Request Summary

Requesters: Huichuan Xia, Syracuse University

Yang Wang, Yun Huang, Syracuse University

Anuj Shah, , Carnegie Mellon University

Title: Privacy Violations in Crowd Work

Abstract:

Crowd work is a major aspect of the rising gig economy. Crowd work platforms such as Amazon Mechanical Turk (MTurk) and CrowdFlower have millions of ordinary people (i.e., crowd workers) around the world performing tasks (e.g., answering a survey, testing a website) to get paid. These platforms are widely used by companies, academic researchers, and other individuals to provide tasks for the crowd workers. While the literature has raised ethical issues in crowd work, little is known about **people's *actual* experiences of privacy challenges and violations in crowd work**.

Using MTurk, the most popular crowd work platform, as a concrete example, we conducted a survey of crowd workers' privacy experiences with 435 MTurk workers from around the world. Our respondents reported their *actual* experiences with a wide range of **privacy violations, such as sensitive information collection, manipulative data aggregation and profiling, unauthorized secondary use and sharing, as well as deceptive practices such as phishing and scam.** To our knowledge, this is the first empirical study that reports actual privacy violations that people experienced in crowd work. We published these results in this year's *ACM Conference on Computer-Supported Cooperative Work and Social Computing (CSCW)*.

Our follow-up work has analyzed actual tasks on MTurk and found that these tasks can **violate crowd workers' privacy but also other people's privacy**. For instance, some tasks asked crowd workers to dig out personal information of other individuals.

***Implications for policy*:** The privacy policies of crowd work platforms such as MTurk are vague from a crowd worker's perspective. We recommend that every crowd work task should be required to have its own privacy policy in which it clearly describes who the task requester is, what a crowd worker needs to do, what data will be collected, shared and used for what purpose, etc. The tasks descriptions on MTurk provide no or vague information about these important points, preventing crowd workers from making informed decisions about whether to perform certain tasks. The crowd work platforms should also enforce measures (e.g., suspend the requester) when the task privacy policy is violated.

***Implications for privacy design*:** First and foremost, there is little or inadequate privacy protection in crowd work platforms. These platforms should do a better job at screening out malicious tasks if they are already doing some screening. These platforms should also warn crowd workers about tasks that might violate their privacy. We believe the platforms can build tools to automatically mark problematic tasks based on user reports/complaints as well as natural language processing and machine learning techniques. Similar tools can be built to inform and remind benign task requesters when they unknowingly or unintentionally design tasks that might violate people's privacy.

Publication:

Xia, H., Wang, Y., Huang, Y., Shah, A. (2017), "Our Privacy Needs to be Protected at All Costs: Crowd Workers' Privacy Experiences on Mechanical Turk," *Proceedings of the ACM: Human-Computer Interaction (PACMHCI): Volume 1: Issue 2: Computer-Supported Cooperative Work and Social Computing (CSCW).*