How Does Bryan’s Online Behavior Impact His Privacy Risk Exposure?

TREO Talk Paper

Janine L. Spears
Cleveland State University
j.l.spears@csuohio.edu

Abstract

Behavioral privacy research has primarily focused on consumers with privacy concern. However, a common refrain in public discourse on online tracking of an Internet user’s browsing activity is “I’ve got nothing to hide,” and thus “have nothing to fear” from such tracking. One prospective explanation for the “nothing to hide” (NTH) view is these Internet users may lack awareness of how online tracking may affect them. In general, more research is needed on defining and illustrating online privacy risk for Internet users so that they can make more informed decisions on privacy risk management. Therefore, the present study conducts an experiment to illustrate online tracking and its potential effects. Secondly, the experiment examines the effectiveness of privacy enhancing technology (PET) to reduce online tracking and its effects.

As a means to provide scope and context, data collection of online browsing is facilitated using a persona; i.e., a fictitious representation of a realistic person. A persona typically captures demographics, technology usage, personality traits, and is typically given a name and often a picture. Personas are commonly used in HCI work as part of system design or user experience analysis for a given system. More recently, personas have been used to facilitate privacy risk and security threat analyses.

The fictitious NTH persona defined for the present research includes a name (e.g., “Bryan”), photo, age, marital status, career, hobbies, technology usage, and “something” to hide. Given Bryan’s NTH view on online tracking, he does not use any PET while browsing online. Based on Bryan’s characteristics, an online browsing script is constructed to simulate the types of web sites that a person like Bryan might visit and the searches he might make. Software is used to collect third-party tracking domain names affiliated with each web page he visits. In addition, the full (i.e., descriptive) URLs are collected per page visited.

Given the third-party trackers and full URLs per web page associated with Bryan’s online browsing, analysis is performed on how much information tracking companies can glean about Bryan. In particular, an inventory of full URLs will be provided to study participants who are not part of the research team. These participants will attempt to infer as much about Bryan’s demographics, career, hobbies, and his “something to hide” as possible. Secondly, the participants will be asked to make various inferences about Bryan based on his online browsing behavior that may have financial or other consequences.

A second set of browsing data will be captured using a PET (e.g., a browser plug-in that blocks tracking). A similar list of online trackers and full URLs per site visited will be captured, followed by similar analysis of inferences made on Bryan. It is expected that Bryan will encounter significantly less trackers, and thus, fewer characteristics may be inferred about him when using a PET.

Output from analyzing Bryan’s online tracking, inferences that may be made, and the degree of PET effectiveness in blocking tracking/inferences will serve as input into vignettes that illustrate privacy risk and the impact of managing such risk. These vignettes will subsequently be used in a survey instrument to test behavioral theory on privacy risk management.