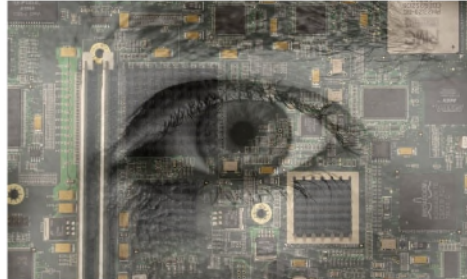


EX1051

A History of the Data-Tracked User

"Cyberspace" was once celebrated as a public, non-tracked space that afforded users freedom of anonymity. How did individual tracking of users come to dominate the web as a market practice?



By: Tanya Kant | Oct 8, 2021

Aa f x e d g

In February 2021, MIT's Schwarzman College of Computing launched a specially commissioned series that, in line with the mission of the school, aims to address the opportunities and challenges of the computing age. The MIT Case Studies in Social and Ethical Responsibilities of Computing (SERC) series features peer-reviewed, freely available cases by topic-area experts that introduce readers to a range of questions about computing, data, and society. Some cases focus closely on particular technologies, others on trends across technological platforms. Still others examine social, historical, philosophical, legal, and cultural facets that are essential for thinking critically about present-day efforts in computing and data sciences and their roles in the wider world.

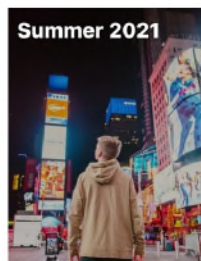
The following article, adapted from Tanya Kant's case study "Identity, Advertising, and Algorithmic Targeting: Or How (Not) to Target Your 'Ideal User,'" maps a brief history of the commercially targeted user, beginning with "identity scoring" in the 1940s and ending with the targeted advertising of today.

—The Editors

Who among us hasn't blindly accepted a cookie notice or an inscrutable privacy policy, or been stalked by a creepy "personalized" ad? Tracking and profiling are now commonplace fixtures of the digital everyday. This stands even if you use tracker blockers, which have been likened to "using an umbrella in a hurricane."

In most instances, data tracking is conducted in the name of offering "personalized experiences" to web users: individually targeted marketing, tailored newsfeeds, recommended products and content. To offer such experiences, platforms such as Facebook and Google use a dizzyingly extensive list of categories to track and profile people: gender, age, ethnicity, lifestyle and consumption preferences, language, voice recordings, facial recognition, location, political leanings, music and film taste, income, credit status, employment status, home ownership status, marital status — the list goes on.

A version of this timeline first appeared in the Summer 2021 issue of the MIT Case Studies series on Social and Ethical Responsibilities of Computing. Tanya Kant's Case Study "Identity, Advertising, and Algorithmic Targeting: Or How (Not) to Target Your 'Ideal User,'" from which the timeline is excerpted, is freely available [here](#).



Of course, there are benefits to being tracked: Cookies streamline site visits by storing user details, auto-filling technologies can quickly complete forms, and filtering systems manage otherwise unmanageable amounts of content, all while the data needed for such user advantages is harnessed to make platform profits.

However, data tracking is not primarily undertaken for the convenience of web users: Instead, the user data mined as part of targeting processes function as the driving

economic resource for the contemporary free-to-use web. This data-for-services market model is extremely successful: In 2020 the world's biggest data tracker, Facebook, made \$31.43 billion in targeted ad revenue in the United States alone. It is not an overstatement to propose that data tracking underpins the online economy as we know it.

Despite (or indeed because of) its monetizable qualities, tracking creates a host of stark ethical problems in relation to identity articulation, collective privacy, data bias, raced and gendered discrimination, and socioeconomic inequality. Beverley Skeggs has found that bids are made by advertisers for access to Facebook users' data on average 50 million times a day in ways that create a kind of data underclass who are then exploited by credit lenders.

Safiya Noble [has found](#) racist categorization practices inherent in Google Search's collaborative data sets. Cathy O'Neil [argues](#) that users' data are used to predict and manage users' future and present socioeconomic positionalities, often with detrimental consequences.

As I explore in [this case study](#), and [as part of my work on algorithmic identity](#), data tracking does not just match the "right" people with the "right" products and services — it can discriminate, govern, and regulate web users in ways that demand close attention to the social and ethical implications of targeting.

It is not an overstatement to propose that data tracking underpins the online economy as we know it.

Commercial platform providers frame data tracking as inevitable: Data in exchange for a (personalized) service is presented as the best, and often the only, option for platform users. Yet this has not always been the case: In the mid-to-late 1990s, when the web was still in its infancy, "cyberspace" was largely celebrated as public, non-tracked space which afforded users freedom of anonymity. How then did the individual tracking of users come to dominate the web as a market practice?

The following timeline outlines a partial history of the data-tracked user. It centers largely on developments that have affected European (and to a lesser extent U.S.) web users. This timeline includes developments in *commercial* targeting in the EU and U.S. rather than global developments in algorithmic policing, spatial infrastructures, medicine, and education, all of which are related but deserve their own timelines. This brief history fits into ongoing conversations around algorithmic targeting by reminding us that being tracked and targeted emerges from a historically specific set of developments. Increased legal scrutiny of targeting means that individual targeting as we know it may soon change dramatically — though while the assumption that profiling web users equates to more profit, it's more than likely that data tracking will persist in some form.

1940s. "Identity scoring" emerges: the categorization of individuals to calculate the benefits or risks of lending credit to certain groups of people.

1969. The U.S. Defense Advanced Research Projects Agency (DARPA) military initiative produces the first iteration of the internet. Its development is driven by "a multitude of actors" — military, state-based, academic and commercial — who are driven by different but divergent aims to make the internet public, commercial and defense-based.[\[1\]](#)

1970-80s. "Niche marketing" is developed in satellite TV and magazine industries, which segments audiences to sell more lucrative advertising slots.

1980s. The internet becomes increasingly privatized by internet service providers (ISPs) who use commercial enterprises to make the internet widely available to the public via pay-to-use services.

1991. The World Wide Web is created.[\[2\]](#) It is championed by its creator Tim Berners-Lee and others such as Howard Rheingold as an anonymous, noncommercial space that exists for the common good.[\[3\]](#)

1993-1997. Business actors urge the commercialization of some elements of the web. Bill Gates (1995) sees potential in online services to automatically suggest personally relevant content, and John Hagel and Arthur Armstrong (1997) envisage a web where user preferences can be monitored to create customized recommendations.[\[4\]](#)

1994. AT&T displays the first ever banner ad on HotWired (now Wired). The development of the "click-through" model allows advertisers to see when an (anonymously defined) user has clicked on their advertisement.

1994. The HTTP cookie is developed.[\[5\]](#) In its development stages, web users could fully restrict what data cookies could collect. However, data privacy measures were quickly removed, and users lost the power to control cookie data before the technology became widespread.[\[6\]](#)

Mid-1990s. Online advertising becomes more prevalent, but most companies struggle to successfully monetize web users' online media consumption.[\[7\]](#)

1996. Ad networks (platforms that serve as brokers between groups of publishers and groups of advertisers) increasingly emerge, including Doubleclick (now owned by Google).

1997. Developer / commentator Ethan Zuckerman hails algorithmic analyses of web pages as a new way form of monetization that works by demographically targeting website visitors for niche marketing purposes.[\[8\]](#) He argues, however, that this form of revenue generation is built on investor speculation rather than data accuracy.

1998. Open Profiling Standard (OPS) is bought and rolled out by Microsoft. OPS could securely store and manage individuals' personal information and credit card details, allowing user profiles to be exchanged between vendors.

2000-2003. Online revenue through advertising actively falls, as "banner ads" fail to compete with TV and print advertising.

2003. Ad networks such as AdSense begin to be adopted, allowing for automated matching of website content to advertising content. This allowed small web publishers to easily sell ad space.

2006. Popular ad-blocking software, Adblock Plus, launches.

Mid-2000s. Real-time bidding^[9] is developed, which allows advertisers to bid in real-time for ad space.

2008. Behavioral targeting begins to be integrated into real-time bidding, marking a crucial shift away from *media content* toward *user behavior* as key to targeting.

2010. Attorney Joseph H. Malley creates the term "zombie cookies" to describe HTTP cookies that are recreated after deletion. These are just one form of cookies designed to be "almost impossible to circumvent". Other variations include "supercookies," "ubercookies," and "evercookies."^[10]

2011. Facebook launches its third-party app system, allowing apps such as Spotify and Candy Crush to collect large amounts of personal data about Facebook users and users' friends.

2013. Edward Snowden reveals that commercial platforms such as Google and Facebook have been aiding the state "dataveillance" of millions of web users in the United States and the United Kingdom.

2013. Nick Nikiforakis *et al.* find that commercial data trackers have a range of "cookieless" methods for identifying and anticipating users, including Flash and canvas fingerprinting, which cannot be easily deleted or detected.^[11]

2014. The World Wide Web Foundation/ Sir Tim Berners-Lee launch The Web We Want campaign, which among other things calls for a return to Berners-Lee's original vision of a noncommercial and public web.

2014. In response to public outcries about data harvesting, Facebook imposes a ban on data collection by apps. Facebook promises to rollout "Anonymous Login" that will allow users to access apps without sharing any data.^[12] However, the ban is soon lifted, and Anonymous Login is never introduced.

2018. Analytics company Cambridge Analytica is found to be exploiting targeting systems that can "target voters with personalised political advertisements," sparking debates that personalized political advertising is unduly influencing elections.^[13]

2018. The European Union (EU) rolls out the General Data Protection Regulation (GDPR) legislation, intended to give EU citizens more security over personal data. However, legal loopholes such as "legitimate interest" allow for collection of otherwise GDPR-protected personal data.

2019. Following a series of lawsuits around discriminatory targeting, Facebook introduces a new nondiscrimination advertising policy. This includes a blanket targeting ban to advertisers in the housing, employment, and credit markets.

2020. The EU proposes a new Digital Services Act, designed among other things to restrict the use of personal data for targeted advertising purposes.

2020. Apple bans third-party cookies and Google pledges to do so by 2022, prompting debates on the so-called "cookie apocalypse." Though welcomed by privacy-concerned users, third-party marketing companies such as Criteo experience a fall in share values and argue that the erasure of third-party cookies gives even more power to monopolistic first-party data trackers.^[14]

2021. The UK government releases the Draft Online Safety bill, which includes legislation designed to restrict how and why individuals are targeted online.

2021. Despite moves to restrict forms of targeted marketing, reports emerge that Facebook is continuing to discriminate against both men and women in their targeted advertising systems,^[15] and that the UK government is using sensitive data to target web users and 'nudge' their behavior.^[16]

Tanya Kant is Lecturer in Media and Cultural Studies (Digital Media) at the University of Sussex, UK. She is Co-Managing Editor of the open access, multimedia publishing platform REFRAME, and author of "Making it Personal: Algorithmic Personalization, Identity, and Everyday Life" (Oxford University Press)

Bibliography

Abbate, Janet. "Privatizing the Internet: Competing Visions and Chaotic

Events, 1987–1995,” *IEEE Annals of the History of Computing* 32, no. 1 (2010): 10–22.

Barker, Alex. “Cookie apocalypse’ forces profound changes in online advertising.” *Financial Times*, February 25, 2020.

BBC. “Facebook to Show Who Buys Political Ads.” *BBC News*, October 28, 2017.

Browne, Simone. *Dark Matters*. Durham, NC: Duke University Press, 2015.

Business Insider. “Facebook Ad Revenue in 2020 Will Grow 4.9% Despite the Growing Number of Brands Pulling Campaigns.” June 23, 2020.

Cellan Jones, Rory. “Facebook accused of allowing sexist job advertising.” *BBC News*, Sept 9, 2021.

Curran, James, and John Seaton. *Power without Responsibility: Press, Broadcasting and the Internet in Britain: Press and Broadcasting in Britain*. London: Routledge, 2010.

Cyphers, Bennett. “Google Is Testing Its Controversial New Ad Targeting Tech in Millions of Browsers.” *Electronic Frontier Foundation*, March 30, 2021.

Egbert, Simon, and Matthias Leese. *Criminal Futures*. London: Routledge, 2020.

Eubanks, Virginia. *Automating Inequality*. New York: St. Martin’s Press, 2017.

Facebook Newsroom. “Introducing Anonymous Login and an Updated Facebook Login.” April 30, 2014.

Gates, Bill. *The Road Ahead*. New York: Viking, 1995.

Greenfield, Peter. “The Cambridge Analytica Files: The Story So Far.” *Guardian*, March 25, 2018.

Hagel, John, and Arthur Armstrong. *Net Gain: Expanding Markets through Virtual Communities*. Boston: Harvard Business School Press, 1997.

Hern, Alex. “Study finds growing government use of sensitive data to ‘nudge’ behaviour.” *Guardian*, Sept 8, 2021.

Kitchin, Rob, and Martin Dodge. *Code/Space: Software and Everyday Life*. Cambridge, MA: MIT Press, 2011.

Nikiforakis, Nick, Alexandros Kapravelos, Wouter Joosen, Christopher Kruegel, Frank Piessens, and Giovanni Vigna. “Cookieless Monster: Exploring the Ecosystem of Web-Based Device Fingerprinting.” *2013 IEEE Symposium on Security and Privacy* (2013): 541–55.

Noble, Safiya Umoja. *Algorithms of Oppression: How Search Engines Reinforce Racism*. New York: New York University Press, 2018.

Ofcom. “Adults’ Media Use & Attitudes report 2020.” Accessed May 17, 2021.

O’Neil, Cathy. *Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy*. London: Penguin, 2016.

Peacock, Sylvia. “How Web Tracking Changes User Agency in the Age of Big Data: The Used User.” *Big Data & Society* 2, no. 1 (2014).

Rheingold, Howard. *The Virtual Community: Homesteading on the Electronic Frontier*. Cambridge, MA: MIT Press, 1996.

Skeggs, Beverley. *You Are Being Tracked. Valued and Sold: An Analysis of Digital Inequalities, 2017*. Lecture, Filmed September 2017 at the London School of Economics, MP4, run time 1:26:42.

Wang, Jackie. *Carceral Capitalism*. Cambridge, MA: MIT Press, 2020.

Zuckerman, Ethan. “The Internet’s Original Sin.” *Atlantic*, August 14, 2014.

[1] Janet Abbate, “Privatizing the Internet: Competing Visions and Chaotic Events, 1987–1995,” *IEEE Annals of the History of Computing* 32 (2010): 10, 10–22.

[2] Invented by Tim Berners Lee, the World Wide Web is a hyperlinked information system that can be used to access the internet.

[3] Howard Rheingold, *The Virtual Community* (Cambridge, MA: MIT Press, 1996).

[4] Bill Gates, *The Road Ahead* (New York: Viking, 1995); John Hagel and Arthur Armstrong, *Net Gain* (Boston: Harvard Business School Press, 1997).

[5] The HTTP cookie is a form of information storage between user and server that exists on a user’s computer and can be later retrieved by the server. Cookie tracking works by storing this text file on a user’s computer

and sending it to either third- or first-party cookie trackers, who then use this data to attribute characteristics to the user in the form of demographic profiling and other profiling mechanisms.

[6] Sylvia Peacock, "How Web Tracking Changes User Agency in the Age of Big Data: The Used User," *Big Data & Society* 2, no. 1 (2014): 6

[7] James Curran and John Seaton, *Power Without Responsibility* (London: Routledge, 2010).

[8] Ethan Zuckerman, "The Internet's Original Sin," *Atlantic*, August 14, 2014.

[9] Real-time bidding is a process used to display personalized advertising on web pages across the internet. Real-time bidding works as an auction process, wherein advertisers bid for an "impression" (ad space) seen by a particular user on the website they are visiting. Bidding, as the name suggests, is in real time and is largely fought and won using a combination of user profiling and content review of the website hosting the advertisement.

[10] Peacock, "How Web Tracking Changes User Agency."

[11] Nick Nikiforakis et al., "Cookieless Monster: Exploring the Ecosystem of Web-Based Device Fingerprinting," *2013 IEEE Symposium on Security and Privacy* (2013): 541–55.

[12] Introducing Anonymous Login," *Facebook Newsroom*, April 30, 2014.

[13] Peter Greenfield, "The Cambridge Analytica Files: The Story So Far," *Guardian*, March 25, 2018.

[14] Alex Barker, "Cookie apocalypse' forces profound changes in online advertising," *Financial Times*, February 26, 2020.

[15] Rory Cellan Jones, Facebook accused of allowing sexist job advertising, *BBC News*, Sept 9, 2021.

[16] Alex Hern, "Study finds growing government use of sensitive data to 'nudge' behaviour," *Guardian*, Sept 8, 2021.

POSTED ON OCT 4, 2021

[Facebook](#) [X Twitter](#) [Reddit](#) [Pocket](#) [Flipboard](#)

The MIT Press is a mission-driven, not-for-profit scholarly publisher. Your support helps make it possible for us to create open publishing models and produce books of superior design quality.

[Donate](#)



60 Years Ago, Congress Warned Us About the Surveillance State. What Happened?

"We must see to it that this agency and all agencies that possess this technology operate within the law and under proper supervision so that we never cross over that abyss. That is the abyss from which there is no return."

Jennifer Holt | Sep 27



Privacy and Paternalism: The Ethics of Student Data Collection

Despite policies that aim to provide both transparency and access, most students and parents are unaware of what data is being stored and who has access to it.

Kathleen Creel & Tara Dixit | Sep 15, 2022



AI Is No Match for the Quirks of Human Intelligence

We may sometimes behave like computers, but more often, we are creative, irrational, and not always too bright.

Herbert L. Roitblat | Oct 4, 2021



Chatting With the Dead

In giving voice to the digital ghosts of the deceased, chatbots are trying to succeed where photography and dreams fail.

Davide Sisto | Jan 4, 2023



illuminating the bold ideas and voices that make up the MIT Press's expansive catalog. We publish thought-provoking excerpts, interviews, and original essays written for a general reader but backed by academic rigor.

[Home](#) [The MIT Press](#) [About](#) [Privacy Policy](#) [RSS](#) [Contact](#)