



Communicative Constitution of Illicit Online Trade Collectives: An Exploration of Darkweb Market Subreddits

K. Hazel Kwon

Walter Cronkite School of Journalism and Mass
Communication, Arizona State University, Phoenix,
Arizona, USA
khkwon@asu.edu

Chun Shao

Walter Cronkite School of Journalism and Mass
Communication, Arizona State University, Phoenix,
Arizona, USA
cshao9@asu.edu

ABSTRACT

This study aims to understand communicative activities of an illicit cyber-trade community on an anonymous social platform, Reddit. Based on the communication-as-constitutive of organization (CCO) perspective, the study identifies participants with different levels of engagement, and examines how their discursive engagement collectively reflects the ways in which illicit market users co-orient themselves to respond to a crisis event (i.e., market shutdown). The empirical case for this study is a subreddit channel dedicated to what was once the largest dark web market: r/AlphabayMarket. We examine two month period's posting activities until market was permanently shut down in July 2017. Our analysis comprises three parts. First, a social network analysis was conducted to identify key and non-key players in the community. Second, a structural topic modeling was computed to inductively infer topic clusters. Third, posts were manually reviewed to articulate the process of co-orientation manifest in the results of topic modeling.

CCS CONCEPTS

• **Security and privacy** → Human and societal aspects of security and privacy.

KEYWORDS

Dark web, illicit online communities, Reddit, communicative constitution of organization, cyber markets, cryptomarket, cybercrime

ACM Reference Format:

K. Hazel Kwon and Chun Shao. 2020. Communicative Constitution of Illicit Online Trade Collectives: An Exploration of Darkweb Market Subreddits. In *International Conference on Social Media and Society (SMSociety '20)*, July 22–24, 2020, Toronto, ON, Canada. ACM, New York, NY, USA, 8 pages. <https://doi.org/10.1145/3400806.3400815>

1 INTRODUCTION

On July 20, 2017, the office of public affairs of the U.S. Department of Justice published a press release announcing a historic victory over cybercrime:

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than ACM must be honored. Abstracting with credit is permitted. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee. Request permissions from permissions@acm.org.
SMSociety '20, July 22–24, 2020, Toronto, ON, Canada

© 2020 Association for Computing Machinery.
ACM ISBN 978-1-4503-7688-4/20/07...\$15.00
<https://doi.org/10.1145/3400806.3400815>

“... the seizure of the largest criminal marketplace on the Internet, AlphaBay, which operated for over two years on the dark web and was used to sell deadly illegal drugs, stolen and fraudulent identification documents and access devices, counterfeit goods, malware and other computer hacking tools, firearms, and toxic chemicals throughout the world. The international operation to seize AlphaBay's infrastructure was led by the United States and involved cooperation and efforts by law enforcement authorities in Thailand, the Netherlands, Lithuania, Canada, the United Kingdom, and France, as well as the European law enforcement agency Europol¹”

Undoubtedly, AlphaBay was the largest cyber-underground market to emerge since the shutdown of the legendary market SilkRoad, with \$600,000 to \$800,000 in daily revenue. The shutdown of AlphaBay attracted mainstream media attention not only because of its scale but also due to an ensuing scandal that involved the death of Alexandre Cazes in a jail in Thailand; Cazes had been a co-founder of the market who was arrested on the same day as the shutdown².

It is important to note that not all dark web activities are bad. Dark web broadly refers to a collection of hidden web services that are usually accessible through special routing and browser systems such as Freenet, TOR or I2P [1, 2]. An essential feature of dark web is its technological ability to conceal user identities and locations [3]. Thanks to its anonymizing capacity, dark web welcomes services designed to protect politically vulnerable individuals, whistleblowers, and high-risk activists [3, 4]. That said, the dark web's affordance of anonymity also provides an ideal environment for many illicit activities to bypass lawful procedures and avoid social ramifications. Illicit economic trade is a prominent illicit activity, with dark web markets serving as a hotbed where buyers and sellers gather to trade in drugs, fraudulent products, hacked information, and weapons. Even after the shutdown of AlphaBay, illicit trades are still active in the dark web, if not thriving ever more. As of this writing in January 2020, for example, there are at least 33 markets known to be operating in the dark web³.

The goal of our project is to understand what makes illicit cyber-trade collectives sustainable even in the face of intense intelligence operations by law enforcement. Among various factors that may

¹See <https://www.justice.gov/opa/pr/alphabay-largest-online-dark-market-shut-down>.

²see <https://www.independent.co.uk/news/world/alphabay-alexandre-cazes-dark-web-dark-net-hanged-thailand-bangkok-narcotics-suicide-drug-a7843626.html>

³see <https://www.thedarkweblinks.com/darknet-market-list/>

contribute to the organizational resilience of dark web market communities, this paper focuses on the characteristics of communicative activities in times of the communities' collective crises. Particularly, we identify key player activities during a crisis period and compare them with non-key players' activities, as well as during a routine period.

Theoretically, this paper is inspired by the “communication-as-constitutive of organization (CCO)” perspective [5], which offers a rationale for examining communicative activities to better understand organizational workings of the illicit trade collectives. In this paper, we analyze a Reddit community (i.e., subreddit) dedicated to AlphaBay market-related discussions (r/alphabaymarket), for two months prior to the official announcement of the market seizure. We use a co-posting network analysis to identify key players, and topic modeling to inductively infer conversational themes that emerged from the subreddit posts. This paper intends to address two questions: (1) how different were communicative activities in the Alphabay subreddit between a routine and crisis period? (2) how different were communicative activities in the Alphabay subreddit between key players and non-key players?

2 THEORETICAL BACKGROUND

2.1 Communicative constitution of illicit cybermarket collectives

While illicit trade collectives tend to be depicted simply as being underground and operated by criminals, such reductionism hinders an understanding of their organizing efficacy and agility [6]. Instead, Stohl and Stohl's (2011) conceptualization of “clandestine organizations” is an adequate starting point from which to examine illicit cyber-trade collectives as a socio-technical organizing system. This frame proposes three characteristics: (1) mutual agreement in protecting all members' anonymity and secrecy of membership; (2) furtive internal operations and governance outside the public view; (3) that the existence of the group eventually becomes known –or at least rumored –to the outsiders [5]. These characters properly illustrate dark web market communities in that (1) OPSEC for anonymity and security is the utmost rule by which members must abide; (2) skills, norms, and know-how for value creation are learned and expanded in the form of “tacit knowledge” [7], most of which is transferred and managed in shadowy modes; and (3) the existence of dark web markets is known to the public through media reports, and partial information is available even in the clear web. In particular, the third characteristic is an important condition to expand the collectives and continue to generate values (e.g., profits).

Considering that knowledge flows in illicit cyber-trade communities are heavily dependent on discursive interactions in online forums, communicative-as-constitutive of organization (CCO) theory offers an insightful framework to understand how individual members negotiate and construct shared identity as a community of practice and build internal knowledge. A main thesis of CCO contends that conversation is the core mechanism of organizing. Importantly, conversation intertwines two dialectic components of an organizing system: agency and text. *Agency* refers to members' goal-oriented discursive activities for the purpose of “*co-orientation*,” which itself is defined as a process to negotiate and coordinate

members' “ways of dealing with the objective world while simultaneously situating themselves within a social world.” *Text* broadly refers to functional language used and created to manifest “the sensemaking practices and habits of interpretation” of the community [[8], p. 410]. In the process of co-orientation, conversations not only generate new texts but also are entered by existing texts as reflexive “resources” that assist collective sensemaking [[8], p.397]. That is, “conversation is the site where organizing occurs” [[8], p.410].

While CCO theory's emphasis on conversations lays the theoretical groundwork and justification for examining communicative activities of illicit cybermarket users, one challenge is the theory's uncontested assumption of organizational transparency [5]. Specifically, the original CCO perspective presupposes relational clarity among leaders, other organizational members, clients, and outsiders, as well as the legitimacy of the relationship between organizational members and social institutions in the material environment. CCO's assumption of clear boundary of roles and membership, and legitimate relations within and outside the collectives, called “the embedded transparency principle”, is challenged in our context by the secrecy and tacit nature of a clandestine organization such as dark web market communities [5], p.[6]. In a dark web market community, the definition of membership is opaque as even the core participants are unlikely to be acquainted with one another outside of the online context. Furthermore, participants' roles within the collectives are not prescribed but rather spontaneously emerge by the level of voluntary engagement and commitment they are willing to commit. And of course, their legitimacy is contested by their illicit value creation.

Accordingly, studies of clandestine organizations strive to overcome the “embedded transparency principle” in expanding the CCO theory to their study contexts, by recognizing the far-flung membership structures and opaque nature of conversations that occur in the process of organizing [e.g., 5, 9, 10]. CCO theory is especially useful in theorizing organizational workings in online communities because most of these groups operate in non-traditional, dispersed, and networked forms, for which maintenance and expansion rely heavily on tacit knowledge flow through users' conversations [11]. In line with the CCO theory, our study examines the two organizing forces, agency and texts, and explores how participants with different levels of engagement co-orient their responses toward the community's crisis moment (agency) and how such co-orientation process is manifest in online posting activities (texts). Accordingly, we propose two research questions, each of which addresses an aspect of agency and of texts respectively.

RQ1: How different are the communicative activities between participants of different levels of engagement?

RQ2: Do communicative activities reflect different ways of co-orientation between a routine period and a crisis period?

2.2 Communication ecology of illicit cyber-market collectives

This study examines a subreddit forum as a conversation site. We chose Reddit as a case study due to the importance of Reddit in the communication ecology of illicit cybermarket system. Studies have suggested two pillars of illicit cybermarket system: markets and



Figure 1: “DeepWebSitesLinks” is an information aggregation site currently viewable in the clear web. The information about Empire market (as of 2018) is exemplified in this figure. As seen here, the subreddit URL is introduced along with the main market URL (and mirror links) and the official forum link. Among the links listed, the subreddit is the only one accessible via clear web browsers.

forums [12]. While markets are where the actual economic practices and values are created, forums are where conversations emerge and organizing occurs. There are different types of forums: an official market forum, which is typically embedded in the given market site, and cross-marketplaces forums, which are additionally operated in the dark web to facilitate competing reviews across different marketplaces. Then, there is Reddit. Reddit is an important platform for several reasons. Firstly, while it is a single platform, it offers tens of sub-communities (subreddits) variegating dark web-related topics, ranging from general issues to a single market-related topic. In Reddit, users can easily traverse from one subreddit to another, contributing to enriching the creation of texts. Secondly, and more importantly, while most dedicated users engage with subreddits of illicit market trades furtively via TOR, Reddit is also accessible via regular browsers, and thus the dark web-related conversations that occur in subreddits are visible to outsiders. That is, Reddit functions as a cross-cutting space where not only experienced dark web users actively engage in conversation, but also potential newcomers can have vicarious learning about illicit cyber-trades and prepare their entrance into the actual marketplaces. As Figure 1 illustrates, almost every marketplace has a dedicated subreddit channel, suggesting Reddit’s distinctive status as a conversation site for illicit cybermarket users.

2.3 Key players in online communities

Unlike project group assignments or traditional organizational institution, most illicit cybermarket communities are run as an open and self-organized network of voluntary members [12], where important roles –who we called “key players” in this paper –are communicatively constituted. That is, individual members’ positions in the community are not determined hierarchically, but based on the level of time and effort they willingly spend in the community.

Several studies have attempted to identify key members and role-playing in online communities based on communication activities. Studies have often applied social network metrics such as degree centrality and link similarity [13, 14] to identify network members who have the “most important” positions in the network (i.e., to pinpoint the set of key players) [15]–[17]. For example, Gupta et al. [13] apply network analysis to identify communities of Twitter users and top users in each of these communities during three different crisis events. Interestingly, their results showed that the top users represented the generalized topics and opinions of all the users in the community with 81% accuracy on average.

Topic modeling is another popular technique to explore communicatively emergent role-playing in online communities. For example, one study used topic modeling to represent emergent discussion topics in a mental health Internet support group and the different types of roles played by participants in the community [18]. Their results showed that “super users” (the top one percent) served as emotional supporters, while non-key users engaged in topics reflective of information and experiential knowledge sharing [17].

Likewise, dark network research has also employed computational methods such as topic modeling, social network analysis, and UserRank algorithms to identify key players in illicit cybercommunities [17, 19]–[21]. For example, Marine et al. combined the network analysis metrics and user reputation scores available in forum sites to detect key hackers [19]. Yang et al. investigated the effect of weights and response immediacy on determining user influence ranking on the dark web [17]. Xu and Chen described how several centrality measures can help identify different types of key players on the dark web [22].

Overall, prior literature points to the advantage of combining multiple metrics in determining key players in online communities, such as the number of posts made by members, the number of threads initiated, the number of different threads participated in, the level of connectedness to other members in the forum, and time spent logged in [22]–[25]. Accordingly, this study employed two metrics to designate a set of high-level participants (key players): the top 10 percent of users based on their *total posting frequency*, and top 10 percent of users based on the *degree centrality* calculated from co-posting network data. While posting frequency represents the *volume of texts* a user created, the co-posting based degree centrality indicates how often a user discursively encountered with other users in different topic threads, reflective of the *diversity of texts* with which the user engaged.

3 METHODS

This study investigates the aforementioned market seizure as a crisis event of interest. The subreddit community of interest was

Table 1: Post Frequency Across Users and Contexts

	Routine Period		Shutdown Period	
	Key players	Non-key players	Key players	Non-key players
Number of user	82 (14%)	510 (86%)	106 (15%)	601 (85%)
Number of post	789 (47%)	874 (53%)	822 (52%)	765 (48%)
Total threads participated	295	352	346	242
Mean post word count	211	268	147	139

r/AlphabayMarket, which has been permanently banned since the market seizure. We obtained this community’s historic posts from CRY3CON, a cybersecurity firm that partnered with this project. The obtained data was in a de-identified format, for which the post texts had no user-related information available to the researchers. The project was approved by the IRB. This subreddit community generated an unusually high volume of conversations for ten days from July 5 to July 14, 2017, which we identified as the “market shutdown” period. The total number of topic threads that were created during the shutdown period was 346, and the total number of posts was 1,587. For comparison, we examined a similar number of topic threads and posts made prior to the beginning of the shutdown period, which spanned from May 19, 2017 to July 4, 2017. This time window is defined for our research as a “routine” period, and included the creation of 383 topic threads and 1,663 posts. As a result, a total of 3,250 posts were included in this study. Our analysis entailed three parts. First, a social network analysis was conducted to identify key and non-key players in the forum. Second, a structural topic modeling was computed for all posts ($N = 3,250$). Third, we manually reviewed the posts to understand what types of co-orientation were manifested in each topic cluster generated by topic modeling.

3.1 Key player identification

Degree centrality and total posting frequency were used as criteria to identify “key players” in the community. The two sets of top 10 percent user lists were combined and deduplicated when the same user ID (which was composed of random digits assigned by the database) appeared in both sets. As a result, we identified a total of 170 key players out of 1083 users. Eighteen (10 percent) of these key players appeared in both routine and market shutdown periods, 64 (38 percent) were associated only with the routine period, and 88 (52 percent) uniquely emerged during the shutdown period. In other words, those who emerged as active participants during the market shutdown period were generally *different* users from those active during the routine period. Table 1 shows the descriptive statistics of posts from different user groups.

3.2 Topic modeling

Structural topic modelling (STM) is a computational technique that identifies latent patterns of word co-occurrence using the distribution of words in a collection of documents. STM is an unsupervised machine learning technique for text analysis, usually based on a topic modeling algorithm called latent Dirichlet allocation (LDA) [26, 27]. Topic modeling has been used in studying social media and online communities [e.g., 28, 29], and the output includes a set

of multiple topic clusters that are detected from the distributions of words [30].

Like other computational text analyses, STM requires natural language preprocessing (NLP) of the data. Specifically, each word in each post was preprocessed by stemming and dropping punctuation, symbols, numbers, and stop words. In addition, we manually reiterated the preprocessing procedure to eliminate meaningless words in our study context as much as possible.

Further, we set a threshold of word occurrence to ten such that any word that occurred less than ten times was excluded. The threshold of ten was determined based on the review of the removal trends for documents, words, and tokens according to the threshold levels. As seen in Figure 2, the word removal became somewhat static around the threshold 15 to 20. However, when we set the threshold at such a high level, the modeling results lost many meaningful words. Accordingly, we decided to set the threshold slightly lower, with 10 as a final cut-off point. With the threshold level of ten, approximately 100 posts, 5500 unique words, and a total of 12000 words were excluded from further modeling.

STM also requires setting the number of topics. Determining the number of topical clusters required both quantitative measures and heuristic judgement. We first chose four candidate models—5, 6, 10, and 20 topic cluster models—and ran estimates for each. Based on the quantitative metrics (e.g., exclusivity and semantic coherence) [30] and manual reviews in terms of thematic overlaps across topic clusters, we finally selected the 6-topic model for further interpretation.

4 RESULTS

Topics were identified based on the distribution of 8757 unique words in the *r/AlphabayMarket*. After generating topic clusters, 100 of the most representative posts for each topic cluster were closely reviewed. As a result, the theme of each topic cluster was interpreted as following: *order issue* (Topic 1), *bitcoin and transaction issue* (Topic 2), *vendor promotion* (Topic 3), *sensemaking* (Topic 4), *vendor/product review* (Topic 5), and *emotional outrage* (Topic 6). Table 2 explains the interpretation of the thematic characteristics of each cluster, along with example posts. Specifically, Topic 1 included discussions about issues occurring in the process of ordering; Topic 2 included financial transaction issues such as the withdrawal and deposit of bitcoins; Topic 3 showed promotion and advertisement of vendors, mostly self-promotions; Topic 4 was discussions about scams, security failure, and shutdown, which collectively represents attempts for sensemaking about a given situation or a problem; Topic 5 included largely personal anecdotes that talked about their experiences with vendors and products. Together, these five topics

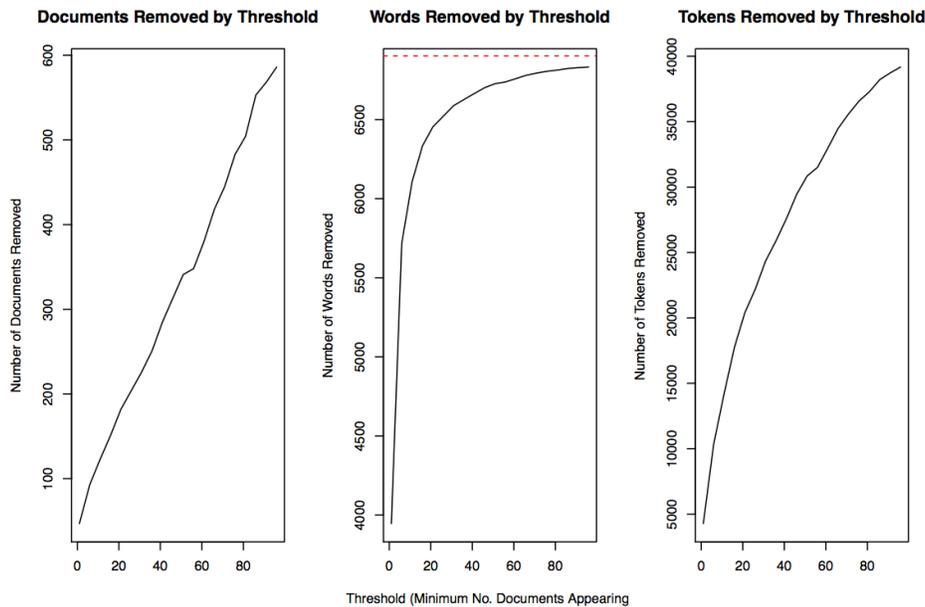


Figure 2: Changes in the number of documents (posts), unique words, and tokens (total word counts) according to the level of threshold.

provided reviews for quality assurance. Topic 6 covered emotional speech that reflected anger, frustrations, and aggressive reactions.

Figure 3 shows the topics that are more closely related than others. The connection between a pair of topics indicates that the two topics are more likely to appear together in the same text. For example, bitcoin and transaction issues (Topic 2) tended to co-occur in the same post that contained sensemaking (Topic 4). As an illustration, a post by one user states, “if bitcoin price remains stable is a good signal that ab [AlphaBay] is not in a [sic] exit scam operation.” Another user posted that “the timing is weird btc [bitcoin] just went up a fuckin hell of a lot now they’re starting to decline.”

To answer RQ1, we compared communicative activities in the Alphabay subreddit between key players and non-key players. First, during the routine period, key players’ posts contained significantly more vendor promotions (Topic 3), and outrage (Topic 6), while non-key players engaged with more conversations reflective of sensemaking (Topic 4) and vendor/product review (Topic 5) (the third graph in Figure 4). This trend suggests that prominent participants in the routine period were more likely to be sellers and therefore have high emotional investment. In contrast, communicative activities between key players and non-key players were *not* significantly distinctive during the shutdown period, with the exception of Topic 1 (the fourth graph in Figure 4). While the high prevalence of ordering issues (Topic 1) among non-key players confirms that the majority of non-key players were indeed buyers, the lack of a difference in distribution for the other topics between key and non-key players implies that participants, regardless of their level of activeness, were co-oriented in a similar manner to deal with the market shut down.

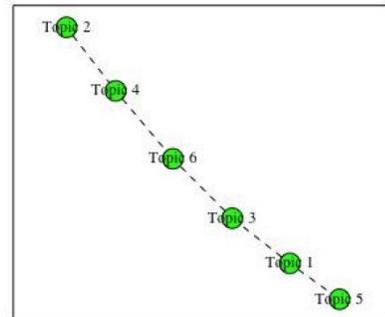


Figure 3: Graphical display of topic correlations.

In response to RQ2, the STM results showed different ways of co-orientation between a routine period and a crisis period (first and second graphs in Figure 4). Specifically, discussions related to order issues (Topic 1) and vendor/product reviews (Topic 5) were more prevalent during the routine period than the shutdown period, suggesting that participants co-oriented one another to build shared knowledge on “legitimate” trades [3] in terms of who were credible actors and determining reliable ordering channels for successful delivery. Then, during the shutdown period, questions and complaints about transaction-related security became abundant (Topic 2) along with sensemaking effort (Topic 4). The conversations in these topic clusters reflected that participants collectively faced the same or similar problems, and thus co-orientation effort was made to reach a common understanding of how severe the situation was (e.g., whether the crisis is temporary) and how to respond to resolve the problem. For instance, one user posted his or her

Table 2: Topic Clusters Generated from Topic Modeling.

Topic Description	Example Posts
<p>Topic Cluster 1 (Order issue) <i>Discussion about order and delivery related issues</i></p>	<ul style="list-style-type: none"> • Yeah, so I ordered 20 xtc pink redbull pills, on april 20th as since his vendor level is so high, I didn't doubt fe, but i still got scammed. order was worth only 70\$. but a scam is still a scam... and he denies that I placed the order.... • order placed. we will see. he had 0 feedback but its free besides shipping so there is huge plausible deniability.
<p>Topic Cluster 2 (Bitcoin and transaction issue) <i>Discussions about bitcoin and money transaction techniques</i></p>	<ul style="list-style-type: none"> • I would appreciate some assistance. I withdrew.0606 btc from the market wallet. this is reflected by the confirmation page on ab however the external wallet only received .01629. this is less than a third of my money. • I have no issues regarding sending btc to ab.i had done it several s and I am level 3 customers of ab.in my experience each i send btc it takes 3 confirmation to appears btc in my ab wets and btc address changes.
<p>Topic Cluster 3 (Vendor Promotion) <i>Self-promotion about vendors ort products</i></p>	<ul style="list-style-type: none"> • 72 hour sale!!!! 1g- \$30.00 (\$5.00) off! http://pwoah7foa6au2pul.onion/listing.php?id=344598. 3.5g- \$90.00 (\$10.00) off! / http://pwoah7foa6au2pul.onion/listing.php?id=344599 • Here is my price list !! cheapest ever !! dutch mdma cristals 84% pure 1gr • We are vendor tanukiman on alphabay. due to the current situation on alphabay we have migrated onto dream and hansa. you can find us under the same user name (tanukiman). we are currently carrying live resin bho, full spectrum terp wax, distillate (the clear), and ready to smoke vape cartridges made from distillate.
<p>Topic Cluster 4 (Sense making) <i>Diagnosis of security failure, market status, and fraudulent signals</i></p>	<ul style="list-style-type: none"> • I'm sure it's coming back this has happened a bunch of times and it's always just maintenance. just the admin for ab are the world's worst communicators despite managing to create the best marketplace on the dn • dude don't cry because they are adding the zcash into the site and fixing bugs. eg they make a decent amount of money everyday I don't think they would do an exit scam...
<p>Topic Cluster 5 (Vendor/Product Review) <i>Personal experiences and commentary about vendors or products</i></p>	<ul style="list-style-type: none"> • I've ordered 4g of his ketamine and the package arrived today, tho I'm pretty sure its mdma hes sent instead. I've sent him a message on ab, but I've read before he gets hundreds of messages so I'm not even sure when or even if he will read it. • what should i do? he seems a pretty solid vendor so should i finalise anyway and sort it out later? the ket cost \$140 dollars, tho the mdma (if that's what it is) would only have cost \$60. not that i want it anyway! • I'm in us but I've gotten landedgypsie's shit before and it's fucking awesome. I've tried a lot of top quality mdma in my day including stuff synthesized by a chemist friend of mine and lg's is up there with the best of them.
<p>Topic 6 (Outrage) <i>Negative emotional statements that display anger, frustration etc.</i></p>	<ul style="list-style-type: none"> • haha tell me about it. I hate feing no I don't give a fuck if they're awesome vendors. fuck that suck my dick lsd vendors be playing games nowadays. • after spending 20 years dealing with stupid, petty, ignorant ass bastards, the markets have taken things to a whole more sensible level. • right on man is a fucking serious thing this vendor is a fucking criminal with our personal information...we need to call the attention of trappy @trappy .. this fucker needs to be stop.

guesswork: “it’s done [sic] this 3 times in the last 6 months to update their servers and they always come back; plus they have no reason to exit scam because their [sic] number 1 at the top making so much money way more money than [sic] if they exit scammed.” Meanwhile, Topic 6 also became prominent in the shutdown period, suggesting that emotional co-orientation occurred with rampant displays of anger, apprehension, and frustration.

5 DISCUSSION AND CONCLUSIONS

Based on the Communicative Constitution of Organization (CCO) theory, this study investigates the qualities that make illicit cyber-trade collectives resilient in the contemporary socio-technical environment. Particularly, this study focuses on the comparison of communication activities of non-key and key players during the routine and market shutdown periods.

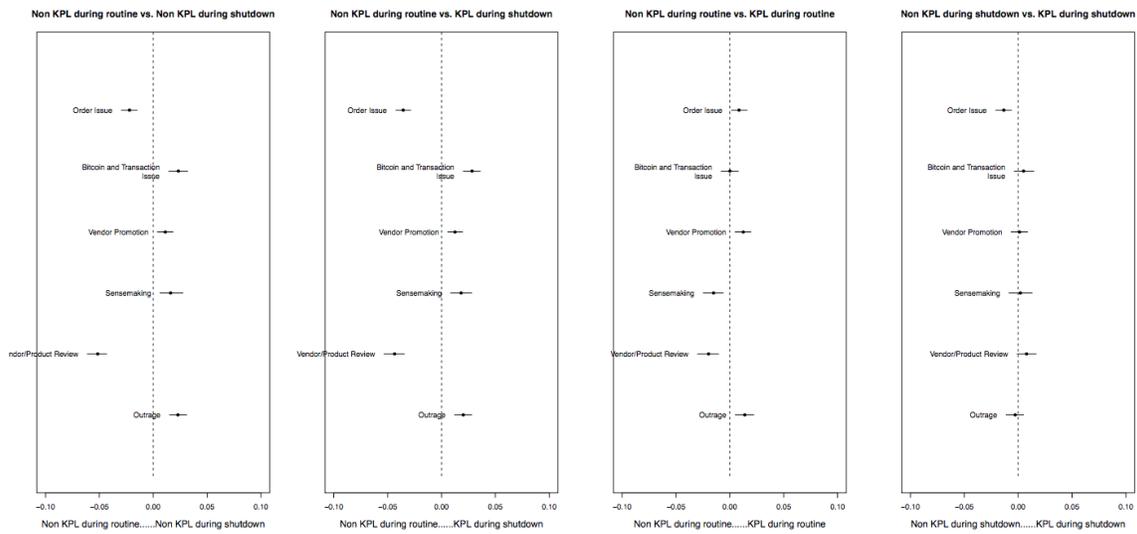


Figure 4: Topical distributions of posts among non-key and key players in different periods, with non-key players during routine period as the reference point. The bars indicate 95 percent confidence intervals.

Based on the results, we conclude that the illicit cyber-trade collectives are communicatively constituted in two ways. First, in ordinary times, users co-oriented to one another to build “legitimacy” of their relationships as sellers and buyers within the community. Specifically, two distinct roles were observed. One the one hand, there were sellers, also identified as the most active participants (key players) during the routine period, who primarily posted self-promotion to signal their legitimacy as credible vendors; and there were buyers, the majority of participants, who shared their personal experiences with vendors and products, and exchanged knowledge about safe practices for ordering and delivery.

Then, when all users simultaneously faced a high-risk problem (i.e., market shutdown), distinctive characteristics of texts generated by sellers and buyers were diluted. Instead, most of communicative efforts were collectively shifted toward “sensemaking” in attempt to interpret the situation as accurately as possible. In this sensemaking process, competing ideas and arguments were proposed (e.g., conspiracy of exit scam, seizure by law enforcement, inadvertent errors), and the discursive atmosphere was not necessarily harmonious, instead it was frequently charged with offensive comments. Nonetheless, participants eventually reached a common interpretation of the situation, and openly shared alternatives to continue their market practices.

Exploring illicit online communities from the CCO perspective is a new approach to dark web research and offers a theory-driven examination of the human side of dark social networks. That said, several limitations remain. First, like many other computational methods, topic modeling is useful for exploratory purposes, but by itself is not sufficient to capture contextual nuances. Also, the results of topic modeling are relatively static, and thus unable to capture the evolution of communicative interactions over time. Combining a rigorous qualitative analysis with topic modeling would enrich discussions. Second, this study dichotomized user

types (key players and non-key players), to which any oversimplification in our identification of roles (sellers and buyers) could be attributed. Future research could consider using more nuanced metrics to classify users and the communicative roles that they play in the community. Last but not least, this study analyzed a relatively small number of posts. The size of the corpus could affect the computational power of topic modelling [31]. While this study is a preliminary work, we intend to expand the time frame and include more data corpora to understand more comprehensively how illicit cyber-trade collectives are communicatively constituted.

In conclusion, illicit market communities in the dark web may be distinctive from other online communities due to their concealed and illegitimate nature, but we emphasize that they are likewise a self-organized community of practice. Efforts to understand the illicit cyber-market system should thus go beyond focusing on its deviancy, by theorizing and analyzing it as an organizing system where individual members collaborate, compete, and coordinate their actions for value creation. Considering that such organizing effort is mainly observed in conversions exchanged in online forums, scholars and practitioners may need to pay more attention to the textual traces of their discursive activities.

ACKNOWLEDGMENTS

This research was sponsored by the Army Research Office and was accomplished under Grant Number W911NF-19-1-0066. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the Army Research Office or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

REFERENCES

- [1] Robert W. Gehl. 2016. Power/freedom on the Dark Web: A digital ethnography of the Dark Web social network. *New Media & Society*, 18, 7, 1219-1235.
- [2] K. Hazel Kwon, J. Hunter Priniski, Soumajyoti Sarkar, Jana Shakarian, and Paulo Shakarian. 2017. Crisis and collective problem solving in dark web: An exploration of a black hat forum. In *Proceedings of the 8th International Conference on Social Media & Society*. Toronto, Canada, 1-5.
- [3] Robert W. Gehl. 2018. *Weaving the Dark Web: Legitimacy on Freenet, Tor, and I2P*. MIT Press.
- [4] K. Hazel Kwon and Jana Shakarian. 2019. Black-Hat hackers' crisis information processing in the darknet: A case study of cyber underground market shutdowns", in *Networks, Hacking, and Media—CITA MS@ 30: Now and Then and Tomorrow*, Emerald Publishing Limited, United Kingdom, pp. 113-135.
- [5] Cynthia Stohl and Michael Stohl. 2011. Secret agencies: The communicative constitution of a clandestine organization. *Organization Studies*, 32, 9, 1197-1215.
- [6] Moises Naim. 2006. *ILLICIT: How Smugglers, Traffickers and Copycats are Hijacking the Global Economy*. Anchor, New York.
- [7] Ikujiro Nonaka and Georg Von Krogh. 2009. Perspective—tacit knowledge and knowledge conversion: Controversy and advancement in organizational knowledge creation theory. *Organization Science*, 20, 3, 635-652.
- [8] James R. Taylor and Daniel Robichaud. 2004. Finding the organization in the communication: Discourse as action and sensemaking. *Organization*, 11, 395-413.
- [9] Leonhard Dobusch, and Dennis Schoeneborn. 2015. Fluidity, identity, and organizationalitaty: The communicative constitution of Anonymous. *Journal of Management Studies*, 52, 8, 1005-1035.
- [10] Hamilton Bean and Ronald J. Buikema. 2015. Deconstituting al-Qa'ida: CCO theory and the decline and dissolution of hidden organizations. *Management Communication Quarterly*, 29, 4, 512-538.
- [11] Faraj, Samer, Georg von Krogh, Eric Monteiro, and Karim R. Lakhani. 2016. Special section introduction—Online community as space for knowledge flows. *Information Systems Research*, 27, 4, 668-684.
- [12] Robertson, John, Ahmad Diab, Ericsson Marin, Eric Nunes, Vivin Paliath, Jana Shakarian, and Paulo Shakarian. 2017. *Darkweb Cyber Threat Intelligence Mining*. Cambridge University Press, Cambridge.
- [13] Aditi Gupta, Anupam Joshi, and Ponnurangam Kumaraguru. 2012. Identifying and characterizing user communities on twitter during crisis events." In *Proceedings of the 2012 workshop on Data-driven user behavioral modelling and mining from social media*, ACM, Hawaii, USA.
- [14] Jennifer Xu and Hsinchun Chen. 2008. The topology of dark networks. *Communications of the ACM*, 51, 10, 58-65.
- [15] Stephen P. Borgatti. 2006. Identifying sets of key players in a social network." *Computational & Mathematical Organization Theory*, 12, 1, 21-34.
- [16] Trevor Van Mierlo. 2014. The 1% rule in four digital health social networks: an observational study. *Journal of Medical Internet Research*, 16, 2, e33.
- [17] Christopher C. Yang, Xuning Tang, and Bhavani M. Thuraisingham. 2010. An analysis of user influence ranking algorithms on dark web forums. In *ACM SIGKDD workshop on intelligence and security informatics*, ACM, Washington D.C.
- [18] Bradley Carron-Arthur, John A. Cunningham, and Kathleen M. Griffiths. 2014. Describing the distribution of engagement in an Internet support group by post frequency: A comparison of the 90-9-1 Principle and Zipf's Law. *Internet Interventions* 1, 4, 165-168.
- [19] Ericsson Marin, Jana Shakarian, and Paulo Shakarian. 2018. Mining key-hackers on darkweb forums. In *2018 1st International Conference on Data Intelligence and Security (ICDIS)*, IEEE, 2018.
- [20] David Décary-Héту and Benoit Dupont. The social network of hackers. *Global Crime*, 13, 3, 160-175.
- [21] Héctor Alvarez, Sebastián A. Ríos, Felipe Aguilera, Eduardo Merlo, and Luis Guerrero. 2010. Enhancing social network analysis with a concept-based text mining approach to discover key members on a virtual community of practice. In *International Conference on Knowledge-Based and Intelligent Information and Engineering Systems*, Springer, Heidelberg, Berlin.
- [22] Jennifer J. Xu, and Hsinchun Chen. 2005. CrimeNet explorer: a framework for criminal network knowledge discovery. *ACM Transactions on Information Systems (TOIS)*, 23, 2, 201-226.
- [23] Nathan K. Cobb, Amanda L. Graham, and David B. Abrams. 2010. Social network structure of a large online community for smoking cessation. *American Journal of Public Health*, 100, 7, 1282-1289.
- [24] Ray Jones, Siobhan Sharkey, Janet Smithson, Tamsin Ford, Tobit Emmens, Elaine Hewis, Bryony Sheaves, and Christabel Owens. 2011. Using metrics to describe the participative stances of members within discussion forums. *Journal of Medical Internet Research*, 13, 1, e3.
- [25] Trevor van Mierlo, Sabrina Voci, Sharon Lee, Rachel Fournier, and Peter Selby. 2012. Superusers in social networks for smoking cessation: analysis of demographic characteristics and posting behavior from the Canadian Cancer Society's smokers' helpline online and StopSmokingCenter.net. *Journal of medical Internet research*, 14, 3, e66.
- [26] David M. Blei, Andrew Y. Ng, and Michael I. Jordan. 2003. Latent dirichlet allocation. *Journal of Machine Learning Research*, 3, 993-1022.
- [27] Carina Jacobi, Wouter Van Atteveldt, and Kasper Welbers. 2016. Quantitative analysis of large amounts of journalistic texts using topic modelling. *Digital Journalism*, 4, 1, 89-106.
- [28] K. Hazel Kwon, Monica Chadha, and Feng Wang. 2019. Proximity and networked news public: Structural topic modeling of global Twitter conversations about the 2017 Quebec mosque shooting. *International Journal of Communication*, 13, 24.
- [29] Christopher Lucas, Richard A. Nielsen, Margaret E. Roberts, Brandon M. Stewart, Alex Storer, and Dustin Tingley. 2015. Computer-assisted text analysis for comparative politics. *Political Analysis*, 23, 2, 254-277.
- [30] Margaret E. Roberts, Brandon M. Stewart, and Dustin Tingley. 2014. stm: R package for structural topic models. *Journal of Statistical Software*, 10, 2, 1-40.
- [31] George C. Banks, Haley M. Woznyj, Ryan S. Wesslen, and Roxanne L. Ross. 2018. A review of best practice recommendations for text analysis in R (and a user-friendly app). *Journal of Business and Psychology*, 33, 4, 445-459.