A Little Antagonism Might Be Nice: Investigation in Information Science

Carter, Daniel[1] Dan Sholler[2] and Amelia Acker[3]

[1] Texas State University, San Marcos, TX, USA

[2] University of California, Berkeley, CA USA

[2] University of Texas at Austin, Austin, TX USA

Abstract. Academic research often claims to investigate phenomena, but we conventionally insist that such investigations take place with the consent of those being studied. In this blue sky paper, we suggest that information science researchers should consider the contexts in which it might be beneficial to violate this norm and pursue what we describe as antagonistic investigation. In relation to illegal and socially harmful activities such as platform manipulation, fraud and the spread of propaganda, we argue that researchers should go against the wishes of those they're studying and possibly, in the process, violate privacy norms, challenge illegal activities and call for accountability as a result of research. While these investigative activities are not conventional in information science research, they draw on core strengths of the field and position researchers to produce impactful work on relevant and pressing topics.

Keywords: methods, methodology, investigation, social media

1 Introduction

While academic research often claims to "investigate" phenomena, this broad use of the term is distinct from more specialized uses, which tend to be associated with the work of detectives, auditors or investigative journalists. One key distinction between academic research and these more forensic forms of investigation is an antagonistic approach to the identification and revelation of activities that are intended to remain hidden. In this blue sky paper we suggest the potential value of antagonistic investigation for researchers seeking to uncover contexts or connections that subjects of the research may not want revealed. In short, we propose that information researchers be more aggressive and more antagonistic. Specifically, we propose the adoption of investigative techniques as a way to provide context to online activities that are illegal or socially harmful and that are likely not fully captured by contemporary research methodologies that rely on, for example, publicly available metadata or statements from individuals engaged in socially harmful activities.

By proposing antagonistic investigation as a potential mode of research or source for inspiration, we do not intend to suggest that current research is naïve or fails to verify claims. Instead, we hope to invite a conversation about how new techniques might enhance current research—a conversation not only about what scholars could do in the future but also about what is currently done but not communicated in formal venues of publication. We suggest that many scholars do incorporate moments of antagonism, friction, opposition, and even enmity in their investigative work but do not explicitly frame their research in this way; we also suggest that one consequence of this pattern is a lack of instruction at the graduate level in techniques that could lead to richer accounts of important online phenomena as well as more compelling forms of public scholarship.

In order to argue for why a little antagonism might be a good thing for empirical research in information science, we briefly review existing approaches for studying activities that are illegal and socially harmful before presenting three brief cases that benefit from an antagonistic investigative approach.

2 Existing Approaches

Models for academic research aimed at activities that are illegal, socially harmful or otherwise hidden include studies that suggest motivation by critiquing interfaces or platforms, studies drawing on data generated through online activities and forms of covert participant observation.

When considering aspects of online platforms or services that might have harmful consequences or that might result from motivations that are purposefully hidden, re-

searchers have sometimes used methods drawn from platform studies or interface criticism as a way to suggest motivation or consequence. Helmond (2015), for example, discusses "platformitization," through an analysis of the technical frameworks of the web (p. 8). Similarly, Gray et al. (2018) describe "dark patterns" in user experience design through an analysis of exemplars. Both of these studies work from observed technical or design features and use these to infer motivations that are unlikely to be shared publicly, for example in a conventional, qualitative interview with an informant.

Another approach to studying activities taking place online is the analysis of metadata or trace data. Approaches focusing on online data made available through social media APIs and similar sources, for example, have described the characteristics of Twitter (e.g., Bastos & Mercea, 2017) and other forms of online propaganda (e.g., Kim et al 2018). While these studies reveal activities that those performing the activities would prefer to remain hidden or misunderstood, they generally do not reveal with any specific information about how the activities were performed. For example, while Bessi and Ferrara (2016) describe the extent to which bots interfered with political communication related to the 2016 US presidential election, but their work does not attempt to draw connections between bots and those who created them or to understand the motivations around their creation to manufacture consensus or manipulate political discourse in social media platforms and beyond.

Using metadata or trace data to study illegal or socially harmful activities is also complicated by current patterns of data access and ownership. While social media platforms such as Twitter and Facebook provide researchers and others with access to some data, this access is limited and increasingly subject to review by the platforms—as Freelon (2018) argues, researchers are now or will soon be working in a "post-API" age. At the same time, researchers working within social media companies have greatly expanded access to data, including greater volumes of data and more sophisticated tools for searching and filtering. While this access—and the non-antagonistic relationships on which it relies—enables large, inductive studies of behavior, it raises questions about scholars' abilities to produce detailed accounts or cases that might reveal undesirable activities occurring on a platform or even ways that those designing or managing the platform are complicit in such activities.

Finally, forms of covert participant observation—in which the researcher embeds with a group or organization while concealing their purpose or goals—provide the most direct access to illegally or socially harmful. As Roulet et al. (2017) note, covert participation conflicts with the conception of informed consent adopted by many ethics guidelines but has value for producing data that would otherwise be unavailable and for providing contexts in which the interpret such data. As Maguire et al.'s (2018) study of heteroactivist groups demonstrates, covert participant observation can also entail an antagonistic relationship between researchers and those studied—however, such methods are unlikely to be available to all researchers or applicable in all contexts.

3   Case Studies

Here we suggest three examples from our own research using investigative approaches that might benefit academic research specifically in relation to online publishing platforms.

3.1 Violating Privacy Norms

In an ongoing research project about online product promotion, Carter has worked over the past year to identify the owner of a Facebook group that posts sweepstakes and contests. While scholarly publications emerging from the project will draw on conventional processes of large-scale data collection, qualitative coding and analysis, he has also attempted to understand the motivations behind the group, the specific ways it generates profit and the relationships between the group's owner and its members. As Carter pursued these questions, the owner of the group sometimes ignored requests for interviews, sometimes instructed people not to speak with him and sometimes responded to emails but gave accounts that contradicted other statements. Looking for alternative approaches, Carter formed relationships with the owners of competing groups and with others familiar with the group's business practices. He combined these sources with domain registration records and property records to identify the owner of the group and to produce evidence of their deceptive practices.

Releasing personal information about someone who does not want to be identified would be considered unethical in many contexts, and the owner of the Facebook group has actively avoided making this information public. Indeed, doing so might directly hurt their business, reveal exploitation, or change the group's interactions in severe ways. However, from a research perspective, any description of the group that does not account, for example, for the discovered wealth disparity between the group's owner and its members risks an incomplete or even misleading portrayal.

3.2 Challenging Illegal Activities

Sholler is conducting a qualitative study of the governance of so-called "Open Science" organizations. Open Science organizations promote and facilitate unfettered access to products of the scientific enterprise, including experiment protocols, datasets, analysis tools and technologies, and scholarly publications. Among the services available for sharing academic publications are illegally-hosted shadow libraries and archives such as Sci-Hub and LibGen, which aim to make paywalled publications available to all via a simple, web-based interface.

Organizations such as Sci-Hub and LibGen operate via illegal activities that violate copyright and other intellectual property laws in the US and around the world. Perhaps due to the illegal nature of the activities, scholars know little about how these organizations develop and sustain the organizational and decision-making structures to support their operations. Some studies (e.g., Himmelstein et al., 2018) have been able to quantify the success of Sci-Hub via analyses of the availability of academic publications, but little is known about the organization's leaders, employees, decision-making structures, and financial sustainability. We might assume that their success is not just a factor of the interface or technical features that we can see from the outside, but might also have something to do with these organizational structures and methods of decision-making. Sholler's study attempts to uncover the avenues through which members of Open Science organizations communicate, collaborate, and make transactions. Members of these organizations have, for the most part, hesitated to speak with researchers and media outlets and often conduct their activities using pseudonyms and encrypted communication. As such, researchers studying in the operation of shadow libraries and archives must find creative, unorthodox ways to access and investigate the phenomena of interest. These methods—such as interviewing former contributors or organizational members, openly challenging the legality of the activities on public and private forums to elicit rationales and explanations, and otherwise antagonizing organizational members to describe day-to-day operations—place the researcher in a precarious and interrogative position relative to the individuals under study.

3.3 Calling for Accountability

The presence of false activity data in platforms has had a number of unpredictable consequences for social media users and society—from election tampering to genocide. Acker has been developing digital methods for tracking disinformation campaigns online by "reading metadata" that has been gamed, falsified, or exploited in social media platforms. Manipulated metadata make fake digital traces look like authentic user behavior. While social media professionals, influencers, and celebrities have been known to inflate engagement activities with inauthentic likes, views, follower counts, and comments for profit, platforms have been slow to identify manipulation that manufactures ideological and political discord. As such, manipulation techniques that leverage platform features to spread disinformation by mimicking authentic platform use without intent to profit are harder to identify with automated moderation techniques.

While many journalists and computational social scientists have leveraged metadata from Twitter APIs to identify the spread of computational propaganda like bots and sockpuppets, few methods for identifying disinformation attempt to locate coordinated inauthentic behavior as disinformation campaigns are seeded and spread. Reading metadata for inauthentic activity signals not only gives us insight into the emerging techniques of manipulators, it is also a way of understanding the power structures of platforms themselves. When we read metadata that's been exploited or gamed in social media platforms, but has not been caught out by the platform's accountability tools, researchers can begin to decode signals from coordinated inauthentic activity found in disinformation campaigns. By investigating such media manipulation tactics that evade

platforms' moderation efforts with metadata of coordinated inauthentic activity, we both learn about how disinformation spread is supported by platform features *and* how they have been slow to correct it. Thus investigating disinformation techniques can position the researcher against manipulators (by cataloging their skullduggery) and hold platforms accountable for their moderation efforts. This antagonistic approach questions narratives put out by social media platforms and can reveal ways that companies are complicit in the activities they claim to police.

4 Conclusion

Each of these cases highlight contexts that information science, as a field, is uniquely positioned to speak to: understanding how social media algorithms and online advertising models encourage unethical business models and create opportunities for the creation of exploitative communities; drawing organizational lessons from illegal libraries, archives, and the dark web; and reading metadata as a way of calling into account both media manipulators and the platforms that quietly enable their schemes. In line with the fundamental methodology of information science, these contexts call for a deeply sociotechnical approach (Bates, 1999). At the same time, however, they also call for something that information science has yet to claim: an antagonistic investigative approach that encourages researchers to pursue cases even when key players are not willing participants. Such an approach would play to many of the core strengths of information science such as a deep understanding of social and technical processes and would place researchers in an advantageous position to speak in different ways, both to academic and public audiences, on pressing topics related to manipulation, fraud, illicit online activities, information sharing and propaganda.

For academic researchers to take more antagonistic and openly investigative approaches would require considerable revision of norms, especially related to ethics and accountability. Many of the questions to be answered relate to the specificity of investigation—rather than producing generalizable knowledge, the models of investigation cited here attempt to nail down facts and, specifically, facts that are purposefully hidden from researchers as a condition of pursuing unethical or antisocial practices. While academic researchers are familiar with producing case studies and using these accounts to further the development of theory (Yin, 2017), the details of how these studies are produced and communicated would require discussion if antagonistic investigative approaches were to become more common.

One challenge for information science researchers is the lack of heuristics or more formal rules for evaluating specific findings when they are not simply observed or communicated in an interview, as would be the case with conventional case study research. The roots of investigation, after all, are in abduction, a form of logical inference that suggests what is probable but does not necessarily prove what must be (as with deduction) or arrive at a general rule through observing particular cases (as with induction). While communities such as investigative journalists have standards that can be applied to findings, academic communities would likely need to develop these for themselves or at least openly articulate those that implicitly exist. Similarly, while investigative journalists share an understanding of ethical norms, this is likely a topic that would require considerable debate among academic researchers who have conventionally worked to ensure that participants have control over how research is conducted and communicated.

Despite these challenges, the broad methodological process for moving from specific cases to theory is firmly established within many academic communities. While the antagonistic nature of the tactics employed might be unusual in that they violate privacy norms, challenge illegal activities and call for accountability, the product of the research is, we suggest, well within the bounds of what many scholars in the social sciences are comfortable working with.

Beginning a discussion on methods of investigation would enable the information science community to pursue topics that are timely and relevant—but that are also only partially revealed by conventional methods. As the reach and consequences of ICTs grow, it is increasingly important to understand how socially harmful practices are conducted, and this should include an understanding of who is involved, how they are organized, what their motivations are and how they profit from their actions. While antagonistic investigation is not a mode that academic researchers are familiar with, we believe that its potential to the study of people, information and technology is worth exploring.

References

Bastos, M. T., & Mercea, D. (2017). The Brexit Botnet and User-Generated Hyperpartisan News. *Social Science Computer Review*, 0894439317734157. https://doi.org/10.1177/0894439317734157

Bates, M. J. (1999). The invisible substrate of information science. *Journal of the American Society for Information Science*, *50*(12), 1043–1050. https://doi.org/10.1002/(SICI)1097-4571(1999)50:12<1043::AID-ASI1>3.0.CO;2-X

Bessi, A., & Ferrara, E. (2016). Social bots distort the 2016 U.S. Presidential election online discussion. *First Monday*, *21*(11). Retrieved from http://firstmonday.org/ojs/index.php/fm/article/view/7090

Freelon, D. (2018). Computational research in the post-API age. *SocArXiv*. https://doi.org/10.31235/osf.io/56f4q

Gray, C. M., Kou, Y., Battles, B., Hoggatt, J., & Toombs, A. L. (2018). The Dark (Patterns) Side of UX Design. In *Proceedings of the 2018 CHI Conference on Human Factors in Computing Systems* (pp. 534:1–534:14). New York, NY, USA: ACM. https://doi.org/10.1145/3173574.3174108

Helmond, A. (2015). The Platformization of the Web: Making Web Data Platform Ready. *Social Media + Society*, *1*(2), 2056305115603080. https://doi.org/10.1177/2056305115603080

Kim, Y. M., Hsu, J., Neiman, D., Kou, C., Bankston, L., Kim, S. Y., … Raskutti, G. (2018). The Stealth Media? Groups and Targets behind Divisive Issue Campaigns on Facebook. *Political Communication*, *0*(0), 1–27. https://doi.org/10.1080/10584609.2018.1476425

Maguire, H., McCartan, A., Nash, C. J., & Browne, K. (2018). The enduring field: Exploring researcher emotions in covert research with antagonistic organisations. *Area*, *0*(0). https://doi.org/10.1111/area.12464

Roulet, T. J., Gill, M. J., Stenger, S., & Gill, D. J. (2017). Reconsidering the Value of Covert Research. *Organizational Research Methods*, *20*(3), 487–517. https://doi.org/10.1177/1094428117698745

Yin, R. K. (2013). *Case Study Research: Design and Methods*. Thousand Oaks, CA: SAGE Publications.