Studying Cybercrime: Raising Questions about Objectivity and Bias

Marie Joan Kristine T. Gloria  
Tetherless World Constellation  
Rensselaer Polytechnic Institute  
110 8th St  
Troy, NY 12180  
glorim@rpi.edu

John S. Erickson  
Web Science Research Center  
Rensselaer Polytechnic Institute  
110 8th St  
Troy, NY 12180  
erickj4@rpi.edu

ABSTRACT
The multidisciplinary nature of Web Science is at once its most unique and most challenging characteristic. In this paper, we pose initial questions on how a Web Scientist may explore the topic of cybercrime and cyberwarfare by offering an integrated study of both the social and technical aspects of the phenomena. With this, however, comes a plethora of questions related to methodological practice, data analysis, narrative creation, etc. As such, the following presents initial discussions related to the specific aspects of - objectivity and bias - in science research. The paper includes one example cybercrime case study, which elaborates on the steps taken by the research team to mitigate biases. The paper concludes with additional questions and potential tools/methods to mediate bias for future Web Science studies.

Categories and Subject Descriptors
K.0 [Computing General]

General Terms
Web Science, Cybercrime, Methodology

Keywords
Web Science, Cybercrime, Methodology

1. INTRODUCTION
The semantics that separate cybercrime from cyberwarfare has generated a great deal of debate amongst security experts, law practitioners, and academic researchers. According to the United Nations Interregional Crime and Justice Research Institute (UNICRI), cyberwarfare “refers to any action by a nation-state to penetrate another nation’s computer networks for the purpose of causing some sort of damage” [1]. This may be broadened to include acts of cybercrime, cybervandalism, cyberterrorism, etc. [1]. Yet, in reflecting on the role of Web Science within this discourse, one aspect surfaces as agnostic to these definitional disagreements. This paper focuses on objectivity and biases within studies of cybercrime and cyberwarfare from the perspective of Web Science and thereby seeks to identify the methodological practices and tools employed by Web Scientists to maintain objectivity. Our aim is to provoke discussion amongst Web Scientists interested in such a complex phenomena wherein notions of power, economics, governance, identity and culture are deeply integrated with the technology. The paper touches upon decades of debate in the social sciences regarding objectivity and the role of the researcher, and assert that to objectively identify instances of cybercrime or cyberwarfare, a researcher must first be transparent about his or her own decision making process as well as the provenance of data collected. To illustrate potential complexities, we offer a specific example of challenges faced in exploring the U.S. anti-surveillance culture. We conclude with a set of recommended practices and tools that may be of interest to Web Scientists studying such socio-technical phenomena. We argue that by inciting such reflection, Web Scientists are primed to present a unique, novel perspective on the study of cybercrime, cyberwarfare, etc.

2. OBJECTIVITY AND BIASES
The discourse of objectivity and research bias has a long, debated history. Philosophy of science literature as well as sociology, history, and anthropology hold varied perspectives on its definition and its influence on the practice of science. For example, Porter (1996) traces objectivity as having multiple interpretations construed to include notions of fairness, mechanical objectivity, and non-subjectivity [2]. Latour’s (2000) critique goes even further suggesting that “objectivity does not refer to a special quality of the mind . . . but to the presence of objects which have been rendered ‘able’ to object to what is told about them” [3]. The natural sciences, which observe natural objects/phenomenon, inherently escape such critique [3]. But, what and how does this critique apply to Web Science; and in particular how should such considerations be made in studying cybercrime or cyberwarfare? We contend that as the Web Science community deliberates how to study cybercrime or cyberwarfare, it must demonstrate its full recognition of the complexities of a socio-technical phenomenon by extending its scope beyond simply the technical or social domains, but towards an integration of the two. Therefore, concerns regarding objectivity and biases in studying social phenomenon and the technology involved in those studies are not mutually exclusive.

3.A MULTI-DISCIPLINARY EXAMPLE
The combining of and collaboration among multiple disciplines is not easy. As Halford et. al (2013) note, the challenge exposes the “different epistemologies, histories and languages of sociology and computer science, which raise questions about the wider
politics of knowledge and dynamics of power and identity that arise in multidisciplinary work” [4]. The study of cybercrime and cyberwarfare provide examples supporting this observation. The following discussion reflects on the methodological process of an early exploratory case study regarding the anti-surveillance culture in the United States. The detailed results and conclusions are not included as they fall outside the scope of this paper.

Fieldwork was conducted in Fall 2013 as part of an exploratory study investigating constructions and practices of online and offline activist behavior. The field site was located in Washington, D.C. and motivated by an anti-surveillance protest rally. From a Web Science perspective, this study required an examination of individual behavior and motivation; identification of authority, power and governance structures; and consideration of technology’s involvement as a propagator and facilitator of information flow. “Cybercrime” in this instance was defined not as an action of a nation-state unto another nation-state, but rather a single agent’s action onto a nation-state, a definition motivated by the U.S. government’s same use of the term in identifying Edward Snowden’s act as a cybercrime.

Given the multi-layered nature of the study and chosen methodological approach, the researcher engaged in a reflexive practice to denote biases that may potentially affect the outcome of the study. Explicit biases such as the construction of the initial interview questions were first examined. These questions focused on gathering information related to individual motivation, modes of information propagation, levels topic comprehension, etc. Next, the researcher mapped out implicit biases such as organizational affiliations. In this instance, the fact of the researcher's membership with one of the protest-rally coordinating organizations and knowledge of non-publicized details of the rally would need to be considered. Upon completion and transcription of the interviews all collected data was reviewed for bias, such as self-selection sampling bias. In addition to the interview narratives, the study was concerned with the role of the Web as the platform of choice to facilitate and coordinate of the protest rally. Social media properties, such as Twitter, Instagram, and reddit, played significant roles. The development and deployment of StopWatchingUs website was also considered. As iterations of this study continue, researchers will explore not just how these platforms were used, but why and by whom. Questions that could be potentially explored include graph structure, influence, information flow, etc.

4. DISCUSSION

The above example presents what we consider to be an attempt at a comprehensive Web Science study, one that includes multiple disciplines, multiple perspectives, and multiple methods. This example touched on the difficulties of combining such diverse analysis and methodological practices. However, in trying to integrate all of these elements into one single, cohesive narrative is difficult. First, identifying intersections between more descriptive, context-dependent interview data with stand-alone graph analysis requires a negotiation of which content is included and excluded in the final analysis. Second, while biases are made explicit by the researcher, it is difficult to express and share these biases with others. Currently, most of these considerations are kept hidden in IRB applications or in individual notebooks. Third, reflexivity is not an explicit practice when considering questions like social network analysis. Growing criticism of unconscious or hidden biases in computing models, analytics and algorithms underscore the potential for such consideration [5] [6].

In conclusion, the goal of this paper was to incite additional conversations related to how the Web Science community can mediate certain disagreements. Specifically, we focus on questions regarding objectivity and biases. We considered how the study of cybercrime and cyberwarfare serves as a prime example for the potential complications that may manifest when studying both the social and technical aspects of a phenomena. Our hope is to learn from the long standing practices and principles of the multiple disciplines associated with Web Science to help answer questions related to core methodological considerations.

5. ACKNOWLEDGMENTS

Our thanks to Joanne S. Luciano of Center for Cognition, Communication, and Culture, Rensselaer Polytechnic Institute for her contributions.

6. REFERENCES