

# Baking Gender Into Social Media Design: How Platforms Shape Categories for Users and Advertisers

Social Media + Society  
October–December 2016: 1–12  
© The Author(s) 2016  
Reprints and permissions:  
sagepub.co.uk/journalsPermissions.nav  
DOI: 10.1177/2056305116672486  
sms.sagepub.com  


Rena Bivens<sup>1</sup> and Oliver L. Haimson<sup>2</sup>

## Abstract

In recent years, several popular social media platforms have launched freeform custom gender fields. This decision reconstitutes gender categories beyond an oppressive binary only permitting “males” and “females.” In this work, we uncover many different user-facing gender category design strategies within the social media ecosystem, ranging from custom gender options (on Facebook, Google+, and Pinterest) to the absence of gender fields entirely (on Twitter and LinkedIn). To explore how gender is baked into platform design, this article investigates the 10 most popular English-speaking social media platforms by performing recorded walkthroughs from two different subject positions: (1) a new user registering an account, and (2) a new advertiser creating an ad. We explore several different spaces in social media software where designers commonly program gender—sign-up pages, profile pages, and advertising portals—to consider (1) how gender is made durable through social media design, and (2) the shifting composition of the category of gender within the social media ecosystem more broadly. Through this investigation, we question how these categorizations attribute meaning to gender as they materialize in different software spaces, along with the recursive implications for society. Ultimately, our analysis reveals how social media platforms act as intermediaries within the larger ecosystem of advertising and web analytics companies. We argue that this intermediary role entrusts social media platforms with a considerable degree of control over the generation of broader categorization systems, which can be wielded to shape the perceived needs and desires of both users and advertising clients.

## Keywords

categorization, social media platform, gender, ad targeting, software programming, identity

## Introduction

As a category, gender materializes throughout society. We readily see gender emerge and become mobilized in cultural spaces such as popular culture, sports, politics, and retail, so it should come as no surprise that computer programmers also bring gender to life when they design software. Design decisions determine where—in the multiple layers of software—gender appears as a category, how it is materialized within code and activated within software processes, and for what purposes it is deployed. Far from neutral, these programming practices bake values and assumptions into technology. As Balsamo (2011) noted, “through the practices of designing, cultural beliefs are materially reproduced, identities are established, and social relations are codified” (p. 3). As a result, design decisions that render gender into a category recursively influence how we identify ourselves and others. As Cheney-Lippold (2011) argues, “we are entering an online world where our identifications are largely made for us”

(p. 165). By programming software to track and categorize users, designers—and software processes on their behalf—enact culture (Kitchin & Dodge, 2011) by inscribing meaning and setting limits on our perception of what is possible for viable social life (Butler, 2004). Alongside the exercise of power through category management, designers must also concern themselves with “the power to cause that categorization to have social and material consequences” (Crenshaw, 1991, p. 1297). The disproportionate rates of violence and discrimination faced by the transgender<sup>1</sup> community (National Coalition of Anti-Violence Programs [NCAVP], 2014) signal

<sup>1</sup>Carleton University, Canada

<sup>2</sup>University of California, Irvine, USA

### Corresponding Author:

Rena Bivens, School of Journalism and Communication, Carleton University, Room 4306, River Building, 1125 Colonel By Drive, Ottawa, Ontario K1S 5B6, Canada.  
Email: rena.bivens@carleton.ca



how the production of gender as a binary category obstructs viable social life for those in our non-binary world who do not fit static and narrow constructions of gender identity.

Since our social lives are now peppered with digitally mediated encounters, software components literally join us as we engage in everyday social practices, whether we are conscious of their presence or not (Beer, 2013; Kitchin & Dodge, 2011). For many, social media sites have become integral platforms for socializing. Indeed, social media companies capitalize on this trend; they are programmed spaces designed for digital identity construction (boyd, 2011). Yet, these platforms often position themselves as neutral or open online spaces designed specifically for users, when their primary goals actually involve generating, capturing, and controlling user data, affording platforms great power over users (Gehl, 2011; Gillespie, 2010; van Dijck, 2013). By emphasizing neutrality, platforms can more easily conceal important choices about what can appear on the site and how data are archived, owned, and monetized (Gehl, 2011; Gillespie, 2010).

With a focus on platform design, this article investigates the construction of gender as a category within the 10 most popular social media platforms. We explore several different spaces in social media software where designers commonly program gender—sign-up pages, profile pages, and advertising portals—to consider (1) how gender is made durable through social media design, and (2) the shifting composition of the category of gender within the social media ecosystem more broadly. Through this investigation, we question how these categorizations attribute meaning to gender as they materialize in different software spaces, along with the recursive implications for society. Ultimately, our analysis reveals how social media platforms act as intermediaries within the larger ecosystem of advertising and web analytics companies. We argue that this intermediary role entrusts social media platforms with considerable control over generating and managing categorization systems, which can be wielded to shape both users' and advertising clients' perceived needs and desires. This work aims to make visible some of the underlying motivations that propel platforms and programmers to bake gender into their design strategies and, following in Bowker and Star's (1999) footsteps, "challenge the silences" surrounding these classification mechanisms (p. 5).

## The Social Cost of Categorization

Sorting and classifying users bolsters digital monetization strategies. We have already established that categorization is a cultural practice that is not neutral, but it is also important to recognize the population control legacies that these systems carry forward, including their racist lineage. After noting human taxonomies dating back to the 18th century onwards, Chun (2013) points out that "'scientific' categorizations of race have been employed to establish hierarchical differences between people, rendering some into mere

objects to be exploited, enslaved, measured, demeaned, and sometimes destroyed" (p. 40). The advertising industry has also capitalized on categorization systems, embracing the capacity of market segmentation to "look for splits in the social fabric [of society] and then reinforce and extend the splits for their own ends" (Turow, 1997, p. 6). Deliberately segmenting society into categories and encouraging each group to consume media and products, with the goal of increasing advertising effectiveness by further separating each category from another (ideologically and otherwise), is a highly divisive strategy. Turow's insights underscore the industry's drive to not simply identify consumers but to actively mold and manufacture categories of consumers.

Historically, the desire to create and cater to demographic categories has been central to market segmentation efforts, beginning with crude measures of inferring class and race such as the use of postcode data. The shift to online marketing has opened doors to many other data types including coveted behavioral data that can be layered on top of demographic data (Battelle, 2005; Turow, 2006). Some of the data now collected were available before, such as purchasing habits. However, cookies that track activities such as browsing habits, clicks, and search strings increase behavioral consumer data's breadth, scope, and accessibility to marketers. There is now "a constant stream of real-time web use that can be matched against existing behavior and identity models—like gender" (Cheney-Lippold, 2011, p. 168). While the gender binary continues to be valued as a dominant audience segmentation device for advertisers (Altstiel & Grow, 2012; Wagner, 2014), offering two "large and profitable" groups (Wolin, 2003), industry experts argue that targeting strategies ought to couple demographic data with behavioral and taste-based data (Churchill, 2013; Smith, 2014).

Yet, Cheney-Lippold (2011) readjusts the critique, reminding us that advertisers' efforts to categorize us do damage beyond our pocketbooks: they reflexively shape how we see ourselves and others. The very definition of gender is filtered through a "marketing logic of consumption" (Cheney-Lippold, 2011, p. 167) and the meaning of that category is often algorithmically determined, operating as a modulating force by constantly shifting in tune with an invisible feedback loop. Web analytics firms in particular—the focus of Cheney-Lippold's (2011) analysis—place us under constant surveillance, re-calibrating categories in response to data collected from our online activities that make up our digital fingerprints. This feedback loop has the effect of perpetually conditioning us via the suggestions and recommendations that populate as we surf and interact online, imperceptibly nudging us toward conformity (Cheney-Lippold, 2011). In other words, instead of disciplining bodies based on societal standards, disciplinary control is based on advertisements' suggestions for how to behave that are perpetually re-calibrated whenever they are deemed unsuccessful or unprofitable (Cheney-Lippold, 2011).

The responsive design of this real-time gender categorization system surprisingly embodies the potential to correct some of the transphobic, binary-centric aspects of the society from which this technical system developed. Gender categories imposed by algorithms are more flexible and fluid than gender categories imposed by state or governmental forces; instead of requiring a doctor's letter to change from female to male, for instance, a classifier algorithm may make this change based on content it has classified as likely to be clicked on by "male" users. At the same time, algorithmically imposed gender categories have no grounding in biological sex, and thus may align more readily with non-essentialist views of gender (Butler, 1999). However, despite the fluidity and non-essentialism of gender categories permitted by data analytic approaches to categorization, the category of gender itself inevitably hardens into something that works for advertisers. Ultimately, the more flexible algorithmic model of gender identity ends up "re-essentializ[ing] gender as a statistically-related, largely market research-driven category" (Cheney-Lippold, 2011, p. 170).

Culture and technology's intimate entanglement (McPherson, 2014) engenders designs that have the capacity to reproduce *or* disrupt hegemonic discourses and social relations of power (Bardzell, 2010; Flanagan, Howe, & Nissenbaum, 2008; Friedman & Nissenbaum, 1997; Wajcman, 2010). To some extent, the advertising industry's consumption-based logic does both: it flattens gender categories while negating user agency, rendering gender self-identification increasingly irrelevant. For many whose self-definition as trans or as a woman is an important and meaningful part of their identity, this lack of agency in online self-presentation is troubling. There is a clear disconnect between platform owners' motivations for *offering* particular gender options and social media users' desires and motivations for *using* particular gender options. The former is almost completely about data collection, advertising, and revenue opportunities. The latter is about online self-presentation. The following analysis charts these tensions as we pursue the materialization of gender as a programmed category within the 10 most popular social media platforms.

## Methods

Google+, Facebook, YouTube, Yahoo, Twitter, LinkedIn, Blogspot, VK, Instagram, and Pinterest make up our sample of the 10 most popular English-speaking social media platforms. This list is culled from Alexa's Global Traffic Rank (Alexa, 2015) over a 30-day period ("monthly unique visitor metrics" from 13 June to 12 July 2015), and each platform meets Ellison and boyd's (2013) widely used definition of a social network site (see Table 1 in the Appendix).

We used the walkthrough method—an immersive, ethnographic approach that seeks thick description (Burgess, Light, & Duguay, 2015)—to record each platform from two different subject positions: (1) a new user registering an account,

and (2) a new advertiser creating an ad.<sup>2</sup> When we encountered a field while registering as a new user, we determined whether it was mandatory by leaving it blank and attempting to move forward in the process. We included Terms of Service and Privacy Policy documents in our analysis, along with other links related to data collection, use, and processing. As a prospective advertiser, we explored demographic data fields and help pages describing how ads can be targeted.

We supplemented our walkthroughs with background research about each platform. This included "About Us" pages, mission statements, advertising terms, each platform's Wikipedia page, and online searches to determine whether any gender field-related changes had been made throughout the years. For our background research, we were most interested in changes to gender fields on sign-up pages and profile pages that took place since the launch of the platforms, public outcry over gender-related data fields and collection, ownership changes, and platform-to-platform alliances. Along with platform names, the following search strings were used to gather this information: redesign, update, sign-up, profile, history, acquisition, gender, binary, advertising, and privacy.

## Sign-Up Your Data

Three characteristics of sign-up pages work in concert to make gender durable. Sign-up pages (1) function as mandatory gateways where non-negotiable data is collected; (2) are transitory spaces, designed to be quickly passed over by users; and, over time, (3) become increasingly immune to change. We begin this section focusing on the first two characteristics, and then explore the third.

All users must eventually visit a sign-up page if they wish to participate in a social media network. Once they do, sign-up pages block access to the site until all mandatory fields are completed. However, designers are also cognizant of the need to advance users past the gateway before they deem the process too onerous and leave. This impetus helps sign-up pages acquire a level of invisibility. They continue to exist within the software, but they are intended as transitory spaces that quickly disappear from view. At the same time, sign-up pages are highly familiar sociotechnical configurations that, similar to Terms and Conditions, users are inclined to hastily pass through without a great deal of reflection (Jensen & Potts, 2004). Yet it is in these moments—when we uncritically accept the categorization systems through which we are asked to identify ourselves—that we are more susceptible to recursive consequences of design decisions, such as the acceptance of the binary as normal and neutral. Gender's categorization as a binary calcifies when it materializes again and again through sociotechnical decision points like social media sign-up pages, which we hastily move through and allow ourselves to be classified by.

For many users, selecting between standard gender classifications (e.g., male, female) while registering for these

platforms is banal; however, for users with complex gender identities, underlying sociotechnical infrastructures can constrain choices. Ultimately, mandatory binary gender fields reify a gender categorization schema that erases everyone who does not fit the binary (Bivens, 2015). Nakamura (2002, p. 120) argued that we must be critical of the ways platforms mediate race and gender identities, and the ways that marginalized and intersectional identities are often “unclickable,” thus making them illegible. When a gender category materializes within sign-up pages as narrow, static, and, ultimately, incomplete, the users who have nowhere to click are faced with technical and even legal obstacles given that they can only bypass the gateway by misgendering themselves (Bivens, 2015). In this software-user configuration, answers given on sign-up pages are swiftly reconfigured into databases as far more permanent artifacts than implied by sign-up pages’ pretense as transitory pit stops. While a social media sign-up page is unlikely to be the first time that a trans or gender non-conforming person has encountered social practices that exclude their identity or reify the gender binary as the only viable option, clicking a box for one binary gender or the other can still be an emotionally harmful or stressful experience (e.g., see Haimson, Brubaker, Dombrowski, & Hayes [2015] for work on gender transition on social network sites).

In the midst of a trans rights movement that is only recently gaining substantial traction, attention, and representation in North America, Facebook’s 2014 decision to add 56<sup>3</sup> gender options beyond the binary was an important signal that advocacy efforts can challenge and affect change in standard, transphobic programming practices. Changes to gender categories on Google+ and Pinterest soon followed, incorporating open text fields for users to enter any label they wished. Yet, as Bivens (2015) details in her investigation of Facebook’s gender coding practices, these software modifications reconfigured gender as more than a binary within only *some* parts of the platform, such as the public-facing profile pages and news feed. The gender binary is maintained, and further entrenched, within less visible spaces (such as the database, where “custom” non-binary genders are reverted back to a binary system based on the pronoun that is selected) and less visited spaces (sign-up pages).

Overall, we found a range of design decisions in the sign-up pages of the most popular social media platforms. Most platforms included gender as a category—only three (of eight<sup>4</sup>) omitted it altogether (see Table 2 in the Appendix). Of those sign-up pages that did incorporate gender, all but one (Google+) conceptualized gender as a binary. When gender did appear, it was almost always programmed as a mandatory field (Pinterest<sup>5</sup> was the only exception). Public criticism followed Google+’s design decision to launch their sign-up page in 2011 with “other” as the only non-binary possibility (Truitt, 2011). This classification system leaves the binary intact—in a privileged, normalized

position—while relegating irrelevant anyone who does identify with a catchall “other” category.<sup>6</sup>

Continuing to bake the binary into sign-up pages, even when gender materializes as a more open category within other parts of the software, suggests that sign-up pages are not designed to serve users. From a programming perspective, sign-up pages are instrumental software spaces because they trigger the process of inserting a new database record for each user. Occurring “under the hood,” this is a vital software process given how frequently platform software retrieves data from databases in order to populate the graphical user interface with the content that users ultimately engage with. At the same time, social media platforms collect identity-based user data through the non-negotiable fields that they program into these spaces. It has been argued, for example, that Google+’s low rates of engagement are largely irrelevant given that Google’s broader business strategy hinges on collecting identity-based user data, which is already accomplished once users complete the sign-up process (Constine, 2012).

Despite the transitory nature of sign-up pages for users, the fields that the system requests or demands within this space play a role in making particular categories durable (see Table 3 in the Appendix). Since start-up ventures in the social media industry tend to be highly invested in the growth of their user-base, early design decisions can be difficult to modify once databases become very large. Facebook is an interesting example since the platform launched in 2004 without a gender field on the sign-up page. Other information, such as “student status,” was considered vital for the database record, but Facebook’s advertising model had yet to emerge. Once platform owners began to value gender as a category, the company’s design strategy turned interventionist, attempting to decrease the number of accounts with an undefined gender field stored in their database (Bivens, 2015). Years later, even after Facebook reconfigured their gender category within profile pages, broadening the definition to accept non-binary gender options, a mandatory, binary gender field remains on the sign-up page (as of this writing). Citing an interview with Facebook, Bivens (2015) argues that modifying the sign-up page to accommodate a non-binary constitution of the category is undermined by the desire to not “break the system” for advertisers, marketers, and developers. Maintaining connectivity with these actors requires a degree of stability within the system to ensure that the many programmed relationships (e.g., between data fields, lines of code, and functions) that materialize in different spaces within the software continue to operate. As a result, identity categories programmed into sign-up pages can become further entrenched because of the unique needs of these software spaces and the limited capacity to modify the meaning of these categories, given that modifications hinge on the relationships that the company has developed with third-party applications and



the meaning that those actors inscribe within their own categorization systems.

## Custom Genders Are Trending

Within the social media ecosystem, the emergence of gender as a category that captures more than a binary is recent. Beginning in 2011 with the launch of Google+ and the materialization of a third option (“other”) on their sign-up page, movement away from the binary stalled until Facebook’s 2014 profile page upgrade. It was at this point—when “custom” gender began to trend among social media designers—that we saw a more explicit break away from a categorization system that exclusively, or primarily, values the binary. Pronouns also emerged as an ancillary arm of custom gender, adding a further layer to the categorization process.

Each platform independently determines how to program gender. In our sample, 3 (of 10) platforms offer custom gender options within profile pages, with Facebook being the first to offer custom gender in February 2014. At that time, upon selecting “custom,” a text field appeared to facilitate a search through possible gender terms, which were limited to 56 predetermined options. To make a gender selection, Facebook’s design also requires users to select a preferred pronoun selection. The options available for mandatory pronoun selection (his, her, or them)<sup>7</sup> reinscribe gender as a three-option data classification and collection system, allowing Facebook to continue to store and process data about gender in effectively the same way they had since their launch in 2004 (when male, female, and undefined constituted gender on profile pages). As Bivens (2015) detailed in her investigation of Facebook, despite the emergence of custom gender on profile pages, the mandatory pronoun was (and remains at the time of this writing) most instrumental in determining how a user’s gender materializes in the database and, subsequently, how gender reappears again through Facebook’s application programming interface (API). In other words, instead of a user’s selected gender (e.g., “gender questioning”), the database reconstitutes a user’s gender according to their pronoun selection (e.g., “she”), which means that “female” appears when that user’s gender is retrieved from the database. Bivens’ (2015) investigation highlights the capacity for software to misgender<sup>8</sup> users under the surface, burying this act of symbolic violence deep in the database. This example demonstrates the complex ways in which the category of gender materializes in different software layers, reconstituted in each location with new limitations, meanings, and constituent parts.

Even though social media platforms act independently, they are not immune to social pressures nor to major changes instituted by other platforms they see as competitors. In December 2014, 10 months after Facebook unveiled custom gender, Google+ revealed its own redesigned gender options for profile pages. In this configuration, gender is categorized

with the help of a freeform text field that permits users to enter any text they wish, as long as they also indicate a preferred pronoun (configured again as a mandatory, three-option layer of gender that can be superimposed over custom gender). This design strategy continues to visually and ideologically separate the gender binary from all other gendered identities. Such programming practices enact a cultural valuation and naturalization of gender categories that idealizes the binary. Our investigation<sup>9</sup> of Google+ tracked the gender category as the system reconstitutes it from profile pages to software spaces designed for advertising clients. Whenever the pronoun “their” is selected, gender simply disappears as though the user has no gender at all. Yet “male” and “female” users materialize whenever a user selects a binary gender *or* the pronoun “his” or “her.” By attributing so much value to gender’s constitution as a pronoun, and collapsing the pronoun category back into a binary, this programming practice reinscribes the category of gender in a way that purports to function more effectively for the perceived needs of advertisers rather than users. In this space, the gender binary is made durable for advertisers, at the expense of some users who are given the illusion of agency in gender self-presentation yet are nonetheless invisibly misgendered.

Two months later, in February 2015, Facebook followed Google+’s lead, replacing the list of custom options with a freeform text field, while again retaining the mandatory pronoun category and reproducing the partition between the binary and all other gendered identities. By June 2015, the freeform text model also appeared on Pinterest’s profile pages. Yet this time, the ancillary pronoun arm of the gender category was not transposed—Pinterest does *not* include a mandatory pronoun categorization layer that collapses complex user gender identifications into binary genders for advertisers. While this is an interesting development, Pinterest’s strategic alliance with Facebook offers access to user data stored on Facebook, which in turn informs Pinterest’s advertising capabilities and bypasses Pinterest’s own gender categorization system, as we consider further in the next section.

Within the profile pages of the seven other social media platforms in our sample, two were programmed to permit a binary (Yahoo and VK), two more afforded the capacity for users to leave their gender unspecified (Blogspot and Instagram), and the final three (YouTube, Twitter, and LinkedIn) omitted gender altogether within their profile pages and avoided pronouns within news feeds.

## Demographically Obsessed With Gender

Based on our walkthroughs of each social media platform’s advertising portal, all 10 platforms offer the ability to target audiences by gender. This finding is in line with research that indicates the continued importance of demographic targeting,

even as behavioral targeting grows in popularity (Smith, 2014; Wagner, 2014). Ad agencies create “consumer profiles” that include demographic characteristics like gender and age to give a face to the group of people they wish to target (Altstiel & Grow, 2012). While these agencies may wish to target behavioral characteristics as well, they often use demographics like gender to reflect targeted behavioral characteristics.

As we have detailed previously, platforms employ a variety of different programming practices to bake gender into different spaces within their software. In the advertising portals, we find only slight differences related to the materialization of the category of gender itself, but greater variation when it comes to determining what data are used to sort users into these categories. The latter occurs through one of the following methods, or a combination thereof: (1) direct sorting through data initially collected from mandatory identity-based fields on sign-up pages, (2) indirect sorting based on data obtained through strategic alliances with other platforms, or (3) inference-based sorting using algorithmically generated data. It is important to note that the outcome of inferring or assigning a binary gender is only visible within the advertising portal, not the user interface. Overall, demand for a binary gender categorization by the advertising industry has ensured its durability within the ad interface.

The category of gender manufactured for advertisers is consistently restricted to three options, with “male” and “female” appearing on every platform. The only difference is the composition of the third option: “all” or “unspecified genders.” While “all” permits advertisers to select everyone who has been programmatically identified as “male” and “female,” the “unspecified” option targets users who have not provided a gender and users who have selected a custom gender with a non-binary pronoun (e.g., “they”). We found that sites that offer non-binary gender options at user sign-up are more likely to allow targeted advertising to those with “unspecified” genders ( $r = .65$ ,  $p = .04$ ), demonstrating how the recent trend toward custom gender has begun to seep outside of profile pages and into software spaces designed for advertisers.

If the shifting categorization of gender within profile pages is beginning to influence the constitution of gender categories presented to advertisers, the recursive implications could instigate an intervention into an industry that plays a prominent role in calcifying gender as both a binary and a segmenting device that naturalizes essentialist stereotypes. At the same time, we have to be cognizant of pressing surveillance concerns that arise when the collection, storage, and processing of identity-based data may inadvertently (or intentionally in response to government and police requests) identify marginalized users who already face disproportionate levels of discrimination, violence, increased surveillance, and harsher penalties in society. These concerns should function to restrict the uninhibited surge of data captured by recent categorization systems that are beginning to make space for trans and gender non-conforming people.

Overall, half of the platforms in our sample programmed a gender category for their advertising clients that permits targeting of male, female, or unspecified users, as well as the option to select all three. This includes Google+, Google partners YouTube and Blogspot, and Yahoo and Pinterest. Interestingly, Yahoo has a mandatory, binary gender field on both its sign-up page and profile pages and yet unspecified genders materialize on its advertising portal. This is because Yahoo advertising—served by Bing—offers clients the ability to target audiences who are not members of its social media platform (and are thus tracked and classified based on their online activities).

Meanwhile, only a binary categorization system materializes within advertising portals programmed by Facebook, Twitter, LinkedIn, VK, and Instagram. While Twitter and LinkedIn maintain genderless sign-up pages and profile pages, both of these platforms use user data and actions to algorithmically infer a binary gender category to satisfy their advertising and marketing clients. Both Facebook and Instagram offer options beyond the binary in their profile pages but their sign-up pages differ: Facebook has a mandatory binary, while Instagram is genderless. Yet, their strategic alliance equates to shared data between the sites, and algorithmic inferences may also occur under the surface. VK, however, has constituted gender as a mandatory binary in both its sign-up page and profile pages. In other words, to be a VK user, one must select a binary gender, and as the category of gender flows through the software, it remains static, materializing again within the advertising portal as a binary.

Based on Terms and Conditions, privacy policies, and other information uncovered in our background research, we discovered that 6 of the 10 platforms populated their advertising portal categories by accessing information from databases that was initially collected through mandatory sign-up fields. Google+, Facebook, Yahoo, and VK use this approach (along with YouTube and Blogspot given their strategic alliance with, and ownership by, Google). Because these platforms (apart from Blogspot and YouTube) permit users to modify their gender on profile pages after registering, both sign-up pages and profile pages can play a role in populating categories for advertisers, as both software spaces afford opportunities to modify the gender stored in a user’s database record.

The indirect strategic alliance approach was most clearly implemented by Instagram and Pinterest. Neither includes a mandatory data field about gender on their sign-up pages: Instagram’s sign-up is genderless, while Pinterest includes a binary gender field that may appear mandatory but is in fact voluntary. Both of these platforms have formal partnerships with Facebook; Instagram is owned by Facebook, while Pinterest is linked through digital delegation.<sup>10</sup> Instagram’s Privacy Policy clearly states that “Instagram and Facebook share data for advertising purposes”: “For example, you might see ads based on the people you follow and things you like on Instagram, your information and interests on Facebook

and the websites and apps you visit” (Instagram, 2015). According to *Pinterest for Business* (2015): “[Gender] is based on Facebook settings and what the Pinner chose when they signed up for Pinterest” (p. 16).

The third approach employs a very different programming strategy to achieve gender datafication: algorithmically inferring gender based on other information the platform has collected about the user. Despite maintaining genderless platforms, Twitter and LinkedIn can still offer gender targeting to advertising clients by algorithmically inferring gender in the background. As the *Twitter for Business* page notes, “Gender is determined via public signals that users share on Twitter, such as @usernames or accounts followed” (Twitter, 2015a). These gender classifications, associated with user accounts, are not directly accessible to users themselves since they are produced for advertisers.<sup>11</sup> Twitter’s Privacy Policy notes an even broader range of data used to make inferences for content customization and ad targeting: “messages you Tweet; the metadata provided . . . lists you create, people you follow, Tweets you mark as favorites or Retweet, and many other bits of information that result from your use of the Twitter Services” (Twitter, 2015b). LinkedIn (2015b) also makes gendered inferences; for example, “your picture or your name may reveal your gender.” Yet this information, noted in the Privacy Policy, contradicts the perhaps more well-read Help Center’s messaging:

When you sign up for LinkedIn, you’re not prompted to identify whether you are male or female, and there is no section on the profile for listing your gender. LinkedIn Customer Service messages are intended to be gender-neutral and we don’t want to make potentially incorrect assumptions regarding the gender of our members. (LinkedIn, 2015a)

To translate, LinkedIn generates a gender for each user, which may be incorrect, and shares it with advertising clients, who use it to send users (stereotypically) gender-targeted ads. However, this system-assigned gender is not visible to users, apart from gender-related advertising suggestions and recommendations, which may misgender users and recursively shape these users’ own sense of identity, as Cheney-Lippold’s (2011) arguments suggested earlier.<sup>12</sup> This design strategy engages two separate programming practices in the hopes of accommodating two different groups of clients/users.

As we discovered earlier from Cheney-Lippold’s (2011) analysis, inferring data about a user is not out of the ordinary, especially for web analytics firms. Indeed, these programming practices are used by social media platforms to classify users into other categories, not just gender. LinkedIn (2015b), for example, also uses “job titles to infer age, industry, seniority, and compensation bracket.” Our exploration of Google+’s advertising portal led us to AdWords, a tool that amalgamates data about users from many sources given

Google’s wide range of services, acquired companies, and third-party relationships. Within the Google Display Network (a large group of sites displaying ads from Google AdWords), gender can be inferred through websites that users visit since cookies<sup>13</sup> remember this data. As yet another testament to the strong reliance on demographic data by advertisers despite the wealth of other behavioral and taste-based data now available, AdWords Help advises advertising clients to “Bear in mind that we aren’t able to gather or infer demographic information from everyone using the web or mobile apps” (Google, 2015).

## Discussion and Conclusion

Social media software has the capacity to enact culture by managing and shaping the construction and meaning of identity categories like gender. It is vital that researchers explore how, where, and for what purposes designers program software to perform this function. Our investigation uncovered a wide range of programming practices within the 10 most popular social media platforms, demonstrating that programmers do not follow an imagined, standard industry practice when making decisions about how and where to code gender. Instead, platforms independently wield control over their own categorization systems. This capacity is further mobilized and entrenched by the emphasis that platforms place on neutrality (Gehl, 2011; Gillespie, 2010). Yet, designers are hardly immune to the cultural forces stemming from the society in which they are embedded. The durability of the gender binary within all of the categorization systems we uncovered is influenced by cultural pressures and perceived needs of advertisers. Even custom gender models that offer non-binary options actively partition these new options away from the normalized binary while also generating an added layer of categorization: mandatory, three-option pronouns. This added layer reinscribes the binary as the gendered data move through the software and materialize as a new category within advertising interfaces, obscured from users.

In addition to being the most recent convert to a free-form custom gender field in our sample, Pinterest’s designers also decided against a mandatory pronoun overlay. Given Pinterest’s advertising-dependent relationship with Facebook, it is possible that Facebook may wield some control over Pinterest’s categorization system. This leads us to question the future capacities of the intermediary role that social media platforms play in categorization practices, particularly if mergers bring about an increasing concentration of ownership within the platform ecosystem more broadly. To explore power structures that emerge around platforms’ intermediary roles as category shapers, future research should attend to the complex relationships between web analytics firms, the advertising industry, and social media platforms, which we have only begun to unravel here.

While our analysis focused on sign-up pages, profile pages, and advertising portals, other software spaces and processes may also reveal important identity-based categorization processes. Our goal was to consider how (and where) the category of gender materializes within these popular platforms, for whom it is deployed, and whether these shifts may engender recursive implications for society more broadly.

From a user's perspective, reporting gender by clicking on a programmed field is a fairly common practice. Thus, sign-up pages are read by users as transitory spaces, yet are actually spaces where categorization systems reify the gender binary and render this narrowly inscribed meaning durable. Existing beyond the binary is possible, but a more rigid gender classification regime often must be bypassed first. From an advertiser's perspective, targeting gender is always a possibility. Because the manufacture and management of gender categories, like all identity-based data categories, is propelled by revenue generation, it is unsurprising that even custom gender is reconstituted into a three-option classification system that partitions "custom" away from the normalized binary. The advertising industry continues to revere the gender binary, yet it may also be the case that social media platforms wield the power of their intermediary role to determine which gender categorization scheme can best serve the advertising industry.

Thus, social media platforms face a double-bind: while platform design recursively shapes the advertising industry's already preferred understanding of gender by making the binary problematically durable, a move beyond the binary on the advertising interface could increase surveillance of marginalized populations in unexpected ways. Retaining binary categories in platform databases contributes to programming practices that bury data about users from users. In our view, each user should have access to all data about themselves, particularly when it is inferred,<sup>14</sup> assigned, or modified (as in the case of a binary gender imposed on a user based on their pronoun selection).

Recently, tech industry professionals in charge of computational predictions for ad targeting, purchasing behaviors, and the like have shifted away from demographic-based recommendations and toward recommendations based on behavior and preference segmentation practices (Churchill, 2013). Instead of targeting women, for example, an advertiser might prefer to target people with purchasing behaviors that parallel those who have already bought their product. This is a promising development, in that it points to a potential future where gender classification would no longer determine what a person would see on a site, negating the recursive implications for how we understand our own identity and that of others. Yet other data merely take the place of gender—acts of cultural consumption, in this case—which may be similarly layered with judgments about users rooted in binary and heteronormative performances, filtered, as Cheney-Lippold (2011) has argued, through a "marketing logic of consumption."<sup>15</sup> Still, given our focus on gender-related programming practices,

industry movement away from demographic personalization could eventually persuade social media platforms to eliminate mandatory and binary fields, allowing users to regain some of the agency in self-presentation they lose when forced to enter a binary gender upon sign-up.

Social media platforms, particularly those platforms that achieve a great deal of popularity and profit, must bear responsibility for their design decisions. Their programming practices may inadvertently advocate for certain groups of people while alienating others. The values they bake into their software have the capacity to influence the next generation of platforms that will go on to play an intermediary role in shaping society's construction of itself. To be clear, however, we do not advocate for genderless online spaces. While some may wish to avoid online gender identification to block gender-related ad targeting or maintain privacy, gender remains a salient and important marker of personal identity for many. While a genderless Internet would force the advertising industry to focus on behavioral and taste-based data, or at least other demographic characteristics, this move would not eliminate gender-based discrimination and marginalization (Kendall, 1998). In the context of race and ethnicity, for instance, online spaces that do not explicitly allow race to be presented or specified do not become raceless utopias; instead, these spaces tend to erase race while positioning whiteness as the default (Kolko, Nakamura, & Rodman, 2013; Nakamura, 2002).

In this work, we have highlighted platform design and the particular capacity of programming practices to generate and manage identity-based categories within multiple software locations, including surface and deep, visible, and invisible software levels. This work should serve as a starting point to encourage more scholarship attending to the values baked within platform infrastructures, and what it means when we allow social media platforms to shape our constructions of self.

## Acknowledgements

Thanks to Mark Handel for organizing the "Facebooking In 'Face': Complex Identities Meet Simple Databases" panel at CSCW 2015 where we first began discussing ideas for this article; Gillian Hayes and Nazanin Andalibi for continued support; and anonymous reviewers for their helpful feedback.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

This work was funded in part by the Social Sciences and Humanities Research Council of Canada, a Banting Postdoctoral Fellowship, a National Science Foundation Graduate Research Fellowship, a Rob Kling Memorial Fellowship in Information and Computer Sciences, and the Kleist Endowment.



## Notes

1. Transgender is a term that refers to “people who move away from the gender they were assigned at birth, people who cross over (trans-) the boundaries constructed by their culture to define and contain that gender” (Stryker, 2008, p. 1). Following (Haimson et al., 2015), we use “trans” for the remainder of this article to refer to the broad transgender population.
2. We performed all walkthroughs in late August 2015, using a browser (Safari) with a cleared cookie history, a fresh private browsing session, and a Canadian Internet Protocol (IP) address. All walkthroughs were web-based apart from Instagram that required a mobile app for registration. We used video recordings of the screen, screenshots, and voice recording for analysis. Pinterest had not yet opened up their advertising opportunities to all users, so we relied on online video tutorials for Pinterest’s advertising clients.
3. This update was launched for English (US) users.
4. Google+ handled YouTube and Blogspot registration during our research period, which is why this section deals with a total sample of eight sign-up pages.
5. While Pinterest’s sign-up page did not explicitly note that gender was a voluntary field, we were able to register without completing the field.
6. Similar options (e.g., including four options: male, female, other, and not known) have been rejected as a gender coding scheme by the Gay, Lesbian, Bisexual, and Transgender (GLBT) round table of the American Library Association on the grounds that it cannot capture the fluidity and complexities of gender (“GLBTRT Task Force on RDA,” 2015).
7. Besides “them” or “their” as a non-binary pronoun, other pronouns are also in popular use (e.g., “ze,” “zir”).
8. Being misgendered is harmful for transgender people, leading to increased stigma and negative affect (McLemore, 2015).
9. This investigation involved manipulating gender fields and pronouns on our new account’s profile page and refreshing Google+’s public-facing ad preferences page (<http://www.google.com/ads/preferences>).
10. Digital delegation practices (e.g., “login to Instagram with your Facebook account”) access, “periodically sync,” and store data from other existing accounts on the new network’s server.
11. However, users may guess which gender the system has inferred based on the results of targeted ads or social media content.
12. Products embody culturally mandated gendered messages about “who we can be” (Churchill, 2010, p. 52), and social media sites do the same. Pinterest, for instance, suggests Home Decor boards to women and Technology boards to Men, while Twitter’s “Who to follow” may similarly enforce gendered stereotypes. Thus, while algorithmic decisions about users’ gender categorizations are made without user input and are often invisible, consequences of those decisions become visible to users in social media sites’ user interfaces.
13. While all 10 platforms in our sample use cookies, and other tracking technologies like beacon and pixel tags were also popular, exploring their use in relation to gender categorization was beyond the scope of this article.
14. However, inferred data are increasingly opaque as much of the data are determined using black box machine learning algorithms.

15. Special thanks to anonymous Reviewer 2 for drawing us toward this insight, and for giving R2s everywhere a good name.

## References

- Alexa. (2015). *The top 500 sites on the web*. Alexa’s Global Traffic Rank. Retrieved from <https://www.alexa.com/topsites>
- Altstiel, T., & Grow, J. (2012). *Advertising creative: Strategy, copy, and design*. Thousand Oaks, CA: SAGE.
- Balsamo, A. (2011). *Designing culture: The technological imagination at work*. Durham, NC: Duke University Press.
- Bardzell, S. (2010). Feminist HCI: Taking stock and outlining an agenda for design. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 1301–1310). New York, NY: Association for Computing Machinery. doi:10.1145/1753326.1753521
- Battelle, J. (2005). *The search: How Google and its rivals rewrote the rules of business and transformed our culture*. New York, NY: Portfolio.
- Beer, D. (2013). *Popular culture and new media: The politics of circulation*. New York, NY: Springer.
- Bivens, R. (2015). The gender binary will not be deprogrammed: Ten years of coding gender on Facebook. *New Media & Society*. Advance online publication. doi:10.1177/1461444815621527
- Bowker, G. C., & Star, S. L. (1999). *Sorting things out: Classification and its consequences*. Cambridge, MA: MIT Press.
- boyd, d. (2011). Social network sites as networked publics: Affordances, dynamics, and implications. In Z. Papacharissi (Ed.), *Networked self: Identity, community, and culture on social network sites* (pp. 39–58). New York, NY: Routledge.
- Burgess, J., Light, B., & Duguay, S. (2015, May). *Towards Hookup apps studies: Sexual cultures, digital queers, locative media*. Paper presented at the ICA Conference, San Juan, Puerto Rico.
- Butler, J. (1999). *Gender trouble*. Hoboken, NJ: Taylor & Francis.
- Butler, J. (2004). *Undoing gender*. New York, NY: Psychology Press.
- Cheney-Lippold, J. (2011). A new algorithmic identity soft biopolitics and the modulation of control. *Theory, Culture & Society*, 28, 164–181. doi:10.1177/0263276411424420
- Chun, W. H. K. (2013). Race and/as technology, or how to do things to race. In L. Nakamura & P. Chow-White. *Race after the Internet* (pp. 38–60). London, England: Routledge.
- Churchill, E. F. (2010). Sugared puppy-dog tails: Gender and design. *Interactions*, 17, 52. doi:10.1145/1699775.1699787
- Churchill, E. F. (2013). Putting the person back into personalization. *Interactions*, 20, 12–15. doi:10.1145/2504847
- Constine, J. (2012). *Why Google+ doesn’t care if you never come back*. Retrieved from <http://social.techcrunch.com/2012/02/28/no-more-no-more-no-more-no-more/>
- Crenshaw, K. (1991). Mapping the margins: Intersectionality, identity politics, and violence against women of color. *Stanford Law Review*, 43, 1241–1299. doi:10.2307/1229039
- Ellison, N. B., & boyd, d. (2013). Sociality through social network sites. In W. H. Dutton (Ed.), *The Oxford handbook of Internet studies* (pp. 151–172). Oxford, UK: Oxford University Press.
- Facebook. (2015). *Data policy*. Retrieved from <https://www.facebook.com/policy.php>
- Flanagan, M., Howe, D. C., & Nissenbaum, H. (2008). Embodying values in technology: Theory and practice. In J. van den Hoven & J. Weckert (Eds.), *Information technology and moral*

- philosophy* (pp. 322–353). Cambridge, UK: Cambridge University Press.
- Friedman, B., & Nissenbaum, H. (1997). Software agents and user autonomy. In *Proceedings of the first international conference on autonomous agents* (pp. 466–469). New York, NY: Association for Computing Machinery. doi:10.1145/267658.267772
- Gehl, R. W. (2011). The archive and the processor: The internal logic of Web 2.0. *New Media & Society*, 13, 1228–1244. doi:10.1177/1461444811401735
- Gillespie, T. (2010). The politics of “platforms.” *New Media & Society*, 12, 347–364. doi:10.1177/1461444809342738
- GLBTRT task force on RDA and gender in authority records. (2008). *GLBT News*. Retrieved from <http://www.glbtrt.ala.org/news/archives/20>
- Google. (2015). *Why Google asks for your information at sign-up*. Retrieved from <https://support.google.com/accounts/answer/1733224?hl=en>
- Haimson, O. L., Brubaker, J. R., Dombrowski, L., & Hayes, G. R. (2015). Disclosure, stress, and support during gender transition on Facebook. In *Proceedings of the 18th ACM conference on computer supported cooperative work & social computing* (pp. 1176–1190). New York, NY: Association for Computing Machinery. doi:10.1145/2675133.2675152
- Instagram. (2015). *How does Instagram decide which ads to show me?* Retrieved from <https://help.instagram.com/173081309564229>
- Jensen, C., & Potts, C. (2004). Privacy policies as decision-making tools: An evaluation of online privacy notices. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 471–478). New York, NY: Association for Computing Machinery. doi:10.1145/985692.985752
- Kendall, L. (1998). Meaning and identity in “Cyberspace”: The performance of gender class, and race online. *Symbolic Interaction*, 21, 129–153. doi:10.1525/si.1998.21.2.129
- Kitchin, R., & Dodge, M. (2011). *Code/space: Software and everyday life*. Cambridge, MA: MIT Press.
- Kolko, B., Nakamura, L., & Rodman, G. (2013). *Race in cyberspace*. London, England: Routledge.
- LinkedIn. (2015a). *Use of gender in site content and communications*. Retrieved from [http://help.linkedin.com/app/answers/global/a\\_id/3566/~/-use-of-gender-in-site-content-and-communications](http://help.linkedin.com/app/answers/global/a_id/3566/~/-use-of-gender-in-site-content-and-communications)
- LinkedIn. (2015b). *Your privacy matters*. Retrieved from <https://www.linkedin.com/legal/privacy-policy>
- McLemore, K. A. (2015). Experiences with misgendering: Identity misclassification of transgender spectrum individuals. *Self and Identity*, 14, 51–74. doi:10.1080/15298868.2014.950691
- McPherson, T. (2014). Designing for difference. *Differences*, 25, 177–188. doi:10.1215/10407391-2420039
- Nakamura, L. (2002). *Cybertypes: Race, ethnicity, and identity on the Internet*. London, England: Routledge.
- National Coalition of Anti-Violence Programs. (2014). *Lesbian, gay, bisexual, transgender, queer, and HIV-affected hate violence in 2014*. Retrieved from [http://avp.org/storage/documents/Reports/2014\\_HV\\_Report-Final.pdf](http://avp.org/storage/documents/Reports/2014_HV_Report-Final.pdf)
- Pinterest for Business. (2015). *A guide to Pinterest analytics*. Retrieved from <https://business.pinterest.com/sites/business/files/pinterest-analytics-guide-en.pdf>
- Smith, M. (2014). *Targeted: How technology is revolutionizing advertising and the way companies reach consumers*. New York, NY: AMACOM Div American Mgmt Assn.
- Stryker, S. (2008). *Transgender history*. Berkeley, CA: Seal Press.
- Truitt, J. (2011). *Google+ has a gender problem*. Retrieved from <https://feministing.com/2011/07/12/google-has-a-gender-problem>
- Turow, J. (2006). *Niche Envy: Marketing discrimination in the digital age*. Cambridge, MA: The MIT Press.
- Turow, J. (1997). *Breaking up America: Advertisers and the new media world*. Chicago, IL: University of Chicago Press.
- Twitter. (2015a). *Gender targeting*. Retrieved from <https://business.twitter.com/en-gb/solutions/gender-targeting>
- Twitter. (2015b). *Privacy policy*. Retrieved from <https://twitter.com/privacy>
- van Dijck, J. (2013). *The culture of connectivity: A critical history of social media*. New York, NY: Oxford University Press.
- Wagner, R. (2014). *Demographic targeting using age and gender data*. Retrieved from <http://www.exactdrive.com/demographic-targeting-using-age-and-gender-data>
- Wajcman, J. (2010). Feminist theories of technology. *Cambridge Journal of Economics*, 34, 143–152. doi:10.1093/cje/ben057
- Wolin, L. D. (2003). Gender issues in advertising—An oversight synthesis of research: 1970–2002. *Journal of Advertising Research*, 43, 111–129. doi:10.2501/JAR-43-1-111-130

## Author Biographies

Rena Bivens (PhD, University of Glasgow) is an Assistant Professor in the School of Journalism and Communication at Carleton University. Her research interests include software design, normativity, gender, and violence.

Oliver L. Haimson is a PhD Candidate in Information and Computer Sciences at the University of California, Irvine. His research interests include social media, online identity, and gender.

## Appendix

**Table 1.** English Language Sites Considered for Inclusion in This Study.

English, fits criteria	English, does not fit criteria
<b>Google (1)</b>	Amazon (6)
<b>Facebook (2)</b>	Wikipedia (7)
<b>YouTube (3)</b>	Reddit (31)
<b>Yahoo (5)</b>	Imgur (43)
<b>Twitter (10)</b>	StackOverflow (55)
<b>LinkedIn (14)</b>	Diply (61)
<b>Blogspot (20)<sup>a</sup></b>	
<b>VK (21)</b>	
<b>Instagram (28)</b>	
<b>Pinterest (40)</b>	
Tumblr (42)	
DailyMotion (88)	
GitHub (95)	
Twitch (139)	

Numbers in parentheses show Alexa Ranking (Alexa, 2015). The 10 most popular sites that fit the criteria (listed in bold) were included in our study.

<sup>a</sup>Blogspot and Blogger appear to direct to the same page yet they are ranked in two different places (20 and 114, respectively).

**Table 2.** Gender Options in Social Media Site Sign-Up Pages, Profiles, and Advertising Targeting.

Site	Sign-up page			Profile		Advertising targeting
	Gender options mandatory?	Gender options	Drop-down list or radio button	Gender options	Pronoun options	Gender options
Google+, YouTube, Blogspot	M	<b>Male, female, other</b>	Drop-down list	<b>Male, female, decline to state, custom</b>	<b>Him, her, them</b>	Male, female, unknown
Facebook	M	Male, female	Radio button	<b>Male, female, custom</b>	<b>Him, her, them</b>	All, male, female
Yahoo	M	Male, female	Radio button	Male, female	—	Male, female, unknown
Twitter	—	—	—	—	—	All, male, female
LinkedIn	—	—	—	—	—	All, male, female
VK	M	Male, female	Drop-down list	Male, female	—	All, male, female
Instagram	—	—	—	<b>Male, female, not specified</b>	—	All, male, female
Pinterest	R	Male, female	Radio button	<b>Male, female, custom</b>	—	Male, female, unknown

M: mandatory; R: requested but not mandatory; —: absent.

**Table 3.** Other Information in Social Media Site Sign-Up Pages.

Site	Name	Age	Email	Username	Password	Location	Anti-bot	Account verification	Log in with Facebook	Other
Google+, YouTube, Blogspot	M	M	R	M	M	M	M	M: phone	–	
Facebook	M	M	M <sup>a</sup>	–	M	–	–	–	–	
Yahoo	M	M	–	M	M	–	–	M: phone	–	R: recover number and relationship to you
Twitter	M	–	M	M	M	–	–	R: email/phone	–	R: option to tailor Twitter based on recent website visits
LinkedIn	M	–	M	–	M	–	–	R: email	–	M: post sign-up: country; postal code; student (y/n); job title; company; industry
VK	M	–	–	–	M	–	–	M: phone	–	
Instagram	R	–	M	M	M	–	–	–	R	R: phone number
Pinterest	M <sup>b</sup>	R	M	–	M	–	–	–	R	

M: mandatory; R: requested but not mandatory; –: absent.

<sup>a</sup>The user can choose between email or phone number, but one is mandatory.

<sup>b</sup>Only the first name was mandatory; the second name was requested.