
Cultural Competence: An Ethical Model for Big Data Research

Amirah Majid

The Information School
University of Washington
Seattle, WA 98105 USA
amirah@uw.edu

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. To copy otherwise, or republish, to post on servers or to redistribute to lists, requires prior specific permission and/or a fee.

Abstract

Data science research *is* human subjects research. The impressive numbers we can generate from big datasets are too often deemed a sufficient indicator of reality. This position paper attempts to promote adopting the cultural competency framework, borrowed from medicine, to big data research. Researchers enact a culturally competent stance when they learn about culture, recognize that the plurality of their work includes their own culture, and actively attempt to minimize harm. This ethical stance encourages using the needs of all stakeholders to shape data gathering and transformation.

Author Keywords

cultural competence, ethics, signaling theory, big data

ACM Classification Keywords

H.3 [Information interfaces and presentation]: Group and Organization Interfaces - *Computer Supported Cooperative Work, theory and models.*

Introduction

Big data research is frequently based around the harvesting and analysis of large scale representations of social interactions. Disembodied datasets are given statistical weight and relevance just by sheer virtue of size. However, the statistical claims to 'truth' are

subject to a number of ethical concerns regarding the collection, analysis and subsequent presentation of the data as a research result.

This position paper proposes that the complex ethical issues around studying people's behaviors and intentions through big data, specifically social media data, necessitates a rigorous and systematic ethical framework. Studying online user behavior *is* human subjects research, however as our distance from the participants becomes greater so do our assumptions. Harvesting millions of lines of 'data' has seemingly given us the opportunity to make broad, sweeping, generalizations about populations irregardless of the cultural differences of those populations. Quantity does not equal quality, nor does quantity erase researchers obligations to accurately represent and deliver back culturally competent findings.

We will begin by describing cultural competency and then provide an example illustrating how coming from a participant centered, culturally competent approach has shaped our research questions and the manner in which we have chosen to gather data. We close by emphasizing a need for greater intentionality and participation at the institutional level in ethical procedures.

Cultural Competency: an ethical stance

Social media is a cross-cultural environment, regardless of the platform, and thus any data harvested from social media is *inherently* cross-cultural. With this in mind researchers are faced by a series of ethical and philosophical questions which must be addressed at all stages of data collection and analysis. "We have no real sense from such data what the motivations behind

engagement and disengagement might be or how socioeconomic and other structural factors might be shaping patterns of activity. Most importantly, there is a basic question of meaning to be applied to so-called 'big social data'" [1]. A culturally competent ethical stance to big data research is a rational framework from which to remove the distance between researcher and the (witting or unwitting) participant because it breaks down the reification of the data and shifts focus back to "...methodological questions concerning data construction, sampling, interpretation and analysis, and the ways in which data 'trends' have themselves been shaped by commercial interests and contingent sociotechnical processes" [1].

Cultural competency encompasses a set of practices and ethical standpoints most frequently explored in medical literature and curriculum. As a model it is mostly absent from CSCW literature even when the study focus is set around examining aspects of culture. After a thorough review of literature Dr. Michael Paaschow-Orlow's, associate professor of medicine at Boston University, summing up of cultural competency into three essential principles presents a good jumping off point: [2]

- *Principle 1: Acknowledgement of the importance of culture in people's lives.*

This principle affirms 'participant' centeredness as the goal of interaction. However, "...detached mastery of particular cultures is never, in and of itself, the goal of cultural competence," because in order "to develop tools for insight and improved working relationships" researchers must acknowledge and attempt to learn how cultural factors influence their subjects.[2]

- *Principle 2: Respect for cultural differences:*
This principle asks that researchers develop a pluralistic understanding of culture and a respect for the autonomy of their participants, witting or otherwise. This demands that CSCW and information science researchers in general gain an explicit self-awareness, and discipline specific awareness of their distinct cultures. They must have, "security with their own viewpoint to not be threatened by alternative points of view." [2]
- *Principle 3: Minimization of any negative consequences of cultural differences:*
Finally, though this principle this can be an expensive and complicated one, "this type of evaluation should explore the range of corrective actions needed, from the removal of frank bias to the discussion of health beliefs and negotiation of a shared model of care." [2]

Ideally these principles, carefully applied, could be especially useful to the field of Computer Supported Collaborative Work (CSCW) because of the naturally interdisciplinary and cross-cultural community found in CSCW, as well as that which is often investigated by CSCW researchers. There is a power-dynamic at work in the harvesting of big data and a culturally competent ethical stance requires a recognition of how researchers become 'cultural brokers' and gatekeepers for certain communities [3]. However, cultural competency is not a proposal for cultural relativism and it goes beyond notions of cultural sensitivity [2]. Culturally competent principles encourage researchers to triangulate on more than just data points. It also demands that we take into consideration cultural influences of the discipline, the researcher, the platform, as well as the participants.

These principles discourage data abuse by minimizing the possibility of giving an inaccurate representation of the data and populations.

Culturally Competent Research Design: an example

Thus far we have established where a culturally competent ethical stance comes from, a general definition, and why it might be useful to apply it as a framework for research using big data. In application a research design guided by culturally competent ethics has wide appeal for publications and policy papers due in large part to how it encourages the researcher(s) to consider a wider variety of standpoints than other more deterministic or relativistic designs. Now what could it look like in practice?

We are currently addressing a research question, using a large Twitter dataset, wherein we are asking *how do interactants interpret certain social signals and cues*. We are attempting to understand how people "read" social signals and cues because users who do not share a culture, and may have less access to established information networks such as mutual friends / shared locations / shared languages, frequently misinterpret cultural signals and cues which can lead to grave consequences. "The starting point is the observation that social signals – aimed at (a) indicating a certain hidden quality (e.g. "would you be a good friend?," "are you smart?") and (b) changing the beliefs or actions of its recipient – are playing a vital role in defining social relations and structuring societies" [4]. Strategies for addressing this question, as suggested by a culturally competent stance, lead us to work in native languages (rather than using translation tools), seek out culturally

appropriate collaborators, and use a mixed methods approach to data collection and analysis.

People make unwitting errors when doing research in languages and cultures not their own. Especially if they have preconceived notions built on political or social information. We are focused on how religiosity, as a form of social signaling, can be misread by casual observers, online translation tools, and even trained academics. Our research motivation is captured in this anecdote shared by a Muslim, male participant who described the following scenario he experienced while attending a prestigious American University:

"I had a professor who was researching terrorist organizations use of social media. He made a node map, which he displayed to the class. It was based on keywords from his analysis and it showed keywords as nodes. His map showed that people writing in Arabic in social media had the central node as the word Allah. Although the professor did not discuss any conclusion based on that map but the picture was surely out of context. Well, you know, anybody who knows something about the culture of many people who speak and write Arabic is that whether they are Muslim or not they use the word Allah a lot. In a lot of different ways. Just like Americans use the word Jesus. I don't think this professor's analysis was very good because he was drawing a conclusion from just showing how keywords linked to each other. Just using the word Allah does not make you Muslim. It also does not mean you are super religious."

Ultimately we intend use our study data to build a tool which will assist in cultural interpretation online. One which goes far beyond the faulty translation tools we have available which cause so much cross-cultural confusion. "One goal for an ideal society is thus not to eliminate social categorization – it's essential for helping us make sense of the world – but to make the prototypes and their triggering cues more closely aligned with actual behavior" [5]. Our pilot study forms the groundwork for identifying how we will be able to begin privileging *actual* intention rather than misinterpretation.

Discussion

To represent the *lived experiences* of the many populations in a 'big data' dataset with accountability and respect is of critical importance in order to advance ethical reliability and accuracy in CSCW research. Currently, Institutional Review Board policies do not meet the needs of big social data researchers. Adopting a culturally competent framework within our own research community demonstrates an ethical commitment to the people whose online social behavior has given us such a rich stream of information. Examining in what ways the ethics of cultural competence are in agreement or disagreement with the current ethics of CSCW and big data research is a discussion our research community should be having. We have the opportunity to set the policies and practices for our institutions in a way which is relevant, supportive and responsible.

Acknowledgements

Thank you to Prof. David McDonald, chair and associate professor of HCDE at the University of Washington for his continuing and gracious guidance.

References

- [1] Hand, M., & Hillyard, S. (Eds.). (2014). *Big Data?: Qualitative Approaches to Digital Research* (Vol. 13). Emerald Group Publishing.
- [2] Paasche-Orlow, M. (2004). The ethics of cultural competence. *Academic Medicine*, 79(4), 347-350.

[3] Shiu-Thornton, S., Balabis, J., Senturia, K., Tamayo, A., & Oberle, M. (2007). Disaster preparedness for limited English proficient communities: medical interpreters as cultural brokers and gatekeepers. *Public Health Reports*, 122(4), 466.

[4] Gasser, Urs, Avishag, *Social Signaling Theory and Cyberspace*. 2007.
<http://blogs.law.harvard.edu/ugasser/2007/02/24/social-signaling-theory-and-cyberspace/#respond>

[5] Donath, Judith,,. *The Social Machine : Designs for Living Online* 2014. (270 – 275)