

Author: Valerie Racine Creation date: February 6, 2017

URL: www.onlineethics.org/40522.aspx

Note: The author wishes to acknowledge the contributions of Karin Ellison, OEC - Life and Environmental Sciences Editor, and Joseph Herkert, OEC Engineering co-Editor. They provided valuable input in selecting topics and crafting the resources.

Big Data & Public Health¹

Zhang Kar-wai, a graduate from Stanford University's department of computer science, has recently accepted his first paid position at a start-up company in San Francisco. The company's main products consist of a line of **smartphone applications designed to track individuals' personal health information**, including their medical records, their seasonal illnesses, their blood pressure and blood glucose levels, their eating habits, their sleep cycles, and even their weight and reproductive health. The products are designed primarily to help individuals reach health related goals, and enable users to manage their overall health and wellbeing.

Zhang was hired to join the team responsible for managing the large data sets generated by the users of these apps. One of his responsibilities will be to **develop algorithms and analytic tools that can track the outbreak and spread of infectious diseases in real-time** using data gathered from individuals using their applications. Their goal is to improve on the traditional methods used by the US's Center for Disease Control and Prevention (CDC) and the UN's World Health Organization (WHO).

Zhang loved developing algorithms as a student and he is looking forward to participating in the team's project. He was excited to tell his father, a professor of epidemiology at Stanford's School of Medicine, about his assignment. However, when Zhang told his father about his responsibilities at his new job, his father's reaction was not what he expected. Zhang's father expressed concern about the project's use of personal data, as well as their ability to design algorithms that could accurately predict and track outbreaks based on their data.

Zhang explained to his father that the data will be aggregated and the algorithms will undoubtedly sometimes fail, as all models do, but they will be continuously tested and upgraded. So, he assured his father that he was up to the task. His father remained hesitant to share in Zhang's enthusiasm and warned him to think more about the ethical implications of his project, and not just about whether his algorithms will succeed or not. Should Zhang take his father's concerns more seriously?

Answer the following three discussion questions. Use complete sentences. All answers should be 1-2 paragraphs and the entire assignment should be a maximum of 2 pages. Upload a separate word file with your responses to the questions.

1. Should users of social media, online applications and/or search engines be notified about the potential use of their personal data (even in aggregated form) for public health measures? Why/why not?
2. What are some of the ethical risks of the proposed big data analytics or algorithms discussed in this case study, specifically focused on outbreak tracking? What might be done to mitigate the risks?

3. What are some of the opportunities and ethical risks of for-profit companies collecting, managing and/or analyzing the health data of large populations? Describe any potential actions or policies that could be taken that would allow the public health community to maximize the opportunities presented by these new technologies while minimizing some of the risks in order to advance public health.

¹This material is based upon work supported by the National Science Foundation under Award No. 1355547, Karin Ellison and Joseph Herkert, Arizona State University sub-award Co-PIs. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.