Facebook History Collector:
A New Method for Directly Collecting Data from Facebook

Rosanna E. Guadagno, National Science Foundation, Arlington VA, USA

Tonio A. Loewald, Loewald New Media, Arlington VA, USA

Nicole L. Muscanell, The University of Alabama, Tuscaloosa, AL, USA

Joan M. Barth, The University of Alabama, Tuscaloosa, AL, USA

Melissa K. Goodwin, The University of Alabama, Tuscaloosa, AL, USA

Yang Yang, The University of Alabama, Tuscaloosa, AL, USA

ABSTRACT

Social networking sites are a rich source of untapped data. While much research has focused on Facebook and other social networks, less has done so by collecting information straight from the source. The purpose of this paper is to present a new and innovative means of gathering raw data from Facebook via a software program the authors call the Facebook History Collector. Initially developed to study Facebook activity in the wake of a destructive tornado that occurred in Tuscaloosa, Alabama on April 27, 2011, this tool allows for retrieving massive amounts of text, photographs, and videos directly from Facebook during a period of time set by the researcher. In this article, the authors argue that the new method has broad implications for use within the research community as it allows capturing real-time social media interactions. The process of software development, data collection, technical details, and ethical considerations are discussed.

Keywords: Data Collection, Facebook, Mediated Communication, Online Social Networks, Research Ethics, Research Methods, Social Media, Social Networking Sites

INTRODUCTION

Social networking sites such as Facebook, MySpace, Twitter, Pinterest, Instagram, and YouTube offer a rich source of data relatable to many of the major research topics in the social sciences. With a few clicks, one can observe user behavior in online social interactions and relationships with others, their personal interests, opinions, and beliefs on a wide variety of subjects, and other activities they engage in via status updates, sharing photographs, and

other posts. A typical Facebook user creates, completes, and publicly displays an extensive profile, referred to as a "Facebook page." Personal pages contain information varying from users' favorite pets, to movies and hobbies, to political affiliation. Facebook thrives because it allows unlimited opportunities for user self-expression and self-disclosure, while making the information available to a handful of friends or to the entire world, depending on individuals' preferences. Facebook is not only the most popular social networking site; it is the most frequently visited one surpassing Google, the world's largest and most powerful search engine (Nuthall & Gelles, 2010). With 1.11 billion monthly active users across the globe,[1] Facebook continues to be the number one choice when it comes to online social networking services. Why do people use Facebook? What about it is so appealing? And what can researchers learn about its use?

While social science research on Facebook and other social networking sites (SNS) is still burgeoning (Wilson, Gosling, & Graham, 2012), most of it relies on self-report, which is susceptible to socially desirable responses and recall errors. This may be due to challenges related to ethical issues and the dynamic nature of SNS themselves, including frequent security and privacy updates. How can researchers study Facebook activity systematically without creating their own SNS, using data crawlers, or employing ethically questionable approaches such as accessing people's personal profiles without their consent? This article presents the *Facebook History Collector*, a new software tool that allows researchers to gather and analyze data on naturally occurring Facebook social interaction. Initially developed to study Facebook activity in the midst of a destructive tornado in Tuscaloosa, Alabama on April 27, 2011, this method allowed for retrieving massive amounts of text, photographs, videos, and other data directly from Facebook. The authors reflect on the process of software development, data collection, technical details, challenges, ethical considerations, and directions for the future.

CURRENT RESEARCH METHODS USED TO COLLECT DATA FROM SNS

A comprehensive review by Wilson, Gosling, and Graham (2012) examined the social science literature on Facebook and other online social networks. Overall, the authors concluded that researchers have studied Facebook from several key perspectives: descriptive analyses of Facebook users, motivations for Facebook use, identity presentation, social interaction on Facebook, and privacy-information disclosure. Existing research has utilized different ways to explore research questions relevant to people's use of Facebook. These include (a) surveys and self-reports, (b) user ratings of mock-up or existing Facebook profiles, and (c) analyses of content retrieved directly from social media platforms. All of these approaches have provided useful data collection techniques. For example, by applying the self-report method, researchers have contributed to a burgeoning body of literature that facilitates our understanding of how and why people use social networking. Successfully employed in experimental studies, the mock-up profile is another viable alternative. Yet, the ability of these methods to capture real-time social media interactions is limited.

Researchers have used the self-report technique in a large number of studies on SNS and self-presentation (Krämer & Winter, 2008; Manago, Graham, Greenfield, & Salimkhan, 2008; Tong, Van Der Heide, Langwell, & Walther, 2008; Wilson, 2008). Using this approach, Krämer and Winter (2008) found that, compared to face-to-face interactions, SNS afford users a greater amount of control in how they appear to others because the user has the capability to regulate what specific information (demographics, photos, hobbies, etc.) is presented to viewers. In addition, researchers have attempted to simply describe social networking site users and their motivations through self-report surveys (Boyd, 2007; Hargittai, 2007; Peluchette & Karl, 2008; Raacke & Bonds-Raacke, 2008). Such research has examined general characteristics of SNS

9 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage: www.igi-global.com/article/facebook-history-collector/84815

Related Content

Seriously Social: Young Adults, Social Media and News
Kelly Kaufhold (2014). *International Journal of Interactive Communication Systems and Technologies (pp. 1-13).*
www.irma-international.org/article/seriously-social/134408

The Future of the Printed Book
Lee San Bao Elizabeth, Noriahni Ismailand Ma Saw Zarchi Tun (2012).
Understanding the Interactive Digital Media Marketplace: Frameworks, Platforms, Communities and Issues (pp. 416-429).
www.irma-international.org/chapter/future-printed-book/60487

Latent Semantic Similarity in Initial Computer-Mediated Interactions: Moderating Effects of Time, Extraversion, and Gender Composition
Vivian P. Taand William Ickes (2020). *International Journal of Interactive Communication Systems and Technologies (pp. 51-71).*
www.irma-international.org/article/latent-semantic-similarity-in-initial-computer-mediated-interactions/250265

Abstraction and Control of Transport Networks
Young Leeand Daniele Ceccarelli (2015). *Handbook of Research on Redesigning the Future of Internet Architectures (pp. 346-363).*
www.irma-international.org/chapter/abstraction-and-control-of-transport-networks/131373

Using Conceptual Models to Implement Natural Language Pedagogic Agent-Student Conversations
Diana Pérez-Marínand Carlos Caballero (2013). *International Journal of Interactive Communication Systems and Technologies (pp. 29-47).*
www.irma-international.org/article/using-conceptual-models-to-implement-natural-language-pedagogic-agent-student-conversations/105655