USA, https://doi.org/10.1145/2145204.2145329

Wasserman S, Faust K (1994) *Social network analysis: Methods and applications.* Cambridge University Press. https://doi.org/10.1017/CBO9780511815478

White HC (1981) Where do markets come from? American Journal of Sociology 87(3):517–547. http://www.jstor.org/stable/2778933

Woo J (2006) The right not to be identified: privacy and anonymity in the interactive media environment [Article]. New Media & Society 8(6):949–967. https://doi.org/10.1177/1461444806069650

Wu CG, Gerlach JH, Young CE (2007) An empirical analysis of open source software developers' motivations and continuance intentions. Information & Management 44(3):253–262. https://doi.org/10.1016/j.im.2006.12.006

Zamani M, Rabbani F, Horicsanyi A, Zafeiris A, Vicsek T (2019) Differences in structure and dynamics of networks retrieved from dark and public web forums. Physica a-Statistical Mechanics and Its Applications 525:326–336. https://doi.org/10.1016/j.physa.2019.03.048

Zhang, L., Han, Y., Zhou, J.-L., Liu, Y.-S., & Wu, Y. (2020) Influence of intrinsic motivations on the continuity of scientific knowledge contribution to online knowledge-sharing platforms. *Public Understanding of Science*. https://doi.org/10.1177/0963662520970782

Author contributions

Zhicong Chen conceived the study, analyzed the data, wrote the first manuscript, and served as the first author. Xiang Meng revised the manuscript, and served as the second author. Cheng-Jun Wang supervised the study, revised the manuscript, wrote the cover letter, and served as the corresponding author.

Ethical approval

This article does not contain any studies with human participants performed by any of the authors.

Informed consent

This article does not contain any studies with human participants performed by any of the authors.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to Cheng-Jun Wang.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.