Available online at www.sciencedirect.com

ScienceDirect

Computers and Composition 35 (2015) 125–140

Computers and Composition

www.elsevier.com/locate/compcom

The Invisible Digital Identity: Assemblages in Digital Networks☆

Estee N. Beck[1]

Bowling Green State University

Abstract

Using tracking cookies and web beacons, online behavioral advertising uses code stored on machines to access users' Internet habits to customize advertisements and better market goods to consumers. This trend of tracking user movements has become concerning because the technologies used reveal personal information about the user to companies. Users can become more informed about the tracking technologies by visiting two websites that provide information about the trackers and give ways to opt-out of tracking technologies. This article provides a historical overview of tracking technologies, analyzes AboutAds.info and the online privacy tool Ghostery <ghostery.com>, theorizes what networked culture means in the 21st century, and closes with a heuristic for educators to use in their classrooms for discussions about invisible digital identity on the web.

Keywords: Digital identity; Surveillance; Cookies; Web beacons; Online behavioral advertising; Object-oriented rhetoric

Strategically hidden inside our computers are files that track our movements on the web. Inside the files are long strings of alphanumeric codes that do not reveal, on the surface, the kinds of personal information they contain. Concealed inside the code, such personal data includes housing type, age, sex, income, spending habits, hobbies and interests, items bought, items you're interested in buying, if you're traveling soon, and other data that may be fairly revealing. We live in an age of invisible digital identities where companies track our demographic information, habits, and online behaviors, and in some cases, sell this information to third-party companies for profit.

Interested in what my invisible digital identity consisted of and what tracking technologies had gathered about my online identity, I turned to BlueKai, a "big data[2]" activation company that brands data for marketing purposes.

☆ I would like to thank Kristine Blair for reading through earlier versions of this manuscript and providing encouragement, feedback, and insight. Additionally, I thank the blind reviewers for their gracious feedback and direction. Lastly, I thank Matthew Bridgewater, Mariana Grohowski, and Megan Adams for reading through a much earlier version of this manuscript and providing helpful directions.

E-mail address: enbeck@bgsu.edu

[1] Estee N. Beck is a PhD Candidate in the Rhetoric & Writing program, Department of English, at Bowling Green State University. Her current research positions computer algorithms as persuasive agents capable of shaping and influencing writers in digital spaces. She regularly tweets from @estee_beck.

[2] It is important to note here that there are several definitions of *big data* that are highly dependent upon the field or organizations' purposes. According to Thomas H. Davenport, Paul Barth, and Randy Bean (2012), data scientists can capture big data from an array of sources, including, "clickstream data from the web, social media content (tweets, blogs, Facebook wall postings, etc.) and video data from retail and other settings and from video entertainment. . . [it] also encompasses everything from call center voice data to genomic and protemic data from biological research and medicine" (p. 43). Big data is a conglomeration of virtual human behavior that involves some form of digital technology that computers and people have the ability to track and analyze. Big data is also pay dirt for organizations who seek advantages over competitors to capture consumer interest and money, especially through custom advertisements based on demographics and online behaviors.

http://dx.doi.org/10.1016/j.compcom.2015.01.005

Within their website, an option reveals consumers' invisible digital identities by threading files from various tracking technologies and illustrates how companies brand customers for customized web experiences.

The BlueKai Registry provided me with telling data about my surfing habits. It correctly identified the type of housing I live in (an apartment), my age range, sex, income level, spending habits, that I contribute to charities, that I will travel in greater than 21 days, the vehicle price-range of my car, and that I may want to buy concert tickets in the near future. Granted, the data provided on the site masked more identifiable characteristics like my name or housing address; however, the data diagrammed enough about my surfing habits to place me in certain behavioral categories, which allowed for the type of segmentation marketing companies create to motivate me to click an advertisement. All of this information came from the various tracking technologies embedded within the Firefox browser I used.

As digital rhetoricians and digital writing instructors, we have been concerned with digital identity representation, and rightly so, because we've focused our efforts on the types of data elements we can control, like developing virtual professional identities and showing our students how to cultivate a healthy online presence. We have, in turn, come to see our *visible* digital identity as something we can somewhat regulate online. On the other hand, with each click of a web page, we also have an *invisible* digital identity constructed through third-party elements and tracking technologies. I find this concerning, as with each click, fragments of our interests, habits, and demographics are stored, collected, and—in some cases—sold. In the Web 2.0 and Web 3.0 eras, our invisible digital identity—constructed through our consent of using web pages—is for sale. This trend of collecting and distributing fragments of our digital identity has led me to think about the ways digital writing teachers ask students to get online, click around, and use certain websites for course projects. Additionally, I wonder how we might better inform our students and each other about the types of online writing environments they and we engage with on a daily basis. Hence, my goal here is to share information about how our *invisible* digital identities develop through third-party elements and sites that track these elements to foster awareness among digital teachers and researchers. It is my hope that through this material, educators can turn to teaching their students about the hidden files stored on their personal machines, how companies use their personal data for online behavior advertising, and what actions students can take to limit the farming of computer files of their surfing habits. This article argues that if educators ask students to dig into digital spaces that use tracking technologies, then they also have some responsibility to teach students about invisible digital identities, how to become more informed about digital tracking, and how to possibly opt-out of behavioral marketing.

After all, discussions about visible digital identity and privacy are important to have with students. Scholarly essays and chapters have provided digital researchers and teachers with a rich understanding of identity online connected with social identity (Blackmon, 2003), developing digital identity with young women (Blair, Dietel-McLaughlin, & Graupner-Hurley, 2010), sexual orientation associated with gaming and literacy (Alexander with McCoy, & Velez, 2007), and the limitations of templates driven by social media sites (Arola, 2010), but there are other issues at stake. Talking about data collection and mining, digital and online surveillance, and various tracking technologies gives teachers points to consider not only when designing a course using the web, but also when asking students to participate in spaces that track their movements and collect user data. Accordingly, Jessica Reyman (2013) argued that we need to educate not only our students, but also our colleagues and ourselves about how data information operates online and how power differentials affect our digital spaces. Altogether, there are several issues at play with web use including *visible* digital identities, but instructors also might turn their attention to the *invisible* identities all web users have and what those in turn might mean for digital writing instruction.

Considering that more people are on the Internet and using mobile technologies, we need further research into surveillance and privacy along with the specific types of information tracking technologies collect about users. According to a Nielson (2012) report on social media use, there were 204 million people connected to the Internet in the U.S. in July 2012 along with over 95 million connected through mobile technologies; while the top social networking sites are Facebook and Twitter, sites like WordPress and Wikia lead the top list in terms of engagement. With this many people connected and with more wired classrooms, there is an obligation to discuss with our students not only the kinds of digital identity and privacy that are somewhat controllable, but also the hundreds of companies that use tracking technologies to capture data to better advertise products and how that shapes interactions on the web. In short, digital surveillance matters. As digital writing instructors, we need to understand how companies use this data, especially when, in the most egregious of cases, large data points reveal inaccurate information about people that in turn affect their legal, medical, and financial lives—which has been a case with American Express negatively adjusting credit