Maryia Šupa
47 Researching online forums

Abstract: The chapter focuses on online forum research in criminology. It discusses existing methodologies and theoretical challenges faced by researchers, and future developments in the area. As socio-technical systems, online hacking forums provide a valuable source of information about the nature of illicit economies and construction of identities and norms in digital societies.

Keywords: online offending, online forums, research methodologies

Online forums are discussion platforms with multiple contributions from online participants, relevant as sources of observational and interactive data for criminologists. They differ from other online services by providing a space for community-building and long-term collection of knowledge that is structured into searchable, easily accessible archives. Many forums exist as grassroots initiatives independent of corporate online service providers such as Facebook or Google, maintained and moderated by volunteers and administered in accordance with community rules. Online forums have developed as modernized web versions of a previous generation of discussion tools such as bulletin boards and newsgroups. As a platform of computer-mediated communication, forums are a hybrid techno-social tool that combines a technological infrastructure with authentic social interactions (see Infrastructures by Grisot and Parmiggiani).

Online forum research is a subset of the broader research on online communities. It started in tandem with the mass adoption of the web in the late 1990s and early 2000s. Disciplines as diverse as anthropology, education, folklore studies, gender studies, healthcare, psychology, sociology, and, last but not least, criminology included online phenomena in their research. It is a versatile and flexible research field offering unique practical insights. It differs from other online research object such as news websites, social networking services, or online shopping platforms. Online forum research provides grounds for interdisciplinary collaboration between different disciplines, methodologies, and topic areas. A single forum may yield data useful for a range of research questions and analytical approaches.

In criminology, online forums provide the opportunity to reach under-researched and hard-to-access groups such as online offenders, participants in illicit online trade, or stigmatized crime survivors (see Accessing Online Communities by Kaufmann). It facilitates the research of sensitive or covert topics such as sexual abuse. Via forums, researchers can reach transnational and intercultural communities from all over the world. They gain access to both openly published content and hidden content (see Darknet by Tzanetakis), thus facing a host of ethical dilemmas.

Online forums provide access to at least four types of communities that are relevant to digital criminology:

- Communities discussing illicit activities that rely on computers and networks, such as distribution of mobile malware (Grisham et al., 2016), sexually explicit scamming (Pastrana et al., 2019), and stolen data sales (Hutchings and Holt, 2015).
- Communities discussing other clearly illicit activities that also occur offline, including illicit drug sales (Ferguson, 2017), paedophilia (Holt et al., 2010), or illegal consumption of copyrighted material such as music (Caldwell Brown, 2016).
- Communities discussing deviant, although not necessarily illicit activities, such as self-injury (Rodham et al., 2016) or incel subculture (Liggett O'Malley and Helm, 2022).
- Stigmatized communities, including support groups for victims of specific crimes, such as rape survivors (O'Neill, 2018).

The main theoretical approaches used by online forum researchers reflect the paradigmatic divide between positivist and critical research agendas. Positivist approaches are most often represented by rational choice theory (e.g., Smirnova and Holt, 2017; Holt et al., 2015) and situational crime prevention (Chavez and Bichler, 2019), while critical approaches, although much less common, include cultural criminology (Alashti et al., 2022). Criminological studies of online forums frequently lack a thorough discussion of how their theoretical approaches have been adapted to an online context, or how online context expands or transforms key theoretical concepts. For instance, while online hacking forums are often identified as subcultures, scant argumentation is provided about their subcultural traits, communal values, or the changes that occur when members of a subculture interact online.

Methodologies

Online forums have been researched using different methodologies, including quantitative, qualitative, and mixed methods. Quantitative approaches most often rely on large datasets such as the Cambridge Cybercrime Centre's 'CrimeBB' database of online hacking forum posts (Pastrana et al., 2018) or other automatically collected (Munksgaard and Demant, 2016) or leaked (Holt and Dupont, 2019; Overdorf et al., 2018) forum databases. Natural language processing (NLP) and machine learning (ML) are used to analyze large datasets. They combine multiple steps and techniques of analysis. Results allow researchers to determine interests and developmental pathways of key actors in online hacking forums (Hughes et al., 2019), to describe specific offenses (Pastrana et al., 2019), and to predict private interactions between forum members based on publicly posted data (Overdorf et al., 2018). Topic modeling has been used for simpler tasks, such as describing the key topics of online forums specializing in payment card scams (Kigerl, 2018), or analyzing and classifying political discourses on cryptomarket forums (Munksgaard and Demant, 2016). Sentiment analysis has been utilized for models predicting the onset of online offences (Deb et al., 2018). Social network analysis has been used to determine the overall social structures of forum communi-

ties (Pete et al., 2020) and characteristics of sub-communities (Garg et al., 2015). Content analysis, grounded in statistics, has also been conducted. An example is the analysis of variables associated with individuals that have requested membership in a closed hacking forum (Holt and Dupont, 2019).

Qualitative approaches may be carried out on purposive samples from the same large datasets. For example, based on data from the CrimeBB database, researchers have conducted a thematic content analysis of offenses targeting the Internet of Things (Bada and Pete, 2020). Custom data may also be collected specifically for the project. Qualitative content analysis of threads and messages posted in online forums has been used for the analysis of information security advice (Chavez and Bichler, 2019), administration of malware (Hutchings and Clayton, 2017) or development of pedophile subcultures (Holt et al., 2010). Online ethnography and participant observations are long-term projects that add interactions with forum members to content analysis. They have been conducted to explain, for instance, online drug trade (Ferguson, 2017).

Mixed methods studies have combined analysis of data from online forums with other sources of data, e.g., interviews (Collier et al., 2021; Alashti et al., 2022), integrated online and offline ethnography (Potter, 2017; Ferguson, 2017), or incorporated quantitative and qualitative methods for analyzing the same data (Bada et al., 2021).

Contributions

A key contribution of online forum research to criminology is the debunking of popular myths about the internet and online offending. One such myth is that communities labeled as online hacking forums support intense illicit activity (see Hacking by Wall). On the contrary, studies have found the majority of content on these forums is unrelated to offending (Siu et al., 2021), or that the majority of the users in online hacking forums posed little actual risk (Holt et al., 2014). Another myth is that communities focused on offending prevail on non-public networks such as Tor (referred to popularly as the darknet). On the contrary, there are many publicly accessible spots on the web where illicit activity occurs including on major platforms like Facebook. There are also specific myths about the mechanisms of the illicit economy, e.g., the myth that anonymity is the key advantage of cryptocurrency in illicit trade (Butler, 2021).

Beyond debunking myths, criminological research of online hacking forums sparks discussions about the broader implications of the illicit economy facilitated by technology (Holt, 2012). One of such frames interprets participation in illicit online activity as an act of resistance to mainstream capitalism (Maddox, 2020). Analysis of online forums also raises questions about the late modern condition, such as issues of a fragmented, mediated postmodern identity apparent on hacking forums. Issues of trust, risk, uncertainty (Yip et al., 2013), or boredom (Collier et al., 2021) may also be connected to global and local social structures, experienced by offenders and non-offenders alike.

There are several benefits of using online forums as a data source. Their anonymous or pseudonymous nature and open nature of discussions provide convenient access to data on a wide variety of topics, including sensitive and taboo topics. The practical value of positivist online forum research for criminal justice and law enforcement lies in the definition of crime scripts (Hutchings and Holt, 2015; Chavez and Bichler, 2019), intelligence, such as identification of illicit supply chains (Bhalerao et al., 2019) or new malware (Grisham et al., 2017), risk modeling and prediction (Deb et al., 2018), and prevention tactics. Since online forums are communities, they represent distinct cultural milieus. Understanding them is critical to practical outcomes of critical research, creating insight into how norms are formed, evaluated, deliberated, and breached. From the point of view of critical theory, it is also possible to observe how external discourses find their expression in online forums and reflect the broader configurations of discursive power.

Forums are not just content publishing platforms but communities that bring together active contributors and a network of committed members. They allow exploring how community members create and deliberate the content, how they express the rationale behind different standpoints, and how they maintain their social ties. Forums are instrumental for the construction and performance of a collective identity (Maltby et al., 2018: 1775) by the community. Their structures facilitate the building of trust, which has special significance for communities focused on offending (Yip et al., 2013: 535). Other community functions include control, coordination, social networking, and management of uncertainties and risks (Yip et al., 2013). All of these help understanding of how norms and transgressions are negotiated and evaluated.

Challenges

Online forum research also presents practical challenges. One challenge, especially for quantitative analysis of online forum data, is its informal and unstructured language. If data is collected automatically, additional tools are needed to filter out spam and unsuitable data. Online forum posts are often short, weakly structured, include non-conventional language, emoticons, and abbreviations that make structural and algorithmic analysis of the text difficult (Ferguson, 2017: 694).

Researchers face a host of ethical dilemmas (see Ethics by Markham). They must delicately balance between disclosing enough information about their research process to make it reproducible, and preserving the anonymity of the studied communities, especially if they are vulnerable or non-public (franzke et al., 2020). Some online forum content may be extreme and harmful to the researcher and requires additional scrutiny and preparation, as well as a situated approach to ethics (Kelley and Weaver, 2020). It is often impossible to request consent from online forum members, and much of the research is carried out covertly, which leads to ethical challenges. While researchers benefit from publishing their findings, there are few opportunities for forum users to familiarize themselves with the research, react to it, or share some of the possible

benefits if research is done covertly. Online forum contents were not initially created as research data, so further questions arise about the ethics of using them.

Beyond researching forums

Online forums, collections of topical posts and discussion threads interactively produced by their users, are not the only platforms where discussions of potentially illicit activities occur. Other platforms should not be overlooked when considering potential data sources. These may include platforms such as Reddit which hosts community discussions, but is less structured than online forums, and provides additional tools such as upvoting or downvoting posts. Applications for chatting and group communication such as Telegram, Signal, or Discord may be accessed. However, their content is unstructured and may quickly disappear, necessitating approaches that allow description, interaction, and analysis as an ongoing process rather than collecting data in a single step. Illicit online marketplaces may provide valuable criminological insight, and should not be confused with online forums in general. While some marketplaces exist within forum ecosystems, the most prominent ones are exclusively focused on trade and provide additional functionality such as escrow services—the payment is held by the marketplace until the seller and buyer complete their transaction.

Summary of key points

- Online forum research is a subset of online community research focusing on online forums—discussion platforms which allow communities to create, deliberate, and organize various forms of knowledge, and to create and maintain social ties via online interactions. They are valued as a source of scientific data for the variety of topics that may be addressed and large diversity of possible approaches.
- Online forum research is significant for digital criminology as a source of knowledge about communities discussing online and offline illicit activities, as well as deviant and stigmatized groups. Although currently positivist theoretical frameworks prevail in criminological works about online forums, critical approaches also exist.
- Criminological research of online forums facilitates the debunking of several popular myths about online and offline offending. It encourages broader discussion about the nature of illicit economies and construction of identities and norms in digital societies.
- Different quantitative, qualitative, and mixed methods designs have been applied in online forum research. They range from automated analysis based on Natural Language Processing of large data samples from multiple forums to in-depth thematic coding by hand of purposively sampled messages. While much research in this area is based on content only, participant online ethnography allows to lever-

age the direct involvement of a researcher with an online forum community (see Online Ethnography by Gibbs and Hall).

– The key benefits of online forum research are convenient access to published data and opportunity to observe how online communities function, including negotiation of norms and transgressions. The key challenges are ethical dilemmas, such as covert research or the use of data not produced for research purposes, linguistic complexities, and competing platforms that may be used to study similar communities.

– The boundary between online and offline life has become blurred and is continuing to do so. Most social practices, including illicit ones, become increasingly represented online, even if the activities themselves occur offline. Therefore, online forums will remain one of the significant sources of data about a large range of offending and deviant behavior (Potter, 2017: 3–4), including newly emerging forms of crime.

Suggested reading

franzke, a. s., Bechmann, A., Zimmer, M., Ess, C., & the Association of Internet Researchers (2020). *Internet Research: Ethical Guidelines 3.0.* Available online at: https://aoir.org/reports/ethics3.pdf

Ferguson, R. H. (2017). Offline 'stranger' and online lurker: Methods for an ethnography of illicit transactions on the darknet. *Qualitative Research*, 17(6), 683–698.

Jemielniak, D. (2020). *Thick Big Data: Doing Digital Social Sciences.* Oxford: Oxford University Press.

References

Alashti, Z. F., Bojnordi, A. J. J., & Sani, S. M. S. (2022). Toward a carnivalesque analysis of hacking: A qualitative study of Iranian hackers. *Asian Journal of Social Science*, 50(2), 147–155.

Bada, M., & Pete, I. (2020). An exploration of the cybercrime ecosystem around Shodan. In *2020 7th International Conference on Internet of Things: Systems, Management and Security (IOTSMS)*, 1–8.

Bada, M., Chua, Y. T., Collier, B., & Pete, I. (2021). Exploring masculinities and perceptions of gender in online cybercrime subcultures. In M. W. Kranenbarg & R. Leukfeldt (eds.), *Cybercrime in Context: The Human Factor in Victimization, Offending, and Policing* (pp. 237–257). Cham: Springer International Publishing.

Bhalerao, R., Aliapoulios, M., Shumailov, I., Afroz, S., & McCoy, D. (2019). Mapping the underground: Supervised discovery of cybercrime supply chains. In *2019 APWG Symposium on Electronic Crime Research (eCrime)*, 1–16.

Butler, S. (2021). Cyber 9/11 will not take place: A user perspective of bitcoin and cryptocurrencies from underground and dark net forums. In: *STAST 2020: Socio-Technical Aspects in Security and Trust*, 135–153.

Caldwell Brown, S. (2016). Where do beliefs about Music Piracy come from and how are they shared? An ethnographic study. *International Journal of Cyber Criminology*, 10(1), 21–39.

Chavez, N., & Bichler, G. (2019). Guarding against cyber-trespass and theft: Routine precautions from the hacking community. *International Journal of Cyber Criminology*, 13(1), 101–116.

Collier, B., Clayton, R., Hutchings, A., & Thomas, D. (2021). Cybercrime is (often) boring: Infrastructure and alienation in a deviant subculture. *The British Journal of Criminology*, 61(5), 1407–1423.

Deb, A., Lerman, K., & Ferrara, E. (2018). Predicting cyber-events by leveraging hacker sentiment. *Information*, 9(11).

Ferguson, R. H. (2017). Offline 'stranger' and online lurker: Methods for an ethnography of illicit transactions on the darknet. *Qualitative Research*, 17(6), 683–698.

franzke, a. s., Bechmann, A., Zimmer, M., Ess, C., & the Association of Internet Researchers (2020). *Internet Research: Ethical Guidelines 3.0.* Available online at: https://aoir.org/reports/ethics3.pdf

Garg, V., Afroz, S., Overdorf, R., & Greenstadt, R. (2015). Computer-supported cooperative crime. In R. Bohme & T. Okamoto (eds.), *Financial Cryptography and Data Security* (pp. 32–43). Heidelberg: Springer Berlin.

Grisham, J., Samtani, S., Patton, M., & Chen, H. (2017). Identifying mobile malware and key threat actors in online hacker forums for proactive cyber threat intelligence. In *2017 IEEE International Conference on Intelligence and Security Informatics (ISI)*, 13–18.

Holt, T. J. (2012). Examining the forces shaping cybercrime markets online. *Social Science Computer Review*, 31(2), 165–177.

Holt, T. J., & Dupont, B. (2019). Exploring the factors associated with rejection from a closed cybercrime community. *International Journal of Offender Therapy and Comparative Criminology*, 63(8), 1127–1147.

Holt, T. J., Blevins, K. R., & Burkert, N. (2010). Considering the pedophile subculture online. *Sexual Abuse: A Journal of Research and Treatment*, 22(1), 3–24.

Holt, T. J., Smirnova, O., Strumsky, D., & Kilger, M. (2014). Case study: Advancing research on hackers through social network data. In G. E. Higgins & C. D. Marcum (eds.), *Social Networking as a Criminal Enterprise* (pp. 45–163). Boca Raton, FL: CRC Press.

Holt, T. J., Smirnova, O., Chua, Y. T., & Copes, H. (2015). Examining the risk reduction strategies of actors in online criminal markets. *Global Crime*, 16(2), 81–103.

Hughes, J., Collier, B., & Hutchings, A. (2019). From playing games to committing crimes: A multi-technique approach to predicting key actors on an online gaming forum. In *2019 APWG Symposium on Electronic Crime Research (eCrime)*, 1–12.

Hutchings, A., & Clayton, R. (2017). Configuring Zeus: A case study of online crime target selection and knowledge transmission. In *2017 APWG Symposium on Electronic Crime Research (eCrime)*, 1–8.

Hutchings, A., & Holt, T. J. (2015). A crime script analysis of the online stolen data market. *British Journal of Criminology*, 55(3), 596–614.

Kelley, B., & Weaver, S. (2020). Researching people who (probably) hate you: When practicing "good" ethics means protecting yourself. *Computers and Composition*, 56, 1–9.

Kigerl, A. (2018). Profiling cybercriminals: Topic model clustering of carding forum member comment histories. *Social Science Computer Review*, 36(5), 591–609.

Maddox, A. (2020). Disrupting the ethnographic imaginarium: Challenges of immersion in the Silk Road cryptomarket community. *Journal of Digital Social Research*, 2(1), 20–38.

Maltby, S., Thornham, H., & Bennett, D. (2018). Beyond 'pseudonymity': The sociotechnical structure of online military forums. *New Media & Society*, 20(5), 1773–1791.

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.

Liggett O'Malley, R., & Helm, B. (2022). The role of perceived injustice and need for esteem on incel membership online. *Deviant Behavior*, 44(7), 1026–1043.

O'Neill, T. (2018). 'Today I speak': Exploring how victim-survivors use Reddit. *International Journal for Crime, Justice and Social Democracy*, 7(1), 44–59.

Overdorf, R., Troncoso, C., Greenstadt, R., & McCoy, D. (2018). Under the underground: Predicting private interactions in underground forums. *arXiv preprint,* arXiv:1805.04494.

Pastrana, S., Thomas, D. R., Hutchings, A., & Clayton R. (2018). CrimeBB: Enabling cybercrime research on underground forums at scale. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web – WWW '18*, 1845–1854.

Pastrana, S., Hutchings, A., Thomas, D., & Tapiador, J. (2019). Measuring e-whoring. In *Proceedings of the Internet Measurement Conference*, 463–477.

Pete, I., Hughes, J., Chua, Y. T., & Bada, M. (2020). A social network analysis and comparison of six dark web forums. In *2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*, 484–493.

Potter, G. R. (2017). Real gates to virtual fields: Integrating online and offline ethnography in studying cannabis cultivation and reflections on the applicability of this approach in criminological ethnography more generally. *Methodological Innovations*, 10(1), 1–11.

Rodham, K., Gavin, J., Lewis, S., Bandalli, P., & St. Denis, J. (2016). The NSSI paradox: Discussing and displaying NSSI in an online environment. *Deviant Behavior*, 37(10), 1110–1117.

Siu, G. A., Collier, B., & Hutchings, A. (2021). Follow the money: The relationship between currency exchange and illicit behaviour in an underground forum. In *2021 IEEE European Symposium on Security and Privacy Workshops*, 191–201.

Smirnova, O., & Holt, T. J. (2017). Examining the geographic distribution of victim nations in stolen data markets. *American Behavioral Scientist*, 61(11), 1403–1426.

Yip, M., Webber, C., & Shadbolt, N. (2013). Trust among cybercriminals? Carding forums, uncertainty and implications for policing. *Policing and Society*, 23(4), 516–539.