

Thoughts on Mirco Musolesi's *Big Mobile Data Mining: Good or Evil?*

Eric Davis

2016

Introduction

Recent advancements in technology have facilitated the rapid expansion of large-scale data mining. Breakthroughs in mobile computing have placed a variety of sophisticated sensors in millions of pockets with an unrivaled ability to collect personal information. A smartphone's camera, accelerometer, microphone, Bluetooth, and GPS signal can all be used to harvest its user's location and personal information (Shilton, 2009, pp. 48-9). Further innovations in machine learning have amplified the value of this data by offering the ability to produce increasingly useful human-behavioral predictions using new technologies in anticipatory computing (Musolesi, 2014, pp. 78-9). With many companies and private organizations realizing the inherent commercial value in accurate behavioral predictions, the field of mass data collection has grown rapidly. The speed of this growth has exceeded the speed of effective regulation and established ethical standards. To maintain a competitive edge, many companies have an intended secrecy for many of their data mining practices. Musolesi believes that this secrecy reinforces ignorance amongst end-users, and therefore promotes fear of all forms of data mining.

Mirco Musolesi's *Big Mobile Data Mining: Good or Evil?* summarizes perhaps the most important ethical discussion of the past decade within its seven-word title. Recently, an increased affordability and therefore, accessibility to mobile computing (namely, smartphones) has encouraged rapid advancements and interest in data mining technology. With data harvesting becoming both progressively pervasive and invasive, many new ethical concerns continue to be raised. While acknowledging the potential to be used with maliciously evil intentions, Musolesi believes that current data mining technologies can be used for "good" and should be viewed as a "triumph for computer science" (2014, p. 80). With a few caveats, Musolesi argues for the necessity of approaching data collection technologies with optimism. By insisting on the importance of education and research, ethical regulation, and transparency in data mining,

Musolesi creates a compelling call to action for widespread education concerning the current data collection systems that already exist – and their ultimate potentials for both good and evil.

Education

Within the global society known as the Internet, an incredible amount of personal data is made public through the use of numerous, now-routine devices. Many of these end-users are unaware of the content of the data being collected, and also, how this data can and will be used. Not entirely out of negligence, many people are ignorant of data mining techniques because an overall lack of regulation has created a shroud of secrecy surrounding current practices (van Wel & Royakkers, 2004, p. 138). Because many organizations desire confidentiality in the name of preserving competitive advantages and trade secrets, the burden of education has been placed on the end-user. Speaking about this general lack of education, ethical theorists L. van Wel and L. Royakkers raise an important philosophical question:

Most people who use the web are not aware of the ways in which their web data can be analysed. Is it fair to say that people choose to give up their privacy when they are not fully aware of the consequences of their actions? (p. 135)

Backing Musolesi's call for increased education, van Wel and Royakkers voice additional ethical concerns supporting Musolesi's worries for the eventual dangers found in ignorance. With the importance of quality education becoming undeniably clear, all of these authors stress the necessity for these new technologies to become "standard body of knowledge for 21st century citizens" (Musolesi, 2014, p. 80).

Ethics

Ethical questions regarding large-scale personal data harvesting practices are evolving nearly as fast as the technology itself. Unfortunately, the level of transparency and disclosure needed to simply begin answering these ethical questions is lagging far behind. As observed by ethical theorists L. van Wel and L. Royakkers, "most web-data mining applications are currently found in the private sector" (2004, p. 130). Immediately, this raises major concerns regarding the

concealed nature of this data's use and storage after collection. If an organization is collecting sensitive personal data (even with user consent), what security mechanisms are in place to prevent theft? Recently, an abundance of headlines around the world suggests that data theft is becoming increasingly profitable. Once a user acquires the education necessary to make the informed decision to share personal data, what guarantees for the safety of this information can be given? Furthermore, as observed by van Wel and Royakkers, "privacy declarations promising not to sell or give information to third parties are often not clear about who those third parties are, and, more importantly, who they are not" (p. 134). How will future changes in these policies impact data that was collected under previous (and perhaps entirely different) pretenses? More importantly, what existing regulations can protect consumers from a company simply changing a policy in an effort to sell data collected under a previous policy? A general lack of disclosure from almost every data mining system has created a seemingly endless number of ethical dilemmas. In agreement with Musolesi, van Wel, and Royakkers, a dramatically increased level of transparency is required to solve many of these current ethical issues.

Transparency

The current level of disclosure from organizations participating in big data collection is often intentionally clouded to preserve trade secrets and competitive advantages. Algorithms can be worth millions of dollars and therefore, companies are rarely given incentives to share proprietary techniques and information. Technology ethics researcher, K. Shilton, believes that developers need to include "data-management tools to put personal data control back in the hands of individuals... and cryptic, fine-print EULAs (end-user license agreements) aren't the answer" (2009, pp. 51-3). Agreeing with these sentiments, van Wel and Royakkers also lend their voices to this cry for transparency: "privacy policies are often difficult to understand, hard to find, take a long time to read, and can be changed without notice" (2004, p. 134). These concerns only become increasingly complicated because, as observed by van Wel and Royakkers earlier, "most people who use the web are not aware of the ways in which their web data can be analysed" (p. 135). Most data mining systems in their current form lack the full disclosure and transparency necessary for consumers to make truly informed decisions on what personal information they share. In full agreement with (and perhaps a little more passionate about)

Musolesi's call for transparency, Shilton, van Wel, Royakkers, and myself find the lack of full disclosure as the driving force behind most ethical concerns.

Conclusion

Mirco Musolesi offers an expertly constructed argument that is approachable to both supporters and skeptics alike. Carefully considering the ability to use personal information for evil, Musolesi believes that ultimately, the potential for good outweighs that of evil. By acknowledging the need to increase current levels of education, address ethical concerns, and establish transparency, Musolesi's argument becomes easy to support. With the backing of many additional contemporaries and scholars, Musolesi's argument is widely supported. Entering this discussion with numerous fears for my own personal information, Musolesi's argument has convinced me that there is a potential for socioeconomic good found in mass data collection. Stubbornly holding on to my key concerns for privacy and transparency, I completely support Musolesi's call for the thorough (and transparent) exploration of all aspects related to big data mining.

References

- Musolesi, M. (2014). Big Mobile Data Mining: Good or Evil? *IEEE Internet Computing*, 18, 78-81. doi: 10.1109/MIC.2014.2
- Shilton, K. (2009). Four billion little brothers?: Privacy, mobile phones, and ubiquitous data collection. *Communications of the ACM*, 52(11), 48-53. doi: 10.1145/1592761.1592778
- Van Wel, L., & Royakkers, L. (2004). Ethical issues in web data mining. *Ethics and Information Technology*, 6(2), 129-140. doi: 10.1023/B:ETIN.0000047476.05912.3d